Air Pollution News Sentiment: Seasonal Trends' Relation with PM10 Levels

Stefani Kulebanova[1], Jana Prodanova[2], Aleksandra Dedinec[1], Trifce Sandev[2], Desheng Wu[3], Ljupco Kocarev[2]

[1] Faculty of Computer Science and Engineering, UKIM, Skopje, Macedonia
[2] Macedonian Academy of Sciences and Arts, Skopje, Macedonia
[3] University of Chinese Academy of Sciences, Beijing, PR China

Abstract: Air pollution remains a major public health concern globally, especially in the Western Balkan countries facing severe challenges with the air quality problem. This study investigates the relationship between air quality, news media sentiment, and public discourse in Macedonia over a ten-year period (2014-2023). We employed sentiment analysis to examine the emotional tone of news coverage related to air pollution, and topic modeling to uncover recurring themes within news articles. Our analysis revealed a distinct seasonal pattern, with negative media sentiment peaking during winter months when PM10 levels were the highest. This finding aligns with the increased reliance on polluting fuels for winter heating. Interestingly, despite a stable number of neutral articles, a rise in positive sentiment articles suggests a potential decrease in pollution levels or the effectiveness of new government policies. We identified recurring topics like air quality concerns in specific cities, public unease regarding factories, and ongoing scrutiny of government policies. Emerging topics included the impact of the COVID-19 pandemic on air quality, public discourse surrounding heating practices, and growing concerns about waste management. This study contributes to a deeper understanding of the complex interplay between air pollution data, public discourse, and media framing, offering valuable insights for policymakers and media outlets in Macedonia.

Keywords: air pollution, Western Balkan, news media, sentiment analysis, topic modeling, PM10

1. Introduction

Air pollution is a pressing global concern, posing significant threats to human health, environmental well-being, and economic stability [1-4]. The air pollution problems in Macedonia stem from earlier years and pose a serious threat [5]. The number of polluting particles is going well into hazardous levels each year, as per the annual analysis from 2017 to 2023 by the World Health Organization (WHO). Tetovo, Strumica and Skopje are in Europe's top 30 most polluted cities, with Strumica in seventh place and Tetovo in ninth place [6]. Fine particulate matter, particularly PM10 particles with a diameter of 10 micrometers or less, is a significant component of air pollution due to its deep penetration into the respiratory system and its potential to cause various health complications, including respiratory illnesses, cardiovascular diseases and even premature mortality [7-11]. Public awareness and concern regarding air quality are crucial for driving policy changes and promoting individual actions to mitigate air pollution [12,13].

Media plays a vital role in shaping public perception and influencing environmental discourse [14]. News coverage of air quality issues can raise awareness, inform the public about health risks, and

hold authorities accountable for pollution control measures [15]. However, the nature and sentiment of media coverage can significantly impact how the public understands and reacts to air pollution concerns.

This research explores the interrelation between PM10 levels and media sentiment. We specifically focus on PM10 particles for several reasons. First, PM10 is widely monitored as a critical indicator of overall air quality by regulatory agencies worldwide [16]. Unlike some other measures, PM10 encompasses a variety of particles from diverse sources, including dust, smoke, and even pollen. This makes it a comprehensive measure of the fine particulate matter that can penetrate deep into the lungs and cause health problems [17]. Additionally, PM10 concentrations are readily available from government monitoring stations, allowing us to establish a clear link between air quality data and media coverage.

Our study chooses news media as the source for analyzing sentiment because it serves as a primary channel for informing the public about environmental issues. News outlets report on air quality data, discuss potential health risks and influence public opinion through the language and tone they employ. By analyzing the sentiment expressed in news articles about air pollution, we can gain insights into how media coverage shapes public understanding of air quality and potentially influences public behavior. By sentiment, we refer to the emotional tone or opinion conveyed in news articles related to air pollution.

This study aims to explore if there is a correlation between rising PM10 levels and a shift towards more negative media sentiment. Understanding this relationship holds significant value. If negative sentiment dominates air quality news during periods of high PM10 levels, it can heighten public anxiety and prompt individuals to take action to protect their health, such as wearing masks or limiting outdoor activities. Conversely, a lack of media attention or overly positive portrayals during poor air quality might lead to complacency and hinder public engagement in pollution mitigation efforts. Investigating the relationship between PM10 and sentiment can shed light on potential media biases in environmental reporting. Understanding these dynamics can inform discussions on media responsibility and encourage a more nuanced approach to environmental reporting.

This research contributes to the ongoing dialogue on media portrayal of environmental issues and its impact on public behavior and policy making. Through a detailed analysis of PM10 levels and media sentiment, we tend to understand the dynamic interactions between air quality data and how it is communicated in the Macedonian media. The findings of this investigation are a valuable asset for media professionals, environmental policymakers, and public health advocates seeking to bridge the gap between air quality data and effective public communication strategies.

2. Literature Review

The impact of air pollution on media coverage has been explored in various contexts and regions, highlighting the interplay between environmental conditions and media narratives. This section reviews several studies that provide a foundation for understanding our findings within the broader literature.

2.1 Relationship Between Air Pollution and Media Coverage

The study by Wang et al. [18] investigates the relationship between air pollution and media slant in China, focusing on publicly listed firms. Using extensive air quality and media data at the city level, the authors found that worsening air quality leads to an increase in negative media coverage. Specifically, as air quality deteriorates from lightly to heavily polluted, the number of negative sentences in news articles increases by about 1%. The effect is consistent across various types of firms and newspaper articles, with a stronger impact observed for firms in heavily polluting industries. These findings suggest that air pollution significantly influences media slant, contributing to a more negative portrayal of firms during high pollution periods. This pattern supports our objective, suggesting a potential influence of air pollution on media sentiment. When air quality deteriorates, media coverage tends to adopt a more critical tone, possibly reflecting public concern and heightened awareness of environmental health risks.

Conversely, the study by He, Zhang & Chen [19] examines how media coverage of haze in China aligns with sustainable development goals and environmental governance. Analyzing articles from China Daily between 2000 and 2018, the authors utilized framing theory and the issue-attention cycle to understand the dynamics of media slant. The results indicate that most haze coverage had a neutral slant and primarily used frames that defined problems and suggested remedies. The study also found significant correlations between the frames used and the types of sources cited, such as officials and specialists. This work highlights the role of media in environmental governance and the cyclical nature of media attention to air pollution issues.

2.2 Focus and Content of Media Coverage

Ramondt & Ramírez [20] focused on how national and regional newspapers in the USA reported on air pollution over five years. The study revealed that nearly 40% of articles mentioned human health risks, yet fewer than 10% provided information on precautionary measures. The findings suggest that media reporting in high pollution areas is not effectively raising environmental health literacy, as it fails to provide actionable information for the public.

Furthermore, Amiraslani [21] explored the role of media in raising public awareness about climate change and air pollution mitigation in urban areas. The research found that news articles predominantly focused on public awareness and alarming messages related to climate change. The study emphasized the importance of researchers in generating scientific news and the need for media to disseminate educational messages more effectively. This underscores the significance of media in educating the public not only on the existence of environmental issues but also on potential solutions.

2.3 Local Context of the Macedonian Case

While the studies mentioned above provide valuable insights from various regions, it is important to consider the specific context of Macedonia. Previous evidence on this subject pinpoint household heating with solid fuels, particularly during winter months, as a major contributor [22,23]. This aligns with national data highlighting residential heating's significant role in PM2.5 and PM10 emissions [24]. Industrial activity, especially from older plants lacking proper emission controls, also plays a part. Traffic emissions, while less impactful nationally, can be more concerning locally due to the prevalence of older vehicles [24,25].

