Assessing Air Pollution in Urban Environments: A Comprehensive Analysis Using Earth Observation and In Situ Measurements

Masterarbeit

zur Erlangung des akademischen Grades

Master of Science

EAGLE - Applied Earth Observation and Geoanalysis

Am Institute for Geography and Geology

Department of Remote Sensing

der Bayerischen Julius-Maximilians-Universität Wuerzburg

Author: Andreas Bury

Peterplatz 3

97070 Wuerzburg

1796909

Supervisor: Prof. Dr. Hannes Taubenböck

Prof. Dr. Tobias Ullmann

Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

Versicherung zur Leistungserbringung

Titel der Abschlussarbeit:

Thema bereitgestellt von (Titel, Vorname, Nachname, Lehrstuhl):

Eingereicht durch (Vorname, Nachname, Matrikel):

Ich versichere, dass ich die vorstehende Arbeit selbstständig und ohne fremde Hilfe angefertigt und mich keiner anderer als der in den beigefügten Verzeichnissen angegebenen Hilfsmittel bedient habe. Alle Textstellen, die wörtlich oder sinngemäß aus Veröffentlichungen Dritter entnommen wurden, sind als solche kenntlich gemacht. Alle Quellen, die dem World Wide Web entnommen oder in einer digitalen Form verwendet wurden, sind der Arbeit beigefügt.

Weitere Personen waren an der geistigen Leistung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe eines Ghostwriters oder einer Ghostwriting-Agentur in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar Geld oder geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Arbeit stehen.

Der Durchführung einer elektronischen Plagiatsprüfung stimme ich hiermit zu. Die eingereichte elektronische Fassung der Arbeit ist vollständig. Mir ist bewusst, dass nachträgliche Ergänzungen ausgeschlossen sind.

Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht. Ich bin mir bewusst, dass eine unwahre Erklärung zur Versicherung der selbstständigen Leistungserbringung rechtliche Folgen haben kann.

Ort, Datum, Unterschrift

Julius-Maximilians-Universität Wuerzburg

- Institute of Geography and Geology –

Earth Observation Research Hub

Master of Science

Assessing Air Pollution in Urban Environments: A Comprehensive Analysis Using Earth Observation and In Situ Measurements

Andreas Bury

Abstract

This master's thesis investigates the feasibility of using Earth Observation and low-cost environmental sensors in urban areas to complement existing government monitoring systems to provide a sharper picture of air quality, which is a growing concern for the health and well-being of urban dwellers. The work uses NO2 tropospheric column data from the Copernicus Sentinel-5P satellite, regulatory PM10 readings and self-constructed low-cost sensor systems measuring PM10, ambient temperature and relative humidity. The environmental monitors were developed with computer-aided design, 3D printed, and their performance evaluated. A network of static sensors and a mobile system were deployed to better cover the urban space in Wuerzburg, Germany, and detect spatio-temporal patterns. Each of the three layers has its own limitations with satellites facing temporal and spatial resolution issues and cloud cover interference, regulatory data providing site-specific measurements but only publishing hourly averages, and low-cost sensors being sensitive to elevated relative humidity and cold ambient temperatures. However, the work also shows that with appropriate data and procedures, calibration of low-cost sensors is possible and that all three levels can be intertwined to make a valuable contribution to research and understanding of urban space and its pollution levels.

Zusammenfassung

Die vorliegende Masterarbeit untersucht, ob Erdbeobachtung und Low-cost-Sensorik im urbanen Raum synergetisch genutzt werden können, um bestehende behördliche Monitoringsysteme zu ergänzen und so ein schärferes Bild über die Luftqualität zu erhalten, die ein wachsendes Problem für die Gesundheit und das Wohlbefinden der Stadtbewohner darstellt. Die Arbeit verwendet NO2-Troposphären Daten des Copernicus Sentinel-5P-Satelliten, behördliche PM10-Messungen und selbst konstruierte Low-cost-Sensor Systeme, die PM10, Außentemperatur und relative Luftfeuchtigkeit messen. Die Umweltmonitore wurden mit einer Software zum computerunterstützten Design entworfen, mit 3D-Druck realisiert und ihre Leistung evaluiert. Ein Netzwerk aus statischen Sensoren und ein mobiles System wurden eingesetzt, um den urbanen Raum in Wuerzburg, Deutschland, besser abzudecken und raumzeitliche Muster zu erkennen. Jede der drei Schichten hat ihre eigenen Einschränkungen, wie zeitliche und räumliche Auflösung sowie Störungen durch Wolkenbedeckung bei Satelliten, ortspezifische Messungen und nur stündliche Durchschnittswerte bei behördlichen Daten und Empfindlichkeit gegenüber erhöhter relativer Luftfeuchtigkeit und kalten Umgebungstemperaturen bei kostengünstigen Sensoren. Die Arbeit zeigt jedoch auch, dass mit geeigneten Daten und Verfahren eine Kalibrierung der Low-cost-Sensorik möglich ist und dass alle drei Ebenen ineinandergreifen können, um einen wertvollen Beitrag zur Forschung und zum Verständnis über den urbanen Raum und seiner Schadstoffbelastung zu leisten.

Acknowledgments

My deepest gratitude goes to my supervisors, Prof. Dr. Hannes Taubenböck and Prof. Dr. Tobias Ullman, for their unwavering support and mentorship. Their expertise has been instrumental in the development of this study.

I must also acknowledge the invaluable contributions from Dr. Martin Wegmann. The technical resources and encouragement provided by him, and the Department of Remote Sensing were crucial in overcoming the challenges I encountered.

A heartfelt thanks to Aline for her consistent support and inspiration. Her presence has been a constant source of strength throughout this journey.

Lastly, I extend my appreciation to everyone who has contributed to this study in any capacity. The collective guidance and encouragement have been essential in fulfilling my research objectives.

Content

List of Figures

List of Tables

For Aline, my inspiration and anchor

I Introduction

1. Background and Motivation

Air pollution remains a critical environmental risk to human health in Europe [1] [2], making the monitoring of exposure levels of critical concern. Despite various policy measures, poor air quality in many places continues to pose a significant health risk to the population [3]. Harmful air pollution can lead to severe respiratory and cardiovascular diseases, which are the most common causes of premature death due to air pollution [4]. The International Agency for Research on Cancer (IARC) has classified air pollution as carcinogenic, with particulate matter identified as the primary component of air pollution mixtures [5]. In 2018, air pollution resulted in a total of 452,400 premature deaths in the European Union, with particulate matter (PM), nitrogen dioxide (NO2), and ozone (O3) pollution being the main contributors [1]. Those most severely affected by the effects of breathing polluted air include children, pregnant women, the elderly, and individuals with pre-existing conditions. Moreover, people with lower socioeconomic statuses are more impacted since they are more likely to reside in environments with subpar air quality [6]. Elevated pollution levels also contribute to increased costs for businesses and healthcare systems due to lost workdays and required medical care [7]. Consequently, the benefits of implementing stricter air quality policies outweigh the associated costs [8].

Efforts to reduce air pollution in Europe are progressing but citizens are still breathing air that widely exceeds the World Health Organization (WHO) guidelines [9] and suggested exposure limits [10] . To better understand the sources of air pollution and its impacts, it is necessary to monitor air quality and pollution levels with high spatial and temporal resolution. Remote sensing technology and low-cost air quality sensors can contribute to this objective and have been used to achieve this goal [11] [12]. However, low-cost sensors have limitations in terms of accuracy, reliability, and consistency [13], while government-grade monitors are expensive and cover only a limited area [14].

Therefore, integrating low-cost sensors, Earth Observation and governmental-grade monitors can provide a more comprehensive and accurate assessment of air quality [15] [16]. My motivation for this study is to contribute to a more profound understanding of urban space and its environmental factors. Healthy air quality is an issue that cannot be overlooked and needs interdisciplinary integration to provide a clear picture and to better plan and adapt our cities. Remote sensing plays a pivotal role due to its unique ability to provide large-scale insights that are invisible to the human eye.

1.1 Research Questions and Objectives

The overarching aim of this research is to critically examine the potential of an integrative methodology for assessing air quality in urban environments. This methodology encompasses data from Earth Observation technologies, conventional in-situ monitoring mechanisms, and emerging low-cost sensor technologies. In the first phase, the research evaluates the extent to which the retrievals of NO2 tropospheric columns from the European Space Agency (ESA) Copernicus Sentinel-5P satellite offer observations that are sufficiently contiguous in both spatial and temporal dimensions. This is critical for generating a robust analysis of air quality in localized urban areas such as Wuerzburg. Subsequently, the study then addresses low-cost particulate matter sensors with an emphasis on comparing these cost-effective alternatives to authorities-operated monitoring sites to assess their reliability and accuracy in data collection. A key component of this research is the investigation of mobile monitoring systems. The objective is to evaluate whether such devices can accurately provide robust data sets compared to stationary instruments. Finally, the study will investigate whether discernible patterns and trends can be identified when contrasting in-situ PM10 measurements with tropospheric NO2 data.

By exploring these avenues of inquiry, the study not only seeks to shed light on the specific air quality dynamics in Wuerzburg but also aspires to contribute to the broader methodological discourse in air quality assessment.

1.2 Limitations

The study grapples with multiple methodological and data-related challenges. Firstly, the spatial resolution of Sentinel 5P Offline Nitrogen Dioxide data is 1113.2 meters, constraining the granularity of results when examining urban-scale phenomena. Budgetary constraints imposed additional obstacles to the use of low-cost sensor systems, which affects data density and completeness. The scope of a master's thesis inherently restricts a more in-depth examination of specific air quality facets.

Data sets for official ambient measurements, essential for calibrating the low-cost sensors, lacked the desired temporal format. Winter conditions posed challenges for sensor performance, which is particularly sensitive to variations in relative humidity and ambient temperature. The segment of the study centred on low-cost particulate matter sensors was methodologically limited to PM10 omitting other pollutants like NO2 and O3. The deployment of mobile monitoring systems faced spatial and temporal constraints due to being a single-person operation.

Lastly, the integration of heterogeneous data sources, including measurement methodologies with various physical units, serves as a distinct limitation in the study. While aiming to create a comprehensive spatio-temporal profile of air quality in Wuerzburg, this diversity of data sources introduces its own set of challenges.

II Literature Review

Building upon the compelling issue of air pollution outlined, this chapter endeavours to provide a comprehensive understanding of urban air quality monitoring in Europe with a focus on Germany. This exploration starts with a historical overview of air quality monitoring, tracing the advancements and shifts in understanding over time. Subsequently, the chapter will evaluate the strengths and limitations of Earth Observation (EO) technologies as mechanisms for monitoring and understanding air quality. Concluding the chapter, an assessment will be presented on the utility and efficacy of low-cost environmental sensors for monitoring air quality. The potential benefits, alongside the inherent limitations of these sensors, will be discussed.

2.1 Historical Overview of Air Quality Monitoring

The pivotal moment that triggered a heightened focus on air quality occurred with the Great Smog of London in 1952 [17]. This environmental disaster led to thousands of premature deaths and severe health issues. This event served as a catalyst for legislative actions, including the UK's Clean Air Act of 1956 [18]. In the European Union, the Air Quality Framework Directive (Directive 96/62/EC [19]) marked the beginning of a concerted effort to assess and control air pollution. This framework established systems for monitoring various pollutants and laid down exposure limits, which were further specified in subsequent daughter directives.

The Ambient Air Quality Directive (Directive 2008/50/EC [3]) represents a significant evolution in European air quality governance, promulgated to emergent scientific evidence highlighting the deleterious impacts of air pollutants. This Directive encapsulates a multi-faceted approach to air quality management, harmonizing advanced analytical methods and scientific techniques for comprehensive air quality evaluation. Additionally, it affords Member States a degree of implementation flexibility to accommodate regional variations in pollution sources and meteorological conditions. The Directive also mandates a stronger framework for public disclosure of air quality metrics, thereby augmenting transparency and public awareness.

The directive accentuates the importance of cross-border cooperation to manage transboundary air pollution, while introducing a legal apparatus to enforce compliance and impose sanctions on non-compliant Member States. These sanctions must be ‘effective, proportionate and dissuasive’ (Directive 2008/50/EC Article 30). As we look towards the future, recent legislative endeavours underscore the European Parliament Environment Committee's commitment to implementing even more stringent regulations on air quality. A report recently ratified by the committee outlines ambitious objectives, including the imposition of tighter emission limits on critical pollutants by 2030. The plan seeks to standardize air quality indices across the European Union and mandates that Member States formulate comprehensive air quality roadmaps [20]. Additionally, the report advocates for an increase in air quality monitoring stations and demands enhanced public information dissemination, including hourly updates to air quality indices. These initiatives are designed to align with the latest World Health Organization Air Quality Guidelines [9], [21] aiming to substantially reduce premature mortality rates attributed to air pollution in the EU. However, it is essential to recognize that despite these advancements in legislative action, significant challenges remain. Even with increasingly stringent regulations, achieving widespread compliance and reducing air pollution to safe levels continues to be an arduous task requiring multidisciplinary efforts [22]. This necessitates the exploration of innovative monitoring technologies, data analytics methods, and policy interventions.

2.2 Monitoring Technologies and Data Sources

In the context of the evolving regulatory landscape and ongoing challenges, this section will review the array of technologies and data sources that are currently being employed for air quality monitoring, highlighting their role as foundational elements in both academic research and policy formulation.

2.2.1 Governmental In-Situ Monitoring

In Germany, where the study area is situated, the 39th Federal Immission Control Ordinance, based on the EU Directive 2008/50/EC establishes legal benchmarks and limit values for diverse pollutants in the air [23]. Both the type of measurements, location criteria and the number of environmental monitoring sites are governed herein. Table 1 outlines the minimum number of stationary sites mandated for PM10 assessment. The study area has a population of approximately 145.000 [24] and belongs in category one. As per this classification, the area is required to have two stationary monitoring sites for PM10 and PM2.5. Notably, if both PM2.5 and PM10 measurements are logged at the same station, they should be counted as two distinct sampling stations. For NO2, the regulations require the inclusion of at least one measuring station for urban background sources and another specifically for traffic-related emissions.

Population of Area (in thousands)	PM (Max Value Exceeds Upper Assessment Threshold)	PM (Max Value Between Upper and Lower Assessment Thresholds)
0 – 249	2	1
250 – 499	3	2
500 – 749	3	2
750 – 999	4	2
1000 – 1499	6	3
1500 – 1999	7	3
2000 – 2749	8	4
2750 – 3749	10	4
3750 – 4749	11	6
4750 – 5999	13	6
≥ 6000	15	7

Table 1: Criteria for determining the number of sampling stations [23]

Air quality monitoring stations are positioned based on macro- and micro-scale siting rules. These stations are typically classified by the dominant emission sources, with traffic stations being located close to major roads, industrial stations near industrial areas or sources, and background stations situated in places where pollution levels best represent the average exposure of the general population or vegetation. In terms of the surroundings, stations are described based on the type of area they are in. An urban station is in a continuously built-up urban area, a suburban station is in a largely built-up urban area, and a rural station is found in areas other than urban or suburban. It is noteworthy that all stations which are officially reported to the EEA are included in Europe's air

quality status briefing [25]. The data quality goals concerning PM10 for air quality assessment stipulate parameters ensuring precision. For stationary measurements, a 25% uncertainty and a minimum 90% data capture are required. Urban background, traffic, and industrial areas have no specified measurement duration for PM10. Indicative measurements have a 50% uncertainty rate but maintain the 90% data capture, with a minimum duration set at 14%, either yearly or over eight weeks. Authorities can choose spot measurements for PM10, provided the uncertainty, considering sample randomness, meets the 25% quality goal. These measurements should spread evenly over the year to avoid bias. For compliance with PM10 limits using spot measurements, the 90.4 percentile value, which shouldn't exceed 50 micrograms per cubic meter, is prioritized over data capture-influenced results. Uncertainty evaluation in measurements adheres to specific standards, aligning with guidelines and ISO norms compliance. These uncertainty percentages relate to single measurements averaged over relevant periods versus the limit value. This uncertainty definition also encompasses stationery and model calculations, considering the model's spatial resolution [23].

2.2.2 Earth Observation for Air Quality Assessment

The capability of remote sensing technology to span large geographical areas offers numerous advantages in evaluating air quality. Satellite-based methods have emerged as indispensable tools in contemporary atmospheric research, going beyond the limitations of traditional in-situ methods that are geographically restricted. Recent advancements in Earth Observation have proven invaluable for assessing air quality, specifically in the quantification of nitrogen dioxide (NO2) levels. Ka Lok Chan et al. focuses on the application of machine learning techniques to estimate surface NO2 levels in Germany. Drawing upon TROPOMI satellite observations coupled with meteorological parameters, the researchers employ a neural network model to predict surface NO2 concentrations. They validate their model against ground-based in situ measurements and regional chemical transport model (CTM) simulations, finding a favourable Pearson correlation coefficient (R) of 0.80. What makes this study particularly noteworthy is that the machine learning-based estimates of surface NO2 show better agreement within situ data than regional CTM simulations.

Therefore, Chan et al. establish that machine learning techniques, when employed alongside satellite data, can provide reliable estimates of air quality. In the study conducted by Inken Müller, Thilo Erbertseder, and Hannes Taubenböck from the German Remote Sensing Data Center and the Julius-Maximilians-Universität Würzburg, high-resolution Sentinel-5P/TROPOMI satellite data is used to explore the spatial variability of nitrogen dioxide (NO2) in Germany [26]. The research identifies 24 significant hot spots for NO2, primarily in urban regions, and investigates the correlation between NO2 levels and non-meteorological factors such as impervious surfaces, population, and road density. The authors demonstrate that TROPOMI data can effectively supplement traditional in-situ measurement networks in understanding the role of road traffic in NO2 emissions. Their findings underscore that spatial variability in NO2 levels is influenced both by local emission sources and broader regional factors, contributing valuable insights for air quality management and climate adaptation strategies. Adding to these observations, a study by D. Oxoli, J. R. Cedeno Jimenez, and M. A. Brovelli from the Department of Civil and Environmental Engineering at Politecnico di Milano focuses on Sentinel-5P's effectiveness in ground-level air quality monitoring. The study compares Sentinel-5P observations with traditional ground measurements in the Lombardy region of Northern Italy. Utilizing data collected during the COVID-19 pandemic, the researchers found a marked 17.5% reduction in NO2 concentrations during the lockdown. The study also reported strong positive correlations between ground-based and satellite observations, with Spearman's rank correlation coefficients (ρs) outperforming Pearson's correlation coefficients (ρp), especially in plain and metropolitan areas. This work sets the stage for future research focused on leveraging machine learning and geostatistics methods to improve the granularity of air quality monitoring [27].

