JULY 2025
No 364

Authors
Thanh T.N Nguyen
Truong X. Ngo
An H. Phan
Linh T.K Nguyen
Thanh Ngo-Duc

Coordination
Cecilia Poggi

Impact of Meteorological Factors and Extreme Weather Events on PM2.5 Pollution in Vietnam

Research papers

ÉDITIONS AFD
AGENCE FRANÇAISE DE DÉVELOPPEMENT

Agence française de développement

Papiers de recherche

Les *Papiers de Recherche de l'AFD* ont pour but de diffuser rapidement les résultats de travaux en cours. Ils s'adressent principalement aux chercheurs, aux étudiants et au monde académique. Ils couvrent l'ensemble des sujets de travail de l'AFD : analyse économique, théorie économique, analyse des politiques publiques, sciences de l'ingénieur, sociologie, géographie et anthropologie. Une publication dans les *Papiers de Recherche de l'AFD* n'en exclut aucune autre.

Les opinions exprimées dans ce papier sont celles de son (ses) auteur(s) et ne reflètent pas nécessairement celles de l'AFD. Ce document est publié sous l'entière responsabilité de son (ses) auteur(s) ou des institutions partenaires.

Research Papers

AFD Research Papers are intended to rapidly disseminate findings of ongoing work and mainly target researchers, students and the wider academic community. They cover the full range of AFD work, including: economic analysis, economic theory, policy analysis, engineering sciences, sociology, geography and anthropology. *AFD Research Papers* and other publications are not mutually exclusive.

The opinions expressed in this paper are those of the author(s) and do not necessarily reflect the position of AFD. It is therefore published under the sole responsibility of its author(s) or its partner institutions.

Impact of Meteorological Factors and Extreme Weather Events on $PM_{2.5}$ Pollution in Vietnam

Authors

Thanh T.N Nguyen
University of Technology and Engineering, Vietnam National University, Hanoi, Viet Nam

Truong X. Ngo
University of Technology and Engineering, Vietnam National University, Hanoi, Viet Nam

An H. Phan
University of Technology and Engineering, Vietnam National University, Hanoi, Viet Nam

Linh T.K Nguyen

University of Science, Vietnam National University Hanoi, Hanoi, Viet Nam

Thanh Ngo-Duc

University of Science and Technology of Hanoi, Vietnam Academy of Science and Technology, Hanoi, Viet Nam

Coordination

Cecilia Poggi (AFD)

Abstract
This study examines how meteorological factors and extreme weather influence $PM_{2.5}$ pollution across six socio-economic regions in Vietnam from 2012 to 2020. Univariate and multivariate analyses, including correlation analyses, Linear Regression, and Generalized Additive Models are used to quantify the independent and the combined effects of meteorological factors on $PM_{2.5}$ fluctuations. The study identifies extreme weather events and uses overlap ratio along with Two-Way Fixed Effect model to analyze and assess their causal impact on $PM_{2.5}$ levels. Two specific cases (drought and typhoon) are quantified using a Difference-in-Differences model. Results show that meteorology strongly affects regional $PM_{2.5}$ variations, with rainfall and wind speed impact the South immediately, while temperature and humidity influence longer-term concentrations mainly in the North; surface pressure and temperature have strongest impact. Together, they explain up to 54% of $PM_{2.5}$ variability in the North but less in the South, indicating additional pollution sources beyond meteorology. Extreme weather impacts $PM_{2.5}$ levels regionally, varying by event type, duration. Cool days, heavy rain, and high humidity reduce pollution, while cool nights, low humidity, low wind speed increase it. A notable spike of 14.2 µg/m³ occurred during prolonged cool nights in the Red River Delta, near the WHO's 10 µg/m³ limit for 24-hour exposure.

Keywords
$PM_{2.5}$; Meteorological factors; Extreme weather events; Vietnam.

Acknowledgements
This research is part of the research programme GEMMES Viet Nam 2 (C3-WP4-A1, 2023-2026), funded by the French Development Agency (AFD).

JEL Classification
Q53; C23; Q54.

Original version
English

Accepted
July 2025

Résumé

Cette étude examine comment les facteurs météorologiques et les conditions météorologiques extrêmes influencent la pollution PM2.5 dans six régions socio-économiques au Vietnam de 2012 à 2020. Les analyses univariées et multivariées, y compris les analyses de corrélation, la régression linéaire et les modèles additifs généralisés sont utilisées pour quantifier les effets indépendants et combinés des facteurs météorologiques sur les fluctuations de $PM_{2.5}$. L'étude identifie les événements météorologiques extrêmes et utilise le ratio de chevauchement ainsi que le modèle à effet fixe bidirectionnel pour analyser et évaluer leur impact causal sur les niveaux de $PM_{2.5}$. Deux cas spécifiques (sécheresse et typhon) sont quantifiés à l'aide d'un modèle de différence dans les différences. Les résultats montrent que la météorologie affecte fortement les variations régionales des PM2,5, avec des précipitations et une vitesse du vent qui ont un impact immédiat sur le Sud, tandis que la température et l'humidité influencent les concentrations à plus long terme principalement dans le Nord ; la pression de surface et la température ont l'impact le plus fort. Ensemble, ils expliquent jusqu'à 54% de la variabilité des $PM_{2.5}$ dans le Nord mais moins dans le Sud, indiquant des sources de pollution supplémentaires au-delà de la météorologie. Les phénomènes météorologiques extrêmes ont des impacts sur les $PM_{2.5}$. niveaux régionalement, variant selon le type d'événement, la durée. Les jours frais, les fortes pluies et une humidité élevée réduisent la pollution, tandis que les nuits fraîches, une faible humidité, une faible vitesse du vent l'augmentent. Un pic notable de 14,2 µg/m3 s'est produit pendant des nuits fraîches prolongées dans le delta du fleuve Rouge, près de la limite de 10 µg/m3 fixée par l'OMS pour une exposition de 24 heures.

Mots-clés

$PM_{2.5}$; facteurs météorologiques ; événements météorologiques extrêmes; Vietnam.

Remerciements

Cette recherche s'inscrit dans le programme de recherche GEMMES Viet Nam 2 (C3-WP4-A1, 2023-2026), financé par l'Agence française de développement (AFD).

Classification JEL

Q53; C23; Q54.

Version originale

Anglais

Acceptée

juillet 2025

Introduction

Air pollution is becoming a serious global issue, with fine particulate matter ($PM_{2.5}$) being one of the most dangerous pollutants (Wei & Semple, 2023). $PM_{2.5}$ particles are commonly produced from household combustion devices, transportation, factories, industrial facilities, and wildfires. Health effects attributable to long-term exposure to $PM_{2.5}$ include ischemic heart disease, lung cancer, chronic obstructive pulmonary disease (COPD), lower-respiratory infections, stroke, type 2 diabetes, and adverse birth outcomes. According to the World Health Organization (WHO), approximately 7 million premature deaths occur annually due to air pollution (World Health Organization, 2014), and 99% of the global population is exposed to $PM_{2.5}$ levels that exceed safe limits (Organization, 2021). In Vietnam, air pollution is becoming increasingly severe, ranking 130th out of 180 countries in the Environmental Performance Index (EPI) (Wolf et al., 2022).

Because of its serious adverse effects, the $PM_{2.5}$ pollution has garnered significant public concern and scientific attention in order to understand its mechanism and to find mitigation solutions. $PM_{2.5}$ concentrations are influenced not only by emissions but also by various natural geographical factors, including topography, vegetation, and climate. Among these, meteorological factors have the most substantial effect on $PM_{2.5}$ levels (Jones et al., 2010). To effectively address air pollution, it is essential to have a complete understanding of the relationship between $PM_{2.5}$ concentration and meteorological factors. In addition, extreme weather events significantly influence fluctuations in $PM_{2.5}$ concentrations by enhancing the formation and dispersion of pollutants, creating favorable conditions for accumulation, and amplifying the effects of harmful substances (T. Chen et al., 2016),(D. Zhao et al., 2018). Deeper insights into these relationships will aid in the development of effective measures and policies to improve air quality and protect human health.

This study analyzes the relationships between $PM_{2.5}$ concentrations and meteorological events on a national scale in Vietnam, taking into account the unique characteristics of each region during the period from 2012 to 2022. The Pearson and Spearman correlations and Linear Regression (LR) and Generalized Additive Models (GAM) are used to assess combined univariate and multivariate effects, respectively. Next, the study examines the impacts of extreme weather events identified from historical data by analyzing the overlap ratio and quantifying their effects with Two Way Fix Effect (TWFE) model. Two specific cases – the 2016 drought in the Mekong River Delta and the 2017 typhoon in Central Vietnam are analyzed using the Difference In Difference (DID) model.

Our findings indicate that meteorology strongly affects regional $PM_{2.5}$ variations. Rainfall and wind speed tend to have immediate effects on $PM_{2.5}$ levels, while temperature and humidity influence concentrations over longer periods in northern regions. In southern regions, all factors show more immediate impacts. Among meteorological factors, surface pressure and temperature have the strongest impacts. When considering the combined effects, all meteorological factors in the GAM explain approximately 53.1% to 54.4% of $PM_{2.5}$ variability in norther regions, which is higher than the 27.5% to 36.7% observed in southern regions. In term of impacts of extreme weather events, cool nights, low pressure, low wind events significantly increase $PM_{2.5}$, while cool days, rainfall, high humidity, high pressure reduce it. Notably, in Red River Delta, extended cool nights up to 7–8 days sharply increase $PM_{2.5}$, peaking at 14.20 *μg/m*^{3}. For the case studies of the 2016 drought in Southern Vietnam and Typhoon Talas (2017), the drought can increase $PM_{2.5}$ by 6.45 μg/m^3 in the Mekong Delta while the typhoon reduced by 2.05 μg/m^3 in Central Vietnam. These study results emphasize the importance of meteorological factors and extreme weather events in regional air quality management, meanwhile air pollution needs to be considered in climate change adaptation and disaster preparedness and resilience, in Vietnam. This is the first study offering a comprehensive overview of how various meteorological factors interact with $PM_{2.5}$ and investigating deeply on the relationships of extreme events and $PM_{2.5}$ levels across all regions in Vietnam. In comparison with other published studies, the quantification impacts of extreme events on $PM_{2.5}$ utilizing the TWFE and DID models is firstly implemented.

The paper is divided into the following section. Section 1 reviews related works. Section 2 depicts the study area, study period, and the specific data used in the analysis. Section 3 presents the methodologies, procedures, and tools employed for data preparation and analysis. Section 4 presents the results of the analysis of the relationships between meteorological factors and $PM_{2.5}$, as well as the impacts of extreme weather events on $PM_{2.5}$. Finally, section 5 concludes the study.

1. Related works

1.1. The relationship between meteorological variables and $PM_{2.5}$

Numerous studies worldwide have shown that meteorological factors such as temperature, humidity, and pressure are closely related to air quality in general and $PM_{2.5}$ concentration variations in particular. Correlation measures are commonly used in analytical problems to assess the relationships between variables. Li, Feng, & Liang (2017) analyzed data from Hong Kong in 2013 using a cross-correlation matrix, revealing a positive correlation between $PM_{2.5}$ concentrations and pressure and negative correlations with temperature, relative humidity, rainfall, and wind speed. Among these factors, temperature, pressure, and rainfall were identified as primary influencers of $PM_{2.5}$ levels Similarly, Li, Ma, et al. (2017) studied the relationship between $PM_{2.5}$ and PM_{10} and meteorological factors in Shenyang, China, noting seasonal and annual variations in PM concentrations, which were generally negatively correlated with wind speed and positively correlated with air pressure, air temperature, and relative humidity. Other studies, including those in Guayaquil, Ecuador (Rincon et al., 2023), and Makkah, Saudi Arabia (Munir et al., 2017), reported negative correlations between $PM_{2.5}$ and relative humidity and wind speed. However, temperature correlations varied: it was positive in Saudi Arabia but exhibited both positive and negative correlations in Ecuador. In D. Zhao et al. (2018), Pearson correlations were used to analyze relationships in five Chinese cities: Beijing, Shanghai, Guangzhou, Chengdu, and Shenyang. The results indicated that pressure positively affected $PM_{2.5}$ concentrations, while temperature and rainfall had negative effects. Stepwise regression showed varying correlations across cities, but pressure and temperature significantly influenced $PM_{2.5}$ levels in most locations, with higher air pressure and lower temperatures linked to increased $PM_{2.5}$ pollution. Spearman R was used in Chen's study in Nanjing (T. Chen et al., 2016) and Yang's research across China (Yang et al., 2017). Chen found that $PM_{2.5}$ inversely correlated with wind speed, relative humidity, and rainfall, while temperature generally showed a positive correlation, with a negative trend noted overall. Yang's study indicated that relative humidity negatively correlated with $PM_{2.5}$ in most regions, except for northern China and Urumqi, while wind speed had a similar negative correlation, except in Hainan Island. Temperature showed a strong negative correlation with $PM_{2.5}$ in most areas, reversing in winter. Pressure positively associated with $PM_{2.5}$ concentrations in northeast and mid-south China but was weaker elsewhere. Wang's report on Nagasaki (J. Wang & Ogawa, 2015) used Spearman R compared with R^2 from unary linear regression and found that temperature negatively correlated and rainfall positively correlated with $PM_{2.5}$, with humidity and wind speed correlations varying based on thresholds. The analysis suggested that westerly winds could transport the most pollutants to Nagasaki.

In order to evaluate combined impacts of meteorological variables on $PM_{2.5,}$ multivariate models are utilized. In Wang's study in China, linear regression models were applied alongside univariate correlation analysis to assess how the impact of meteorological factors changes in a multivariate context. The patterns observed were similar to those from the univariate analysis, showing that relative humidity positively influences $PM_{2.5}$ concentrations in northern China but negatively in southern China, temperature and wind speed have a negative impact. However, the results for pressure showed larger variation. In Amos P.K. Tai's report, a positive relationship between temperature and $PM_{2.5}$ was noted in the U.S. (Tai et al., 2010), while both positive and negative relationships were identified in R. Zhao et al. (2018) in Beijing, China, through the temperature coefficients in multivariate linear regression models. Generalized Additive Mixed Models (GAMM) were developed by Huang et al. in Beijing, China (Huang et al., 2015), demonstrating that meteorological factors were linked to daily $PM_{2.5}$ concentrations, achieving an R^2 of 0.59.