The consequences of air pollution in Macedonia are severe. Studies estimate thousands of premature deaths annually due to air pollution exposure, with cities like Tetovo, Bitola, and Skopje experiencing the highest mortality rates [23]. The gravity of Macedonia's air pollution problem is undeniable; however, recent research offers guarded optimism for the future. Previous studies suggest the possibility of gradual air quality improvements, offering a cause for cautious optimism [25]. This positive outlook is further strengthened by the temporary decrease in pollution observed during COVID-19 lockdowns [23]. These observations highlight the potential impact of stricter regulations or a shift towards cleaner energy sources, suggesting a path towards a healthier future for Macedonian citizens.

3. Materials and Methods

3.1 Data Sources

We have meticulously compiled two distinct sets of quantitative data spanning from January 2014 to December 2023, measures of PM10 particle data and news data. The measures of PM10 particle data were sourced from the official website of the Ministry of Environment and Physical Planning (MOEPP) of Macedonia, ensuring reliability and accuracy [26]. At the time of our extraction process, there were provided measurements from 24 official monitoring stations in Macedonia from the official website of MOEPP. More specifically, there were eight monitoring stations in Skopje, two in Bitola, two in Veles,

and one each in Kichevo, Lazaropole, Tetovo, Gostivar, Prilep, Kochani, Kavadarci, Kumanovo, Strumica, Gevgelija, Berovo and Ohrid (Figure 1).

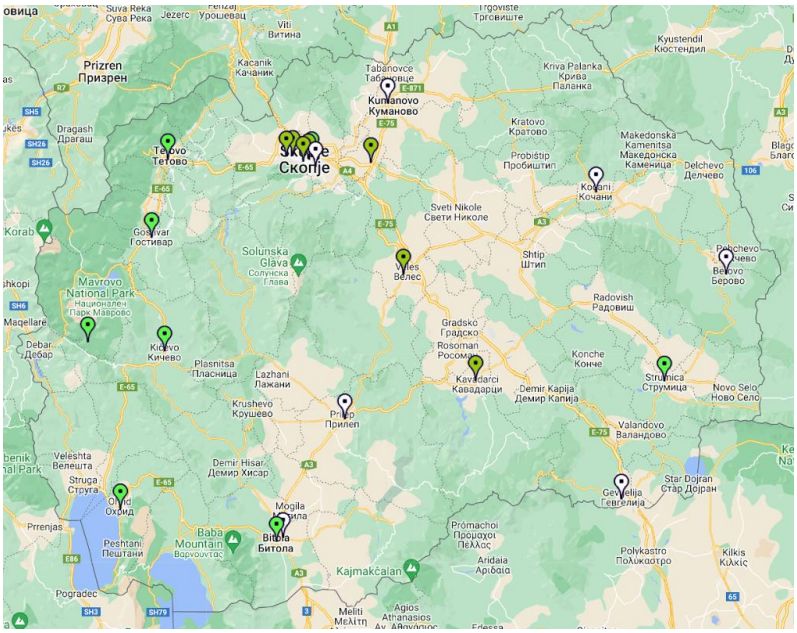

Figure 1. Locations of official air monitoring stations in Macedonia

Meanwhile, the news data was collected from Time.mk, a reputable news outlet. Time.mk is a news aggregator that uses 120 distinct news media outlets as a source each day [27]. Despite employing diverse methodologies to extract information from these sources, we maintained consistency in the timeframe, ensuring the comparability and integrity of our dataset.

3.2 Data Gathering and Preprocessing

3.2.1 Air Data

The air data consists of hourly measures of PM10 particles over the 10-year period. We used a Python script, utilizing the requests library. This library allows for sending a HTTP request to a website, while abstracting the process for sending a request behind an API that simplifies the work [28]. After the hourly data was extracted, it went through careful examination and transformation. The data fields were transformed to the appropriate format, we cleansed it of duplicates, null rows and as a final step - aggregated the hourly measures of PM10 particles on a daily basis. This prepared the dataset for subsequent analysis and interpretation.

3.2.2 News Data

The extraction of the news media was performed with a web-crawling tool called Octoparse. Octoparse is a free cloud-based web data extraction solution that helps users extract relevant information from various types of websites [29]. It is an automated scraping tool that allows for visiting a page and extracting the data based on the HTML tags without using code. It has various visual tools for the user to add and personalize them to the user's needs.

We focused on extracting the news media teasers with their title, time, and source, from the Time.mk website. A media teaser is a brief preview of an article, enticing readers to explore specific news items [30]. The data was scraped according to a list of keywords targeting air pollution: "аерозагадување" (air pollution), "загадување" (pollution), "загаден" (polluted), "пм10" (pm10). After the collection, we transformed the data to the appropriate format, cleansed it from NULL rows, removed duplicate news (based on the same title) and manually processed it to remove irrelevant news. The teasers were translated into English language using Google translate [31] to enable the application

of sentiment analysis and topic modeling techniques. Furthermore, the news teasers went through tokenization, lowercasing, removal of punctuation, stop words, and special characters for the next steps of the methodology.

3.3 Sentiment Analysis

After the collection and preprocessing of the news data, we performed sentiment analysis on each of the news teasers with a specialized lexicon-based tool called VADER (Valence Aware Dictionary and Sentiment Reasoner). VADER is part of the library NLTK, (Natural Language Toolkit), which is a suite of libraries and programs for symbolic and statistical natural language processing for English written in the Python programming language [32]. Specifically, VADER is used for analyzing sentiments in social media texts. VADER employs a lexicon that assigns pre-defined sentiment scores to individual words. As a first step, VADER breaks down the text into its individual words to identify its sentiment. Each word in the text is assigned with a polarity score, ranging from -1 (indicating extreme negativity) to +1 (representing strong positivity), where 0 indicates neutrality. These scores stem from a human judgment. On top of the individual word scores, VADER considers context, punctuation, capitalization and certain words and phrases called valence words that alter the sentiment of adjacent words. For example, 'I had a good day' will have a lower positive score than 'I had a very good day', which will have lower positive score from 'I had a VERY good day'. VADER aggregates all the individual word scores and normalizes them to a scale between -1 (most negative) and +1 (most positive). This normalization takes into consideration the length of the text and by doing this, it ensures the sentiment score is proportional to the density of sentiment-bearing words in the text. The result from the normalization is called a compound score, which provides a comprehensive assessment of the texts' sentiment. Research suggests VADER achieves an F1 score of 0.96, a metric combining precision and recall, for sentiment classification on tweets [].

$$\mathrm{X} = \frac{x}{\sqrt{x^2 + \alpha}} \tag{1}$$

In (1), x is the sum of Valence scores of constituent words and α is a normalization constant with a default value equal to 15.

For severalizing the tweets into positive, negative, and neutral sentiment groups, the default threshold value of - 0.05 and + 0.05 was used. Each news teaser went through sentiment analysis and was scored for positive, negative, and neutral sentiment, as well as an overall compound score representing the aggregated sentiment.

3.4 Cross-correlation

To explore the relationship between air pollution dynamics and media discourse, we conducted time series statistical analysis using cross correlation. Cross correlation measures the similarity between two time series as a function of the displacement of one relative to the other [33]. It is analogous to a sliding dot product that finds how well one signal matches another when it is moved back and forth in time. In signal processing, it measures similarity without normalization (values not limited between -1 and 1). The result (correlation function) shows how much one signal needs to be shifted to best match the other. For continuous functions, such as our time series data, integrals are used involving the product of one function and the complex conjugate of the other shifted by a lag (τ) [34].