While EO technology offers several advantages, there are critical limitations. One of the most significant challenges is the data gaps arising from cloud cover and other retrieval limitations. A study by Philipp Schneider and his team from the NILU—Norwegian Institute for Air Research, highlights these shortcomings [28]. According to their study, depending on the site and the season, only 20% to 50% of the retrievals were found to be valid, significantly limiting the utility of EO technologies during high pollution periods. These data gaps also impede the operational assimilation of local-scale dispersion models.

The study suggests that data availability could be improved by lowering the acceptable retrieval quality threshold, albeit at the cost of data quality. Another limitation pertains to the temporal constraints of the technology. Sentinel-5P's TROPOMI sensor collects data at specific acquisition times, rendering it incapable of providing continuous, 24/7 monitoring of air quality. The temporally sparse data is particularly problematic given the high volatility of gases such as NO2, which can fluctuate substantially over short periods [29]. This temporal limitation restricts the efficacy of EO technologies in real-time air quality management and necessitates the integration of complementary data sources for a comprehensive air quality assessment.

2.2.3 Advantages and Challenges of Low-Cost Sensors

The advent of low-cost sensors has brought a revolution in the realm of air quality monitoring. These affordable, easy-to-use, and accessible tools offer a means for various stakeholders - including citizens, non-governmental organizations, and local authorities - to monitor air quality in different contexts[29]. One way these low-cost sensors enable in situ data collection is by forming networks that provide near-real-time air quality observations with high spatial and temporal resolution. This fine-grained data can fill the gaps left by existing official regulatory networks that usually operate at lower spatial and temporal scales [30]. Peer Nowack et al. investigated the potential of machine learning algorithms to calibrate low-cost sensors for NO2 and PM10. Conducted in the urban area of London, their study evaluated multiple regression methods and found that machine learning algorithms like Gaussian Process Regression (GPR) and ridge regression could significantly improve the sensor performance. The study also shed light on the limitations of these machine-learning-based calibration techniques, such as the difficulty in transferring calibrated sensors to new locations with different air pollution patterns [31]. In a multi-city epidemiological study, Marina Zusman et al. evaluated the performance characteristics of the Plantower PMSA003 and Shinyei PPD42NS sensors for measuring fine particulate matter. Their findings indicate that region-specific calibration models significantly improve the performance of these sensors. For example, Plantower PMS A003 sensors calibrated with regional models showed high precision and accuracy compared to regulatory instruments.

The study emphasizes the need for localized calibration and warns against using manufacturer-provided general calibration factors, particularly when the data is intended for use in environmental epidemiological studies [32]. Expanding on sensor-based approaches, a research study by Amin Anjomshoaa et al. introduced the concept of drive-by sensing in smart cities. Utilizing portable sensors mounted on urban vehicles, this dynamic method allows for comprehensive and efficient monitoring of the urban environment. The study posited that this strategy offers lower deployment and maintenance costs compared to traditional stationary sensor networks. Furthermore, the research highlighted the influence of street network topology on data sampling frequency and spatial granularity, suggesting the need for custom models that take into account the complexity of urban parcels [33].

Despite these advancements, there are still challenges to overcome. Alice Cavaliere et al. further emphasized the importance of calibration in a study focused on O3 and NO2 sensors. They found that while machine learning models like multiple random forests (MRF) offered high accuracy, they were less reliable in generalizing to new datasets. In contrast, parametric models like multiple linear regression (MLR) were more robust and easily adjustable over time. This is particularly relevant for NO2 and O3 sensors, as these require unique parameter values due to intersensory variability [30]. However, even with improved calibration techniques, low-cost air pollution sensors and drive-by sensing methods often still fall short of the performance metrics set by state-of-the-art air quality monitoring stations. The quality of the data produced may not be suitable for legal sanctions, but these emerging technologies offer a rapid and flexible approach to air quality monitoring, filling in gaps that traditional methods may miss.

III Study Area

Wuerzburg, the focal point of this study, is situated in the federal state of Bavaria in Germany. The city's geographical coordinates range from 49.7105°N to 49.8456°N in latitude and 9.8716°E to 10.0144°E in longitude. Set within a valley and framed by surrounding hills, the Main River winds its way through the city, adding to the complexity of its topography. Wuerzburg is divided into 13 administrative districts, covering a total land area of 87.63 km² and sustaining a population density of about 1486 inhabitants per square kilometre [24]. Based on records from 1991-2020, the average annual temperature in Wuerzburg fluctuates from 1.2°C in January to 19.7°C in July [34].

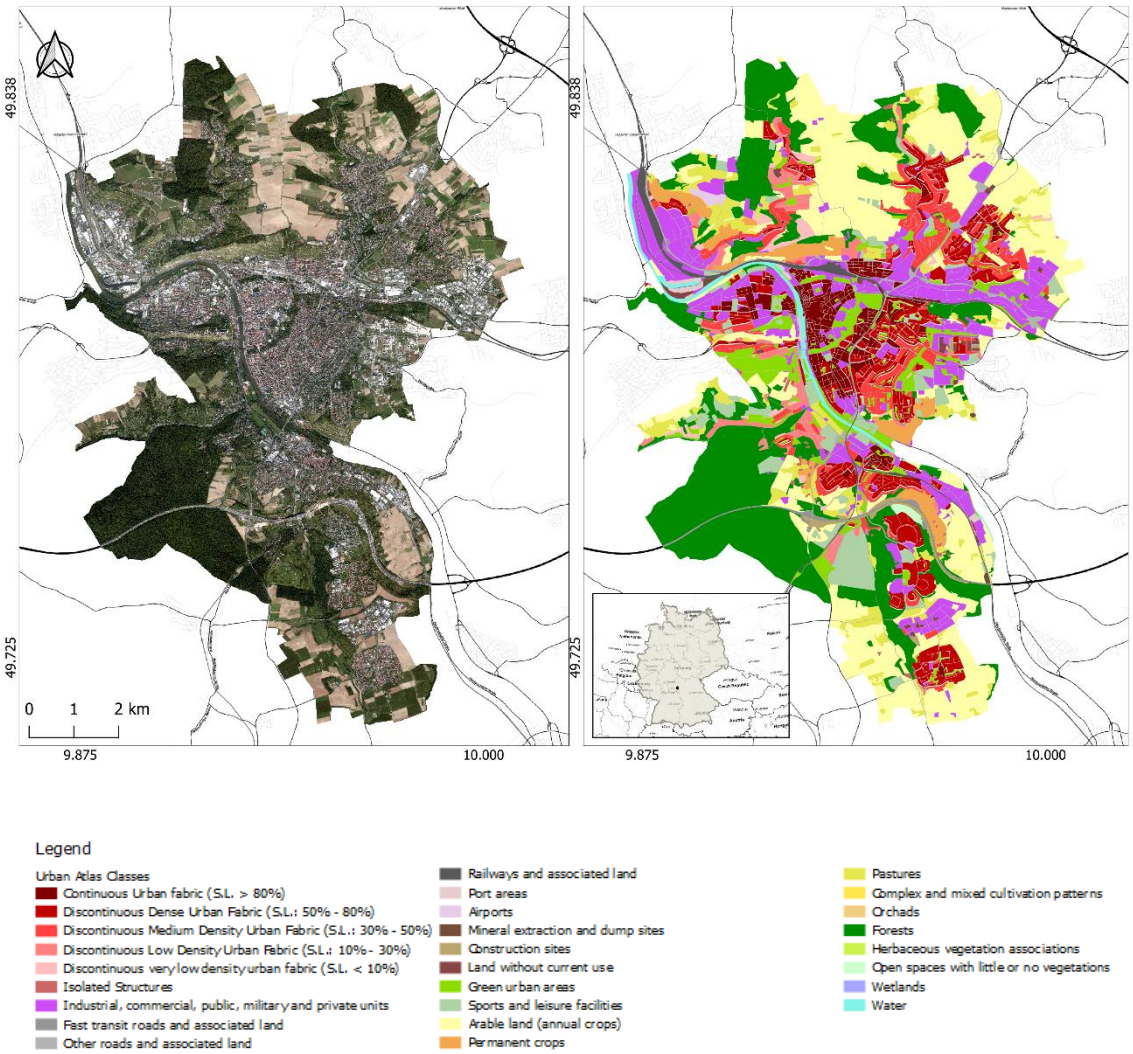


Figure 1: Research area – Wuerzburg.

Left: Digital Orthophoto [35] Right: Urban Atlas [36]

Figure 1 offers a comprehensive visual representation of Wuerzburg, showcasing both high-resolution RGB imagery and Copernicus Urban Atlas 2018 classifications. According to the Copernicus Urban Atlas 2018 [36], the city primarily consists of discontinuous dense urban fabric, accounting for approximately 16.19% of the total area. This is followed by industrial, commercial, public, military, and private units, making up about 16.07%. Continuous urban fabric contributes to 14.05%, whereas green urban areas comprise approximately 10.89% of the land. Other notable land classes include discontinuous medium-density urban fabric (6.85%), pastures (5.83%), sports and leisure facilities (5.42%), arable land (5.18%), and forests (4.76%). The least prevalent is discontinuous low-density urban fabric, occupying only around 3.27% of the city's total area.

IV Materials and Methods

This section outlines the materials and methods utilized in this master's thesis, as detailed in Figure 2, which illustrates the workflow and key components of the study.

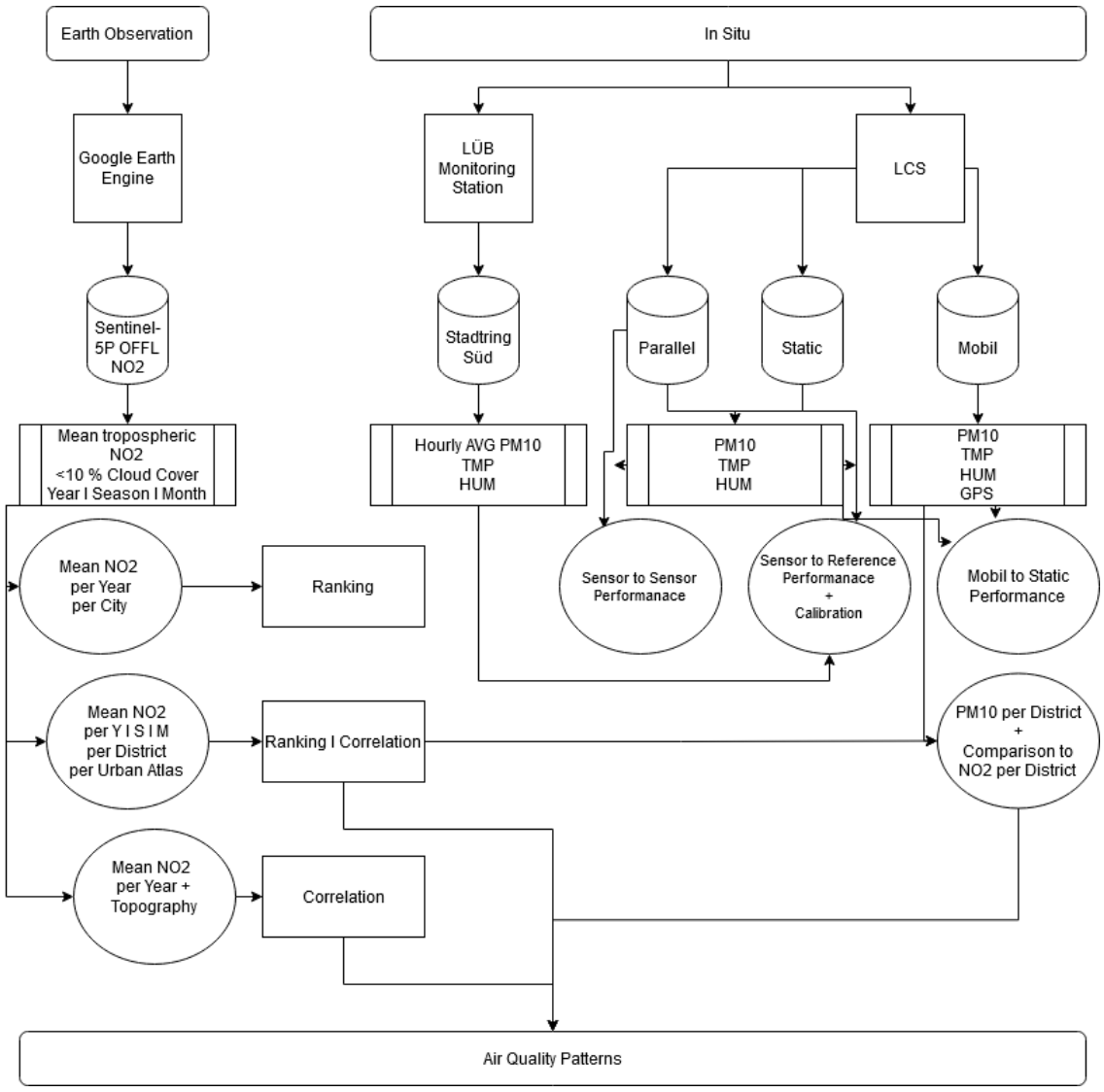


Figure 2: Applied workflow

4.1 Data

4.1.1 Sentinel-5 Precursor

The present thesis focuses on the analysis of Earth Observation data obtained from the TROPOspheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 Precursor satellite [2]. The aim is to examine the patterns of tropospheric nitrogen dioxide (NO2) concentrations in Wuerzburg, Germany, over a period of January 2019 to March 2023. Sentinel-5 Precursor, launched on 13 October 2017, is a low Earth orbit satellite designed to provide daily atmospheric data products for a period of seven years. The TROPOMI instrument achieves high spatial resolution (0.01 arc degrees) and uses its UV-VIS spectrometer to measure backscattered solar radiation within the 405–465 nm wavelength range. [37]. This allows for precise measurements of tropospheric NO2 columns, a significant indicator of air quality produced by both natural and anthropogenic sources, including fossil fuel combustion and forest fires. The study leverages the Sentinel-5P Offline Nitrogen Dioxide (OFFL NO2) dataset, with a spatial resolution of 1113.2 meters, which is accessible via Google Earth Engine [38], a cloud-based platform for satellite imagery and geospatial data. The dataset covers the period from 28 June 2018 to 03 February 2023 [39] and is updated periodically. The acquired data collection underwent multiple layers of filtering to ensure the reliability of the study's findings. First, the data was restricted by time parameters, encompassing records from January 2019 through March 2023, and segmented further by year, seasons, and months. Additionally, the dataset was subjected to a cloud fraction filter to maintain the maximum effective cloud fraction at 0.1 (10%) for the entirety of the image collection [40]. Following this, the data was spatially clipped to conform to the geographical boundaries of the city of Wuerzburg and its districts. Finally, the unit of measurement was converted from µmol/m^3 to µmol/m^2 by multiplying the values by 10^6. These pre-processing steps culminated in datasets representing the mean NO2 tropospheric column for individual years, seasons, and monthly intervals, which were subsequently stored for further analysis.

4.1.2 Urban Atlas 2018

To establish a robust spatial framework for correlating land use patterns with air quality metrics in Wuerzburg, this study utilized the Urban Atlas 2018 [36]. This dataset serves as a pivotal source of standardized land use and land cover information for major cities within the European Union and European Free Trade Association (EFTA) countries. Under the leadership of the Copernicus program [41], a land monitoring initiative, the Atlas is collaboratively produced by the Directorate-General for Regional Policy, the Directorate-General for Enterprise and Industry, and supported by the European Environment Agency. The Atlas primarily uses high-spatial-resolution imagery from SPOT 5 & 6 and Formosat-2, with 2 to 2.5m spatial resolution, to derive its land cover classes, which are further enriched by ancillary data such as local city maps. Introduced initially in 2006 covering 305 Functional Urban Areas (FUAs), the 2018 iteration has expanded its scope to nearly 700 cities in the EU28 and EFTA countries. It has also evolved in its nomenclature, expanding from the initial 17 classes to include a broader array of land use categories. The atlas now includes detailed rural fringe classes and an additional 'street tree layer.' In this study, the Urban Atlas 2018 provided invaluable land cover data for Wuerzburg, which will be analysed in conjunction with tropospheric NO2 data to offer a multi-dimensional view of the city’s air quality.

4.1.3 Governmental Permanent Monitoring Sites

To evaluate the accuracy and calibration performance of the low-cost Air Quality Sensor (LCS) deployed in this study, benchmark data was sourced from the Bavarian Air Quality Monitoring System (LÜB) [23]. Specifically, datasets from the Measuring Station Wuerzburg Stadtring Süd, located at coordinates 49°47.430 N and 9°56.858 E, served as the reference for comparative analyses. This site continuously monitors various air pollutants and meteorological indicators, encompassing carbon monoxide (CO), nitrogen monoxide (NO), nitrogen dioxide (NO2), as well as particulate matter with diameters less than 10 micrometres (PM10). The NO2 concentration is determined using a method described in DIN EN 14211:2012 that employs Chemiluminescence. This technique leverages a chemical reaction between O3 and NO that results in the emission of light.