Studies conducted in Vietnam have utilized the aforementioned methods to investigate the relationships between meteorological factors and air quality, particularly focusing on $PM_{2.5}$ concentrations. Dung et al. (2019) focus on the relationship between hourly meteorological factors (temperature, humidity, and wind speed) in Hanoi and PM_{10} levels in the year 2018. Spearman coefficients are calculated for each variable and an independent T-test is performed to observe differences between the two seasons (dry and rainy). The results indicate that temperature, humidity, and wind speed all exhibit negative correlations with PM_{10} concentrations in both seasons, with wind speed and temperature being the most significant influencing factors compared to humidity. In the study of Tran et al. (2020) for Hanoi, a log-linear regression model was used to examine the relationship between $PM_{2.5}$ concentrations and meteorological factors—air pressure, wind speed, temperature, and surface pressure—using hourly data from 2017 and 2018. The results revealed a positive relationship between air pressure and $PM_{2.5}$ levels, while wind speed, temperature, and surface pressure showed negative trends. In contrast, Ly et al. (2021) construct a random forest model to evaluate the partial effects of individual meteorological factors on $PM_{2.5}$ levels across three sites in Hanoi. Their findings reveal a negative correlation with wind speed and temperature, whereas humidity and atmospheric pressure show positive correlations.

The relationships between $PM_{2.5}$ concentrations and meteorological factors vary across geographical, climatic, and seasonal contexts as shown in the related studies worldwide. However, research on $PM_{2.5}$ in Vietnam remains limited, with most studies focusing on single cities or districts, failing to comprehensively address these relationships. Unlike previous studies focused on local ground data and therefore limited in regional-level data and analysis, this study uses modeled data from multiple sources to examine these relationships at a national scale, offering a broader perspective previously unexplored. This study offers a comprehensive overview of how various meteorological factors interact with $PM_{2.5}$ across the Vietnam, supporting more informed air quality management and climate adaptation.

1.2. Impacts of extreme weather events on $PM_{2.5}$

The most common extreme weather events are typically related to temperature and rainfall factors. These events are identified using percentiles based on historical data, allowing for the classification of weather phenomena that exceed normal levels over a specified time period and specified location (Sheridan et al., 2020), (Jiang et al., 2016), (Hong & Ying, 2018). Other meteorological factors such as relative humidity and wind speed, which are also extreme events, play a significant role in affecting air quality (Zhang et al., 2017). The study by Henian Zhang et al. for the United States (Zhang et al., 2017) identified extreme events using the 5th and 95th percentiles based on a 21-day window of data over 30 years from 1980 to 2009. The results showed that annual extreme ozone and $PM_{2.5}$ days in the eastern U.S. strongly correlated with maximum temperature (Tmax), minimum relative humidity (Rhmin), and minimum wind speed (Vmin). Specifically, the number of annual extreme $PM_{2.5}$ pollution days positively correlated with extreme Rhmin and Tmax days at 0.722 and 0.559, respectively. In contrast, the number of extreme $PM_{2.5}$ days negatively correlated with extreme Vmin days at -0.504, indicating that as extreme Vmin days increased, extreme $PM_{2.5}$ days tended to decrease, especially in rural areas. In the study by Wang et al. in China (W. Wang et al., 2019), it was found that between 1990 and 2014, severe droughts were linked to an average increase in surface ozone and $PM_{2.5}$ during the rainy season (from March to October) by 3.5 ppm (8%) and 1.6 $\mu g/m^3$ (17%), respectively. The variations in ozone and $PM_{2.5}$ showed a negative spatial correlation with the SPEI (Standardized rainfall Evapotranspiration Index) for drought, meaning they were proportional to drought severity (as indicated by a negative SPEI).

This study should be the first investigating deeply on the relationships of extreme events and $PM_{2.5}$ levels in Vietnam. In comparison with similar studies in the world, our study considered a larger number of extreme events for weather factors, including minimum, and maximum temperature, rainfall, humidity, pressure, and wind speed. Besides, quantification impacts utilizing the TWFE and DID models is firstly applied on this issue.

2. Study area and data

2.1. Study area

Vietnam is a country located in the eastern part of the Indochinese Peninsula, near the center of Southeast Asia, bordered by the East Sea and the Gulf of Tonkin, with a land area of approximately 331,344 km^2. The country has an S-shaped form, stretching around 1,650 km from north to south. Vietnam shares borders with China, Laos, and Cambodia, with a total border length of up to 4,639 km. It is a nation with diverse natural features and is undergoing rapid economic and social development.

With a population of approximately 99.47 million, according to the 2022 census (General Statistics Office, 2022), Vietnam is one of the fastest-growing economies in Southeast Asia, with a steady GDP growth rate. The country has seen a significant shift in its economic structure, with the industry and service sectors taking a larger share, alongside the continued significance of agriculture, particularly for ensuring food security. Vietnam is divided into six socio-economic regions according to Decree 92/2006/ND-CP issued by the Government in 2006, including the Northern Mountainous and Midlands (NMM), the Red River Delta (RRD), the North Central and Central Coastal Region (NC&CC), the Central Highlands (CH), the Southeast Region (SE), and the Mekong River Delta (MRD), as illustrated in Figure 1.

Geographically, Vietnam is predominantly mountainous, covering about three-quarters of the total area, while plains account for one-quarter. This complex topography, shaped by Neo-tectonic activities, creates a stepped elevation from the northwest to the southeast. Arable land for agriculture makes up less than 20% of the total area. Regarding climate, Vietnam lies within the tropical and subtropical climate zone, with monsoons, abundant sunlight, and plentiful rainfall. The climate varies depending on the changes in topography from north to south. The northern region experiences four distinct seasons, with cold and dry winters from November to March due to the Northeast monsoon. The summer, from May to September, is humid and influenced by the Southwest and Southeast monsoons. Meanwhile, the southern region has two main seasons: the rainy season from May to November, characterized by Southwest monsoons, and the dry season from December to April, affected by Northeast monsoons. Climate variations from north to south significantly influence the environment and $PM_{2.5}$ pollution in different regions.

Figure 1. Study area and six socio-economic regions

Source: https://thuvienphapluat.vn/hoi-dap-phap-luat/83A4D2B-hd-ban-do-6-vung-kinh-te-viet-nam-chi-tiet-nhat-nam-2024.html

2.2. Data

In analyzing the relationship between $PM_{2.5}$ and meteorological factors, the primary data utilized includes $PM_{2.5}$ data and meteorological data. This data encompasses both spatial and temporal variations of $PM_{2.5}$, including temperature, rainfall, relative humidity, surface pressure, and wind speed.

2.2.1. $PM_{2.5}$

To have information on $PM_{2.5}$ levels, we utilize data from the study by Ngo et al. (2023), which provides daily $PM_{2.5}$ concentration maps for Vietnam with a spatial resolution of 3x3 km from 2012 to 2020 (the study has been updated to include data through 2022). The research employed a Mixed Effect Model on a dataset that includes $PM_{2.5}$ measurements from monitoring stations, satellite AOD (Aerosol Optical Depth) data, and meteorological, and land-use data. The AOD products used in the study include MOD04 3K and MYD04 3K, which are AOD products at 550 nm from Aqua MODIS, and Terra MODIS (Collection 6.1, Level 2), combined with VIIRS AOD data. MODIS is a key instrument aboard two satellites: Terra and Aqua. MOD04 3K, derived from the Terra satellite, provides AOD measurements at a spatial resolution of 3 km, primarily for terrestrial applications. In contrast, MYD04 3K, generated by

Aqua, delivers similar data with a focus on oceanic and atmospheric processes, also at 3 km resolution. VIIRS (Visible Infrared Imaging Radiometer Suite), on the other hand, is another instrument within the Suomi-NPP (National Polar-orbiting Partnership) satellite system. Its AOD data provides high-quality atmospheric data and enhances the accuracy of $PM_{2.5}$ estimations. The daily average $PM_{2.5}$ maps generated through this method show high accuracy, with a Pearson correlation coefficient (R) of 0.87, R^2 of 0.75, RMSE (Root Mean Square Error) of 11.76 $\mu g/m^3$, and MRE (Mean Relative Error) of 36.57% over 13,886 data samples. Monthly and annual average maps from 2012 to 2020 also demonstrate excellent quality when compared to other global $PM_{2.5}$ products.

2.2.2. Meteorological data

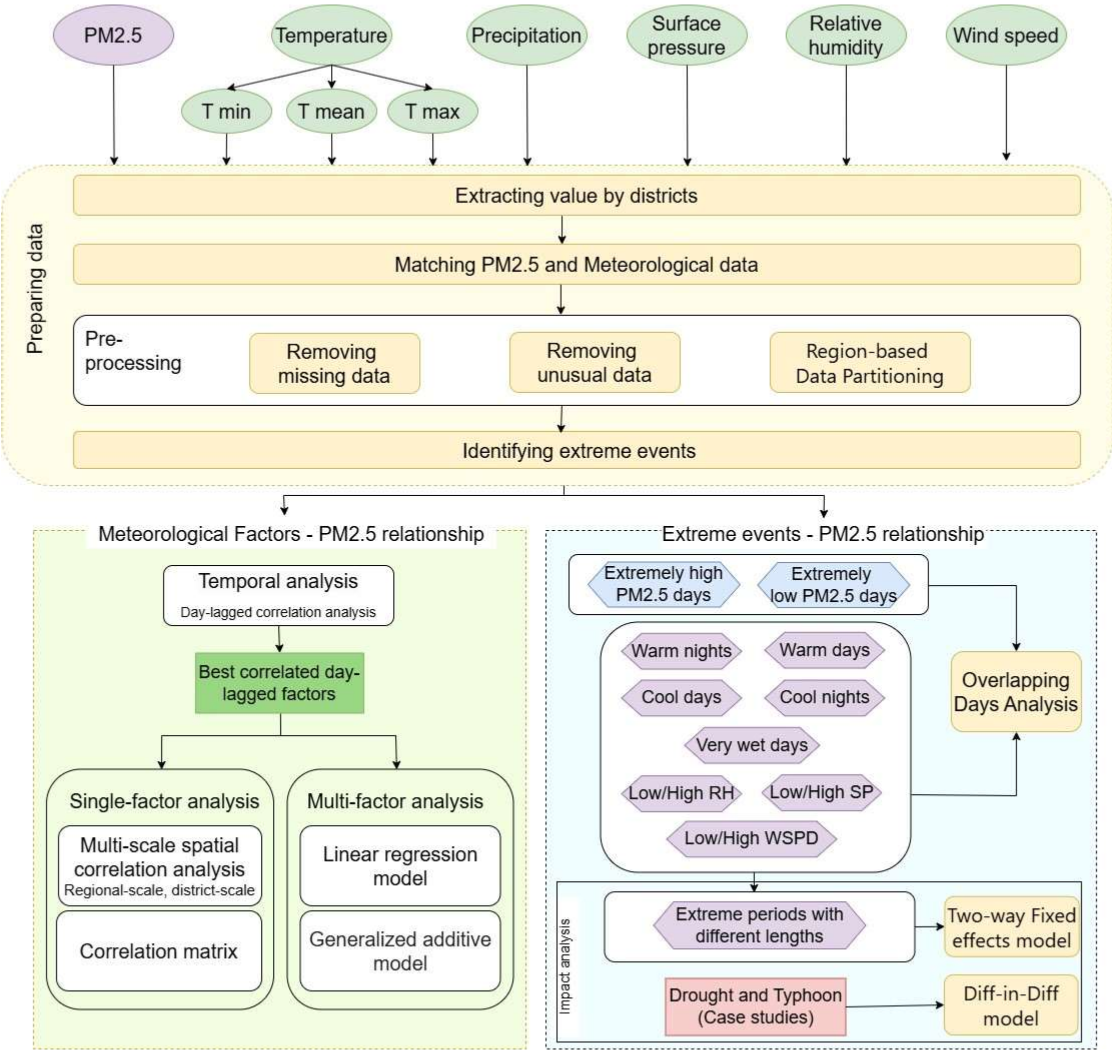
To gather temperature and rainfall data, we utilize average temperature, maximum temperature, and minimum temperature maps for Vietnam from 2012 to 2018 provided by Tran-Anh et al. (2023). Daily observed rainfall and temperature data are collected from 481 and 147 stations across Vietnam, respectively. A three-sigma (and five-sigma) rule was applied to identify suspect values in the data, with each case being re-examined. The observed station data are then interpolated into a gridded dataset at a resolution of 0.1° × 0.1° (referred to as OBS) using Spheremap interpolation for rainfall and Kriging interpolation for temperature. For data from 2019-2022, we utilized ERA-5 Land, which provided hourly that had been aggregated into daily averages.

Other meteorological data are collected from two sources, ERA-5 and ERA-5 Land (https://cds.climate.copernicus.eu). Specifically, u/v wind components from 2012 to 2022 are obtained from ERA-5 Land, while surface pressure and humidity data for the same period are sourced from ERA-5. Both datasets are global weather and climate reanalysis products developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA5 is ECMWF's fifth-generation product, providing data with a spatial resolution of 0.25° x 0.25° (about 25 km) and an hourly temporal resolution, offering a detailed and comprehensive view of global weather and climate. ERA5-Land is an extended version of ERA5, reanalyzed at a higher spatial resolution of 0.1° x 0.1° (about 10 km), focusing on land-related features. These data are also aggregated into daily averages on an hourly basis. All meteorological variables obtained from ERA-5 and ERA-5 Land are measured at a height of 2 m, reflecting the climatic conditions close to the surface.

3. Methodology

The research process and analysis are illustrated in Figure 2, comprising three main phases. First, after collecting $PM_{2.5}$ and meteorological data, we enter the data preparation phase, which includes extracting values by district, integration, preprocessing and identifying extreme events. Following this, we conduct the analysis, focusing on two primary relationships: the relationship between meteorological factors and $PM_{2.5}$ concentrations, and the relationship between extreme weather events and $PM_{2.5}$ levels.

Figure 2: Method for analyzing the relationship between $PM_{2.5}$ and meteorological factors, extreme weather events



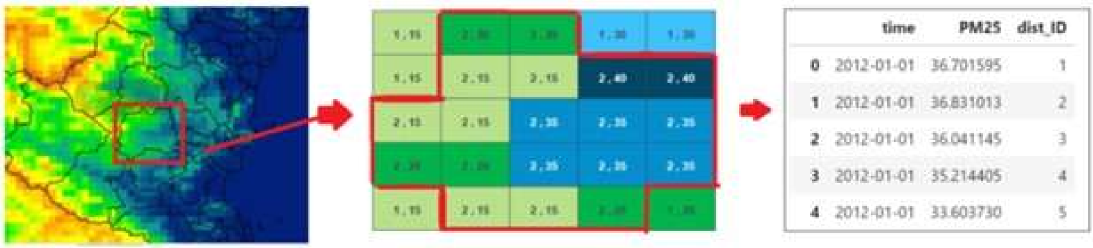
Source: Authors' own calculations. Original.