Consider that g(k) and p(k), with k integer, are two sampled sequences of time series. We assume, for simplicity, $-\infty < k < +\infty$. A standard, unnormalized definition of cross-correlation is:

$$\Phi_{pg}(\tau) = \sum_{k=-\infty}^{+\infty} p(k-\tau)g(k), \tag{2}$$

where the integer τ is the relative delay [33]. High positive values at a specific τ indicate a strong match, with the larger value signifying a better alignment. This suggests a feature in f(t) aligns with a similar

feature in the shifted g(t + τ). Conversely, low values indicate a poor match, and negative values, potentially arising when peaks in f(t) coincide with troughs in the shifted g(t + τ) or vice versa, reflect the strength of this mismatch based on the magnitude. In essence, the continuous cross-correlation function provides a quantitative measure of how similar two signals are by calculating the product of one function and the shifted version of the other across all possible lags and integrating the result.

We calculated cross correlation coefficients between daily PM10 concentrations and sentiment scores derived from news articles over corresponding time periods. This analysis helped identify any potential temporal relationships or lag effects between air pollution trends and media sentiment.

Furthermore, we performed the Mann-Kendall (MK) test on a yearly basis for both time series, to explore whether the time series exhibit an increasing or decreasing trend over time. The Mann-Kendall (MK) test is a statistical method used to assess whether a time series exhibits a monotonic trend (increasing or decreasing) over time [35,36]. It uses hypothesis testing with a null hypothesis (H0) claiming there is no monotonic trend in the data and alternative hypothesis (Ha) claiming a monotonic trend is present in the data. The test initially assumes the null hypothesis (H0) is true. It needs strong evidence to reject H0 and accept the alternative of a trend. It analyzes all possible pairs of data points. For each pair, it calculates the difference and assigns a sign to this difference: +1 if the later observation is larger, 0 if there is no change, or -1 if the later observation is smaller. Then sums the signs of all difference calculations:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sign\left(x_j - x_i\right), \quad (3)$$

where sign(y_j -y_i), is equal to +1, 0, or -1 as indicated above. The S (the result) is the number of positive differences minus the number of negative differences and n is the number of data points. If *S* is a positive number, observations obtained later in time tend to be larger than observations made earlier. A negative S suggests decreasing trends (later values tend to be smaller). A significant positive or negative S value indicates a rejection of the null hypothesis (no trend) and supports the presence of a monotonic trend in the data (increasing or decreasing, respectively).

3.5 Topic Modelling

Latent Dirichlet Allocation (LDA), developed by David Blei, Andrew Ng, and Michael Jordan in 2003, is a generative probabilistic model for discovering the hidden topics in a collection of documents [37]. LDA assumes that each document is a mixture of a set of topics and each topic is a distribution over words. The process involves two Dirichlet distributions: one that models the distribution of topics in a document and another that models the distribution of words in a topic. For each document, a topic distribution is drawn from the first Dirichlet distribution, and for each word in the document, a topic is chosen according to this distribution, followed by choosing a word from the corresponding topic's word distribution. Inference and parameter estimation in LDA are performed using techniques like Variational Bayes and Gibbs Sampling. The probability distribution for each topic z in document d can be computed as:

$$P(z|d) = \frac{P(z)P(d|z)}{P(d)} \quad (4)$$

LDA has applications in topic modeling, text classification, and collaborative filtering, providing interpretable topics from large text corpora. However, it requires pre-specifying the number of topics and is sensitive to hyperparameter settings and dataset size.

In this study, LDA was utilized to analyze pre-processed news teasers on a yearly basis. For each year, five topics were identified, along with the eight most prominent words associated with each topic. The extracted topics for each year were compiled into a Pandas DataFrame and saved in a comma-separated values (CSV) file. A DataFrame is a two-dimensional data structure commonly used in data analysis and statistical computing. It is akin to a table in a database or an Excel spreadsheet, where data is organized in rows and columns. [38] Since LDA topic labels are not inherently meaningful, human intervention was required to assign descriptive titles to the thirteen topics identified across all years.

Based on this thematic analysis, a valuable insight was gained into the recurring themes discussed throughout the study period in media coverage of air pollution.

4. Results

4.1 Sentiment Analysis

Our analysis revealed distinct seasonal patterns in news sentiment regarding air pollution. As provided in Figure 2, even though the number of news articles fluctuates each year, the increase of the negative sentiment during the winter months is extremely noticeable across all 10 years. This seasonal trend, spanning from November to February, persisting across all years in the dataset, indicates a consistent pattern of heightened negative sentiment during high pollution periods. Positive sentiments follow a similar trend but remain lower than negative sentiments. Neutral sentiments are consistently the lowest each month. Significant spikes in negative sentiment are observed around the end of 2015, late 2017, late 2019, early 2020, and again in late 2023.

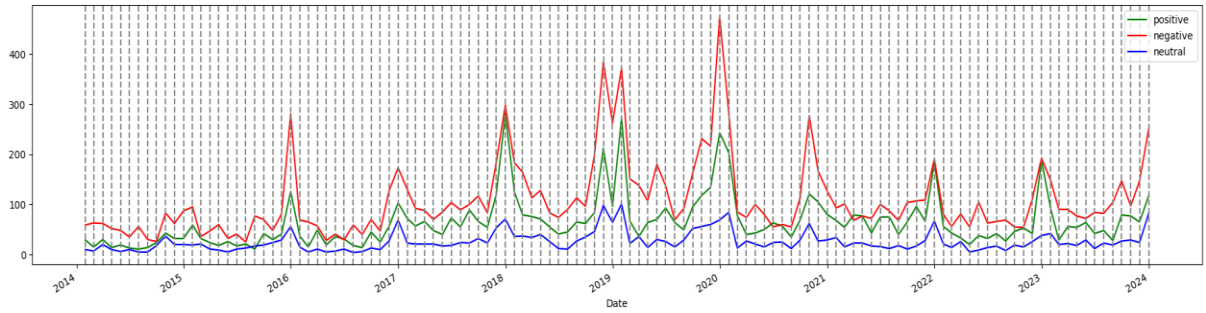

Figure 2. Monthly trend of number of positive, negative and neutral Time.mk teasers from 2014 to 2023.

Aside from Figure 2, which shows the number of positive, negative and neutral news teasers each month, further analysis was provided on the yearly number of positive, negative and neutral news (Figure 3). This yearly analysis of the number of positive, negative and neutral news articles gave us insights into the fluctuations of the number of total news articles throughout the years. From Figure 3, we observed an extreme peak of the number of news articles in 2019. From 2014 until 2016, the number of news articles doesn't vary significantly, while from 2016 until 2019 there is a notable increase in the number of articles (Figure 3). From 2019-2022 there is a steady decline, while after 2022 there seems to be a gradual increase.