The spectrum of this emitted light is characteristic of the substances being measured and serves as a highly reliable detection method [42]. Other detection methods include colorimetry, coulometry, and IR and microwave spectroscopy. More parameters recorded include benzo(a)pyrene, lead, cadmium, arsenic, nickel, air temperature, and relative humidity. The particulate pollutants are captured at inlet heights of 3.8 m and 6.0 m above ground level, situated 6.0 m from the edge of the adjacent roadway. The particulate pollutants are captured at inlet heights of 3.8 m and 6.0 m above ground level, situated 6.0 m from the edge of the adjacent roadway. The monitoring station is geographically located to the east of Wuerzburg's city centre, adjacent to the four-lane federal road B 4 [24]. The local terrain exhibits a downward incline toward the west and is bordered by open residential zones. For the purposes of this study, datasets of PM10 concentrations, ambient temperature, and relative humidity were collected for the months of January and February 2023. These datasets were subjected to pre-processing steps to ensure data quality and integrity. Available in CSV format, the datasets include date, time stamp, and an hourly average value for PM10.

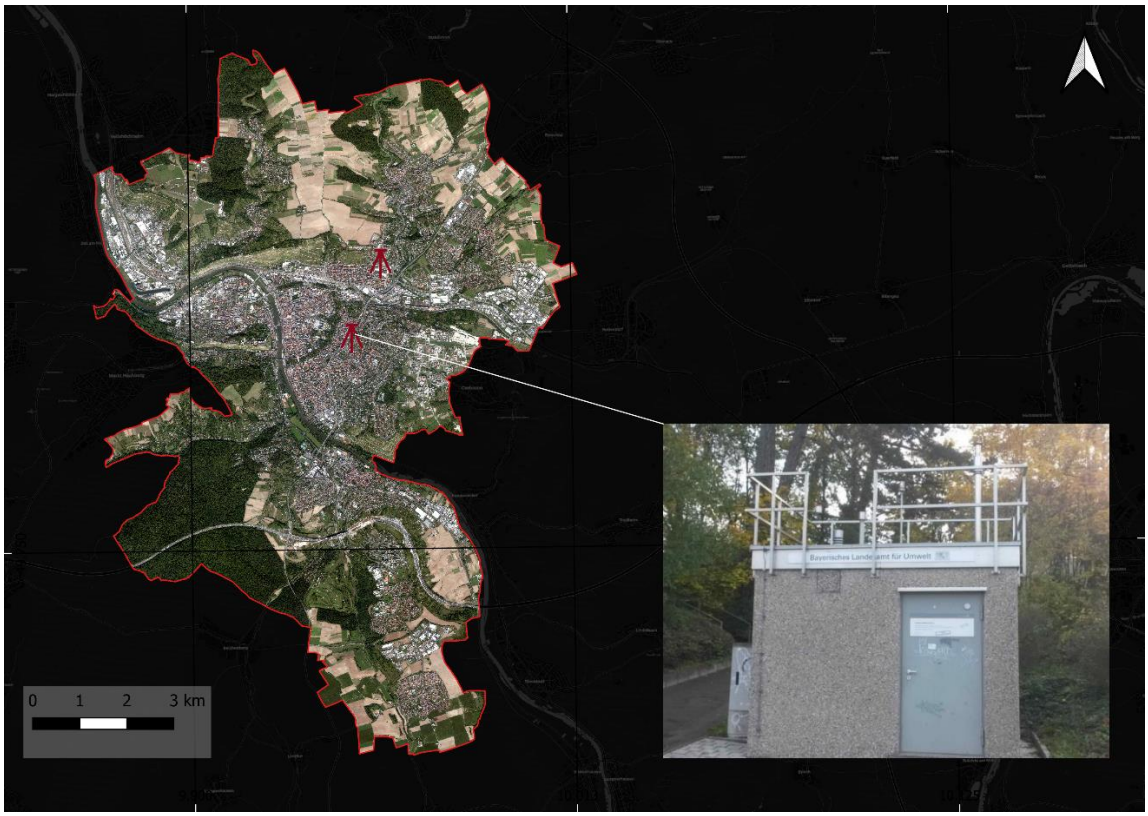


Figure 3:Locations of the two governmental measurement sites, featuring Wuerzburg Süd [43].

Base-Map: CartoDB Inc.[44]

4.2 Software and Programming Languages

In the conduct of this research, a range of software applications and programming languages were employed to support geospatial data management, sensor design, and data analytics. QGIS functioned as an open-source Geographic Information System (GIS) [45], providing a comprehensive suite of tools for the manipulation, analysis, and visualization of geospatial data. It was pivotal in computing zonal statistics of the average NO2 tropospheric column for Wuerzburg and in generating maps for data interpretation. Concurrently, FreeCAD [46], an open-source Computer-Aided Design (CAD) platform, was utilized for the conceptualization and prototyping of weather-resistant enclosures tailored for the environmental monitoring stations. To facilitate the physical realization of these prototypes, Prusa Slicer was used to prepare 3D STL files for additive manufacturing on a Prusa 3D printer [47]. Google Earth Engine, a cloud-based platform, was used to acquire and pre-process the Earth Observation data [38]. The platform granted access to Sentinel-5P OFFL NO2: Offline Nitrogen Dioxide datasets [39], forming a core component of the research data pool. Java [48] was implemented as the primary programming language within the Google Earth Engine environment to conduct complex data filtering and modelling tasks. These included constraining the datasets to a maximum cloud cover of less than 10% and spatially confining them to the geographical boundaries of Wuerzburg. Moreover, the data were temporally categorized to capture annual, seasonal, and monthly fluctuations for later analysis. R, a robust open-source language for statistical computing and graphical representation, was invoked for processing, filtering and merging datasets [49]. It was utilised for statistical analysis and data visualisation. Furthermore, MicroPython [50], an open-source programming language optimized for microcontrollers and other embedded systems, was employed for the programming of microcontrollers that manage the environmental sensors. Data storage was executed both locally and, in a cloud,-based database, with data transmission facilitated via the Message Queuing Telemetry Transport (MQTT) protocol [51] to the Amazon Web Services (AWS) IoT Core [52]. Collectively, the strategic utilization of these diverse software and programming tools enabled a comprehensive and multi-faceted approach to achieving the research objectives.

4.3 Sensor Design

This section outlines the conceptualization and realization of cost-effective environmental sensors employed in this research project. The primary objective was to ensure the reliable protection of sensor electronics against external environmental influences. To achieve this goal, customized weather-resistant enclosures were designed using Computer-Aided Design (CAD) methodologies. These prototypes were subsequently produced through additive manufacturing techniques with the aid of a 3D printer. Following the assembly stage, the environmental sensors were integrated with a microcontroller unit. This unit was programmed to methodically collect and record data from the connected sensors at predetermined intervals.

4.3.1 Platforms and Sensors

In the present study, the Raspberry Pi Pico W microcontroller board [33] was designated as the central unit responsible for sensor management and environmental data logging. The RP2040 microcontroller chip, which serves as the computational core of the Pico W, facilitates efficient data acquisition and storage. This is achieved through its capability to communicate with attached sensors via the Inter-Integrated Circuit (I2C) protocol and to record data onto a Secure Digital (SD) Card through a Serial Peripheral Interface (SPI). The Pico W boasts several advantageous features, including a 2.4 GHz wireless interface for data transmission, 2 MB of onboard flash memory for code storage and execution, and a Micro USB B port for both power provision and data transfer. The boards power supply architecture is both straightforward and flexible. Additionally, it is equipped with an array of 40 digital pins, providing a suite of digital peripherals such as dual UARTs, I2Cs, and SPIs, along with 16 PWM channels and an integrated real-time clock.

Optimized for both cost-efficiency and performance, the Pico W was selected for its modest price point of approximately seven Euro and its high-quality construction, establishing it as an ideal platform for development and prototyping within this research context.

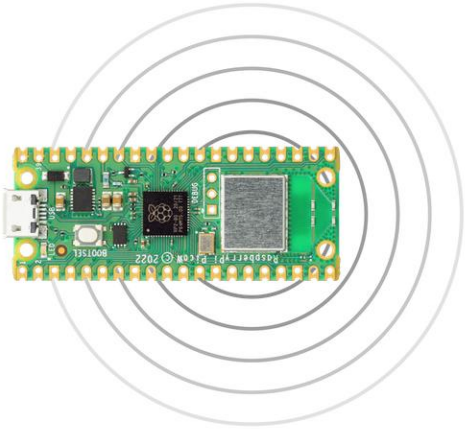


Figure 4: Raspberry Pico W microcontroller [53]

4.3.2 Computer Aided Design

Three distinct instrument enclosures were engineered to accommodate the array of components deployed in this research project. Utilizing the Prusa i3 MK3S+ printer, these housings were fabricated through an additive manufacturing method (AM) with polylactides filament (PLA), a synthetic polymer, as the printing material [54]. Particular attention was afforded to the design to ensure rapid and secure assembly, incorporating threaded nuts compatible with the 3D-printed parts. The design also took into account the necessity to shield the sensors and other constituents from harsh external environmental conditions, while still ensuring their optimal functionality.

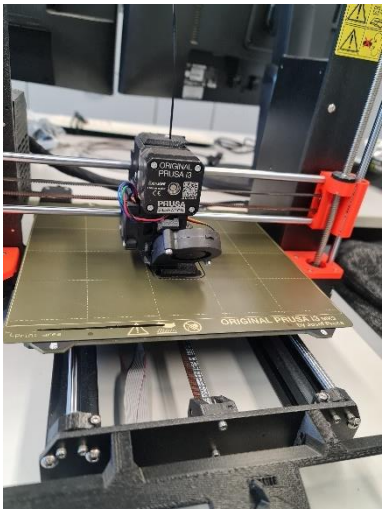


Figure 5:Photograph of Prusa i3 MK3S+ 3D printer

4.3.3 Enclosure Design for Simultaneous Multi-Sensor Monitoring

To enable synchronized air quality assessments via four identical Plantower PMSA003I particulate matter sensors [55], a custom enclosure was designed. This configuration seeks to examine sensor-to-sensor variability and assess the performance of the low-cost sensors relative to the reference station. To mitigate potential biases in the data, the design incorporates air inlets at the front and exhaust outlets at the bottom, thereby ensuring that air sampled is not contaminated by prior measurements. The enclosure dimensions are 24 cm in length, 10 cm in height, and 8 cm in width, with an approximate additive manufacturing time of 25 hours for all essential components.

In addition to accommodating the control unit and the four air quality sensors, the enclosure is designed to house two battery units (one with a 5200 mAh capacity and another with 3800 mAh) providing a dependable power source for uninterrupted data collection. To facilitate data storage, a MicroSD card adapter (HW-125 [56]) is integrated into the system architecture. Enhancing the data acquisition capabilities further, an 8-channel multiplexer, TCA9548A [57], has been integrated to resolve I2C address conflicts inherent in the PMSA003I sensors and enables multiple I2C connections for efficient data collection. Given that ambient temperature and relative humidity are already monitored by the proximately co-located Stadtring Süd measuring station, the inclusion of sensors for these parameters was deemed superfluous.

Figure 6: Photographs of the 3D-printed case for four PMSA300I particulate matter sensors - showing the instruments, the closed housing, and the co-location with the reference station

4.3.4 Enclosure Design for Static Monitoring Configurations

To establish a compact and cost-effective network of monitoring stations, a small instrument housing was designed with the aim of enhancing the spatial and temporal availability of PM10 values within the research area. The enclosure, with a diameter of 10 cm, is fabricated through additive manufacturing processes, requiring approximately 18 hours for each unit. The assembly is comprised of three parts: a base that hosts the control unit and SD card for data storage, a top part that contains the PMSA300I and the Bosch BMP/BME 280 sensors for accurate measurement of ambient temperature, relative humidity, and atmospheric pressure, and a divider that separates the electronics and sensors into two distinct chambers. Power is supplied to the microcontroller via a micro-USB cable connected to a standardized electrical socket.

To ensure the durability and reliability of the monitoring system, two safety measures were implemented. The first involves the physical separation of the electronic circuitry from the sensor array, a feature designed to mitigate the risks associated with potential moisture ingress. The second involves the strategic routing of cable connections through a sealed cable gland, thereby ensuring a hermetic interconnection between the enclosure's compartments. Additionally, a lid designed in the style of a Stevenson Screen was mounted to ensure proper ventilation and protection of the monitoring chamber, while accurately measuring the ambient air.

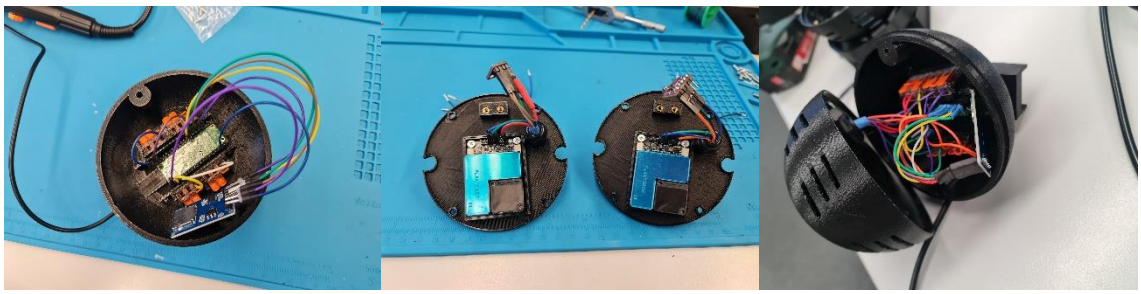


Figure 7: Photographs of the 3D-printed case for the static monitor for one PMSA300I particulate matter sensors - showing the instruments and separated chambers

4.3.5 Enclosure Design for Mobile Monitoring Applications

To complement the capabilities of the static monitoring units, a mobile variant of the sensor system was developed to collect data in motion. This dynamic configuration mirrors the architectural attributes of the static setups, but introduces an auxiliary bracket situated at the base of the enclosure for facile attachment to a vehicular platform. A mountain bike was selected as the vehicular platform for data collection, providing the mobility needed to travel long distances and collect data at diverse locations. The bracket was securely fastened to the handlebars of the bike and further stabilised with a pair of magnets on both the mount and the enclosure, ensuring robust attachment even in rough terrain.

A supplementary pocket was added atop the spherical case to accommodate a 5200 mAh batterie, thereby providing an autonomous energy source for the mobile monitoring unit. The total printing time required for the completion of this mobile configuration was roughly 25 hours.

Figure 8: Photographs of the 3D-printed case for the mobile monitor system - showing the closed case with mount and attached to bike

4.3.6 Sensor Suite: Core Sensors Employed

This section details the sensor modules used for data acquisition. In order to obtain reliable and comparable air quality metrics across diverse urban settings in Wuerzburg and its districts, the study leveraged an array of cost effective sensors. Serving next to the EO products as the analytical foundation of the research, these sensors were integral for generating a comprehensive dataset for the evaluation of atmospheric conditions. A standardization was maintained across all monitoring units by equipping them with homogenous sensor components, thereby ensuring dataset integrity suitable for comparative analysis.

Components	**Simultaneous**	**Static**	**Mobil**
Pico W	1	1	1
PMSA003I	4	1	1
BME280	0	1	1
MicroSD Adapter	1	1	1
Multiplexer	1	0	0
Plug	0	1	1
Batterie	9000 mAH	0 mAH	5200 mAH

Table 2: Components used for each type of monitoring system

4.3.6.1 PMSA003I Air Quality

To capture air quality measurments the study employed multiple PMSA003I dust sensors from Plantower, priced at approximately 50 Euros each. This low-cost sensor uses an optical measurement method that determines the number of suspended particles in the air by laser scattering. The scattered light is then collected and analysed to determine the equivalent particle diameter and distribution of particles of different sizes (PM1.0, PM2.5 and PM10.0) per unit volume. This data is recorded at 2.3-second intervals and expressed in units of mass concentration ($\mu g/m^3$). Communication with the Pico W is established via I2C communication protocol. For data stability, the sensor requires a minimum of 30 seconds after waking from sleep mode to stabilize its internal fan. It operates within a temperature range of -10 to 60 °C and is effective in environments with relative humidity levels up to 99% relative humidity (RH). Given its compact dimension, the PMSA003I is adaptable for both stationary and mobile air quality monitoring systems. Figure 9 provides a top and side view of the sensor, detailing the fan opening and air outlet, while Figure 10 shows its operational mechanism.

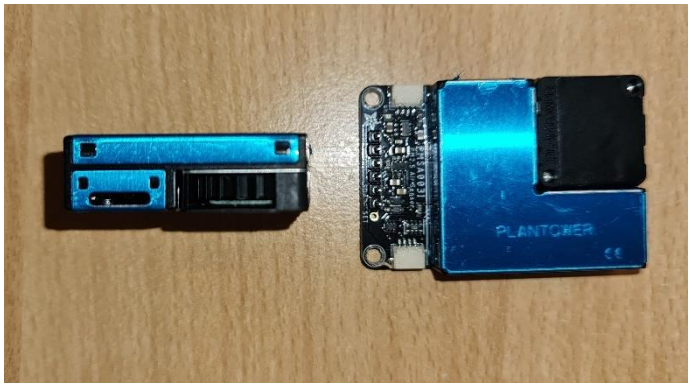


Figure 9: Photograph of a Plantower PMSA033I sensor - top and side view

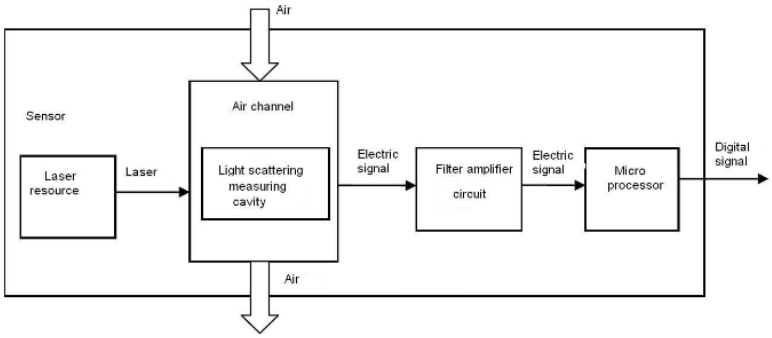


Figure 10: Functional block diagram of PMSA003I [55]

4.3.6.2 Bosch-Sensor BME280

For precise measurement of ambient temperature, relative humidity, and atmospheric pressure, the Bosch BME280 sensor [21] was used. Priced at approximately 15 euros per unit, this multifunctional sensor integrates both a digital humidity sensor and a high-precision temperature and pressure sensor. It operates over an extensive temperature range of -40 to 85°C and can accommodate relative humidity levels from 0 to 100% RH. With a data acquisition interval in the millisecond range, the sensor assures rapid and reliable readings. The BME280 uses a capacitive humidity sensing element for humidity measurement and a thermistor for gauging temperature and pressure. It communicates with the controller via the I2C protocol, ensuring high accuracy and stability in its data. Given its compact form factor and low energy consumption, the sensor is highly suitable for both stationary and mobile environmental monitoring applications. The data collected by the BME280 not only enriches the study's dataset but also serves as a pivotal element for understanding the relationship between air quality and various environmental factors.