3.1. Data preparation

3.1.1. District-level data extraction

The map data will be reprojected and resampled to ensure the images have a common grid. Specifically, the grid covers the entire territory of Vietnam using the EPSG:32648 - WGS 84 / UTM zone 48N geographic coordinate system, with each cell measuring 100x100 meters. Each pixel is then assigned an ID corresponding to the district ID through the district boundary map. The pixel data is grouped by district ID, and then we compute the average value by district and day. Finally, the aggregated results are exported to a CSV file, including columns for observation time, district ID, and the average values of $PM_{2.5}$ or other meteorological data. This process is illustrated in Figure 3.

Figure 3: Illustrates the process of data extraction by districts



Source: Authors' own calculations. Original.

3.1.2. Data integration

To construct the analysis dataset, $PM_{2.5}$ and meteorological data are integrated based on temporal (time) and spatial (dist ID) constraints. The integrated dataset is stored in CSV format and includes the following fields (Table 1).

Table 1: Meteorological factors and PM2.5 concentration in dataset

PARAM	DESCRIPTION
TP	Daily total rainfall (mm)
T2M	Daily Average Temperature (oC)
T2M MAX	Daily Maximum Temperature (oC)
T2M MIN	Daily Minimum Temperature (oC)
WSPD	Wind speed (daily average) (m/s)
SP	Surface pressure (daily average) (hPa)

RH	Relative Humidity (daily average) (%)

Source: Authors' own calculations. Original.

3.1.3. Data preprocessing

In the data pre-processing step, first, we remove missing data and days without data. Additionally, outlier data will be reviewed and removed to avoid affecting the accuracy and reliability of the data and analysis results. Specifically, days with abnormal temperatures, such as Tmin>Tmean or Tmean>Tmax, will be excluded. Similarly, days with abnormal rainfall (tp<0), unreasonable humidity (rh>100 or <0), and invalid pressure (sp<0) will also be removed. Finally, to better understand pollution patterns and influencing factors in each region, the dataset will be divided into six socio-economic regions: Northern Midlands and Mountainous Areas, Red River Delta, North Central and Central Coastal Areas, Central Highlands, Southeast, and Mekong River Delta. This approach enhances analytical capabilities and allows for more accurate conclusions about air pollution in each region, rather than generalizing across the whole country, as each region has its unique climate characteristics.

In total, there are 42,959 missing samples (1.506% of the dataset), along with 19,767 instances of abnormal data. After cleaning, we retain 2,790,054 samples, which is 97.801% of the original dataset.

The statistics of the input variables after preprocessing are presented in Table 2.

Table 2: Descriptive statistics of input data

	T2M	T2M MAX	T2M MIN	WSPD	TP	RH	SP	$PM_{2.5}$
NO. SAMPLES	2,790,054	2,790,054	2,790,054	2,790,054	2,790,054	2,790,054	2,790,054	2,790,054
MEAN	24.50	28.72	21.77	1.73	5.27	79.93	98555.89	19.18
STD	4.50	5.08	4.47	1.21	9.50	9.68	3269.70	13.53
MIN	0.44	0.76	-0.97	0.00	0.00	27.42	86243.51	2.55
MAX	35.75	55.80	31.91	13.50	530.87	99.99	103829.19	207.28

Source: Authors' own calculations. Original.

The $PM_{2.5}$ concentration and rainfall values indicate that the maximum values are significantly higher than the average, which may point to certain pollution events and unusual climatic conditions. In contrast, the pressure, humidity, and temperature show fluctuations within a relatively stable range, consistent with normal meteorological conditions.

3.1.4. Identifying extreme events

According to the IPCC Data Distribution Center (DDC), an extreme weather event is defined as "An event that is rare at a particular place and time of year. Definitions of rare vary, but an extreme weather event would normally be as rare as or rarer than the 10th or 90th percentile of a probability density function estimated from observations". (Edenhofer et al., 2013)

In this study, extreme weather events related to meteorological factors, including temperature, rain- fall, humidity, pressure, and wind speed, are analyzed for their relationship with air quality. These extreme indices are constructed based on a set of 27 core climate indices established by the World Meteorological Organization's (WMO) Commission for Climatology (CCI) and the Expert Team on Climate Change Detection and Indices (ETCCDI) of the CLIVAR Project. We identified extreme events for both $PM_{2.5}$ and weather conditions using the 95th or 5th percentile threshold. Specifically, extreme rainfall is defined as rainfall values exceeding the 95th percentile of all events with more than 1 mm over the study period for the entire study area. For other meteorological events and $PM_{2.5}$, extremes are identified using a 5-day window centered on each calendar day over 11 years and for specific districts. This approach allows for a detailed temporal and spatial analysis of extreme events and their potential impact on $PM_{2.5}$ concentrations. The day is considered extreme if its value exceeds the 95th percentile or falls below the 5th percentile of historical data. A detailed description of these extreme events can be in Table 3.

Table 3: Description of extreme weather events

METEOROLOGICAL	INDICES	DESCRIPTION
TEMPERATURE	TNi < TNin5 (Cool nights)	Let TNi be the daily minimum temperature on day i and let TNin5 be the calendar day 5th percentile centered on a 5-day window for the base period 2012-2022.
	Tni > TNin95 (Warm nights)	Let TNi be the daily minimum temperature on day i and let TNin95 be the calendar day 95th percentile centered on a 5-day window for the base period 2012-2022.
	Txi < TXin5 (Cool days)	Let TXi be the daily maximum temperature on day i and let TXin5 be the calendar day 5th percentile centered on a 5-day window for the base period 2012-2022.

	Txi > TXin95 (Warm days)	Let TXi be the daily maximum temperature on day i and let TXin95 be the calendar day 95th percentile centered on a 5-day window for the base period 2012-2022.
RAINFALL	Rri > R95p (Very wet days)	Let RRi be the daily rainfall amount on a wet day i (RR ≥ 1.0mm) and let R95p be the 95th percentile of rainfall on wet days in the 2012-2022 period
RELATIVE HUMIDITY	Rhi $<$ RHin5	Let RHi be the daily relative humidity on day i and let RHin5 be the calendar day 5th percentile centered on a 5-day window for the base period 2012-2022.
	Rhi $>$ RHin95	Let RHi be the daily relative humidity on day i and let RHin95 be the calendar day 95th percentile centered on a 5-day window for the base period 2012-2022.
SURFACE PRESSURE	Spi $<$ SPin5	Let SPi be the daily pressure on day i and let SPin5 be the calendar day 5th percentile centered on a 5-day window for the base period 2012-2022.
	Spi $>$ SPin95	Let SPi be the daily pressure on day i and let SPin95 be the calendar day 95th percentile centered on a 5-day window for the base period 2012-2022.
WIND SPEED	WSPD $<$ WSPDin5	Let WSPDi be the daily wind speed on day i and let WSPDin5 be the calendar day 5th percentile centered on a 5-day window for the base period 2012-2022.
	WSPDi $>$ WSPDin95	Let WSPDi be the daily wind speed on day i and let WSPDin95 be the calendar day 95th percentile centered on a 5-day window for the base period 2012-2022.
$PM_{2.5}$	$PM_{2.5}$i $<$ $PM_{2.5}$i n5	Let $PM_{2.5}$i be the daily $PM_{2.5}$ concentration on day i and let $PM_{2.5}$i n5 be the calendar day 5th percentile centered on a 5-day window for the base period 2012-2022.

	$PM_{2.5}i > PM_{2.5}i$ n95	Let $PM_{2.5}i$ be the daily $PM_{2.5}$ concentration on day i and let $PM_{2.5}i$ n95 be the calendar day 95th percentile centered on a 5-day window for the base period 2012- 2022.

Source: Authors' own calculations. Original.

After defining the extreme weather events for each region, the number of extreme events is summarized in Table 4. In general, the total number of extreme events across all districts is approximately 5% of the total data samples for that region.

Table 4: Statistical of the Number of Extreme Events in Each Region

	HIGH $PM_{2.5}$	LOW $PM_{2.5}$	COOL NIGHTS	COOL DAYS	WARM NIGHTS	WARM DAYS	VERY WET DAYS	LOW RH	HIGH RH	LOW SP	HIGH SP	LOW WSPD	HIGH WSPD
NMM	29,548	29,134	29,377	29,304	29,229	29,899	16,769	29,546	29,895	29,698	29,007	30,440	29,956
RRD	27,077	26,279	26,564	26,676	26,098	26,883	12,358	26,276	26,869	26,644	26,056	27,429	27,001
NC&CC	35,266	34,804	35,269	35,176	35,558	35,177	19,899	35,142	35,833	35,383	35,246	35,957	35,295
CH	12,981	12,850	12,928	13,088	13,292	13,104	7,303	13,118	13,137	13,600	13,256	13,082	13,033
SE	14,841	14,692	15,180	15,378	15,435	14,990	8,688	14,824	15,261	15,268	15,650	15,264	15,149
MRD	27,709	27,415	28,290	28,417	28,254	28,512	16,965	27,855	28,038	28,378	28,793	27,911	27,390

Source: Authors' own calculations. Original.

3.2. Analysis of the Relationship Between Meteorological Factors and $PM_{2.5}$

3.2.1. Analysis of Cumalative Effects

In each economic region, $PM_{2.5}$ is compared with meteorological parameters at different time lags (current $PM_{2.5}$ compared to current meteorological data and 1 day, 2 days, up to 15 days prior). The aim is to determine whether the relationship between $PM_{2.5}$ and meteorological factors is influenced by temporal factors (lags). The correlation is assessed using Pearson R (linear correlation) and Spearman R (non-linear correlation). After identifying the relationship between $PM_{2.5}$ and meteorological factors with respect to lags, for each economic region, the optimal lags for each meteorological variable will be averaged to create a dataset for that v variable, which will then be used in subsequent univariate and multivariate analyses.

3.2.2. Univariate Analysis

This involves constructing Pearson R and Spearman R to identify the relationships between $PM_{2.5}$ and meteorological factors. Correlation coefficients for each district are also mapped to examine spatial variations in correlation.

3.2.3. Multivariate Analysis

This section focuses on developing linear and nonlinear regression models using meteorological parameters as inputs and $PM_{2.5}$ data as the output (target variable). In reality, many meteorological factors have high correlations with one another, influencing the overall variability of $PM_{2.5}$ rather than just individual factors. Therefore, in addition to univariate correlation analysis, we perform analyses using models such as multiple linear regression (LR) and generalized additive models (GAM). The model indicates how well $PM_{2.5}$ variability can be explained based on different meteorological parameter combinations. To avoid multicollinearity issues, we include only one of the three temperature factors that has the highest correlation with $PM_{2.5}$. Additionally, we analyze the importance (weight) of each meteorological parameter in the model. Two models are used including the Linear Regression Model (LR) and the Generalize Additive Model (GAM). Linear regression is a fundamental statistical method used to model the relationship between one dependent variable and one or more independent variables. It assumes a linear relationship, meaning that changes in the independent variables lead to proportional changes in the dependent variable. The Generalized Additive Model (GAM) is a flexible statistical model that extends traditional linear models by allowing for nonlinear relationships between the dependent variable and independent variables. It does so by modeling the relationship as a sum of smooth functions of the predictor variables, rather than assuming a strict linear form.

3.3. Analysis of the Relationship Between Extreme Weather Events and $PM_{2.5}$

After analyzing the relationship between meteorological factors and $PM_{2.5}$ levels, we also perform an analysis to explore how extreme phenomena of these factors impact $PM_{2.5}$ concentrations.

3.3.1. Analysis of Overlap Ratios of Extreme Days

To investigate the relationship between these phenomena, we conduct an analysis of overlap ratios between extreme weather days and extreme $PM_{2.5}$ days. This approach enables us to assess the distribution and frequency of extreme events over time and space, providing deeper insights into how meteorological factors influence air quality.

The overlap ratio is calculated as follows:

$$\text{Overlapratio} = \frac{\text{Number of extreme PM2.5 days overlapping with extreme weather days}}{\text{Total number of extreme weather days}}$$

Additionally, the overlap between current extreme weather events and previously lagged extreme $PM_{2.5}$ events, as well as subsequent led extreme $PM_{2.5}$ events are examined. This offers a more comprehensive view of the relationship between extreme weather events and $PM_{2.5}$ over time.

3.3.2. Analysis impact of Extreme Weather Events on $PM_{2.5}$

In the next phase, we investigate the specific impacts of extreme weather events on $PM_{2.5}$ concentrations, examining whether they lead to increases or decreases, and quantifying the extent of these changes. We evaluate the impacts of extreme weather events of varying lengths, from at least 1 day to at least 8 days, to assess the extent of the influence of long-term extreme events. These events include cool nights, cool days, warm nights, warm days, very wet days, low RH, high RH, low SP, high SP, low WSPD, and high WSPD. For each extreme period, we identify key information such as the start and end dates, duration, and affected districts, and the districts that did not experience extreme events during that period are used as controls. We aim to identify completely unaffected districts during the extreme period. If no suitable districts can be found, we do not consider that extreme event. The control periods are defined as the time before the extreme events occurred, with a duration three times longer than that of the corresponding extreme period. Finally, we concatenate the data for each extreme period into the final dataset for analysis.

Since weather events occur multiple times throughout the study period, the treatment can switch on and off, making the Two-Way Fixed Effects (TWFE) method suitable for assessing the causal impact of extreme weather events. This model is applied to regress "$PM_{2.5}$" concentration on "Event", which represents the occurrence of an extreme event in district i on day t. The regression is performed at the district i-day t level as follows:

$$PM_{2.5i,t} = \alpha + \beta * Event_{i,t} + C(District) + C(time) \quad (2)$$

Where:

- $PM_{2.5i,t}$ is $PM_{2.5}$ concentration of each district i on day t.
- Event is a binary indicator that takes the value of 1 if district i experiences extreme events on day t and 0 otherwise.
- C(district) and C(time) are group-fixed effects and time-fixed effects respectively. We implement fixed effects for each **district**, which will absorb the impact of any

time-invariant district characteristics, and each **day** to control for unobserved heterogeneity.

- α is the intercept.

Our estimate (coefficient β) assesses the causal impact of the "Event" on $PM_{2.5}$ concentrations.