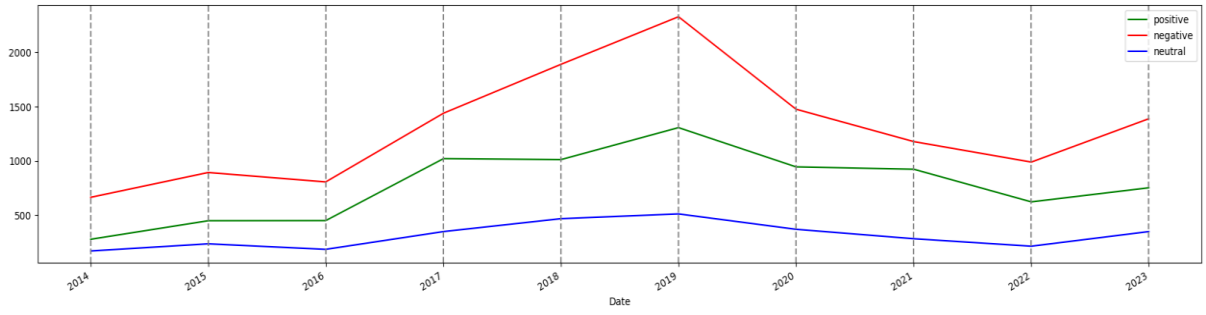

Figure 3. Yearly trend of number of positive, negative and neutral Time.mk teasers from 2014 to 2023

To understand the extent of the involvement of each type of sentiment, we analyzed the yearly percentage of number of positive, negative and neutral news (Figure 4). In 2014, negative sentiments accounted for around 60% of the total, but this gradually decreased to just above 50% by 2017 (Figure 4), then another gradual increase to 2019 at about 60%, decreased in 2021 and followed an upward trend

until 2023. Positive sentiments increased from about 30% in 2015 to around 40% by 2024. Neutral sentiments remained between 10% and 15% throughout the period. The most significant drop in negative sentiments occurred around 2017 and 2018, while positive sentiments increased during the same period. In recent years (2021-2024), there's a notable increase in both positive and negative sentiments, with neutral sentiments remaining largely unchanged.

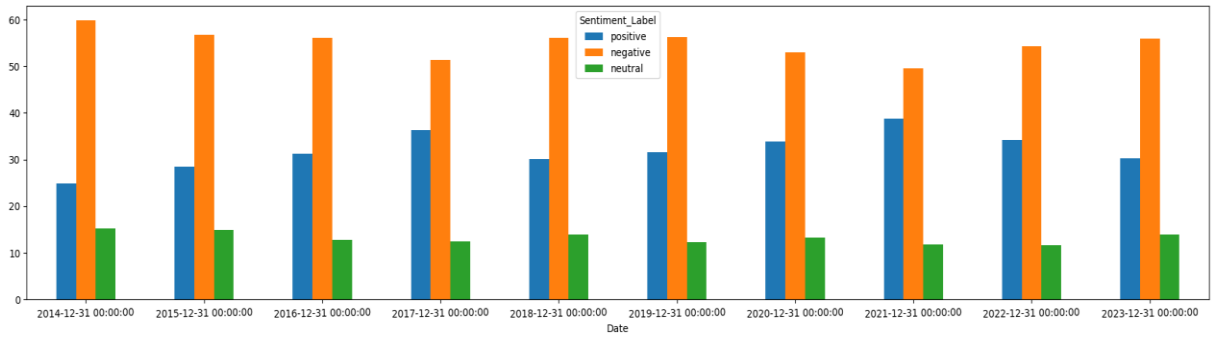

Figure 4. Yearly trend of percentage of positive, negative and neutral Time.mk teasers from 2014 to 2023

4.2 Cross-correlation

A cross-correlation analysis was conducted to investigate the potential relationship between daily air pollution levels and the negative sentiment of news teasers. After the cross-correlation analysis we were able to quantify the relationship between the two variables. The analysis was conducted between the daily average PM10 measures and daily average negative news media teasers. This analysis revealed a moderate positive correlation (coefficient = 0.23) between these time series, with a one-day lag between the air pollution levels and the media sentiment scores. In simpler terms, as air pollution levels increased, the media sentiment tended to become more negative with a one-day delay. This finding is portrayed in Figure 5 where we observe a visual alignment between months with higher PM10 concentrations and increased media coverage with negative sentiment. There is a clear seasonal pattern in both PM10 levels and the frequency of Time.mk teasers. PM10 levels peak during the winter months each year. The peaks in negative teasers often coincide with or slightly lag behind the peaks in PM10 levels, suggesting a relationship where higher pollution levels lead to higher negative sentiment of news teasers.

The frequency of the monthly average number of teasers with positive sentiment shows a less pronounced seasonal pattern compared to negative teasers and PM10 levels. Positive teaser frequency tends to remain relatively low and stable with slight fluctuations, indicating that positive sentiment of the news is less sensitive to changes in air pollution levels. The lag can also be noted on the graph, where there appears to be a slight delay in the increase of negative news frequency after a rise in PM10 levels.

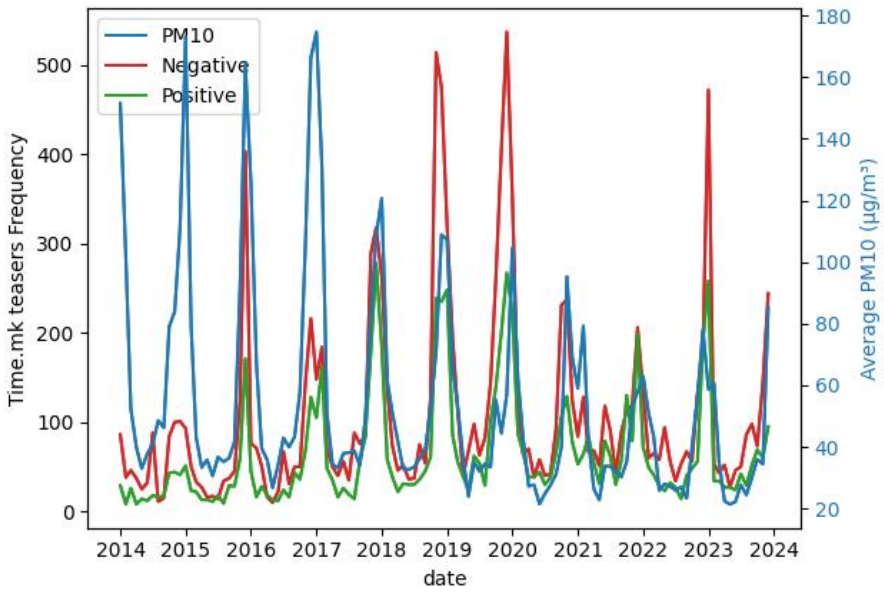

Figure 5. Cross-correlation of monthly frequency of PM10 particles and Time.mk teasers from 2014 to 2023

Additionally to the monthly cross-correlation analysis (Figure 5), we provided investigation in the yearly seasonal trend of the average number of PM10 concentrations and Time.mk teasers (Figure 6). The yearly analysis of both average PM10 particles and average news articles elucidated the fluctuating trends during longer time periods. In Figure 6, the yearly PM10 levels show a decreasing trend from 2014 to 2021, indicating an overall improvement in air quality over this period. There is a slight increase in 2022 and 2023. Negative teaser frequency also shows a general decline from 2014 to 2021, which aligns with the decreasing PM10 levels, suggesting that better air quality leads to less negative news coverage. Positive teaser frequency has an increasing trend from 2014 to 2019, peaks around 2021, and then starts to decline.