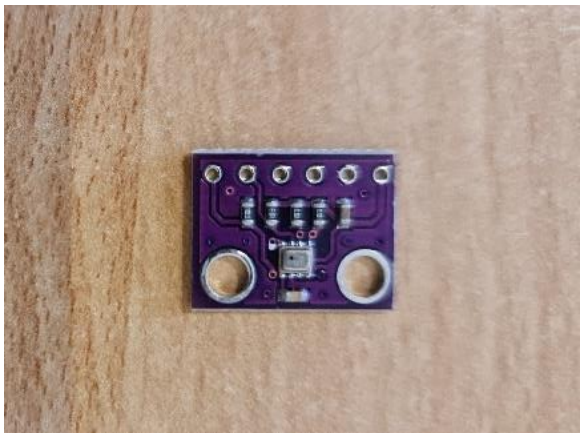


Figure 11: Photograph of a BME280 sensor – top view

V Methodology

The methodology adopted for this research is structured into three main stages. The first stage leverages Earth Observation techniques to examine average tropospheric NO2 levels across five cities in Bavaria, namely Munich, Augsburg, Regensburg, Nuremberg, and Wuerzburg. This approach aims to place the study areas NO2 levels in a broader geographic context. To further understand the exposure, the study investigates the annual, seasonal, and monthly mean levels of tropospheric NO2 for the city and its districts. Sentinel 5P datasets, extracted via Google Earth Engine Cloud Platform, serve as the basis for this temporal analysis. Further, correlation techniques are applied to assess the temporal stability of NO2 levels within the studied period. A focal correlation between pixels is then utilized to examine spatial correlation in NO2 levels across the city.

For a more nuanced analysis, zonal statistics are employed to calculate district-wise mean NO2 levels, thereby ranking Wuerzburg's districts based on their tropospheric NO2 concentrations. The study also incorporates the Urban Atlas 2018 dataset to categorize various regions of the city into clusters of comparable land use types. This information is then correlated with the tropospheric NO2 data to discern any trends. Additionally, a comparison is made between tropospheric NO2 values and a Digital Elevation Model (DEM) to assess how topography might affect tropospheric NO2 levels.

The second stage of the methodology focuses on the evaluation and calibration of low-cost PM sensor systems. Four sensors measuring simultaneously are assessed for their performance relative to the reference system at Wuerzburg Stadtring Süd, using statistical metrics such as the coefficient of determination (R^2) and the root mean square error (RMSE). A random forest-based approach is employed to calibrate these low-cost sensors using hourly temperature and relative humidity data. Subsequently, three fixed stations are deployed in Wuerzburg's districts along with a mobile unit that collects data over a five-day period, sampling three times per day. Correction factors derived from the calibration are applied to generate time-series data for these sensor systems. This data is then correlated with the reference measurements from Wuerzburg Stadtring Süd.

5.1 S5P Timeseries

The following section outlines the methodology employed to investigate tropospheric NO2 concentrations over Wuerzburg and selected cities in Bavaria. Utilizing data from the Sentinel-5P satellite, the study employs a multi-tool approach, incorporating both R for statistical analysis and QGIS for geospatial evaluations. Sub-sections detail the analysis, including district-level assessments, land-use correlations, and topographical influences.

5.1.1 NO2 Levels in Selected Bavarian Cities

This section of the study utilizes Earth Observation methods to examine average tropospheric NO2 concentrations across five Bavarian cities: Munich, Augsburg, Regensburg, Nuremberg, and Wuerzburg. The overarching goal is to contextualize Wuerzburg's tropospheric NO2 levels within a broader geographic and environmental framework. The initial stage of the methodology entailed gathering Sentinel-5P OFFL NO2 tropospheric column number density data via Google Earth Engine (GEE). This raw dataset was confined to a specific time frame, spanning the years 2019 to 2022, with each year considered as a separate entity for subsequent analyses. To enhance data quality, a filter was applied, eliminating scenes with more than 10% cloud cover and thus increasing the reliability of the tropospheric NO2 measurements. After this quality filtering, the dataset's spatial boundaries were adjusted to match those of the selected Bavarian cities. Within these confines, average monthly tropospheric NO2 levels were calculated for each city and compiled into distinct yearly datasets. The final step utilized the R programming language to summarize and rank the cities based on their calculated average monthly tropospheric NO2 concentrations. This ranking offers a comparative metric, allowing Wuerzburg's tropospheric NO2 levels to be evaluated in relation to those of the other cities examined.

5.1.2 Tropospheric NO2 Variability in Wuerzburg

5.1.2.1 Geospatial Visualization of Annual NO2 Trends

Initial insights into the fluctuations in tropospheric NO2 pollution levels over the study period were garnered through a Time Series map, generated using R and QGIS. This visual representation offered a preliminary understanding of the trends and variations in pollution levels across the years under examination.

5.1.2.2 Temporal Correlation of NO2 Concentrations

A Pearson correlation analysis was executed using the R programming language to examine the relationship between the mean tropospheric NO2 concentrations for the years 2019 to 2022. The mean tropospheric NO2 levels were pre-processed using Google Earth Engine and stored as CSV files. In addition to this initial analysis, a focal correlation approach was implemented using a square window of size 5x5 pixels, each having a resolution of 1113.2 m. This method allowed for the incorporation of local spatial context in the temporal correlation analysis. This data was imported into R using the read.csv function, and further transformed and cleaned using the tidyr [58] and dplyr [59] packages. The Pearson correlation coefficient (r) was calculated using the cor() function in R, according to the formula:

$$r = \frac{\Sigma[(x - \bar{x})(y - \bar{y})]}{\sqrt{[\Sigma(x - \bar{x})^2 \Sigma(y - \bar{y})^2]}}$$

Where:

- r is the correlation coefficient between x and y.
- x represents the mean tropospheric NO2 concentrations for a given year.
- y represents the mean tropospheric NO2 concentrations for another year.

5.1.3 District-wise NO2 Analysis in Wuerzburg

5.1.3.1 Ranking and Correlation Analysis Across Districts

To investigate the spatial distribution of tropospheric NO2 concentrations across Wuerzburg, the zonal statistics tool in QGIS was used to calculate annual mean levels for each district from 2019 to 2022. The objectives were to generate a ranking based on these tropospheric NO2 concentrations and to employ Analysis of Variance (ANOVA) [60] to assess both seasonal and monthly fluctuations. This statistical tequnique serves as an effective means for identifying significant disparities in average concentrations across districts.

The equation for the ANOVA is given by:

$$F = \frac{MSbetween}{MSwithin}$$

Where:

- F represents the F-statistic, which serves as an indicator of the statistical significance between group means.
- $MS_{between}$ denotes the mean square between groups, essentially the variance in means across the different districts.
- MS_{within} refers to the mean square within groups, essentially the variance within each respective district.

5.1.4 Urban Atlas Classes and NO2 Correlation Analysis

5.1.4.1 Temporal and Spatial NO2 Patterns in Different Land-Use Classes

To systematically assess the spatial distribution of tropospheric NO2 concentrations across varying land use classes, the analytical strategy delineated in Section 5.1.3.1 for district-level scrutiny was adapted to conform to the specificities of the 2018 Urban Atlas data. The adjustment was necessary to accommodate the high spatial resolution of the dataset. Classes with similar land uses were aggregated into broader categories. The cluster "Vegetation" includes the classes of Green urban areas, Arable land (annual crops), Sports and leisure facilities, Permanent crops (vineyards, fruit trees, olive groves), Pastures, and Forests. The "Urban/Built-up" category comprises Discontinuous low density urban fabric (S.L.: 10% - 30%), Discontinuous dense urban fabric (S.L. : 50% - 80%), Continuous urban fabric (S.L. : > 80%), Discontinuous medium density urban fabric (S.L. : 30% - 50%), and Discontinuous very low density urban fabric (S.L. : < 10%). The "Industrial Areas" category encompasses Industrial, commercial, public, military and private units, Mineral extraction and dump sites, and Construction sites. Lastly, the "Transport" category includes Fast transit roads and associated land, Other roads and associated land, Railways and associated land, Airports, and Port areas. In preparation for the analysis, the Urban Atlas dataset was manipulated in R by first adding an empty column called "cluster". Finally, the new clusters were created by grouping and combining the original land use classes into the categories mentioned. Analysis of Variance (ANOVA) was applied to these clustered Land-Use Classes with the aim of generating rankings based on mean tropospheric NO2 concentrations from 2019 to 2022.

Cluster	Classes
Vegetation	Green urban areas, Arable land (annual crops), Sports and leisure facilities, Permanent crops (vineyards, fruit trees, olive groves), Pastures, Forests
Urban/ Built-up	Discontinuous low density urban fabric (S.L. : 10% - 30%), Discontinuous dense urban fabric (S.L. : 50% - 80%), Continuous urban fabric (S.L. : > 80%), Discontinuous medium density urban fabric (S.L. : 30% - 50%), Discontinuous very low density urban fabric (S.L. : < 10%)
Industrial Areas	Industrial, commercial, public, military and private units, Mineral extraction and dump sites, Construction sites
Transport	Fast transit roads and associated land, Other roads and associated land, Railways and associated land, Airports, Port areas

Table 3: Cluster based on Urban Atlas 2018 classes

5.1.5 Topographical Influence on NO2 Levels

5.1.5.1 Correlation Between Elevation and NO2 Concentrations

To examine the impact of topography on tropospheric NO2 concentrations, a systematic workflow was employed. A Digital Elevation Model with a one meter raster cell size ("DGM1 – GeoTIFF" Bayerische Vermessungsverwaltung – www.geodaten.bayern.de) for Wuerzburg and mean tropospheric NO2 levels with a 1113,2 meter raster cell size spanning the years 2019 to 2022 was used as the data sources for this analysis. In the first step, the DEM was segmented into elevation zones with vertical intervals of five meters using QGIS reclassification tool. This interval was chosen to capture small-scale topographic changes in an urban environment. The zonal statistics tool was then applied to calculate the average concentrations for each elevation zone. The resultant data was exported to R, where it was filtered to remove missing values. A linear regression model was established, formulated as follows:

$$lm = \left(x_{meaη} \sim DN\right)$$

Where:

- X_{mean} represents the mean tropospheric NO2 concentrations for 2019 to 2022.
- DN denotes the elevation zones.

5.2 Sensor Placement

In determining the optimal locations for the three stationary air quality sensor systems, multiple criteria were evaluated. As delineated in Section 6.1, urban districts manifest elevated levels and greater variability of tropospheric NO2, thereby warranting a more nuanced air quality analysis. To align the stationary sensor data with mobile data collection routes for performance evaluation, the sensors were strategically located in proximity to the mobile sampling paths, as well as near power sources and Wi-Fi access points. Given the logistical constraints, mobile tours were capped at a one-hour duration. To capture a comprehensive picture of diurnal air pollution trends, the districts of Altstadt, Steinbachtal, and Frauenland were selected as the sites for the stationary air quality sensors. Specifically, Monitor 2 was situated in Frauenland, adjacent to the high-traffic B4 city ring road. Monitor 3 was installed in Steinbachtal in a residential setting, perched on a balcony along Mergentheimer Straße. Lastly, Monitor 4 was positioned in the Altstadt district, overlooking a parking area near the Government of Lower Franconia. Figure 12 depicts the geographical distribution of the sensor locations.

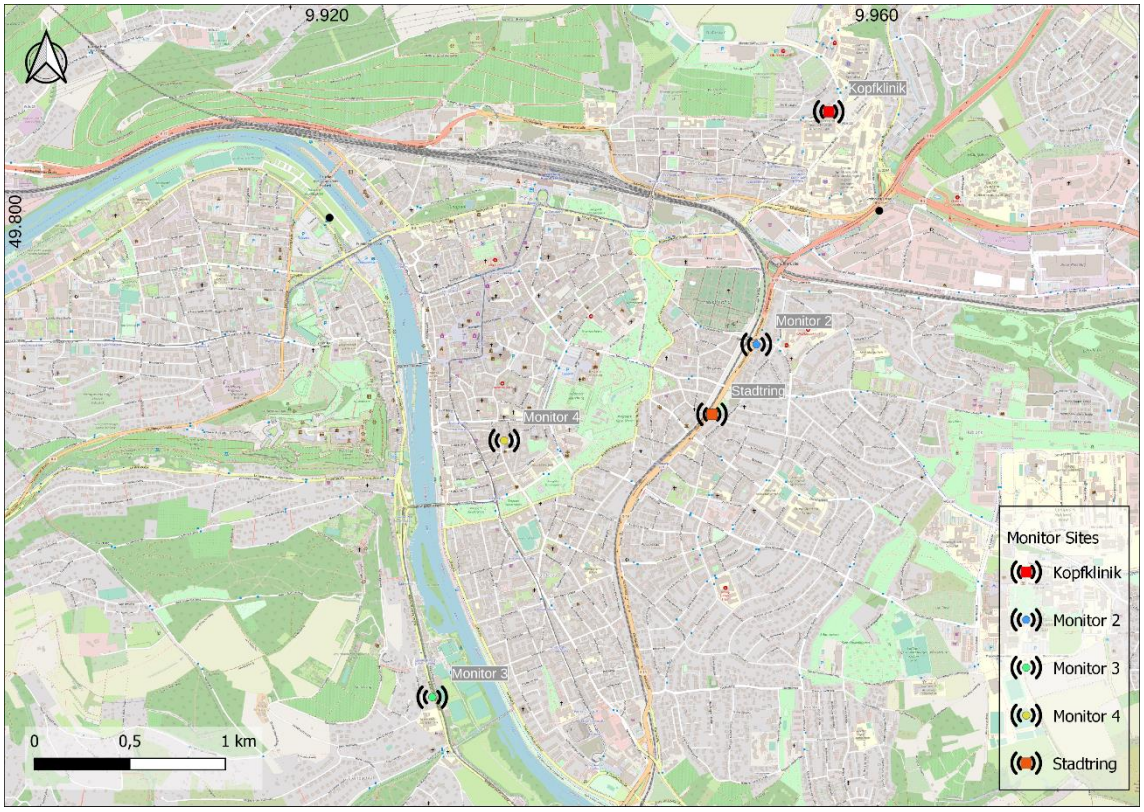


Figure 12: Locations of static Air Quality monitors. Base-Map: Open Street Map [61]

5.3 LCS Data Measurement and Validation

5.3.1 Simultaneous Monitor Method

5.3.1.1 Data Acquisition Parallel System

To assess the performance of the four low-cost particulate matter sensors, they initially were co-located adjacent to the reference station at Wuerzburg Stadtring Süd. Data was collected from February 9, 2023, to February 10, 2023. Figure 13 illustrates the environmental context of the monitoring site. Positioned on the heavily trafficked B4 roadway, the sensor array was strategically placed about five meters from the official measurement station and was adequately protected from environmental elements.

Figure 13: Photographs of the area around the monitor site Stadtring Süd

Data acquisition and storage for measuring particulate matter levels commenced with the connection of a micro-USB port to a battery source, powering on the Pico W controller. The microcontroller, which stores the managing code, initiates operations automatically. Upon powering up, an integrated LED illuminates while the device attempts to establish a Wi-Fi connection and synchronize its real-time clock with network time. Successful synchronization is indicated by the LED extinguishing, and the Wi-Fi chip is subsequently disabled to conserve energy.

Following a 30-second initialization period, the Pico W controller engages its main data collection loop. An LED signal briefly activates at the start of each loop, allowing for visual confirmation of ongoing data collection. Initial PM10 values for each sensor are provisionally recorded as "NA." Utilizing a multiplexer, the system cycles through each sensor to obtain five readings at 2.3-second intervals, subsequently calculating an average value. These averaged PM10 values are stored in distinct variables and replace the initial "NA" placeholders. Timestamping occurs immediately after PM10 value calculation. Data are preserved in CSV format on an SD card, with the date appended to the file name for daily differentiation. Error handling is incorporated through Except options to ensure the data collection loop continues uninterrupted until either the battery is depleted, or the power source is disconnected.

5.3.2 Sensor to Sensor Performance

The examination of the PM10 measurements entailed an in-depth statistical analysis executed within the R programming environment. The initial phase was devoted to data cleaning and pre-processing tasks to assure the integrity of column order and data format. Seasonal time adjustments were made by adding an hour to the timestamp column to align it with wintertime, and records with 'NA' values were excluded. Following this, minute-based data was aggregated to an hourly average, accompanied by the addition of an hourly log. For the purpose of comparative assessment, datasets containing hourly averaged PM10 readings from the Wuerzburg Stadtring Süd data logger were sourced from the publicly accessible database of the Bayerisches Landesamt für Umwelt [62]. These datasets were subsequently filtered and restructured to exclusively retain Wuerzburg station data corresponding to the co-location period. The final analytical dataset was formed by merging these government-sourced readings with data from the low-cost sensors.

To gauge the accuracy of the PMSA003I device, the coefficient of determination, R^2, was computed using the formula:

$$r = \frac{\Sigma[(x - \bar{x})(y - \bar{y})]}{\sqrt{[\Sigma(x - \bar{x})^2 \Sigma(y - \bar{y})^2]}}$$

Here, X represents one PM $sensor_i$ and Y another $sensor_j$. This calculation was iteratively applied to all possible sensor pairings, generating a range of R^2 values between 0 and 1. Values approaching 1 indicate higher sensor-to-sensor consistency.