3.3.3. Analysis impact of some specific extreme events on $PM_{2.5}$

In this subsection, we estimate the impacts of two different extreme meteorological events a drought and a typhoon on $PM_{2.5}$ concentrations. We perform two additional analyses, following the same methodology as above, using two specific events that occurred in Vietnam: the drought in southern Vietnam in 2016 (Vietnam News Agency, 2016) and Typhoon Talas in central Vietnam in 2017 (Vietnam News Agency, 2017).

The **drought** in Southern Vietnam in 2016 was one of the most severe droughts in the past 100 years, occurring in the Mekong Delta in March 2016. The dataset also shows that March 2016 had the lowest average rainfall for the Mekong Delta in the dataset from 2012 to 2022, with an average rainfall of 0.048 mm. This event lasted from March 1 to March 31, 2016, while the comparison control period was set from February 1 to February 29, 2016. The Red River Delta, which was not affected by this drought, served as the control region.

The **typhoon**, known as Typhoon Talas in Vietnam, was a tropical storm that affected the country in mid-July 2017. The storm made landfall in Nghe An, Vietnam at around 1 AM on July 17. According to the National Center for Hydro-Meteorological Forecasting of Vietnam, strong winds reached 100 km/h (62 mph), resulting in significant damage to property and communities in the provinces of Nghe An, Thanh Hoa, Ha Tinh, and Quang Binh. For the analysis, we identify the duration of the event that spanned from July 16 to July 31, 2017, while the control period for comparison was set from July 1 to July 15, 2017. The areas most affected by the typhoon included Nghe An, Thanh Hoa, Ha Tinh, and Quang Binh, where strong winds and heavy rainfall caused substantial damage. The control areas are selected as Vinh Phuc, Thai Binh, Bac Ninh, and Hanoi, which were not impacted by the typhoon.

This setting allows us to apply a DID method, comparing outcomes from districts affected by the "event" defined as the extreme event ("affected" districts), with outcomes from districts unaffected by the event ("control" districts), before and after the event occurred.

The equation regression is defined as follows:

$$PM_{2.5} = \alpha + \beta 1 * Event + \beta 2 * Post + \beta 3 * Event * Post \quad (3)$$

Where

- $PM_{2.5}$ is daily $PM_{2.5}$ concentration of each district.
- The Event variable is a binary indicator that denotes whether a district is impacted by the event. It takes the value of 1 for districts that are affected and 0 for those that are not.
- Post is a binary variable indicating the timing of the event: a value of 1 is assigned to data collected during the event, while a value of 0 is assigned to data collected before the event.
- Event*Post is calculated as the product of the Event and Post variables. This means that it is set to 1 only when both conditions are met specifically, when the data belongs to the treatment group and was collected during the intervention period. In all other cases, this value is set to 0.
- Intercept α represents the average $PM_{2.5}$ concentration before the extreme event for the control district group.
- $\beta 1$ denotes the difference in $PM_{2.5}$ concentration between the extreme and control district groups before the extreme event.
- $\beta 2$ indicates the difference in $PM_{2.5}$ concentration for the control district group before and during the extreme event.
- $\beta 3$ reflects the difference in $PM_{2.5}$ concentration before and during the extreme event for the extreme district group compared to the difference for the control group. This is the core result of the DID model, measuring the actual impact of extreme weather events on $PM_{2.5}$ fluctuations.

4. Results

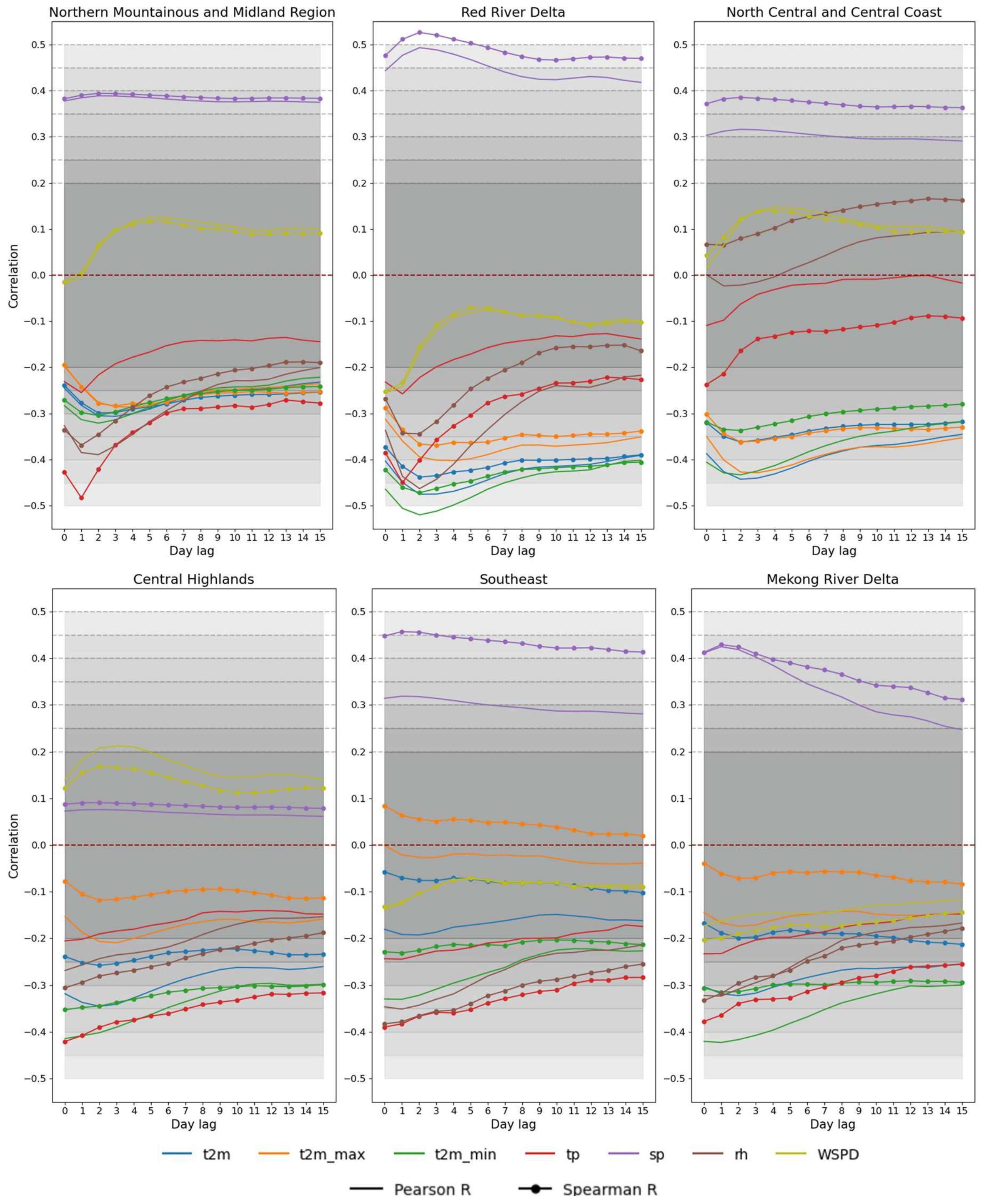
4.1. Analysis of the Relationship Between Meteorological Factors and $PM_{2.5}$

4.1.1. Analysis of Cumulative Effects

After calculating the correlation of different lags of each meteorological variable with $PM_{2.5}$, we categorize the results into eight ranges of correlation levels, as visualized in Figure 4.

The influence of different lags of meteorological factors demonstrates the different response times of $PM_{2.5}$ to changes in these factors. Some factors may have an immediate impact, while others may require a longer duration to affect $PM_{2.5}$ concentrations. Furthermore, the variation in the influence of each factor across regions highlights spatial effects on $PM_{2.5}$.

Figure 4: Correlation between day-lagged meteorological factors and $PM_{2.5}$ in different regions. The correlations are categorized into eight groups: very strong correlation for $|R|>0.5$, strong correlation for $0.45<|R|<0.5$, moderate to strong for $0.4<|R|<0.45$, and moderate for $0.35 <|R|<0.4$. Additionally, we classified weak to moderate correlation for $0.3<|R|<0.35$, weak correlation for $0.25<|R|<0.3$, very weak correlation for $0.2<|R|<0.25$, and negligible correlation for $|R|<0.2$



Source: Authors' own calculation. Original.

In the NMM regions, rainfall from the previous day shows the strongest correlation with $PM_{2.5}$ concentrations, indicated by a Spearman correlation coefficient of 0.45 to 0.5. Surface pressure exhibits no significant variation across different lags, with both the current day and data from 15 days prior yielding similar correlations range ($0.35<|R|<0.4$) with current $PM_{2.5}$ levels. Humidity from one to three days prior has the most substantial impact on $PM_{2.5}$ concentrations. Following this, temperature influences are observed, with the minimum temperatures affecting $PM_{2.5}$ levels primarily from one to four days prior, while the maximum temperatures begin to have significant effects from two days prior and continue through earlier days. Average temperature, as measured by Pearson R, indicates a strong influence from two to four days prior, whereas Spearman R suggests relevant impacts from one to 15 days prior. Wind speed shows minimal influence, with all lags exhibiting low correlations ($|R|<0.2$).

In the RRD regions, minimum temperature and pressure are two factors that significantly impact on $PM_{2.5}$ concentrations ($|R|>0.5$). Specifically, pressure from one to five days prior has the greatest influence, while minimum temperature affects concentrations primarily from one to three days prior. Following these, other factors such as average temperature show optimal lag effects of two to five days prior, relative humidity has a two-day lag, maximum temperature affects concentrations from three to four days prior, and rainfall impacts from one to two days prior. Similar to the NMM regions, wind speed has the least influence on $PM_{2.5}$ concentrations ($0.25<|R|<0.3$). The results indicate that wind speed at lag 0 has the highest correlation, meaning that the wind speed on the same day directly affects $PM_{2.5}$ concentrations.

In the NC&CC regions, temperature and pressure are the two most significant factors influencing $PM_{2.5}$ concentrations. The average temperature from one to six days prior has the largest impact, while the optimal lag of maximum temperature is one to five days prior. The minimum temperature from the current day to four days prior significantly affects $PM_{2.5}$ levels ($0.4<|R|<0.45$). Pressure displays minimal variation across lags, with all showing Spearman R between 0.35 and 0.4. Other variables such as rainfall, humidity, and wind speed demonstrate weaker effects on $PM_{2.5}$ ($|R|<0.2$).

In the CH regions, the minimum temperature and rainfall significantly influence $PM_{2.5}$ concentrations, with optimal lags of 0 to 2 days and 0 to 1 day, respectively ($0.4 <|R|<0.45$). Additionally, average temperature impact $PM_{2.5}$ levels across a broader range of time lags, from 0 to 5 days prior. Relative humidity has an immediate effect on $PM_{2.5}$, with an optimal lag of 0 days. On the other hand, maximum temperature, wind speed, and pressure have minimal effects. Specifically, maximum temperature (t2m max) shows some influence at lags of 2 to 3 days, while wind speed (wspd) affects $PM_{2.5}$ concentrations at lags of 2 to 4 days, both with correlations between 0.2 and 0.25. Pressure has a negligible effect, with all Pearson R and Spearman R values less than 2.

In the SE regions, among these meteorological factors, pressure has the strongest influence on $PM_{2.5}$ concentrations, with optimal lags of 1 to 2 days ($0.5>|R|>0.45$). Relative humidity affects $PM_{2.5}$ levels with optimal lags ranging from 0 to 4 days, while rainfall shows an impact with lags of 0 to 5 days. Minimum temperature influences $PM_{2.5}$ levels with optimal lags of 0 to 3 days. In contrast, average temperature, maximum temperature, and wind speed do not have significant effects, with |R| values less than 0.2.

In the MRD regions, minimum temperature and pressure are the most influential factors, with optimal lags of 0 to 3 days and a correlation range of $0.4 < |R| < 0.45$. Rainfall on the current day and the previous day affects $PM_{2.5}$ levels the most. Average temperature shows an influence with lags from 0 to 4 days, while relative humidity impacts $PM_{2.5}$ with lags of 0 to 2 days. In contrast, maximum temperature and wind speed have negligible effects. In general, meteorological factors in the southern regions tend to influence $PM_{2.5}$ concentrations more rapidly compared to the northern regions. Rainfall and relative humidity in the south show immediate effects, while in the north, their impacts become evident after a delay of 1-2 days. Similarly, temperature in the south affects $PM_{2.5}$ levels within 0-2 days, whereas in the north, this influence is observed after 2-4 days. Importantly, minimum temperature impacts $PM_{2.5}$ concentrations more rapidly than both average temperature and maximum temperature.

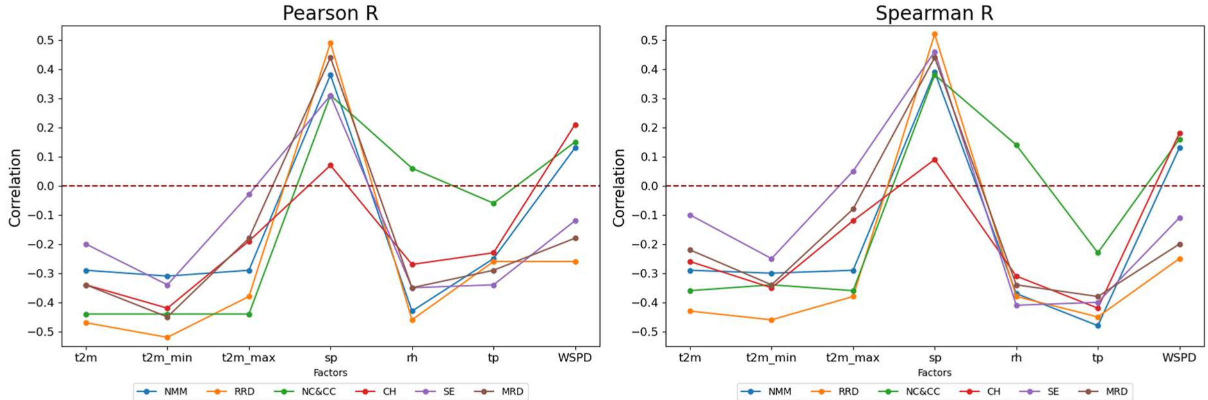
In general, temperature requires more time to impact $PM_{2.5}$ fluctuations than TP and RH. Wind speed has an immediate effect, directly impacting $PM_{2.5}$ concentrations on the same day in the RRD, SE, and MRD regions, as indicated by the highest correlation at lag 0. However, in other regions, the influence of wind speed takes 3-5 days to manifest. Surface pressure exhibits minimal variability across lag times, except in the RRD and MRD regions, where significant effects are seen 1-2 days after changes. These spatial differences indicate that $PM_{2.5}$ concentrations also depend on various factors such as terrain, meteorological conditions, and emission sources.