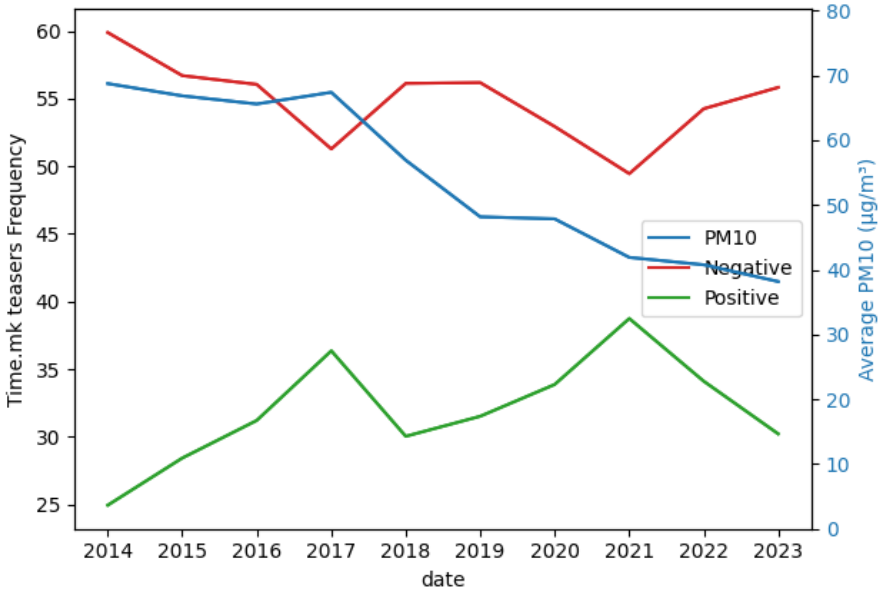

Figure 6. Trend of yearly average PM10 particles and yearly percentage of Time.mk teasers from 2014 to 2023

From this graph (Figure 6), we can discover as well that the peaks of PM10 measures tend to go lower each year, pointing to a decreasing tend over the years. The Mann-Kendall test supports this claim, utilized for the yearly average particles over the 10-year period, with a p-score of 0.0001 and S value of -43.0, reveals that there is a certain decreasing trend. Alternatively, the yearly average numbers

of news teasers with negative sentiment fall under no particular trend with a p-value of 0.28 and S value of 13.0. It is evident from these results that negative news articles do not follow the same trend as PM10 concentrations over longer periods, as can also be seen in the graph, which shows peaks in negative news articles during years of decreased air pollution.

4.3 Obtained Topics

To understand how the Macedonian air pollution crisis has been covered in the media over the past ten years, we analyzed the collection of news teasers through Latent Dirichlet Allocation (LDA). This approach helped us identify the key themes and topics that have dominated the conversation during the researched time period. Thirteen topics we extracted are displayed in Table 1, where the rows represent the years from 2014 to 2023 and the columns represent the five identified topics from the LDA model, with varying prominence each year.

Table 1. Topics from LDA topic modelling of news teasers for each year from 2014 to 2023

Years/Topics	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
2014	Factories and public concern	Environmental issues in Tetovo and Skopje	Skopje and Air Quality	Public protest in Bitola	
2015	Factories and public concern	Environmental issues in Tetovo and Skopje	Government and policies (Tetovo)	Government and policies (Skopje)	
2016	Factories and public concern	Environmental issues in Tetovo and Skopje	Skopje and Air Quality	Air Quality in Bitola and Skopje	
2017	Skopje and Air Quality	Government and policies	Public health	Level of PM10 particles	Water pollution
2018	Skopje and Air Quality	Government and policies	Air Quality in Bitola and Skopje		
2019	Environmental issues in Tetovo and Skopje	Government and policies	Public health	Air Quality in Kumanovo	Air Quality in Bitola and Skopje
2020	Factories and public concern	Government and policies	Level of PM10 particles	Coronavirus	
2021	Skopje and Air Quality	Government and policies	Water pollution		
2022	Skopje and Air Quality	Environmental issues in Tetovo and Skopje	Air Quality in Bitola and Skopje	Heating	
2023	Skopje and Air Quality	Government and policies	Heating	Landfill and waste	

The frequency of each topic provides a window into the evolving dynamics of news coverage over time. An increase in a specific topic's frequency suggests an intensified media focus within a particular year. This heightened attention could be driven by relevant events or a rise in public concern surrounding that specific aspect of air pollution. To facilitate a more granular understanding, we will present a dedicated discussion for each identified topic. These discussions will highlight the most prominent

keywords associated with each theme, offering insights into the specific sub-topics covered by the news media:

- **Factories and Public Concern**: With keywords such as environmental, factory, state and protest, this topic appears consistently from 2014 to 2016, highlighting ongoing issues related to factories and their impact on public concern. It appears again in 2020 exclusively, while its absence from 2021 onwards suggests a potential resolution or shift in focus.
- **Environmental Issues in Tetovo and Skopje**: This topic is prominent in 2014, 2015, 2016, 2019, and 2022, indicating persistent environmental challenges in these cities, stemming from the words: Skopje, Tetovo, particle, environment, measure and high
- **Skopje and Air Quality**: Mentioned frequently (2014, 2016, 2017, 2018, 2021, 2022, 2023), this topic underscores a recurrent concern about air quality in Skopje, with the recurring keywords: Skopje, measure, reduce, environmental and vehicle.
- **Public Protest in Bitola**: This topic appears only in 2014 with the keywords: Bitola, protest, today, problem, car and march, suggesting specific events or a series of events related to public protests in Bitola during that year.
- **Government and Policies**: Featured prominently from 2015 onwards, this topic reflects ongoing governmental efforts and policies addressing environmental issues. Main keywords associated with this topic are: government, authority, health, public, measure, environmental and certain political figures and organizations at the time.
- **Air Quality in Bitola and Skopje**: Present in 2016, 2018, 2019 and 2022, this topic indicates concerns about air quality in these cities. Keywords: Skopje, Bitola, measure, public and high.
- **Public Health**: Emerging in 2017 and 2019, this topic suggests an increasing awareness of public health issues linked to environmental conditions. Keywords: world, health, people, year and child.
- **Level of PM10 Particles**: Mentioned in 2017 and 2020, this topic points to concerns regarding particulate matter and its impact on air quality. Keywords: particle, measuring, station, pm, high and day.
- **Coronavirus**: Only appearing in 2020, this topic reflects the global focus on the COVID-19 pandemic and its environmental implications. Keywords: world, people, coronavirus, health, year and death.
- **Heating**: Appearing in 2022 and 2023, this topic suggests newer concerns related to heating and its environmental impact. Keywords: environment, heating, gas, oil, fuel and use.
- **Landfill and Waste**: Mentioned in 2023, this topic indicates emerging concerns about waste management and landfill issues. Keywords: landfill, waste, health, environmental and state.
- **Water pollution**: Appears as a notable issue in 2017 and 2021, reflecting ongoing concerns about water quality. Keywords: water, contaminated, quality, landfill and lake.
- **Air Quality in Kumanovo**: Emerges as a specific concern in 2019, pointing to regional air quality issues, with keywords such as: particle, environment, Kumanovo, concentration and high

5. Discussion

This study investigated the sentiment of news coverage and its potential association with air pollution levels over a ten-year period, as well as key topics related to air pollution reported by the news. Our findings highlight distinct seasonal patterns and potential relationships between air quality and public sentiment as reflected in media coverage.

As we observed, one of the main interesting and expected points were the peaks of negative sentiment in news coverage consistently during the winter months (November to February) over the ten-year period, aligning with higher air pollution levels (PM10). These findings were also provided by the government of Macedonia in 2017, where they conducted research about the air pollution and found that the highest levels of air pollution are during the heating seasons in Macedonia [39]. According to this governmental study and also the Statistical department of Macedonia around 60% of the population is using wood, pellets, coal or oil as a source for indoor heating [40]. This seasonal trend aligns with previous research by Szép et al. [41] and Zsolt Bodor et al. [42], who attributed the increased air pollution to a combination of factors such as increased heating and atmospheric stability. Furthermore, specific studies in Macedonia have also linked elevated air pollution levels during the winter months

to the increased use of heating [23-25]. These factors, working together, explain the observed seasonal pattern of elevated PM10 levels during winter.