The second evaluative criterion employed was the Root Mean Square Error (RMSE), utilized to measure each low-cost sensor's performance against the reference governmental station. The RMSE was calculated as follows:

$$RMSE_{ij} = \sqrt{\frac{\sum(LCS_i - RFS_j)^2}{n}}$$

Where LCS_i represents a PMSA003I sensor and RFS_j serves as the reference of Stadtring Süd monitor.

5.3.2.1 Calibration

The third phase of analysis entailed calibration of the low-cost sensor dataset. This was accomplished by generating a linear regression model [63] that employed the hourly-averaged PM10 readings from the four PMSA300I sensors. The model also incorporated ambient temperature and relative humidity data obtained from the reference station. The objective was to derive corrected PM10 values based on these sensor and environmental parameters, using the PM10 readings from Stadtring as the dependent variable. After linear regression modelling, a machine-learning approach was implemented by creating a random forest model [64]. This was applied to the set of instruments engaged in simultaneous data logging with the aim of determining a correction factor.

This correction factor was subsequently calculated by dividing the predicted PM10 values by the mean PM10 values in the test dataset. The last step of the calibration process involved applying the derived correction factor to the raw PM10 mean readings. The initial linear regression model served as the calibration tool for this final adjustment.

5.3.3 Static Stations

5.3.3.1 Data Acquisition of the Static System

As described in section 5.2 a constellation of three of the low-cost sensors was strategically deployed across the named districts within the scope of the research area to monitor PM10 levels within the focused area. These sensors engaged in a continuous five-day data collection regimen. The Pico W controller was powered up by connecting the micro-USB port to an electrical source, and it initiated a code sequence pre-stored in its microcontroller to manage the instrument operations. Upon initialization, the device sought to establish a Wi-Fi connection to a predefined network and synchronized its Real-Time Clock with the network time. A 30-second preparatory phase was allocated to prime the PM sensor for data collection. Unlike the mobile sensors, these static stations featured an additional functionality of live data streaming. Utilizing MQTT Protocol, the control unit initiated an SSL connection to the AWS IoT core for this purpose. As a result, LED indicators were unnecessary and thus excluded from the design. The Pico W controller initiated the primary data collection loop. For each iteration of the loop, placeholders were created to store the values for PM10, ambient temperature, relative humidity, and pressure, initially marked as "NAN." A series of 25 PM sensor readings were taken at intervals of 2.3 seconds, after which the BME280 sensor recorded the ambient conditions. All placeholder variables were then updated, assuming no error emerged during this data acquisition cycle. The current date and time were also stored. The acquired data was saved in CSV format on an SD card, appending the current date to the filename for differentiation. Following the local storage, the dataset was transmitted to a database via MQTT. Any exceptions or errors encountered during this process were addressed through error-handling mechanisms, ensuring the uninterrupted flow of the data collection loop until either the power supply was exhausted or manually terminated.

5.3.3.2 Data Visualisation

For an enhanced comprehension of the spatial variability in PM10 levels across the selected districts, time-series plots were generated for each monitoring station within the designated study period. To provide a contextual baseline, these plots were augmented with corresponding PM10 measurements sourced from both governmental monitoring stations located in Wuerzburg.

5.3.4 Mobile System

To secure high-resolution spatio-temporal data, the mobile data collection system was configured for high-frequency readings. This system functioned in a manner analogous to its stationary counterpart, albeit with a modification: a quicker sampling interval comprised of two readings within 2.3 seconds for averaging PM10 levels, followed by immediate logging of BME280 environmental variables and data storage.

Due to the constraints of the Pico W microcontroller in interfacing with multiple external peripherals and software libraries, a supplementary approach was devised to capture location data. The GPS functionality of a Samsung Galaxy Note 20 Plus, utilized in conjunction with the Geotracker GPS Logging application [65], enabled the recording of geographical coordinates at one-second intervals. Subsequent preprocessing and filtering of these datasets occurred in R, where they were merged based on date and time. A snapping function facilitated the alignment of timestamps, accommodating minor discrepancies by correlating entries within a five-second buffer.

Structured to encompass a range of districts, the mobile data collection tours had an approximate length of 11 km as shown in Figure 15. Commencing at Monitor 4 in Altstadt, each tour proceeded to Monitor 3 in Steinbachtal, then to Monitor 2 in Frauenland, before retracing back to the point of origin at Monitor 4. To guarantee the integrity of data for comparative analysis with static sensors, intermittent stationary periods were instituted in proximity to each static sensor.

Tri-daily data collections - morning, noon, and evening - were performed to enrich the dataset, offering insights into the diurnal and spatial variations of PM10 levels. Figure 14 details the described data collection and processing steps in a Flowchart.

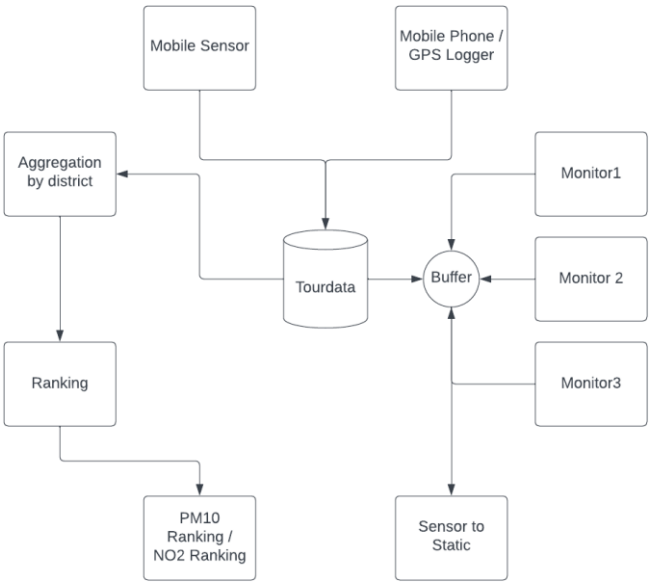


Figure 14: Flowchart of mobile system

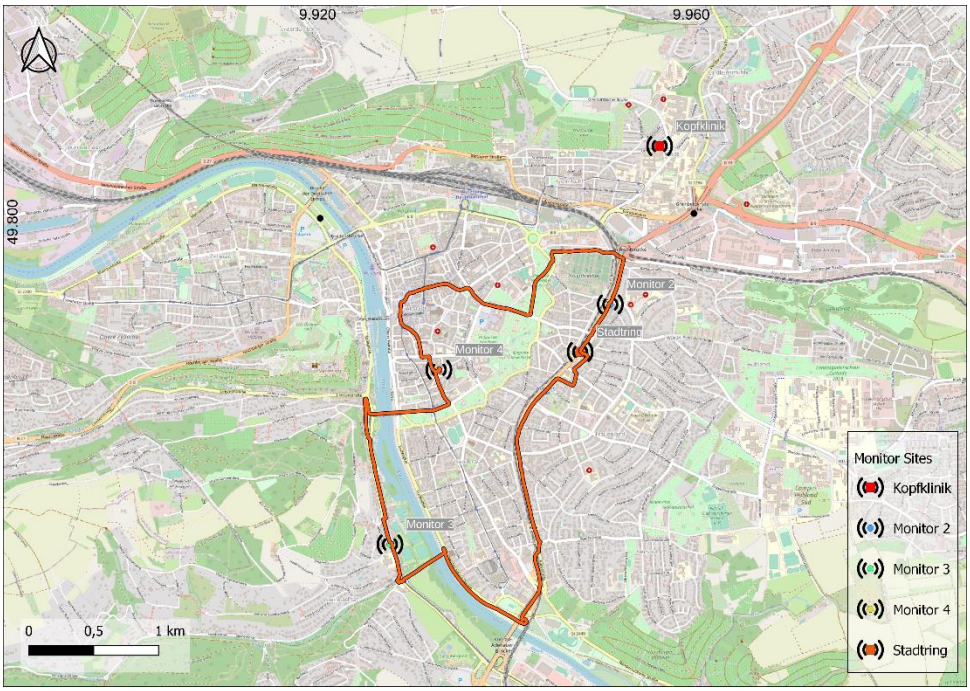


Figure 15: Mobile path. Base-Map: Open Street Map [61]

5.3.4.1 Mobile Analysis

The evaluation of the mobile datasets was conducted through two principal methodologies. The first approach entailed a spatio-temporal comparison with the fixed monitoring stations. A 50-meter spatial buffer was defined around each stationary sensor, capturing instances when the mobile sensor entered this vicinity. Data captured by both monitoring systems within this area were partitioned into discrete subsets. The arithmetically mean of these subsets were calculated to yield a singular PM10 value for each of the mobile and stationary systems. To enhance the robustness of the analysis, these individual PM10 values were subjected to further aggregation across the trio of daily tours, thereby yielding a comprehensive perspective on the spatio-temporal discrepancies between the mobile and fixed monitoring systems. The second analytical technique focused on comparing the mobile sensor data against district-level tropospheric NO2 rankings.

IV Results

6.1 Sentinel 5P NO2 Analysis

6.1.1 Comparative Analysis of NO2 Levels Across Selected Bavarian Cities

Figure 16 depicts the monthly variations in NO2 concentrations for the years spanning 2019 to 2022. Seasonality in NO2 levels is observed across all cities, with elevated concentrations typically seen during the colder months. The concentrations range from a minimum of approximately 20 to a maximum of about 144 µmol/m². More specifically, the winter months show tropospheric NO2 levels ranging from 20 to 144 µmol/m², whereas the summer months show a more constrained range from about 25 to approximately 60 µmol/m².

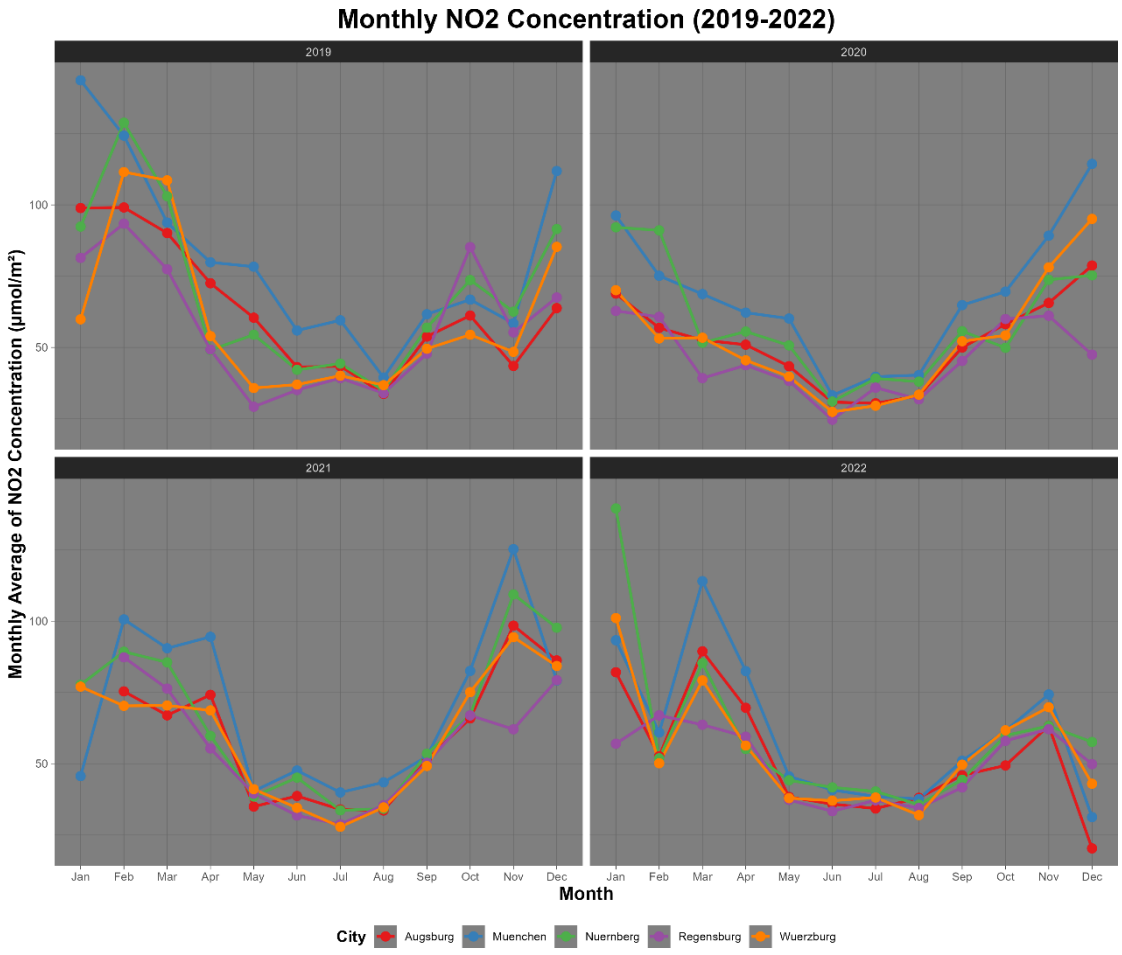


Figure 16: Monthly tropospheric NO2 levels (2019-2022)

Figure 17 illustrates that Munich emerges as the city with the consistently highest average monthly tropospheric NO2 concentrations for the studied years 2019 to 2022 among the selected Bavarian cities. Nuremberg frequently ranks second.

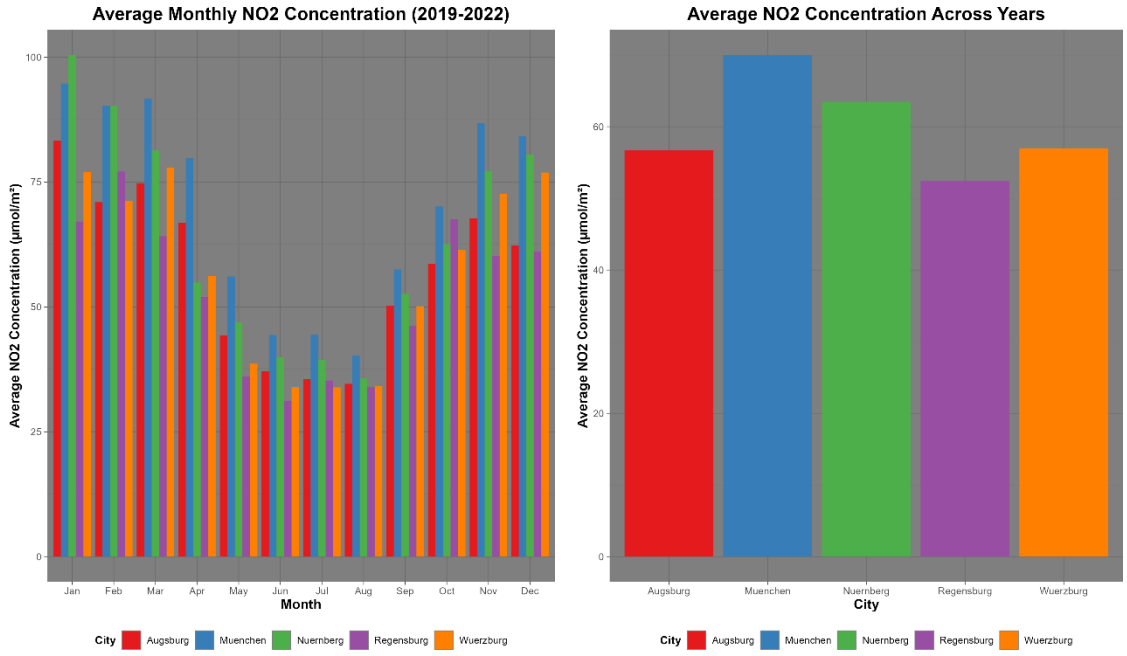


Figure 17: Monthly and yearly average tropospheric NO2 levels (2019 -2022)

Table 4 further validates these observations, highlighting that Munich and Nuremberg have the highest average tropospheric NO2 concentrations over the four-year period, recorded at 70 µmol/m² and 63.5 µmol/m², respectively. They are followed by Augsburg and Wuerzburg, with averages of 56.7 µmol/m² and 57 µmol/m², while Regensburg exhibits the lowest four-year average concentration of 52.4 µmol/m².

City / Year	2019	2020	2021	2022	AVG
Augsburg	63,6	51,6	59,9	51,6	56,7
Munich	81,1	67,8	70,2	60,9	70
Nuremberg	69,5	58,7	65,9	59,9	63,5
Regensburg	57,9	45,9	55,7	50,1	52,4
Wuerzburg	60,1	52,7	60,6	54,6	57

Table 4:Annual citywide average tropospheric NO2 levels (2019-2022)

6.1.2 NO2 Variability in Wuerzburg

6.1.2.1 Visualizing Annual NO2 Trends in Wuerzburg

Figure 18 displays the annual average tropospheric NO2 concentrations in Wuerzburg for the years 2019 to 2022. The raster cell size is 1113.2 meters. To ensure comparability across the four years, a consistent range from 49 to 62 µmol/m² is applied. A visual assessment reveals that areas with elevated concentrations are primarily situated within the city's urban core. In contrast, the outskirts of Wuerzburg tend to show lower concentrations, underlining a noticeable spatial differentiation in tropospheric NO2 levels across the city. It is worth noting that the average values presented in Table 5 are slightly different from those obtained through Google Earth Engine Table 4, due to the larger analysis area covered in the R-based calculations. However, these values are characterized by narrow minimum and maximum ranges for the analysed years 2019 to 2022. The year 2020 specifically showed the lowest average concentrations across all the cities when compared to the other years in the dataset.