From there, the data for each meteorological factor will be recalculated using the average of the lags with the highest correlation to continue with the subsequent analysis.

4.1.2. Univariate Analysis

The Pearson correlation coefficient and Spearman rank correlation for meteorological variables and $PM_{2.5}$ concentrations at the regional scale are summarized in Figure 5. To describe the spatial variation more intuitively, we presented the correlation coefficients for each district on the map shown in Figure 6. Analysis of correlation at the district level allows us to capture localized trends and variations that may be obscured when looking at broader regional averages.

Figure 5: Pearson R (Spearman R) correlations between the meteorological factors and $PM_{2.5}$ concentration by region.



Source: Authors' own calculation. Original.

Generally, temperature, rainfall, and humidity are all negatively correlated with $PM_{2.5}$ concentration, except for pressure, which has a positive correlation. The impact of wind speed on $PM_{2.5}$ varies by region.

The average temperature (t2m) has an inverse effect on $PM_{2.5}$ concentrations across all regions. The Pearson R values are higher than the Spearman R values, indicating a stronger linear correlation between these variables and $PM_{2.5}$. This suggests that a decrease in temperature leads to an increase in $PM_{2.5}$ concentrations. The RRD and NC&CC regions exhibit the highest correlations, with Pearson R values of -0.47 and -0.44 respectively, followed by the CH and MRD regions. Lastly, the NMM and SE regions show the lowest correlations. (Figure 6a)

The minimum temperature factor (t2m min) also shows a negative impact across all regions, with |R| values greater than 0.3. Similar to the average temperature, the Pearson R values are also higher than the Spearman R values. The highest correlation is observed in the RRD region (-0.52), followed by MRD, NC&CC, and CH (ranging from -0.45 to -0.42). Lastly, the SE and NMM regions exhibit the lowest correlations, with values of -0.34 and -0.31, respectively. (Figure 6b)

The maximum temperature factor (t2m max) generally shows a negative impact across most regions, except for the SE region, which has a positive Spearman value of 0.05; however, this is not significant. See Figure 6c for a clearer visualization, as some districts in the SE region demonstrate a positive relationship, indicated by the red color plot. The Pearson R values are also higher than the Spearman R values, consistent with the average temperature and minimum temperature factors. The region with the highest correlation is NC&CC (-0.44), followed by RRD (-0.38) and NMM (-0.29). The remaining regions exhibit negligible correlations (|R|<0.2).

Among the three temperature factors considered, the minimum temperature has the most significant effect on $PM_{2.5}$ concentrations across all regions. The negative correlation of temperature factors can be explained by the fact that higher temperatures promote air convection, which dilutes and disperses air pollutants. Conversely, lower temperatures may lead to increased emissions from electricity generation and heating activities. This result is supported by several studies, including those by Tran et al. (2020) and Ly et al. (2021) in Hanoi, Vietnam, as well as Li, Feng, and Liang (2017), Rincon et al. (2023), and Yang et al. (2017) for other regions in the world. The positive correlation between temperature and $PM_{2.5}$ is also not uncommon in analyses, as noted in the study by Li, Ma, et al. (2017), T. Chen et al. (2016), J. Wang and Ogawa (2015). In the summer, high temperatures lead to secondary particles being formed through photochemical processes.

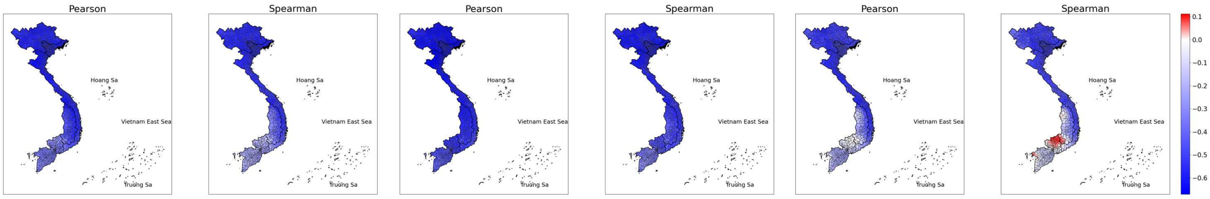
The relationship between pressure (sp) and $PM_{2.5}$ is positive in all regions. When pressure is high, air movement decreases, leading to the accumulation of $PM_{2.5}$ and increased concentrations. The Spearman correlation for pressure is higher than the Pearson correlation in all regions, suggesting a potentially nonlinear relationship with $PM_{2.5}$. However, the impact of pressure on $PM_{2.5}$ varies across regions. In the CH region, pressure is not a significant factor affecting fluctuations in $PM_{2.5}$ concentrations ($R < 0.1$). While, other regions exhibit a clear relationship, with the RRD region showing the highest correlation ($R = 0.52$), followed by SE ($R = 0.46$), MRD ($R = 0.44$), NMM ($R = 0.39$), and NC&CC ($R = 0.38$) (Figure 6d). This finding is similar to results from other studies, such as those by Tran et al. (2020) and Ly et al. (2021) for Hanoi, Vietnam. Yang et al. (2017) and Li, Feng, and Liang (2017) in China.

Rainfall (tp) shows a negative impact across all regions, with the Spearman R values being higher than the Pearson R values. This suggests that rainfall plays a crucial role in cleansing the air by removing dust particles, including $PM_{2.5}$. The degree of correlation is relatively similar across regions, with the NMM and RRD regions having the strongest negative correlations (-0.48 and -0.45, respectively), followed by CH (-0.42), SE (-0.40), MRD (-0.38), and the lowest correlation in NC&CC at -0.23. This negative correlation has been reported in the Nanjing region of China, according to T. Chen et al. (2016), and in Hong Kong, as noted by Li, Feng, and Liang (2017). Nevertheless, when analyzing correlations at a more detailed level, such as by district, the coastal districts in the Central region showed a positive correlation with $PM_{2.5}$ concentrations. (Figure 6e). Most studies have analyzed an inverse relationship, with the washout effect of rainfall on $PM_{2.5}$ concentrations. However, the quantitative impact of light rainfall on $PM_{2.5}$ concentrations is rarely studied. A review of the effects of meteorological conditions on $PM_{2.5}$ concentrations Z. Chen et al. (2020) identified that heavy rainfall has a significantly reduced effect on $PM_{2.5}$ concentrations, while light rain and mist can increase $PM_{2.5}$ concentrations. Light rain leads to increased air humidity and a continuous increase in hygroscopicity, resulting in higher $PM_{2.5}$ concentrations.

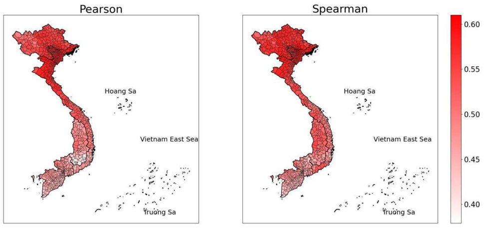
Relative humidity (rh) generally has an inverse impact on $PM_{2.5}$ concentrations in most regions, except for the NC&CC. When humidity increases, $PM_{2.5}$ concentrations tend to decrease because higher moisture levels promote the settling of particulate matter and facilitate the formation of larger droplets that can more easily fall out of the air. The NMM and RRD regions exhibit slightly higher correlations compared to other areas (-0.46 and -0.43, respectively). The three regions: CH, SE, and MRD showed relatively similar correlations, ranging from -0.31 to -0.41. In contrast, the NC&CC region displays a positive correlation (Spearman R = 0.14). When analyzing at the district scale, some districts in the Central Coast region demonstrate a moderate correlation ($|R| \approx 0.4$) (see Figure 6f). Relatively high humidity promotes the partitioning of semi-volatile substances into the aerosol phase. Additionally, humid atmospheric conditions are often associated with a lower boundary layer height, which increases $PM_{2.5}$ concentrations near the ground. This indicates that the influence of these factors on $PM_{2.5}$ concentrations varies by region, geographic location, terrain, and climate. The negative correlation between relative humidity and $PM_{2.5}$ has also been reported in studies by T. Chen et al. (2016) and Li, Feng, and Liang (2017). In contrast, a positive correlation was observed in Shenyang (Li, Ma, et al. ,2017), while in Nagasaki, Japan, humidity's impact varied by season (J. Wang and Ogawa, 2015).

Finally, with wind speed (wspd), there are differences in both the direction and the degree of impact on $PM_{2.5}$ concentration across regions (See Figure 5 and Figure6g). In the NMM, NC&CC, and CH region, wind speed has a positive correlation with $PM_{2.5}$. This may be because these regions have lower pollution levels, and strong winds can blow dust particles from other areas, increasing $PM_{2.5}$ concentration. Conversely, in other regions, wind speed has a negative impact on $PM_{2.5}$ concentration, meaning stronger winds help disperse fine dust particles in the air. In other studies, wind speed also shows a negative effect such as Tran et al. (2020), Ly et al. (2021), Li, Feng, and Liang (2017), T. Chen et al. (2016), and Li, Ma, et al. (2017). However, in (J. Wang and Ogawa (2015)), Munir et al. (2017), wind speed had negative and positive relationships with $PM_{2.5}$.

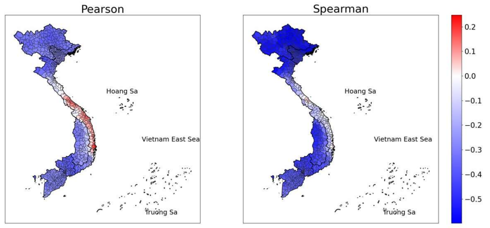
Figure 6: Correlation distribution map of meteorological factors and $PM_{2.5}$ by district



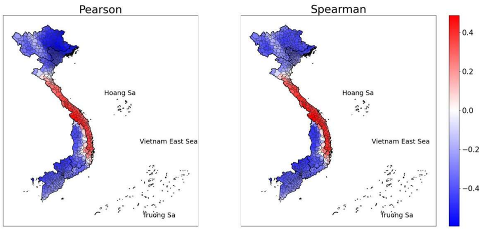
(a) Average temperature (b) Minimum temperature (c) Maximum temperature



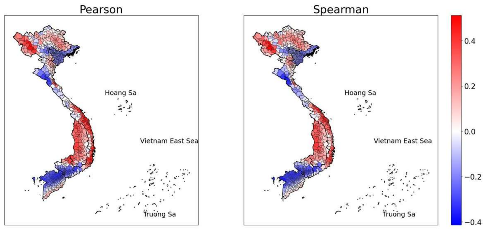
(d) Surface pressure



(e) Total rainfall



(f) Relative humidity



(g) Windspeed

Source: Authors' own calculation. Original.

The effects of meteorological factors on $PM_{2.5}$ concentrations in related studies from Vietnam and other regions in the world are summarized in Table 5.

Table 5: Summary of the Correlations between Meteorological Factors and $PM_{2.5}$ Concentrations Found in Different Study Areas from Previous Studies

STUDY AREA	AUTHOR	TEMPERATURE	SURFACE PRESSURE	RAINFALL	RELATIVE HUMIDITY	WIND SPEED
VIET NAM	This study	both	positive	negative	both	both
NORTHERN MOUNTAINOUS AND MIDLAND REGION, VIETNAM	This study	negative	positive	negative	negative	both
RED RIVER DELTA, VIETNAM	This study	negative	positive	negative	negative	negative
NORTH CENTRAL AND CENTRAL COAST, VIETNAM	This study	negative	positive	both	both	both
CENTRAL HIGHLAND, VIETNAM	This study	both	positive	negative	both	both
SOUTHEAST, VIETNAM	This study	both	positive	negative	negative	both

MEKONG RIVER DELTA, VIETNAM	This study	both	positive	negative	negative	both
HANOI, VIET NAM	Tran et al. (2020)	negative	positive	-	-	negative
	Ly et al. (2021)	negative	positive	-	positive	negative
CHINA	Yang et al. (2017)	negative	positive	-	both	negative (Hainan island positive)
SHENYANG, CHINA	Li, Ma, et al. (2017)	positive	positive	-	positive	negative
HONG KONG	Li, Feng, and Liang (2017)	negative	positive	negative	negative	negative
NANJING, CHINA	T. Chen et al. (2016)	both	-	negative	negative	negative
BEIJING, CHINA	R. Zhao et al. (2018)	both	-	-	both	positive
GUAYAQUILM ECUADOR	Rincon et al. (2023)	both	-	-	negative	positive
NAGASAKI, JAPAN	J. Wang and Ogawa (2015)	positive	-	negative	both	both
UNITED STATES	Tai et al. (2010)	positive	both	negative	both	negative
MAKKAH, SAUDI ARABIA	Munir et al. (2017)	positive	-	-	negative	positive

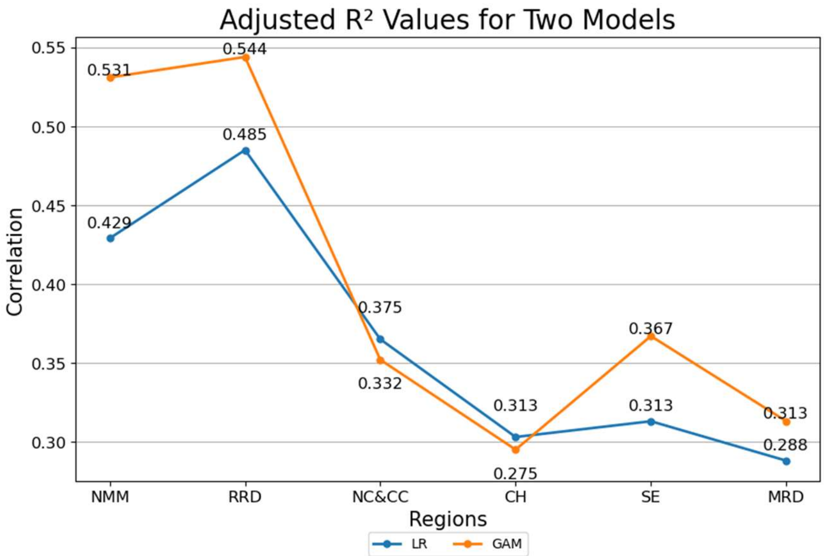
Source: Authors' own calculations. Original.

In terms of the influence of meteorological factors, pressure is the most significant factor, showing a high correlation with $PM_{2.5}$ in 5 out of 6 regions, except for the CH region (|R| ranging from 0.38 to 0.52). Closely following is minimum temperature, which also has a notable impact (|R| ranging from 0.3 to 0.52). Rainfall and relative humidity have a lesser influence, while wind speed has minimal impact across most regions (R <0.2).