The number of neutral articles remains the same throughout the years, which may be a result of the media trying to deliver more engaging articles by using words that increase the sentiment. The positive number of articles increased by 10%, which may be a result of the air pollution decreasing or focusing on the new government policies, as per our topic modeling. The Macedonian government has delivered new solutions to air pollution during this 10-year period. There are attempts to lower the air pollution from the plan of 2017 [39], giving subsidies for bikes from 2014, as well as indoor air conditioners and pellet stoves for heating – the drop in negative sentiments we observed around 2017 and 2018 and the simultaneous increase in the positive sentiment can be contributed to these new regulations. The plan corroborates our findings, as they clearly state the problem of air pollution during the winter months and observed increase in the PM10 measurements. In our results we noticed an increase in the positive sentiment in 2017 as well, indicating that the news media most probably responded with positive acknowledgement of the government engagement.

The peak in year 2017 can also be linked to the new regulations and policies by the government presented in that year by the government plan. We can note that the 'Government and policies' topic is present in that year and remains persistent in the following years. These policy changes could potentially explain the increase in the yearly number of news articles from 2017 to 2020, where the news media was focusing on the ongoing efforts by authorities to address environmental challenges This alignment with policy changes suggests that governmental actions are a key driver of public and media attention to environmental matters. The emergence of public health as a topic in 2017 and 2019 highlights the growing awareness of the link between environmental conditions and health outcomes. This is reinforced by keywords such as "world," "health," "people," and "child," indicating that the health impacts of air pollution are becoming more prominent in public discourse.

The highest peak of negative news articles is noted between late 2019 and early 2020. This could be linked to the ongoing protests in Bitola and Skopje in late December 2019 [43,44], which is as well supported with the topic 'Air Quality in Bitola and Skopje' appearing in 2019. Concerns about air quality in Bitola and Skopje, indicated by recurring mentions in 2016, 2018, 2019, and 2022, emphasize the need for effective pollution control measures in these cities. The frequent mention of Skopje and air quality across numerous years underscores a consistent concern about pollution levels in the capital city. This aligns with efforts to address pollution sources, particularly vehicular emissions, as indicated by keywords like "measure," "reduce," and "vehicle." Public protests in Bitola in 2014 highlight specific events that drew significant media attention, reflecting heightened local activism over environmental issues.

The recurring theme of factories and public concern from 2014 to 2016, with a resurgence in 2020, highlights persistent public unrest regarding the environmental impact of factories. The peak of number of negative articles in 2015 can be related to heightened public awareness during this time about the factories and lack of government regulations. This topic's absence from 2021 onwards suggests a potential resolution or a shift in media focus. Similarly, environmental issues in Tetovo and Skopje, appearing consistently in multiple years, underscore ongoing challenges in these cities. Keywords such as "Skopje," "Tetovo," "particle," and "environment" indicate that air pollution remains a critical issue, necessitating continued attention and action. This is expected, as Tetovo and Skopje are one of the 30 most polluted cities in Europe [6].

The exclusive appearance of the coronavirus topic in 2020 reflects the global focus on the COVID-19 pandemic and its environmental implications. This intersection of public health crises and environmental issues shows how pandemics can shift media and public attention.

New concerns related to heating and its environmental impact, emerging in 2022 and 2023, suggest a growing recognition of the need for sustainable heating solutions. The mention of landfill and waste in 2023 highlights emerging issues in waste management, reflecting a shift towards the environmental and health impacts of waste disposal practices.

Water pollution, appearing notably in 2017 and 2021, indicates ongoing concerns about water quality, suggesting that this remains a significant environmental challenge. The specific focus on air quality in Kumanovo in 2019 points to regional air pollution issues, emphasizing the need for broader geographical attention and action.

The cross-analysis provided hard evidence that there is a relationship between the sentiment of media coverage and PM10 levels. The data indicates a strong correlation between PM10 levels and the frequency of negative news teasers. High pollution levels correlate with increased negative coverage, reflecting public concern and media focus on environmental issues. According to Ye, Binbin et al. [45] and Samuel Agyei-Mensah et al. [46], both studies provided an analysis of media data and air pollution and both found that correlations between air pollution and media articles were also stronger when there was lower air quality and it generally led to a more negative media slant.

Our study further proves that there is a correlation between these sources especially in the winter months, when the media coverage reports the changes in PM10 levels with a lag of one day. This result is expected, as recent studies confirm that there is a correlation between the media reporting and air pollution levels.

One unexpected finding emerged from our analysis: an increase in negative news teasers about air pollution despite potentially decreasing PM10 readings. This intriguing phenomenon suggests a disconnect between objective air quality data and public perception as reflected in media coverage. Several possible explanations exist. Increased media scrutiny and heightened public awareness could lead to a perception of worsening air quality despite objective improvement. Alternatively, negative media attention could be targeting a lack of progress in implementing effective air pollution policies. Sensationalized reporting or the influence of social media trends on news framing could also be contributing factors.

6. Conclusions

This study provides a comprehensive analysis of the sentiment of news coverage and its association with air pollution levels over a ten-year period in Macedonia, leveraging sentiment analysis and topic modeling to uncover significant insights. A consistent seasonal pattern emerged: negative news coverage peaked during winter months when PM10 levels, a key air pollution indicator, were highest. This finding aligns with the increased use of heating fuels during winter and is further supported by existing research. Despite a stable number of neutral articles, a rise in positive sentiment articles suggests a potential decrease in pollution levels or the effectiveness of new government policies.

The overall number of news articles peaked in 2019, coinciding with public protests regarding air quality. Similarly, the peak in 2017 aligns with the introduction of new environmental policies, highlighting a persistent public and media focus on governmental actions.

Topic modeling revealed the evolving nature of public discourse. Recurring topics like air quality in specific cities, factories, and government policies underscore longstanding environmental concerns. Emerging topics like the COVID-19 pandemic's impact, heating practices, and waste management reflect an evolving understanding of environmental issues and shifting media focus.

An unexpected finding was the increasing trend of negative news teasers despite potentially decreasing PM10 levels. This warrants further investigation but could be attributed to factors such as heightened public awareness of long-term health impacts, or broader socio-political considerations influencing media narratives.

Our cross-analysis confirmed a strong correlation between high PM10 levels and negative news coverage, emphasizing the media's crucial role in highlighting environmental challenges.

6.1 Academic implications

Our study contributes to the literature by providing a comprehensive analysis of air pollution-related media coverage, using a large dataset of news articles published in a period of a decade, from 2014 to 2023, complementing the existent research focusing primarily on social media data. This approach allows for a more structured and reliable analysis of public sentiment and media focus, enabling us to understand the influence that media might have on public awareness of the subject. The methodology, including sentiment analysis and topic modeling, offers robust tools for examining the intersection of environmental issues and media representation. By employing topic modeling, we identified key themes and their prevalence over the investigated time of 10 years. Moreover, the cross-correlation analysis exploring the alignment of the sentiments identified in the media news with the actual air

pollution evaluated through the PM10 measures, adds value to the body of knowledge by providing evidence-based implications related to different sources of data.

Our results align with the study by He, Zhang & Chen [19], which emphasizes the importance of investigating news content beyond sentiment analysis. They coincide with several studies, specifically about the heightened levels of air pollution during winter months [18, 22-24]. Additionally, our results emphasize the broader impact of media coverage in promoting conservation efforts and proactive measures, as highlighted by Ramondt & Ramírez [20]. However, our study differs from these researches with an analysis of a significantly long time period.

This study contributes to the field by providing empirical evidence linking media sentiment with actual PM10 levels. The findings suggest that negative sentiments in news teasers correlate with higher PM10 levels, indicating that media sentiment can indeed reflect the real air quality situation, being at the same time a valuable asset that might shape public awareness on the subject. Moreover, the research underscores the importance of considering media news sentiment in environmental policy-making and public health interventions, aligning with the World Bank's emphasis on comprehensive approaches to air quality management [47].