Metric / Year	**2019**	**2020**	**2021**	**2022**
Min	53.4	43.2	50.5	49.2
Max	61.3	49.7	61.8	56.4
Avg	57.4	46.5	56.2	52.8

Table 5: Summary of annual minimum, maximum and average NO2 concentrations in Wuerzburg (2019-2022)

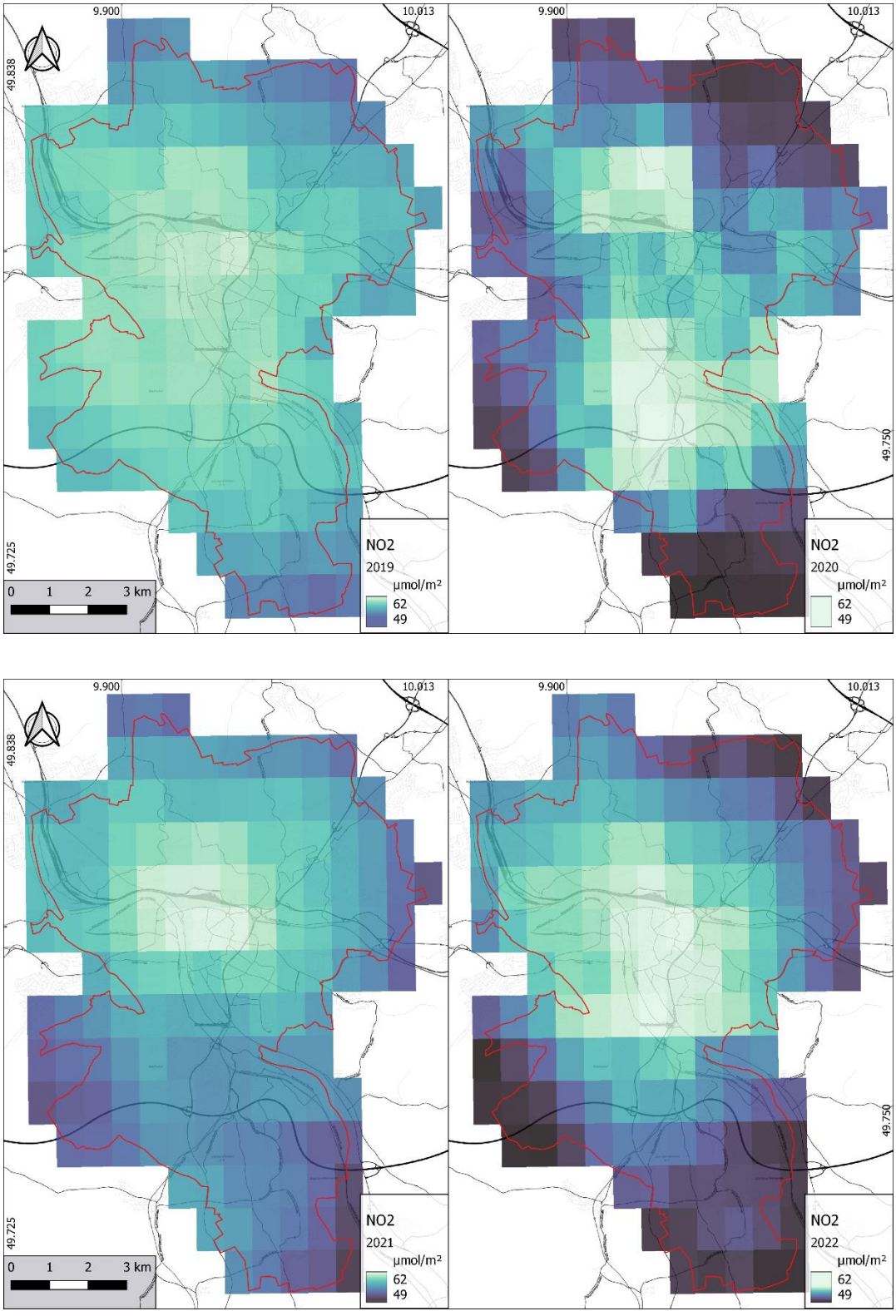


Figure 18: Annual average tropospheric NO2 concentrations for Wuerzburg (2019-2022)

6.1.2.2 Temporal Consistency of NO2 Concentrations in Wuerzburg

The Correlation Analysis was conducted using the Pearson method, and the results are displayed in Figure 19. The matrix employs a colour gradient that ranges from -1, represented in red, to +1, depicted in green. This matrix serves to illustrate the correlation between the average NO2 values for the years 2019, 2020, 2021, and 2022. The analysis reveals a mixed pattern of correlations among the datasets; notably, the NO2 levels in 2022 and 2021 have a correlation coefficient of -0.135, the levels in 2022 and 2020 show a correlation coefficient of -0.223, and the levels in 2022 and 2019 exhibit a correlation coefficient of -0.798. Meanwhile, the correlation coefficient for NO2 levels in 2021 and 2020 is 0.996, the coefficient for levels in 2021 and 2019 is -0.489, and the coefficient for levels in 2020 and 2019 is -0.410. These results suggest a lack of consistent positive correlation between the average NO2 levels across the years studied.

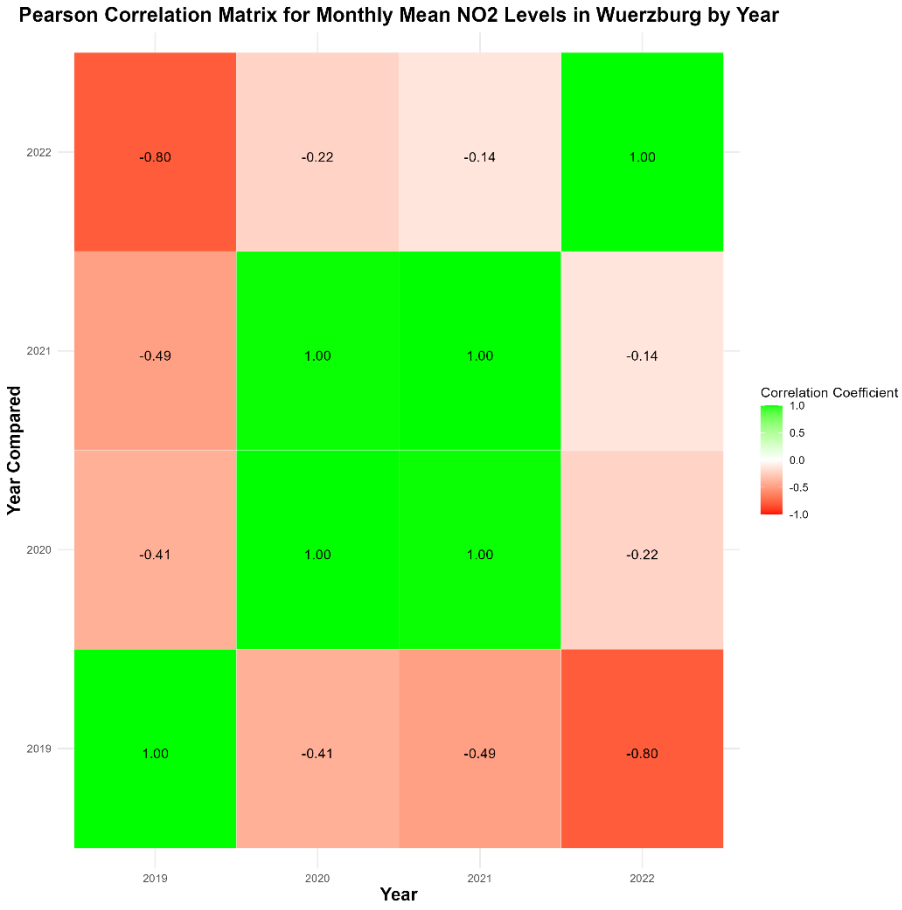
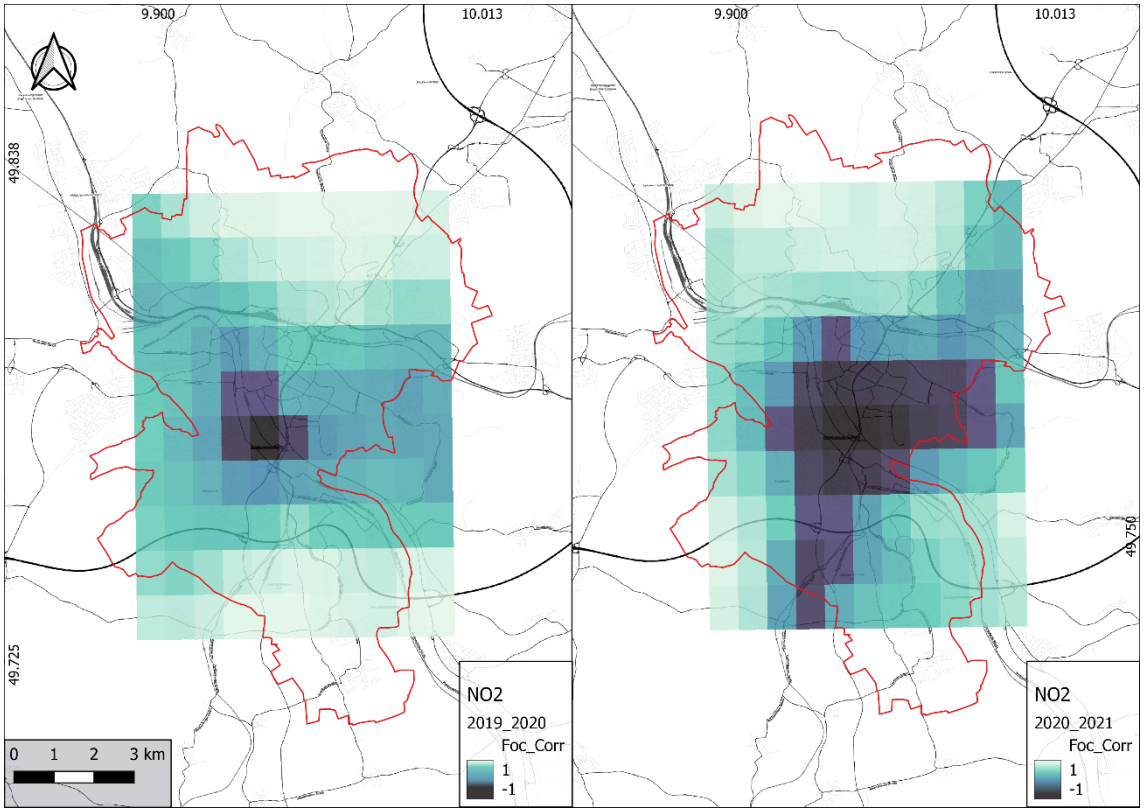
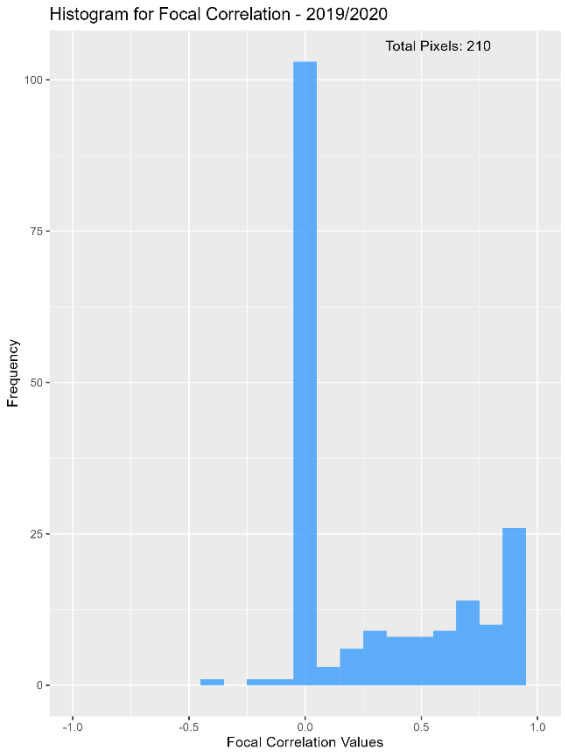
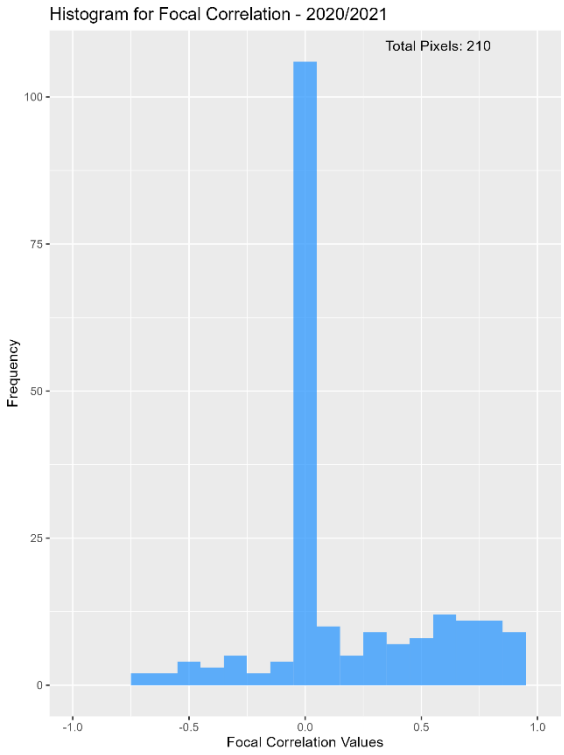


Figure 19: Result of the Pearson Correlation analysis (2019-2022)

The focal correlation analysis shown in Figure 20 provides insight into the relationship between the tropospheric NO2 concentrations in the years 2019 to 2022. The correlation coefficient calculated using a 5x5 moving window, ranges from -1 to 1, where -1 indicates a perfect negative correlation, one indicates a perfect positive correlation, and 0 indicates no correlation. The maps display a spatial representation of the relationship between the tropospheric NO2 levels, highlighting areas where the concentrations tend to change in similar or opposite directions. It is crucial to note that the resultant correlation raster maps exhibit a smaller spatial footprint compared to the original NO2 datasets. This reduction is due to the averaging of values over the 5x5 pixel area by the moving window, thereby eliminating some peripheral data and resulting in a compact dataset. The results of the analysis reveal distinct spatial patterns in correlation coefficients. For the 2019-2020 map, negative correlations are evident in the areas of Steinbachtal, Frauenland, Sanderau, and Altstadt. The 2020-2021 map displays the highest density of negative correlation pixels, particularly concentrated in centralized districts of the city. Lastly, the 2021-2022 map shows negative correlations exclusively in the City Centre Altstadt and in Heuchelhof. In all three maps, the northern and southern outskirts, which are more rural in character, consistently exhibit positive correlations. Within the city, negative and near-zero values predominate. Importantly, near-zero correlation coefficients, indicative of little to no relationship between NO2 concentrations, are the most frequent values across all maps, covering over 100 of the total 206 pixels.


9.900
10.013
49.838
49.725
49.750
0 1 2 3 km
NO2
2019_2020
Foc_Corr
1
-1
NO2
2020_2021
Foc_Corr
1
-1


Histogram for Focal Correlation - 2019/2020
Total Pixels: 210
Frequency
Focal Correlation Values
100
75
50
25
0
-1.0
-0.5
0.0
0.5
1.0


Histogram for Focal Correlation - 2020/2021
Total Pixels: 210
Frequency
Focal Correlation Values
100
75
50
25
0
-1.0
-0.5
0.0
0.5
1.0

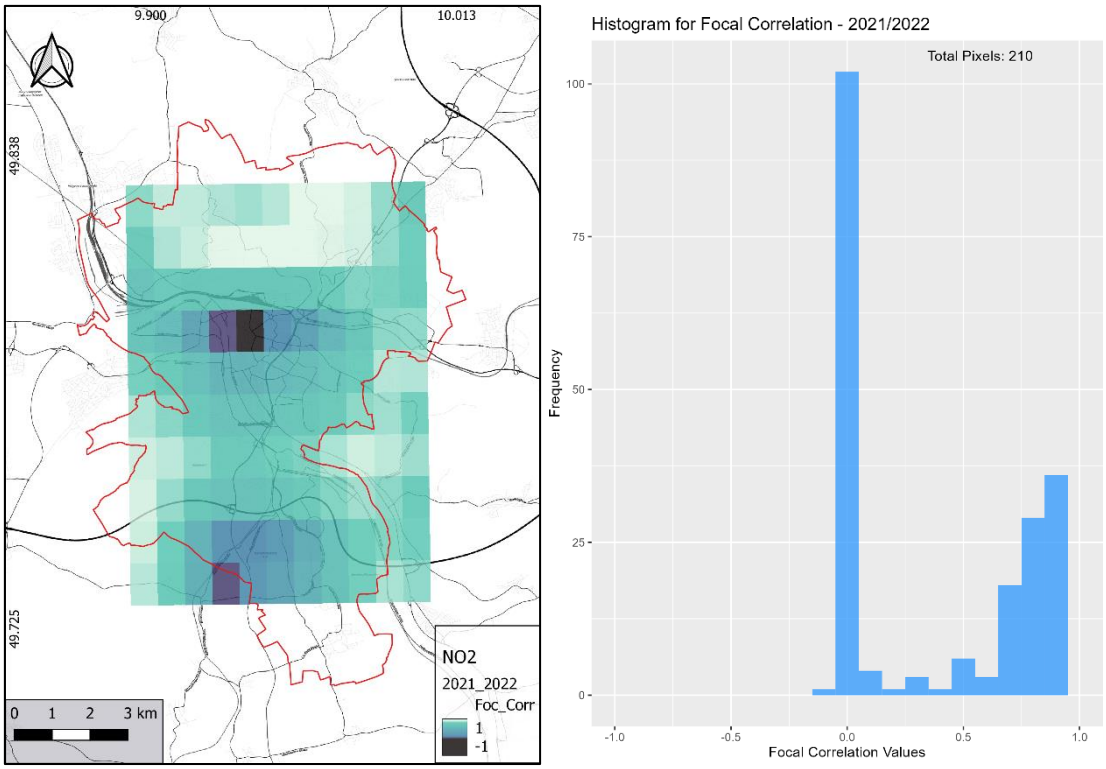


Figure 20: Result of the Focal Correlation analysis (2019-2022)

6.1.3 Zonal Statistical Analysis: Wuerzburg's Districts

6.1.3.1 District Ranking

Figure 21 shows a cartographic representation of the 13 Würzburg districts, providing the spatial information for the following statistical analyses. The districts are subjected to a ranking procedure as well as an analysis of variance (ANOVA) to explore both the spatial and seasonal variations in tropospheric NO2 concentrations.