4.1.3. Multivariate Analysis

Of the three temperature factors analyzed, the minimum temperature exerts the greatest impact on $PM_{2.5}$ concentrations across all regions, and is therefore included in the model. The variables used in the model are: t2m min, sp, tp, rh and WSPD. The results of the multivariate analysis conducted for each region are summarized through the Adjusted R^2 values of two models presented in Figure 7.

Figure 7: Adjusted R^2 Values for Two Models

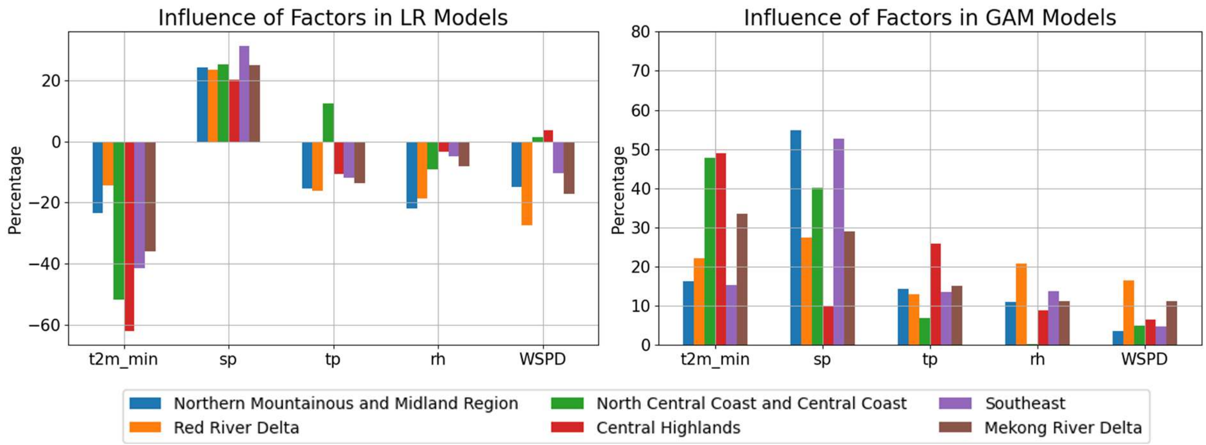


Source: Authors' own calculation. Original.

In most regions, the GAM model provides a better explanation of the relationship between $PM_{2.5}$ concentrations and meteorological factors, significantly improving the Adjusted R^2 compared to the standard linear regression model (with R^2 values ranging from 0.295 to 0.544). This indicates that the relationship between meteorological factors and $PM_{2.5}$ levels is relatively complex and nonlinear. Exceptfor the NC&CC and CH regions, the Adjusted R^2 for the linear regression model is slightly higher than that of the GAM model.

In Northern regions like NMM and RRD, the higher Adjusted R^2 values suggest that meteorological factors significantly influence $PM_{2.5}$ concentrations. In contrast, the other regions in Central and Southern Vietnam show that meteorological variables account for approximately 30% of the variability in $PM_{2.5}$ levels. This implies that, in addition to meteorological factors, other influences, such as emissions from transportation, industry, agriculture, etc, also play a role in determining $PM_{2.5}$ concentrations.

Figure 8: Influence of factors in two models.



Source: Authors' own calculation. Original.

Figure 8 illustrates the strength of each meteorological variable's impact on $PM_{2.5}$ concentrations. The coefficients of the meteorological factors in the LR model were normalized to represent percentages, ensuring that the total contribution of all factors sums to one. The signs of these coefficients indicate the direction of their impact, whether positive or negative. Additionally, the figure illustrates the percentage contributions of each factor to the overall R^2 in GAM models, highlighting the varying degrees of influence that each meteorological variable has on $PM_{2.5}$ levels. All variables exhibit small p-values, approximately 0, in the models, indicating statistical significance.

The direction of influence of the factors in the LR model largely aligns with those observed in the univariate analysis. In the NMM region, all meteorological factors have negative coefficients, except for surface pressure, which is consistent with the univariate analysis; however, wind speed show a positive correlation but has a negative coefficient. In the NC&CC region, rainfall has a positive coefficient while relative humidity has a negative one, contrary to the univariate analysis where rainfall displayed a negative correlation and relative humidity a positive one, although both findings were not significant ($|R|<0.1$). The other factors maintain the same directional influence as indicated in the univariate anal- ysis. In the remaining regions—RRD, CH, SE, and MRD—all factors, including temperature, humidity, rainfall, and wind speed, exhibit signs consistent with the trends observed in the univariate analysis.

In terms of impact, there are differences among the regions; however, minimum temperature and surface pressure consistently show the most significant influence on $PM_{2.5}$ concentrations across nearly all areas. Rainfall and relative humidity have a lesser impact, while wind speed generally shows minimal influence. The results are quite consistent with the univariate analysis; however, some differences exist, as wind speed demonstrates a significant effect on $PM_{2.5}$ concentrations in the LR model.

4.2. Analysis of the Relationship Between Extreme Weather Events and $PM_{2.5}$

4.2.1. Analysis of Overlap Ratios of Extreme Days

Figure 9 visualizes the overlap ratio between extreme weather events and extremely high/low $PM_{2.5}$ concentration across six regions. By visualizing the overlap ratios, we can gain a clearer understanding of how each extreme weather event correlates with high or low $PM_{2.5}$ levels. This analysis revealed the differences among regions and highlighted variations between different types of weather events.

With extremely low minimum temperature events, cool nights exhibit a higher overlap ratio with extremely high $PM_{2.5}$ days in most regions, ranging from 5.99% to 14.72%, compared to 3.68% to 11.74% (Figure 9a). In contrast, in the RRD region, the overlap ratio between cool nights and extremely high $PM_{2.5}$ events did not differ significantly from the ratio with extremely low $PM_{2.5}$ events, with values of 11.34% and 11.74%, respectively. This may indicate a link between cool nights and an increase in $PM_{2.5}$ concentration.

On the other hand, extremely low maximum temperature events, cool days show a significantly higher overlap ratio with extremely low $PM_{2.5}$ events, ranging from 7.39% to 24.97%, compared to extremely high $PM_{2.5}$ events across all regions. Notably, in the two northern provinces, the overlap ratios are particularly high, with the NMM region at 22.89% and the RRD region at 24.97%. (Figure 9b)

The overlap ratio between warm days and both extremely low and high $PM_{2.5}$ events is insignificant, ranging from 1.71% to 6.35%. (Figure 9d)

Similarly, warm nights also showed negligible overlap with both extremely low and high $PM_{2.5}$ events, consistently staying below 7% across all regions. (Figure 9c)

Very wet days showed a significantly higher overlap with extreme low $PM_{2.5}$ events compared to extreme high events, ranging from 9.27% to 20.01%. This may indicate the air-cleansing effect of rain, which is consistent with the results of the negative correlation analysis presented in Section 4.1. In the two northern provinces, the overlap ratios are the highest, with the NMM region at 19.07% and the RRD region at 20.01%. According to the results of the previous analysis, these two regions also demonstrated the strongest correlation between rainfall and $PM_{2.5}$ levels among the six regions. (see Figure 9e).

The difference in the overlap ratio between extremely low relative humidity and extremely high $PM_{2.5}$ events is quite pronounced compared to extremely low $PM_{2.5}$ events across all regions (see Figure 9f). The overlap ratios across the regions are relatively consistent, ranging from 7.33% to 12.48%, with the NMM region showing a slightly higher value (12.48%). This may suggest a connection between low humidity and increased $PM_{2.5}$ concentrations, as dry conditions make it easier for dust particles to remain suspended in the air.

In contrast, extremely high humidity shows a greater overlap with extremely low $PM_{2.5}$ events across all regions, ranging from 6.77% to 12.44%, with the NMM region again having the highest overlap ratio (12.44%). (see Figure 9g)

Extremely low sp events showed a greater overlap with low $PM_{2.5}$ events compared to high $PM_{2.5}$ events, although the difference is not significant. Additionally, the overlap ratio is relatively low across all regions, remaining below 7%. (Figure 9h)

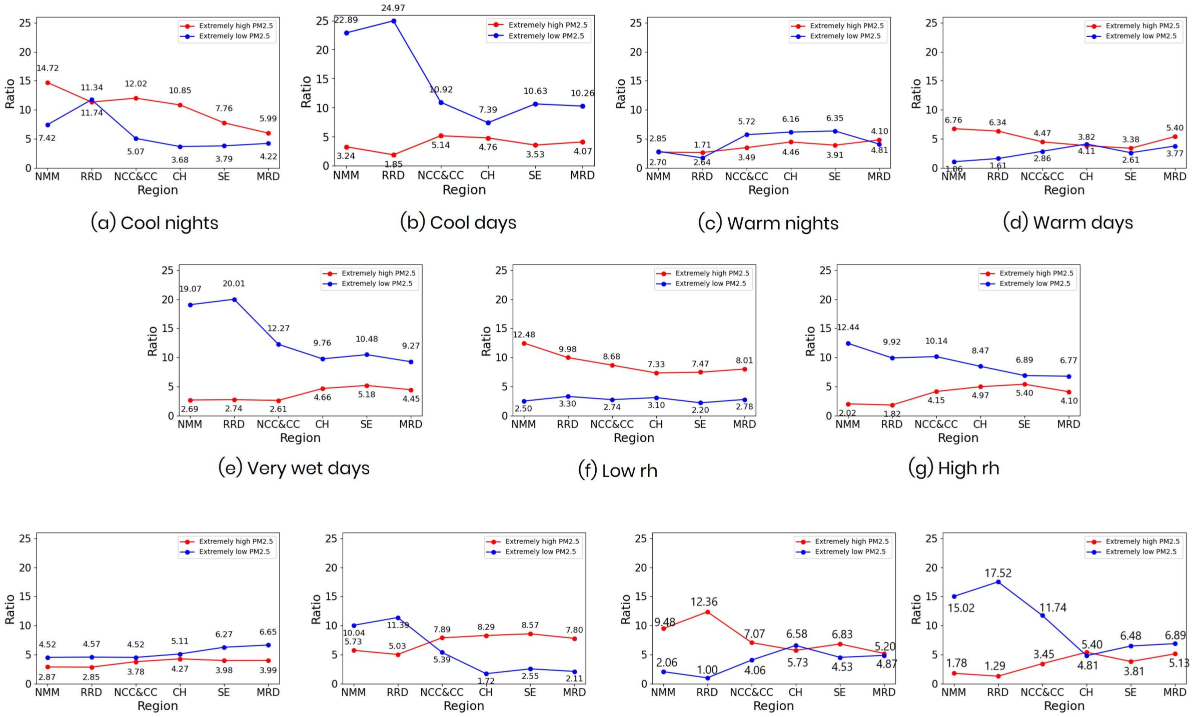
The overlap ratio between high sp and extreme high $PM_{2.5}$ is generally higher compared to extreme low $PM_{2.5}$ across most regions, ranging from 7.8% to 8.57%. However, the exceptions are the NMM and the RRD regions. In these two northern regions, the overlap ratio between high sp and extreme high $PM_{2.5}$ is lower than that with extreme low $PM_{2.5}$, at 10.04% and 11.39% respectively. While these overlap ratios for high sp are higher than the case of low sp, they are also moderate. (Figure 9i)

Regarding extreme wind speed events, the overlap ratio between low WSPD events and extremely high $PM_{2.5}$ is notably higher compared to the ratio with extremely low $PM_{2.5}$ in northern regions. The NMM and RRD regions exhibit the highest overlap ratios at 12.36% and 7.07% respectively. In contrast, the other regions do not show significant differences between the overlap ratios of low WSPD with extremely high $PM_{2.5}$ versus extremely low $PM_{2.5}$, remaining below 7% and not particularly notable. (Figure 9j)

Similarly, the overlap ratio between high WSPD and extreme low $PM_{2.5}$ is significantly higher com- pared to the ratio with extreme high $PM_{2.5}$. In the NMM, RRD, and NC&CC regions, the overlap ratios are 15.02%, 17.52%, and 11.74% respectively for high WSPD and extreme low $PM_{2.5}$. The other regions, including CH, SE, and MRD, do not show significant differences between the overlap ratios of high WSPD with extremely high $PM_{2.5}$ versus extremely low $PM_{2.5}$, ranging from 3.81% to 6.89% and not particularly notable. (Figure 9k)

In general, when comparing different extreme weather events, the overlap ratio for warm nights, warm days, and low sp events is consistently low across all regions, with rates of less than 7%. In contrast, cool days, very wet days and high WSPD exhibit the highest overlap, particularly in the NMM and RRD regions, where rates exceed 15%. This trend suggests that cool, wet and strong wind conditions are more likely to coincide with a significant decrease in $PM_{2.5}$ levels. Other events exhibit moderate overlap ratios ranging from 7% to 15%, including cool nights, low RH, high RH, high SP, and low WSPD. On the other hand, most northern areas show a higher overlap ratio than their southern counterparts. Notably, the CH, SE, and MRD regions have lower overlap rates for all extreme weather events, generally remaining below 10%.