By supporting the findings of previous studies with data on media news sentiment, this research strengthens the argument for addressing environmental health concerns through integrated strategies. It also improves upon previous conclusions by highlighting the critical role of media in influencing public opinion and behavior towards environmental issues. Specifically, the study's findings challenge the notion that media accurately reflects the severity of air pollution, suggesting that media outlets may sometimes underestimate or overestimate the extent of the problem.

6.2 Managerial implications

For policymakers and environmental managers, our findings highlight the importance of managing air pollution, particularly during the winter months. The recurring themes identified through topic modeling can inform targeted communication strategies and policy interventions. By understanding the specific concerns and topics that dominate media coverage, authorities can better address public concerns, enhance environmental regulations, and improve public health outcomes. Understanding the cyclical nature of media attention, as observed in the study by He, Zhang & Chen (2020) [19], can help stakeholders maintain consistent focus on environmental issues, ensuring long-term public awareness and action. Additionally, the insights into seasonal patterns and the impact of heating practices on air pollution can guide efforts to promote cleaner, more sustainable heating solutions. Furthermore, fostering collaboration between media outlets and environmental agencies could help ensure exact reporting of air quality data that can offer a more precise tracking of the air quality, which will then service as a basis for improving the related laws and regulations.

For media outlets, the study suggests a need to move beyond solely reporting peak pollution levels and delve deeper into solutions. The unexpected rise in negative news coverage despite potentially decreasing PM10 levels highlights a potential disconnect between air quality data and related interpretations as portrayed by the media. Exploring topics like clean energy initiatives, public education campaigns, and the effectiveness of government policies could improve the correct reporting about the polluters, challenges and solutions, and will empower the media to influence the public to engage in mitigating air pollution.

The persistent air pollution over the ten-year period underlines the need for continuous media attention, and not just during winter months when concerns are heightened. By adopting a solutions-oriented approach and fostering collaboration with policymakers and environmental experts, media outlets can play a crucial role in promoting long-term strategies for managing and preventing air pollution.

6.3 Limitations and future research lines

This research presents several limitations that open new avenues for exploration of the subject of air quality challenges. Firstly, the data was collected only for the case of Macedonia, restricting the

possibility to broaden the conclusions to a wider region. Future research can explore air quality concerns in other Western Balkan countries and involve into a deeper observation of air quality challenges and the public discourse in specific regions, comparing the trends and identifying possible similarities among them.

Secondly, this study primarily focused on news media coverage, but social media platforms have become a powerful force in shaping public discourse on environmental issues. Thus, future research could incorporate data from social media, and additionally surveys or public opinion polls or interviews with policy makers, to understand how the different sources might influence or be influenced by media coverage and gain a broader understanding of the public and institutional perception of air quality issues.

Thirdly, the unexpected increase in negative news teasers despite potentially decreasing PM10 levels warrants further investigation, which could involve in-depth analysis of these news articles to understand the specific narratives or framing techniques used. Interviews with journalists or media editors could also be conducted to explore the decision-making process behind such coverage.

Finally, this study focused on PM10 levels as a key indicator of air pollution. However, air pollution encompasses a range of pollutants with varying health effects, that could be a subject of research future investigation of the public awareness and media focus on the long-term health consequences.

References

1. Lelieveld, J.; Evans, J.S.; Fnais, M; Giannadaki, D.; Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **2015**, *525*, 367-71. https://doi.org/10.1038/nature15371.
2. Safiri, S; Carson-Chahhoud, K.; Noori, M.; Nejadghaderi, S.A.; Sullman, M.J.M.; Ahmadian Heris, J.; Ansarin, K.: Mansournia, M.A.; Collins, G.S.; Kolahi, A.A.; Kaufman, J.S. Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990-2019: results from the Global Burden of Disease Study 2019. *BMJ* **2022**, 378. https://doi.org/10.1136/bmj-2021-069679
3. Singh, A.; Singh K. K. An Overview of the Environmental and Health Consequences of Air Pollution. Iranica Journal of Energy & *Environment* **2022**, *13*, 231-237.
4. Safo-Adu, Godfred, et al. Outdoor and indoor particle air pollution and its health consequences in African cities: New evidence and an exhortation. Sustainable Environment 2023, 9. https://doi.org/10.1080/27658511.2023.2265729
5. Martinez, G.S.; Spadaro, J.V.; Chapizanis, D.; Kendrovski, V.; Kochubovski, M.; Mudu, P. Health Impacts and Economic Costs of Air Pollution in the Metropolitan Area of Skopje. Int. J. Environ. Res. *Public Health* **2018**, *15*, 626. https://doi.org/10.3390/ijerph15040626
6. IQAIR. World's most polluted cities. Available online: https://www.iqair.com/world-most-polluted-cities?continent=59af92ac3e70001c1bd78e52&country=&state=&sort=-rank&page=1&perPage=50&cities= (accessed on 25 May 2024)
7. World Health Organization. Ambient (outdoor) air pollution, 2022. Available online: www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health (accessed on 1 January 2024)
8. NSW Government. Particulate matter (PM10 and PM2.5), 2020. Available on: www.health.nsw.gov.au/environment/air/Pages/particulate-matter.aspx (accessed on 21 March 2024)
9. Adebayo-Ojo, T.C.; Wichmann, J.; Arowosegbe, O.O.; Probst-Hensch, N.; Schindler, C.; Künzli, N. Short-Term Joint Effects of PM10, NO2 and SO2 on Cardio-Respiratory Disease Hospital Admissions in Cape Town, South Africa. Int. J. Environ. Res. *Public Health* **2022**, *19*, 495. https://doi.org/10.3390/ijerph19010495
10. Bodor, K.; Micheu, M.M.; Keresztesi, Á.; Birsan, M.-V.; Nita, I.-A.; Bodor, Z.; Petres, S.; Korodi, A.; Szép, R. Effects of PM10 and Weather on Respiratory and Cardiovascular Diseases in the Ciuc Basin (Romanian Carpathians). Atmosphere 2021, 12, 289. https://doi.org/10.3390/atmos12020289
11. Saygın, M.; Gonca, T.; Öztürk, Ö.; Has, M.; Çalışkan, S.; Has, ZG.; Akkaya, A. To Investigate the Effects of Air Pollution (PM10 and SO2) on the Respiratory Diseases Asthma and Chronic Obstructive Pulmonary Disease. *Turk Thorac J* **2017**, *2*, 33-39. https://doi.org/10.5152/TurkThoracJ.2017.16016.
12. Orru, K.; Nordin, S.; Harzia, H.; Orru H. The role of perceived air pollution and health risk perception in health symptoms and disease: a population-based study combined with modelled levels of PM10. *Int Arch Occup Environ Health* **2018**, *91*, 581-589. https://doi.org/10.1007/s00420-018-1303-x.
13. Mahajan, S.; Chung, M.; Martinez,J.; Olaya,Y.; Helbing, D.; Chen, L. Translating citizen-generated air quality data into evidence for shaping policy. *Humanities and Social Sciences Communications* **2022**, *9*, 1-18.