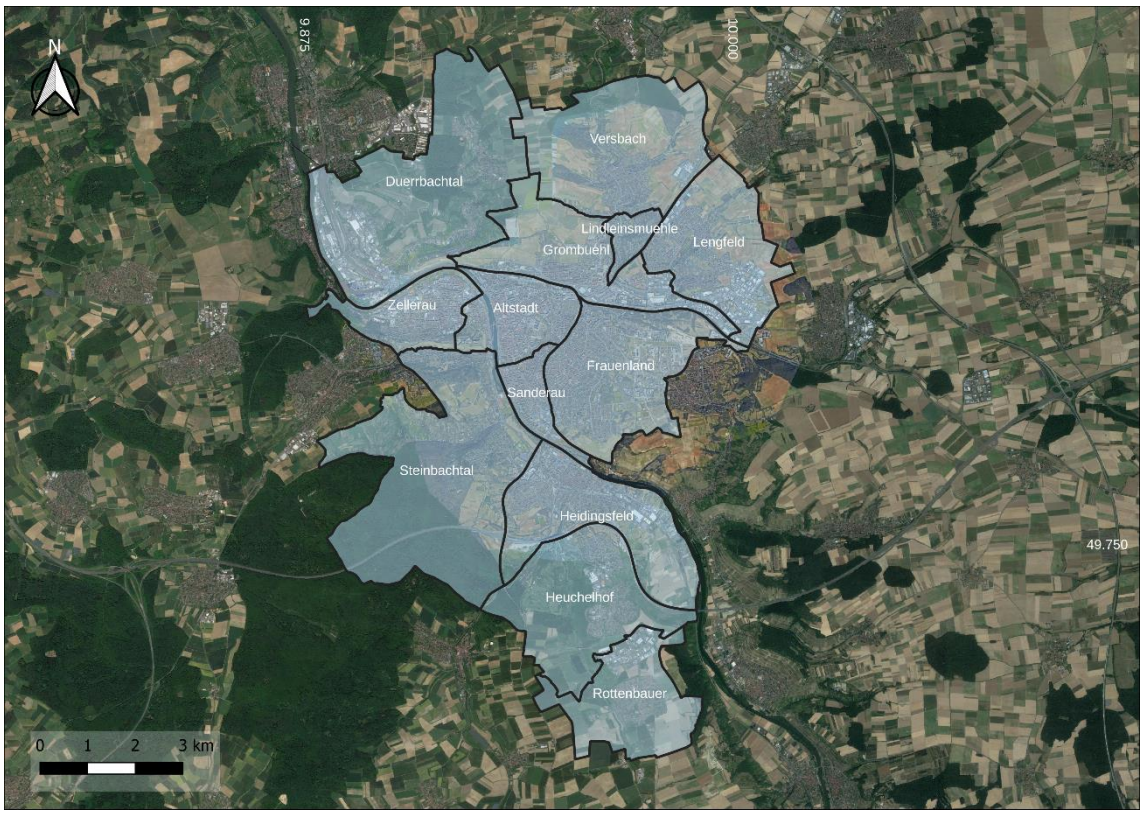


Figure 21: Districts of Wuerzburg [66]. Base-Map: Google Satellite 2023 [67]

Figure 22 reveals minimal intra-annual and inter-annual variation in mean tropospheric NO2 levels across all thirteen districts of Wuerzburg. This consistency in data points to a stable NO2 presence over the observed period. An exception to this pattern is observed in the year 2020, where mean tropospheric NO2 were noticeably lower.

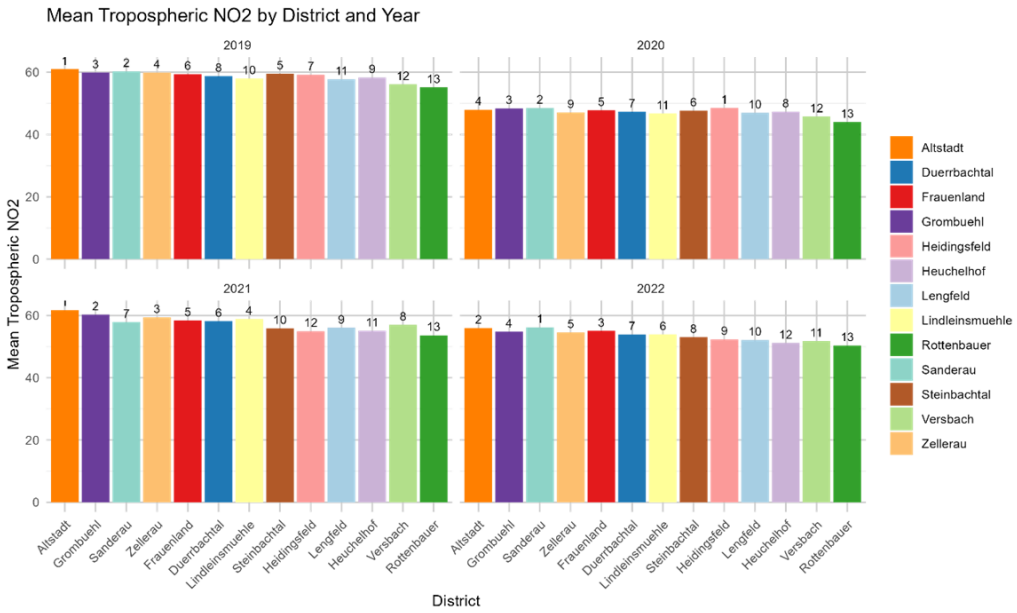


Figure 22: Mean tropospheric NO2 by district and year (2019-2022)

Figure 23 presents the aggregated data for all years to facilitate the ranking of Wuerzburg's districts based on their overall mean tropospheric NO2 levels. The figure reiterates the minor variations in NO2 concentrations across districts, confirming a pattern of consistency over the study's entire time frame. Spatially, it is notable that districts with the highest mean NO2 levels are concentrated in the central part of Wuerzburg.

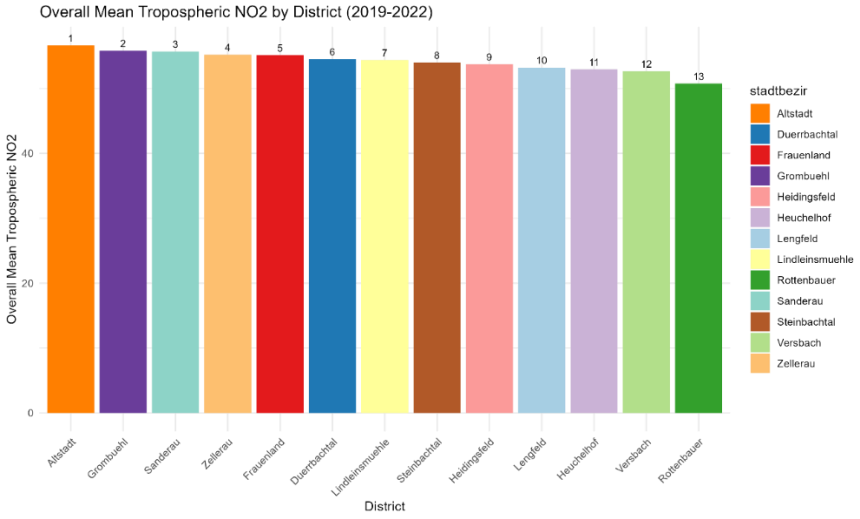


Figure 23: Mean tropospheric NO2 by district (Aggregated 2019-2022)

Figure 24 displays the aggregated seasonal data, providing a more detailed look at the spatiotemporal variations in tropospheric NO2 levels across Wuerzburg. The data reveal seasonal fluctuations, with summer levels ranging from 30 to 37 µmol/m² and winter levels spanning from 50 to 90 µmol/m² over the years studied. The spatial distribution of these concentrations aligns with the patterns observed in Figure 23 that districts situated in the central part of the city consistently exhibit higher tropospheric NO2 levels, while those further from the city centre show lower concentrations.

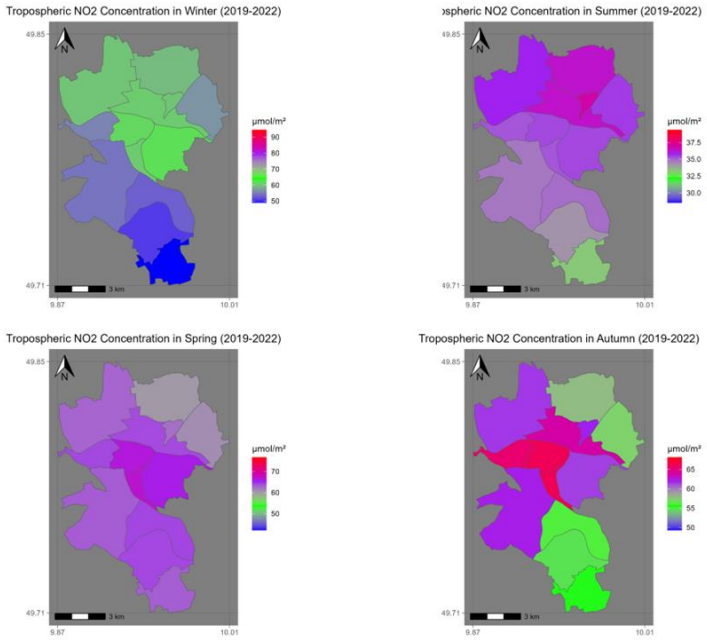


Figure 24: Seasonal Mean tropospheric NO2 by district (Aggregated 2019-2022)

6.1.3.2 Correlation Based on NO2 Exposure

6.1.3.2.1 Seasonal Differences

Table 6 displays the Analysis of Variance (ANOVA) results used to quantify seasonal variations in mean tropospheric NO2 levels. The F-Statistic is 230.4 and the p-Value is less than 2^{-16}.

Parameter	Value
F-Statistic (F)	230.4
p-Value (Pr(>F))	$<2^{-16}$

Table 6: ANOVA results for seasonal differences in tropospheric NO2

6.1.3.2.2 District-Level Seasonal Differences

Table 7 illustrates the findings of the ANOVA conducted to investigate if mean NO2 concentrations differ by district during different seasons. The table reveals an F-Statistic of 1.865 and a p-Value of 0.173.

Parameter	Value
F-Statistic (F)	1.865
p-Value (Pr(>F))	0.173

Table 7: ANOVA results for seasonal and district wise differences in tropospheric NO2

6.1.3.2.3 Monthly Differences

Table 8 outlines the ANOVA results for exploring monthly fluctuations in mean NO2 levels. The data shows an F-Statistic of 86.609 and a p-Value less than 2^{-16}.

Parameter	Value
F-Statistic (F)	86.609
p-Value (Pr(>F))	$<2^{-16}$

Table 8: ANOVA results for monthly differences in tropospheric NO2

6.1.4 Urban Atlas Classes and their NO2 Profile

6.1.4.1 NO2 Patterns across different Land-Use Classes

Figure 25 visualizes the Urban Atlas land-use classes from 2018, along with the mean tropospheric NO2 levels measured in µmol/m² (averaged for the years 2019 to 2022) overlaid on top. The range of tropospheric NO2 levels falls between 50 and 57 µmol/m². Each pixel, with dimensions of 1113.2 m x 1113.2 m, encompasses a mix of different land-use classes. Elevated NO2 levels are observed in the central area of the city.

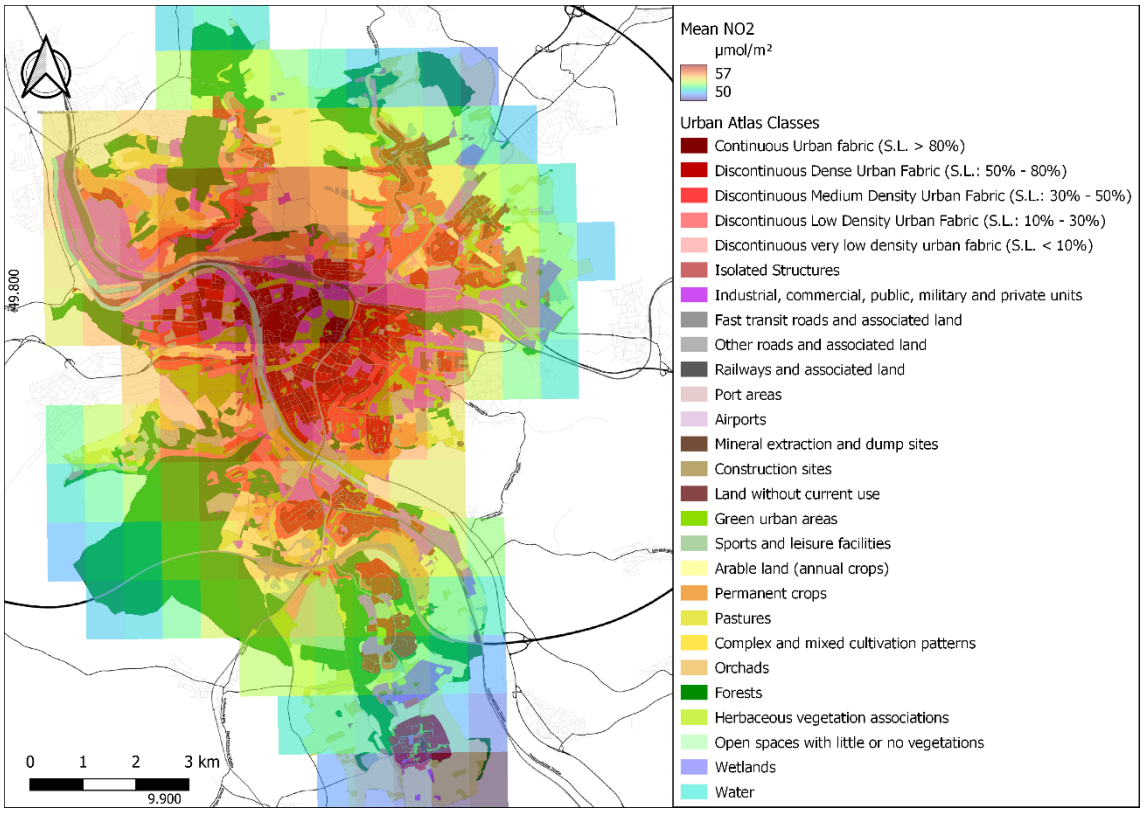


Figure 25:Yearly mean tropospheric NO2 concentration for Wuerzburg (Aggregated 2019-2022) overlayed by Urban Atlas classes. Base-Map: CartoDB Inc. [44]

To address the challenges of mixed land-use categories within the NO2 pixels, clusters were formed as outlined in Table 3. These clusters were then sorted based on their mean tropospheric NO2 measurements from 2019 to 2022. Table 9 outlines the mean, median, and standard deviation of tropospheric NO2 levels across different Urban Atlas clusters. The data shows minimal variation in average NO2 levels among the aggregated land-use categories.

Cluster	Mean	Median	SD
Urban / Built-up	54.9	55.2	1.30
Industrial Areas	54.5	54.7	1.42
Transport	54.5	54.5	1.07
Vegetation	54.3	54.5	1.45
Other / Uncategorized	54.2	54.5	1.29

Table 9: Tropospheric NO2 by Urban Atlas Cluster (Aggregated 2019-2022)

Figure 26 provides a violin plot illustrating the minimum, mean, and maximum tropospheric NO2 concentrations for each of the four examined clusters. The plot shows that Industrial Areas have a minimum tropospheric NO2 level of 49.7 µmol/m², a mean of 54.5 µmol/m², and a maximum of 56.9 µmol/m². The Transport cluster exhibits a minimum concentration of 51.0 µmol/m², a mean of 54.5 µmol/m², and a maximum of 56.9 µmol/m². Lastly, Urban/Built-up areas have a minimum of 49.7 µmol/m², a mean of 54.9 µmol/m², and a maximum of 56.9 µmol/m². Similarly, the Vegetation cluster shows a minimum level of 49.5 µmol/m², a mean of 54.3 µmol/m², and a maximum of 56.9 µmol/m².

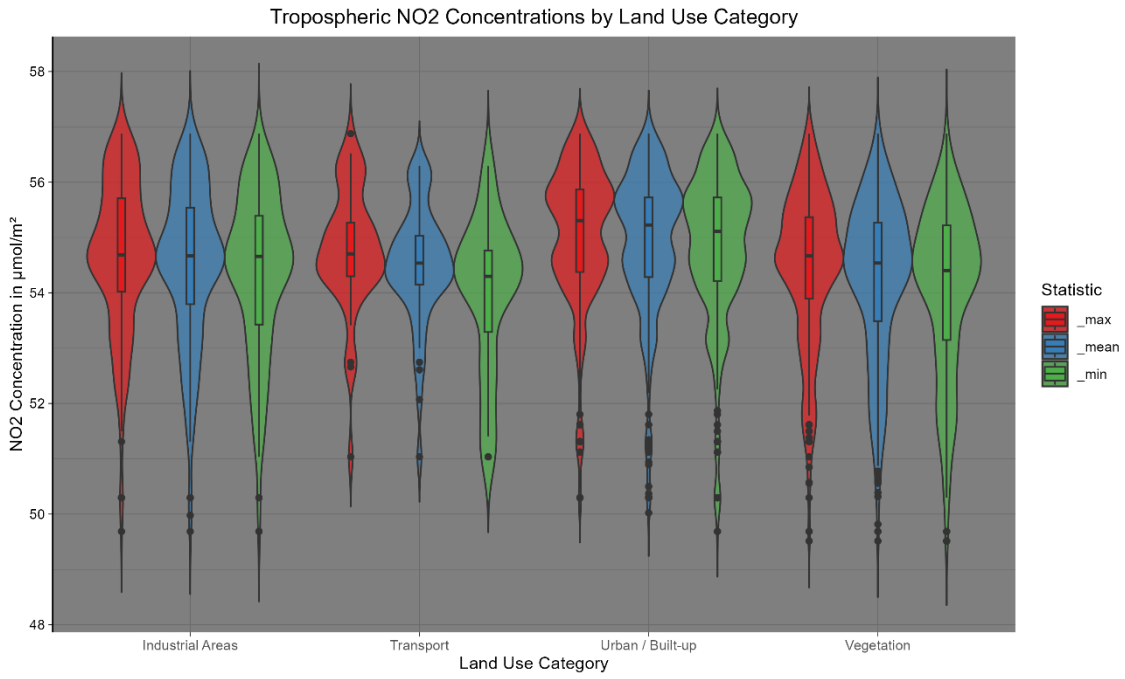


Figure 26: Tropospheric NO2 by Urban Atlas Cluster (Aggregated 2019-2022) - Violin Plot

6.1.5 Topographical Influence on NO2 Distribution

6.1.5.1 Elevation-NO2 Correlation Analysis

The elevation range of the processed DEM spans approximately from 165 to 360 meters above sea level. Over the four-year period from 2019 to 2022, mean tropospheric NO2 concentrations were observed to vary within a narrow range, approximately 49.5 to 57 µmol/m². Figure 27 offers a three-dimensional representation of Wuerzburg's elevation zones, reclassified at 5-meter vertical intervals. These zones are color-coded according to NO2 concentration, transitioning from low (purple) to high (red). The visualization shows that higher tropospheric NO2 levels are predominantly associated with lower elevation zones.