Figure 9: The overlap ratio between extreme weather events and extreme $PM_{2.5}$ levels by region



(a) Cool nights (b) Cool days (c) Warm nights (d) Warm days

(e) Very wet days (f) Low rh (g) High rh

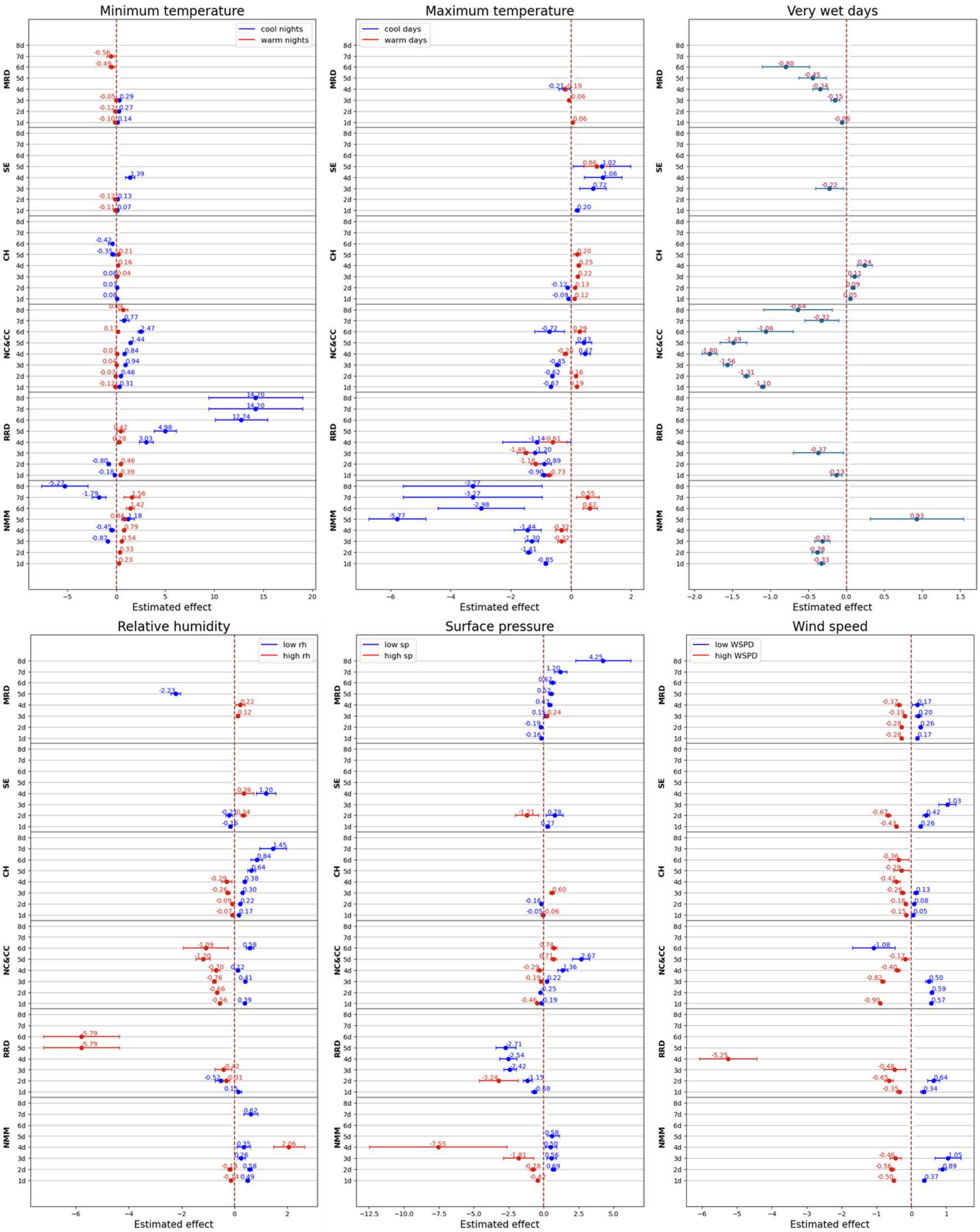
(h) Low sp (i) High sp (j) Low WSPD (k) High WSPD

Source: Authors' own calculation. Original.

4.2.2. Impact of Extreme Weather Events on $PM_{2.5}$

The above overlap ratio results alone do not directly indicate the magnitude or impact of these extreme weather events on $PM_{2.5}$ concentrations. To better understand the specific effects, we quantify the impact through the coefficients of the Event variable in the TWFE model, as summarized in Figure 10. The results in the figure only display statistically significant effects and events that have matched clean control districts.

Figure 10: Impact of Extreme Events with Varying Durations in Two-way Fixed Effect Models



Source: Authors' own calculation. Original.

The results of the TWFE model across varying durations of extreme weather events reveal notable regional disparities in the response of $PM_{2.5}$ concentrations. Regarding minimum temperatures, cool nights generally lead to an increase in $PM_{2.5}$ in most regions, except for the NMM region, where $PM_{2.5}$ decreases significantly, up to 5.27 μg/m³. In contrast, regions such as RRD, NC&CC, and SE experience moderate to strong increases, especially as the duration of extreme events extends. In RRD, short term cool nights (1–2 days) slightly decrease $PM_{2.5}$ (–0.18 to –0.8 μg/m^3), but from 4 days onward, the concentrations rise sharply, reaching as high as 14.20 μg/m^3. The NC&CC and SE regions show increasing trends as well, though to a lesser extent. Meanwhile, the MRD and CH exhibit negligible responses. Warm nights also tend to raise $PM_{2.5}$ levels, though their effects generally remain weaker than those of cool nights. Notably, SE and MRD show slight reductions in $PM_{2.5}$ during warm nights. In most regions, the influence of warm nights does not intensify with longer durations. Only in NMM is there a clear increasing trend, with $PM_{2.5}$ rising from 0.23 μg/m^3 with one-day extremes to 1.56 μg/m^3 for events lasting at least eight days. In contrast, the other regions mostly record changes below 1 μg/m^3.

Regarding maximum temperatures, cool days consistently reduce $PM_{2.5}$, particularly in NMM, where the effect strengthens with longer extreme durations (–0.85 to –5.77 μg/m^3). Other regions such as RRD, NC&CC, CH, and MRD also experience reductions, though to a lesser extent. RRD shows a range of -0.9 to -1.14 μg/m^3, while NC&CC, CH, and MRD exhibit negligible reductions of less than 1 μg/m^3. Conversely, in the SE region, cool days lead to an increase in $PM_{2.5}$, gradually rising with the length of extreme conditions, ranging from 0.2 to 1.02 μg/m^3. Warm days, in contrast, typically cause minor increases in $PM_{2.5}$ across most areas (<1 μg/m^3), except in RRD, where they lead to notable reductions, peaking at –1.49 μg/m^3 for three-day events. In the RRD and SE regions, the effects of cool days and warm days are not distinctly observed. In the RRD, both extreme events of maximum temperature lead to a decrease in $PM_{2.5}$ levels, while in the SE, both events result in an increase in $PM_{2.5}$. The impacts tend to intensify as the duration of the extremes extends.

Extreme rainfall events, represented by very wet days, predominantly act as natural air cleansers, lowering $PM_{2.5}$ concentrations. However, the degree of impact varies by region and event duration. Both RRD and MRD experience gradual decreases, with effects trending upwards as the duration of extreme conditions increases. (RRD: –0.13 to –0.37 μg/m^3; MRD: –0.06 to –0.8 μg/m^3). The NC&CC region shows the most pronounced reductions, from –1.1 to –1.8 μg/m^3 for four-day extremes, before the impact diminishes to about -0.32 μg/m^3. In NMM, short-term wet periods (1–3 days) also contribute to the reduction of $PM_{2.5}$ by around 0.3 μg/m^3, but a five-day event unexpectedly results in a slight increase of 0.93 μg/m^3. In contrast, CH experiences small increases (0.05–0.24 μg/m^3) as wet periods extend which is opposite to the trend observed in the other regions.

Regarding relative humidity, Low RH increases $PM_{2.5}$ levels in the NMM, NC&CC, and CH regions, with CH experiencing the most significant impact, rising from 0.17 to 1.45 μg/m^3 as the duration of extremes increase. The effects in the other two regions are negligible (<1 μg/m^3). In contrast, low RH in the MRD leads to a relative decrease in $PM_{2.5}$, with a significant reduction of -2.23 μg/m^3 during the at least five-day extreme events. The RRD and SE regions do not show a clear trend regarding the impact of low RH. High RH decreases $PM_{2.5}$ levels in RRD, NC&CC, and CH regions, with all regions showing increasing effects as the duration of extremes lengthens. In RRD, high RH significantly reduces $PM_{2.5}$, with the most pronounced decrease ranging from -0.31 to -5.79 μg/m^3. In NMM, high RH for 1–2 days leads to a slight reduction in $PM_{2.5}$ (-0.14 to -0.18 μg/m^3); however, when this condition persists for at least five days, it unexpectedly results in a significant increase of 2.06 μg/m^3. High humidity can cause water to condense on the surfaces of dust particles, making them heavier and causing them to fall to the ground, thereby reducing $PM_{2.5}$ concentrations in the air. However, frequent high humidity can lead to fog formation, which reduces the dispersion of pollutants, causing them to accumulate in the air and worsen pollution levels in that area. In contrast, in SE and MRD regions, high RH tends to increase $PM_{2.5}$ levels, although the effects are not significant and remain unchanged over time (ranging from 0.34 to 0.36 and 0.12 to 0.22 μg/m^3). The NC&CC and CH regions clearly demonstrate the contrasting effects of low and high RH, while the other regions do not show distinct impacts from either condition.

Regarding surface pressure, the impacts of extremes vary widely across regions. High SP tends to lower $PM_{2.5}$ in NMM, RRD, and SE, especially in NMM, where reductions increase sharply with longer durations (–0.4 to –7.55 μg/m^3). For NC&CC, CH, and MRD, the effects remain minor (<1 μg/m^3) and lack clear trends. Low SP, on the other hand, often raises $PM_{2.5}$ in NMM, NC&CC, SE, and MRD, with the extent of the increase becoming more pronounced during prolonged extreme conditions, particularly in NC&CC and MRD where increases reach 2.67 and 4.25 μg/m^3, respectively. In contrast, low sp in the RRD decreases $PM_{2.5}$ levels, with reduction ranging from –0.58 to –2.71 μg/m^3. The NMM, CH, and SE regions show minimal impacts, each below 1 μg/m^3. Notably, the NMM exhibits a clear distinction between the effects of high and low sp, while the other regions do not show as clear a differentiation.

Finally, wind speed extremes exhibit the clearest directional effects. High WSPD consistently reduces $PM_{2.5}$ across all regions, while low WSPD increases it. This reflects the fundamental role of wind in pollutant dispersion: strong winds facilitate dilution, whereas calm conditions promote accumulation. The impact of both extremes intensifies with longer durations in most regions. Low WSPD raises $PM_{2.5}$ from 0.37 to 1.05 μg/m^3 in NMM and from 0.26 to 1.03 μg/m_3 in SE. Conversely, high WSPD reduces $PM_{2.5}$ dramatically in RRD (–0.35 to –5.25 μg/m^3). Notably, NC&CC shows an interesting pattern where the influence of high WSPD diminishes over time (–0.90 to –0.17 μg/m^3), and extended low WSPD events eventually reverse the effect, leading to a reduction of –1.08 μg/m^3.

Overall, the impact of extreme weather events on $PM_{2.5}$ concentrations varies significantly across regions, with some events demonstrating particularly strong effects. In the NMM region, cool nights, cool days, and high SP consistently lead to substantial reductions in $PM_{2.5}$

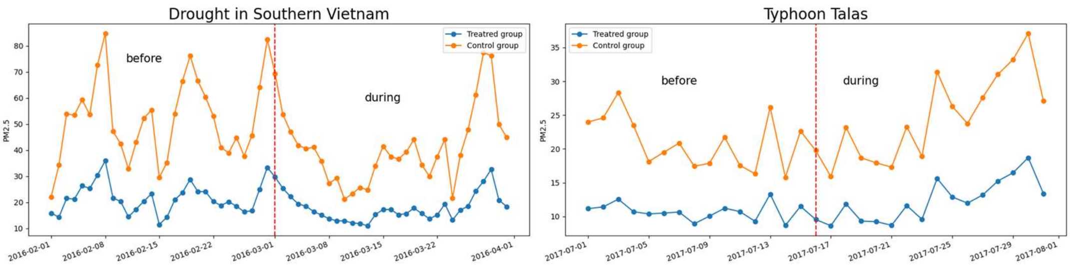
levels. These effects become more pronounced with longer durations of extremes, reaching decreases of approximately –5 to –7 μg/m³. In contrast, in the RRD, extended cool night events (lasting 7–8 days) result in a sharp increase in $PM_{2.5}$, peaking at 14.20 μg/m³. This region also benefits from significant pollution reductions under high RH and high WSPD events, with declines of up to –5 μg/m³. In the MRD, low SP exerts the most notable effect, contributing to an increase in $PM_{2.5}$ concentration of 4.25 μg/m³. Aside from these highlighted cases, the remaining extreme events across other regions tend to have relatively minor impacts on $PM_{2.5}$, typically staying within ±1 μg/m³, suggesting more localized or less sensitive responses.

4.2.3. Impact of specific extreme events on $PM_{2.5}$

Besides identifying the impacts of extreme event ranges on the dataset in the previous section, we analyze the impact of two significant events: drought in Southern Vietnam in 2016 and Typhoon Talas in Central Vietnam in 2017. These events provide valuable insights into how weather conditions can influence air quality. The drought led to severe water shortages, reduced moisture in the air, and increased dust and particulate matter from dry soil. In contrast, Typhoon Talas brought heavy rainfall and strong winds, significantly altering air quality.

In Figure 11, the parallel trend test for the treated and control groups before the intervention shows a relatively similar pattern. Both groups before Southern drought and typhoon Talas exhibit comparable increases and decreases, supported by Pearson correlation coefficients of 0.94 and 0.88, respectively. These results indicate that the $PM_{2.5}$ trends of both groups are consistent in the pre-intervention period, reinforcing the parallel trends hypothesis necessary for the Difference-in-Differences model.

Figure 11: Trends in $PM_{2.5}$ for the treatment and control groups during two specific events.



Source: Authors' own calculations. Original.

Table 6 presents the results of equation (3), which assesses the impacts of a drought and a typhoon separately, on $PM_{2.5}$ concentrations.

Table 6: Summary of DID Model Results for two Extreme Weather events

EVENTS	INTERCEPT	EVENT	POST	EVENT POST
DROUGHT	52.2 ***	-30.65 ***	-10.05 ***	6.45 ***
TYPHOON	20.96 ***	-10.22 ***	3.56 ***	-2.05 ***

Source: Authors' own calculations. Original.

The results indicated that the drought in Southern Vietnam in March 2016 had a negative impact, increasing $PM_{2.5}$ concentrations in the Mekong Delta by 6.45 $\mu g/m^3$. Conversely, Typhoon Talas led to a reduction in $PM_{2.5}$ concentrations by 2.05 $\mu g/m^3$ across the four provinces it affected. These findings align with the impact of extreme meteorological events on $PM_{2.5}$ levels discussed in the previously analyzed section. Both extremely low humidity and rainfall contribute to increased $PM_{2.5}$ concentrations; thus, during droughts, the lack of rainfall and low humidity leads to higher levels of $PM_{2.5}$. In contrast, extremely high wind speeds and rainfall result in lower $PM_{2.5}$ concentrations. During storms, heavy rainfall and strong winds promote the dispersion and washing away of pollutants, effectively decreasing $PM_{2.5}$ levels. These results confirm that extreme weather events can significantly influence air quality, with droughts exacerbating pollution levels while typhoons help to mitigate them.