14. Marukhovskyi, O.; Marukhovska-Kartunova, O.; Kolybabiuk, S.; Pavlova, L.; Karchevska, O. The Role of Online Media in Shaping Public Opinion about the Prospects for Peace and Military Conflict. Pakistan Journal of Criminology 2024, 16.
15. Liao, CH. Exploring the Influence of Public Perception of Mass Media Usage and Attitudes towards Mass Media News on Altruistic Behavior. *Behavioral sciences* **2023**, *13*, 621. https://doi.org/10.3390/bs13080621
16. Jafri, I.M.M.; Noor, N.M.; Rahim, N.A.A.A.; Saufie, A.Z.U; Habil, G.D.. Prediction of PM10 Level During High Particulate Event in Malaysia Using Modified Model. *EDP Sciences* **2023**, *437*, 8. https://doi.org/10.1051/e3sconf/202343701001
17. Jahn, H.J.; Schneider, A.; Breitner, S.; Eissner, R.; Wendisch, M.; Krämer, A. Particulate matter pollution in the megacities of the Pearl River Delta, China - a systematic literature review and health risk assessment. *Int J Hyg Environ Health* **2011**, *214*, 281-295. https://doi.org/10.1016/j.ijheh.2011.05.008.
18. Wang, X.; Wu, G.; Xiang, Z.; Zhang, J. Air pollution and media slant: evidence from Chinese corporate news. *Emerging Markets Finance and Trade* **2022**, *58*, 2880-2894.
19. He, Y.; Zhang, G.; Chen, L. Analysis of News Coverage of Haze in China in the Context of Sustainable Development: The Case of China Daily. *Sustainability* **2020**, *12*, 386. https://doi.org/10.3390/su12010386
20. Ramondt, S.;Ramírez, A.S. Media reporting on air pollution: health risk and precautionary measures in national and regional newspapers. International Journal of Environmental Research and Public Health 2020, 17, 6516.
21. Amiraslani, F. Climate change and urban citizens: the role of media in publicizing the conservation of green spaces and mitigation of air pollution. *Conservation* **2022**, *2*, 219-232.
22. Mirakovski, D.; Boev, B.; Boev, I.; Nikolova, M.H.; Zendelska, A.; Shijakova, T. Wintertime urban air pollution in Macedonia–composition and source contribution of air particulate matter. *18th World Clean Air Congress*, Istanbul, Turkey , 23-27 September 2019, 492-500.
23. Dimovska, M.; Gjorgjev, D. Assessing Health Impact of Air Pollution in Macedonian Cities. *Biomedical Journal of Scientific & Technical Research* **2018**, *10*. https://doi.org/10.26717/BJSTR.2018.10.001887
24. Ministry of Environment and Physical Planning. Macedonian Air Quality Assessment Report for the Period 2005-2015. Available online: https://air.moepp.gov.mk/wpcontent/uploads/2017/07/AirQualityReport_EN.pdf (accessed on 20 February 2024)
25. Avis, R. Causes and consequences of air pollution in North Macedonia. *Environ. Sci. Pollut. Res* **2022**, *21*. Available online: https://mvdsi.seeu.edu.mk/papers/K4D_NM2.pdf (accessed on 25 May 2024)
26. Ministry of environment and physical planning. Air Quality Portal. Available on: https://air.moepp.gov.mk/?page_id=175, (accessed on 1 March 2024)
27. Trajkovski, I. How does TIME.mk work?. Time.mk 2008. Available online: https://time.mk/info/site, (accessed on 12 MArch 2024)
28. Sarker, MO Faruque, and Sam Washington. Learning Python Network Programming. Packt Publishing Ltd, 2015. 86-91
29. Almaqbali, I. S. H., Al Khufairi, F. M. A., Khan, M. S., Bhat, A. Z., & Ahmed, I. Web Scrapping: Data Extraction from Websites. *Journal of Student Research 2019*. https://doi.org/10.47611/jsr.vi.942
30. Karn, S.K.; Buckley, M.; Waltinger, U.; Schütze, H. News Article Teaser Tweets and How to Generate Them. arXiv preprint arXiv:1807.11535. 2018.
31. Herdawan, D.; Sari, D.; Setyoningrum, A. Google Translate: Strategies to Optimize Its Academic Writing Translation. *English Education: Jurnal Tadris Bahasa Inggris* **2023**. *16*. 108. https://doi.org/10.24042/ee-jtbi.v16i2.18593.
32. Hutto, C.; Gilbert, E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Eight International AAAI Conference on Web and Social Media*, Ann Arbor, Michigan USA, 1-4 June 2014. 8, 216-225. https://doi.org/10.1609/icwsm.v8i1.14550
33. Poletto, F.; Miranda, F. Chapter 8 - Preprocessing of SWD data, Seismic While Drilling (Second Edition); Elsevier: Amsterdam, The Netherlands, 2004; Chapter 8, 325-399, ISBN 9780128231456, https://doi.org/10.1016/B978-0-12-823145-6.00020-1
34. Pang, Z.; Wang, G.; Wang, B.; Wang, L. Comparison Between Time Shifting Deviation and Cross-Correlation Methods. *Journal of Lightwave Technology* **2022**, *40*, 3003-3009. https://doi.org/10.1109/JLT.2022.3146718.
35. Hamed, K.H.; Rao, A.R. A modified Mann-Kendall trend test for autocorrelated data. *J. Hydrol.* **1998**, *204*, 182-196. https://doi.org/10.1016/S0022-1694(97)00125-X
36. Tech, Tetra. Statistical analysis for monotonic trends. 2011.
37. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. Journal of machine Learning research 2003, 3, 993-1022.
38. Pandas. Available on: https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.html#pandas.DataFrame (accessed on 01 June 2024)
39. Government of the Republic of North Macedonia. Available online: https://vlada.mk/node/15965?ln=en-gb (accessed on 03 June 2024)

40. State Statistical Office. Households by heating method and type of settlement, by municipalities. Census 2021. Available online: https://makstat.stat.gov.mk/PXWeb/pxweb/mk/MakStat/MakStat__Popisi__Popis2021__NaselenieVkupno__Domakinstva/T2006P21.px/ (accessed on 25 March 2024)
41. Szép, R.; Keresztes, R.; Korodi, A.; Szende, T.; Niculae, A; Birloiu, A. Dew point - Indirect particulate matter pollution indicator in the Ciuc Basin - Harghita, Romania. *Rev. Chim* **2016**, *67*. 1914-1921.
42. Bodor, Z.; Bodor, K.; Keresztesi, Á; Szép, R. Major air pollutants seasonal variation analysis and long-range transport of PM10 in an urban environment with specific climate condition in Transylvania (Romania). Environ Sci Pollut Res Int. 2020, 27, 38181-38199. doi: 10.1007/s11356-020-09838-2.
43. Slobodna Evropa. Available online: https://www.slobodnaevropa.mk/a/30348933.html (accessed on 02 June 2024)
44. Radio Mof. Available online: https://www.radiomof.mk/srednoshkolski-marsh-za-chist-vozduh-i-vo-bitola/ (accessed on 02 June 2024)
45. Ye, B.; Krishnan, P.; Jia, S. Public Concern about Air Pollution and Related Health Outcomes on Social Media in China: An Analysis of Data from Sina Weibo (Chinese Twitter) and Air Monitoring Stations. *Int J Environ Res Public Health* **2022**, *19*, 16115. https://doi.org/10.3390/ijerph192316115.
46. Agyei-Mensah, S.; Kyere-Gyeabour, E.; Mwaura, A.; Mudu, P. Between Policy and Risk Communication: Coverage of Air Pollution in Ghanaian Newspapers. *Int. J. Environ. Res. Public Health* **2022**, *19*, 13246.
47. The World Bank. Western Balkans Regional AQM - Western Balkans. Available online: https://documents1.worldbank.org/curated/zh/116521576516981237/pdf/Air-Quality-Management-in-North-Macedonia.pdf (accessed on 02 June 2024)