5. Conclusions

This study highlights the varying impacts of meteorological factors and extreme weather events on $PM_{2.5}$ concentrations across different regions in Vietnam. In the NMM, tp has the greatest impact on $PM_{2.5}$, followed by rh and sp. sp and wspd have positive correlations, though wspd's influence is minimal ($R<0.2$) The others show negative correlations. Together, these factors explain 53.1% of the $PM_{2.5}$ variability in GAM. Extreme weather events, cool nights coincide with high extreme $PM_{2.5}$ levels at 14.72%, while cool days, very wet days, high WSPD, high RH, and high SP link to low extreme $PM_{2.5}$ events (10.% to 22.89%) and decrease $PM_{2.5}$ levels. Cool nights, cool days, and high SP have the strongest impact, especially during prolonged extremes, reducing $PM_{2.5}$ from 5.27 to 7.55 $\mu g/m^3$.

In the RRD, all meteorological factors significantly influence $PM_{2.5}$ concentrations ($0.26<|R|<0.52$). sp and minimum temperature have the strongest effects, and together, all meteorological factors explain 54.4% of the variability in $PM_{2.5}$ levels. All variables exhibit negative correlations with $PM_{2.5}$, except for sp. Extreme weather events in this region typically exhibit the highest overlap ratios compared to other areas. Cool days, very wet days, high WSPD, and high SP associate with low extreme $PM_{2.5}$ events (11.39% - 24.97%). Low WSPD correlates with extreme high $PM_{2.5}$ (12.36%). In terms of quantity, cool nights increase $PM_{2.5}$

concentrations up to 14.20 μg/m^3 during longer-lasting events. The high RH and high WSPD reduce $PM_{2.5}$ up to 5.25 – 5.79 μg/m^3, respectively.

In the NC&CC, temperature had the most substantial negative impact on $PM_{2.5}$ levels (R = -0.44), while sp exhibited a positive correlation (R = 0.38). Both regression models accounted for approximately 35-36% of the variability in $PM_{2.5}$. The overlap ratios of cool nights and extreme high $PM_{2.5}$ levels at 12.02%. Cool days, very wet days, high RH, and high WSPD correlated with low $PM_{2.5}$ levels (10.14%-12.27%). The very wet day only reduces $PM_{2.5}$ at 1.8 μg/m^3, while other extreme weather events have almost no significant impact.

In the CH, $PM_{2.5}$ is negatively correlated with temperature, rh, and tp, positively with sp and wspd. Minimum temperature and rh are the primary factors (R = -0.34 and -0.31, respectively), while sp, only at the region, has a minor effect (|R|<0.2). The LR and GAM models have similar R^2 values of about 30%. Only cool nights overlap with extremely high $PM_{2.5}$ levels at 10.85%. Low RH significantly increases $PM_{2.5}$ up to 1.45 μg/m^3 by extreme lengthen, while other impacts are generally minimal (less than 1 μg/m^3).

In the SE, all variables show negative correlations with $PM_{2.5}$, except for sp, which is the most influential factor (R = 0.46). The GAM explains slightly more variability (36.7%) than the LR model (31.3%). All extreme weather events exhibit less than 10.68% overlapping with extreme $PM_{2.5}$ levels. The impact of extreme events in this region also has minimal effects on air quality.

In the MRD, all variables show negative correlations with $PM_{2.5}$, except sp. Minimum temperature and sp are key factors influencing $PM_{2.5}$ levels (R = -0.45 and 0.4 respectively), with the GAM and LR models showing small differences in explanatory power (31.3% vs 28.8%). All extreme weather events show low overlap with extreme $PM_{2.5}$ levels, less than 10. 26%. Notably, low SP contributes to an increase in $PM_{2.5}$ concentration of 4.25 μg/m^3, while low RH results in a reduction of 2.23 μg/m^3.

Finally, two major events show how extreme weather can impact air quality. The 2016 drought in Southern Vietnam raised $PM_{2.5}$ by 6.45 μg/m^3 in the Mekong Delta, reflecting how dry spells intensify pollution in vulnerable lowland areas. In contrast, Typhoon Talas (2017) reduced $PM_{2.5}$ by 2.05 μg/m^3 in Central Vietnam, likely due to strong winds and rainfall dispersing pollutants. These cases highlight how different extreme events can worsen or improve local air quality depending on their nature and location. However, the findings are subject to methodological and data-related constraints. The quality of meteorological and $PM_{2.5}$ data may affect accuracy. The TWFE model assumes stable district- and time-specific effects, which may oversimplify real-world dynamics. The DID method depends on parallel trends between treatment and control groups, which may not always hold. Also, defining extreme events, selecting data for TWFE and DID models, and the limited number of case studies may limit the generalizability of the results.

The findings of this study provide several recommendations for policy dialogue in Vietnam. Firstly, region specific air quality management policies should be implemented due to significant regional differences in factors influencing $PM_{2.5}$ levels. Meteorological factors have a stronger impact on $PM_{2.5}$ concentrations in northern regions than in southern regions, suggesting that these factors should be considered carefully in air quality management in the northern and interregional collaboration could be effective. Secondly, current and lag-day values of meteorological variables, such as surface pressure, temperature, rainfall, relative humidity, wind speed and extreme event patterns, should be integrated into regional air quality models to enhance their accuracy. Thirdly, the strong link between meteorological factors and air pollution in the Northern regions suggests that future climate change could worsen air pollution levels, with adverse effects on human health and the environment. Air pollution should be integrated into climate change adaptation plans to mitigate these long-term effects. Finally, extreme weather events such as cool nights, low humidity, and low wind speed often coincide with high $PM_{2.5}$ concentration. Prolonged extreme events, like cool nights in the Red River Delta, low surface pressure in the Mekong River Delta, and droughts, can significantly increase $PM_{2.5}$ levels. Strengthening disaster preparedness and resilience is essential to mitigate these combined impacts.

Bibliography

Chen, T., He, J., Lu, X., She, J., & Guan, Z. (2016). Spatial and temporal variations of pm2. 5 and its relation to meteorological factors in the urban area of nanjing, china. *International journal of environmental research and public health*, 13 (9), 921.

Chen, Z., Chen, D., Zhao, C., Kwan, M.-p., Cai, J., Zhuang, Y., . . . others (2020). Influence of meteorological conditions on pm2. 5 concentrations across china: A review of methodology and mechanism. *Environment international* , 139 , 105558.

Dung, N. A., Son, D. H., Tri, D. Q., et al. (2019). Effect of meteorological factors on pm10 concentration in hanoi, vietnam. *Journal of Geoscience and Environment Protection*, 7 (11), 138.

Edenhofer, O., Seyboth, K., & Shogren, J. (2013). Intergovernmental panel on climate change (ipcc). In Encyclopedia of energy, natural resource, and environmental economics (volume 1: Energy) (pp. 48–56). Elsevier.

Hong, Y., & Ying, S. (2018). Characteristics of extreme temperature and precipitation in china in 2017 based on etccdi indices. *Advances in Climate Change Research*, 9 (4), 218–226.

Huang, F., Li, X., Wang, C., Xu, Q., Wang, W., Luo, Y., . . . others (2015). Pm2. 5 spatiotemporal variations and the relationship with meteorological factors during 2013-2014 in beijing, china. *PloS one*, 10 (11), e0141642.

Jiang, C., Mu, X., Wang, F., & Zhao, G. (2016). Analysis of extreme temperature events in the qinling mountains and surrounding area during 1960–2012. *Quaternary international*, 392 , 155–167.

Jones, A. M., Harrison, R. M., & Baker, J. (2010). The wind speed dependence of the concentrations of airborne particulate matter and nox. *Atmospheric Environment*, 44 (13), 1682–1690.

Li, X., Feng, Y., & Liang, H. (2017). The impact of meteorological factors on pm2. 5 variations in hong kong. In Iop conference series: *Earth and environmental science* (Vol. 78, p. 012003).

Li, X., Ma, Y., Wang, Y., Liu, N., & Hong, Y. (2017). Temporal and spatial analyses of particulate matter (pm10 and pm2. 5) and its relationship with meteorological parameters over an urban city in northeast china. *Atmospheric research*, 198 , 185–193.

Ly, B.-T., Matsumi, Y., Vu, T. V., Sekiguchi, K., Nguyen, T.-T., Pham, C.-T., . . . others (2021). The effects of meteorological conditions and long-range transport on pm2. 5 levels in hanoi revealed from multi-site measurement using compact sensors and machine learning approach. *Journal of Aerosol Science*, 152 , 105716.

Munir, S., Habeebullah, T. M., Mohammed, A. M., Morsy, E. A., Rehan, M., Ali, K., et al.(2017). Analysing pm2. 5 and its association with pm10 and meteorology in the arid climate of makkah, saudi arabia. *Aerosol and Air Quality Research*, 17 (2), 453–464.

Ngo, T. X., Pham, H. V., Phan, H. D., Nguyen, A. T., To, H. T., & Nguyen, T. T. (2023). A daily and complete pm2. 5 dataset derived from space observations for vietnam from 2012 to 2020. *Science of The Total Environment*, 857 , 159537.

Organization, W. H. (2021). Air pollution. Retrieved from https://www.who.int/health-topics/air-pollution#tab=tab 1

Rincon, G., Morantes, G., Roa-L´opez, H., Cornejo-Rodriguez, M. d. P., Jones, B., & Cremades, L. V. (2023). Spatio-temporal statistical analysis of pm1 and pm2. 5 concentrations and their key influencing factors at guayaquil city, ecuador. *Stochastic Environmental Research and Risk Assessment*, 37 (3), 1093–1117.

Sheridan, S. C., Lee, C. C., & Smith, E. T. (2020). A comparison between station observations and reanalysis data in the identification of extreme temperature events. *Geophysical Research Letters*, 47 (15), e2020GL088120.

Tai, A. P., Mickley, L. J., & Jacob, D. J. (2010). Correlations between fine particulate matter (pm2. 5) and meteorological variables in the united states: Implications for the sensitivity of pm2. 5 to climate change. *Atmospheric environment*, 44 (32), 3976–3984.

Tran, C. C., Ta, T. D., Duong, A. T., Phan, O. T., & Nguyen, D. A. (2020). Analysis on temporal pattern of fine particulate matter (pm2. 5) in hanoi, vietnam and the impact of meteorological conditions. Journal of Environmental Protection, 11 (3), 246–256.

Tran-Anh, Q., Ngo-Duc, T., Espagne, E., & Trinh-Tuan, L. (2023). A 10-km cmip6 downscaled dataset of temperature and precipitation for historical and future vietnam climate. Scientific Data, 10 (1), 257.

Vietnam News Agency. (2016). Mekong delta struggling with drought, saline intrusion. Retrieved from https://vietnam-vnanetvn.translate.goog/english/tin-tuc/mekong-delta-struggling-with-drought-saline-intrusion-112398.html? x tr sl=vi& x tr tl=en& x tr hl=en& x tr pto=sc (Accessed: 2024-11-26)

Vietnam News Agency. (2017, July 19). Storm talas claims seven lives, leaves four missing. Retrieved from https://en.vietnamplus.vn/storm-talas-claims-seven-lives-leaves-four-missing-post115091.vnp (Accessed: 2024-11-26)

Wang, J., & Ogawa, S. (2015). Effects of meteorological conditions on pm2.5 concentrations in nagasaki, japan. *International journal of environmental research and public health*, *12* (8), 9089–9101.

Wang, W., Sun, L., Pei, Z., Chen, Y., & Zhang, X. (2019). Analysis of temporal and spatial variation of growing season drought in jiling province based on standardized precipitation evapotranspiration index. In 2019 8th international conference on agro-geoinformatics (agro-geoinformatics) (pp. 1–5).

Wei, S., & Semple, S. (2023). Exposure to fine particulate matter (pm2. 5) from non-tobacco sources in homes within high-income countries: a systematic review. *Air Quality, Atmosphere & Health*, 16 (3), 553–566.

Wolf, M. J., Emerson, J. W., Esty, D. C., de Sherbinin, A., Wendling, Z. A., & et al. (2022). 2022 Environmental Performance Index. New Haven, CT: Yale Center for Environmental Law & Policy. (Available at https://epi.yale.edu/)

World Health Organization. (2014). 7 million premature deaths annually linked to air pollution. Retrieved from https://www.who.int/news/item/25-03-2014-7-million-premature-deaths-annually-linked-to-air-pollution

Yang, Q., Yuan, Q., Li, T., Shen, H., & Zhang, L. (2017). The relationships between pm2. 5 and meteorological factors in china: seasonal and regional variations. *International journal of environmental research and public health*, 14 (12), 1510.

Zhang, H., Wang, Y., Park, T.-W., & Deng, Y. (2017). Quantifying the relationship between extreme air pollution events and extreme weather events. *Atmospheric Research*, 188, 64–79.

Zhao, D., Chen, H., Sun, X., Shi, Z., et al. (2018). Spatio-temporal variation of pm2. 5 pollution and its relationship with meteorology among five megacities in china. *Aerosol and Air Quality Research*, 18 (9), 2318–2331.

Zhao, R., Gu, X., Xue, B., Zhang, J., & Ren, W. (2018). Short period pm2. 5 prediction based on multivariate linear regression model. PloS one, 13 (7), e0201011.

List of acronyms and abbreviations

PM	Particulate matter
TP	Total Precipitation
T2M	Average Temperature
T2M_MAX	Maximum Temperature
T2M_MIN	Minimum Temperature
RH	Relative Humidity
SP	Surface Pressure
WSPD	Windspeed
LR	Linear Regression
GAM	Generalized Additive Mixed models
DID	Difference-in-Differences
TWFE	Two-Way Fixed Effects

Agence française
de développement
5, rue Roland Barthes
75012 Paris l France
www.afd.fr

What is AFD?
Éditions Agence française de développement publishes analysis and research on sustainable development issues. Conducted with numerous partners in the Global North and South, these publications contribute to a better understanding of the challenges faced by our planet and to the implementation of concerted actions within the framework of the Sustainable Development Goals.
With a catalogue of more than 1,000 titles and an average of 80 new publications published every year, Éditions Agence française de développement promotes the dissemination of knowledge and expertise, both in AFD's own publications and through key partnerships.
Discover all our publications in open access at editions.afd.fr.
Towards a world in common.

Publication Director Rémy Rioux
Editor-in-Chief Thomas Melonio

Legal deposit 3rd quarter 2025
ISSN 2492 - 2846

Graphic design MeMo, Juliegilles, D. Cazeils
Layout PUB
Printed by the AFD reprography service

To browse our publications:
https://www.afd.fr/en/ressources-accueil