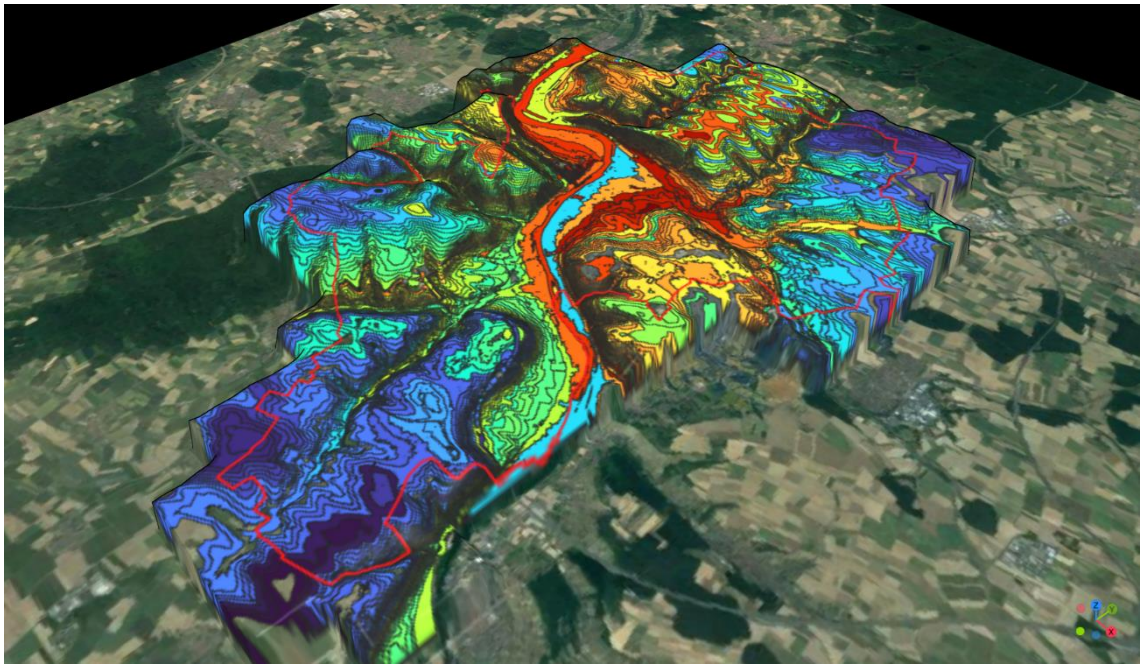


Figure 27: 3-Dimensional view of tropospheric NO2 concentration by height (Aggregated 2019-2022)

In Figure 28, a scatterplot visualizes the results of the linear regression model. The intercept is approximately 55.10, serving as the estimated mean NO2 level at an elevation of 165 meters. The slope for the elevation variable is -0.0813, indicating that as elevation rises by five meters, the mean tropospheric NO2 levels correspondingly decrease by about 0.0813 µmol/m². The multiple R-squared value stands at 0.2262, revealing that elevation accounts for 22.62% of the variance in the mean NO2 levels.

High F-statistic and a p-value less than 2.2^{-16} affirm the statistical significance of the model. Additionally, the scatter of residuals around the regression line is minimal, suggesting a good fit of the model to the data.

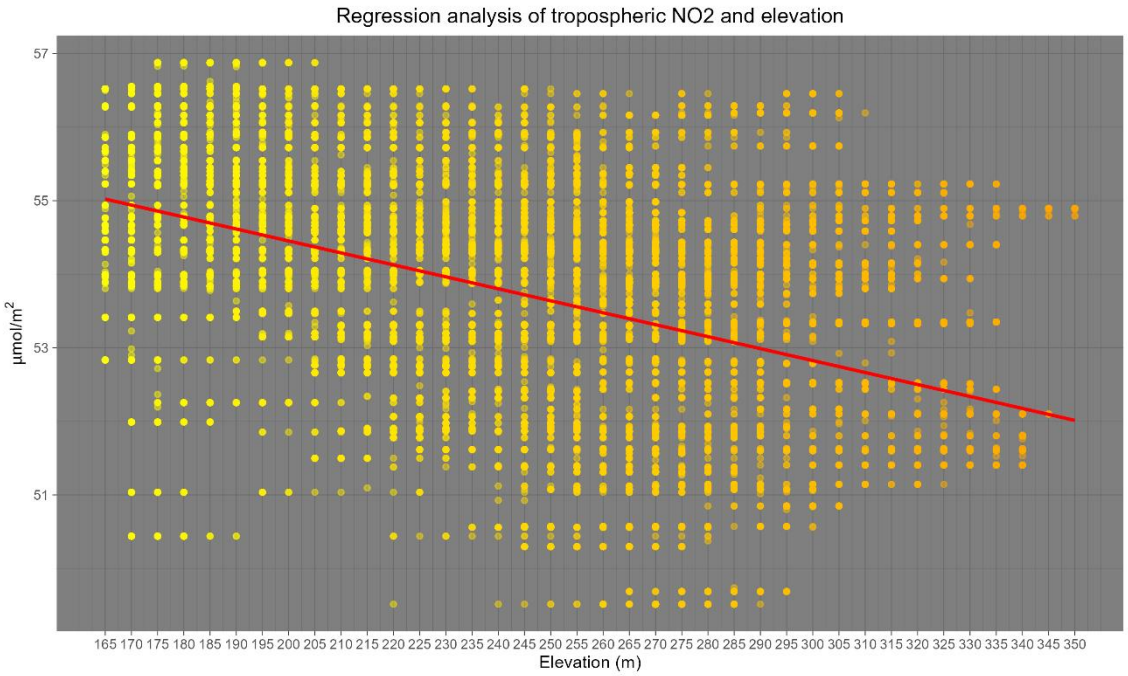


Figure 28: Regression Analysis of tropospheric NO2 and elevation

6.2 Low-cost Sensor Systems

This section presents an analytical overview of the performance metrics obtained from the self-developed low-cost sensor systems. Several metrics, such as root mean square error (RMSE) and coefficient of determination (R^2), were evaluated to assess the sensors' reliability and accuracy.

6.2.1 Sensor to Sensor Performance:

Figure 29 displays the calculated R^2 results for the four PMSA300I low-cost sensors that operated simultaneously. Data was gathered on February 9, 2023, over a span of 24 hours. Calculations for all possible sensor pairings were conducted. The R^2 values ranged between 0.9109 and 0.9251, indicating low variability between the sensors during the study period.

PM10_1 and PM10_2: R-squared = 0.9177

PM10_2 and PM10_3: R-squared = 0.9161

PM10_1 and PM10_3: R-squared = 0.9251

PM10_2 and PM10_4: R-squared = 0.9109

PM10_1 and PM10_4: R-squared = 0.9176

PM10_3 and PM10_4: R-squared = 0.9203

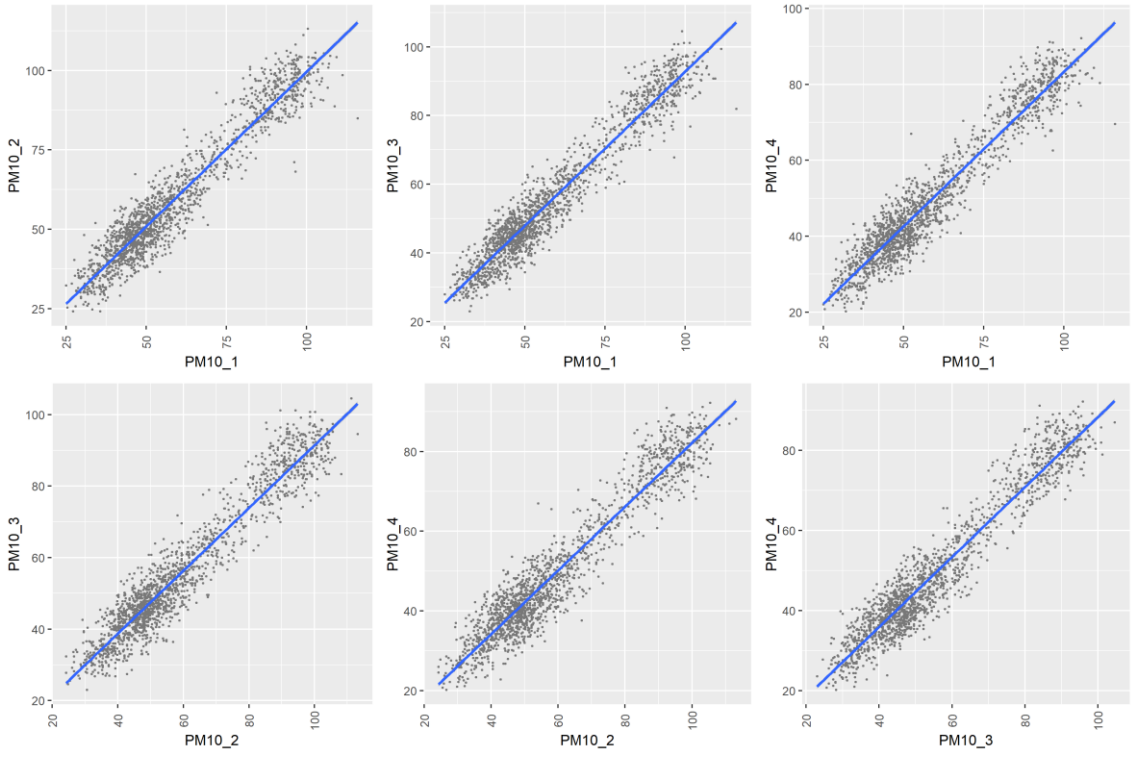


Figure 29. Sensor to sensor performance

6.2.2 Sensor 2 Reference Performance

In this paragraph, the performance of the low-cost sensors is evaluated by comparing it with the values of the reference station Satdtring Süd. The root mean square error (RMSE) was calculated to quantify the average difference between each low-cost sensor and the reference station on an hourly basis, after the sensor data was averaged hourly to align with the reference station data. The RMSE scores for the air pollutant PM10 from each sensor vs. the reference station are as follows:

RMSE PM10_1_Mean = 19.23607 RMSE PM10_2_Mean = 17.94657

RMSE PM10_3_Mean = 19.42284 RMSE PM10_4_Mean = 16.16234

High RMSE values indicate poor agreement between the low-cost sensors and the reference station. A separate analysis was conducted to evaluate the relationship between the Low-cost Sensors and the reference station Wuerzburg Stadtring using a linear model. Despite the sensors showing high consistency among themselves, the relationship between the low-cost sensors and the reference station is only moderate aligned. This is indicated by an R^2 value of 0.441 and visualised in Figure 30.

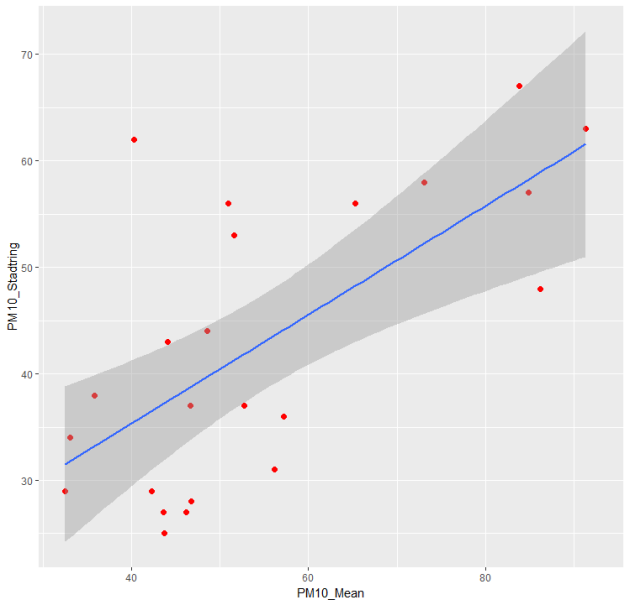


Figure 30: Sensor to reference station performance without calibration

6.2.3 Sensor Calibration

A calibration model was implemented utilizing a random forest algorithm. The model integrated PM10 concentrations, ambient temperature, and relative humidity data from the reference station with PM10_Mean readings from all four low-cost sensors. The result of the corrected dataset is visualized in Figure 31 and Figure 32. Post-calibration, the coefficient of determination (R^2) improved to 0.83.

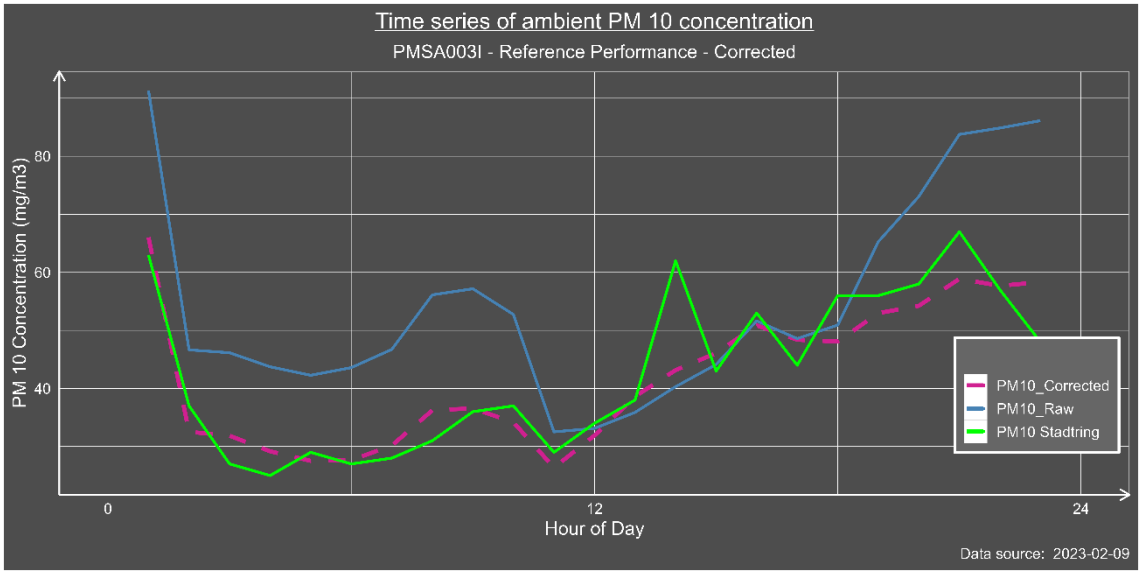


Figure 31:Comparison of raw and. corrected sensor readings

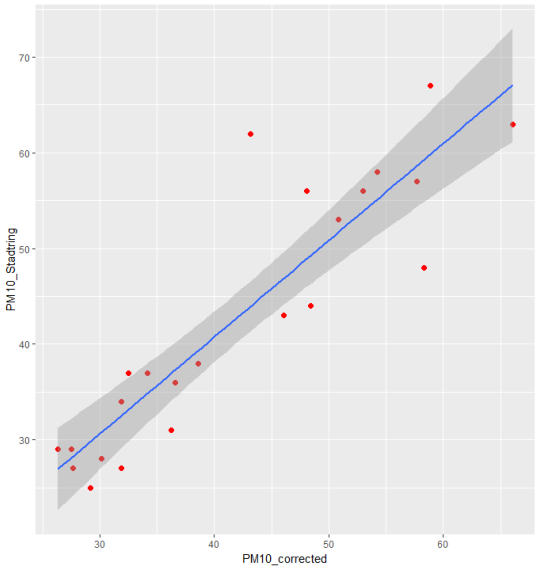
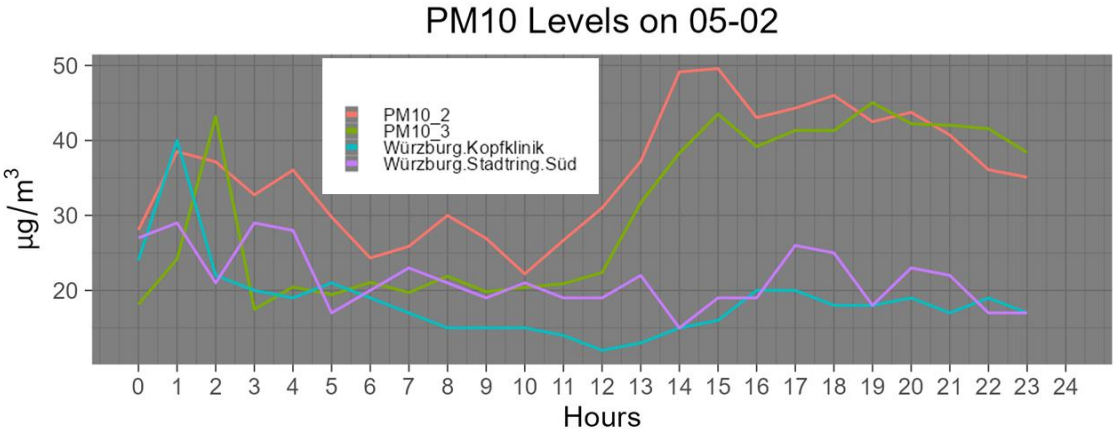
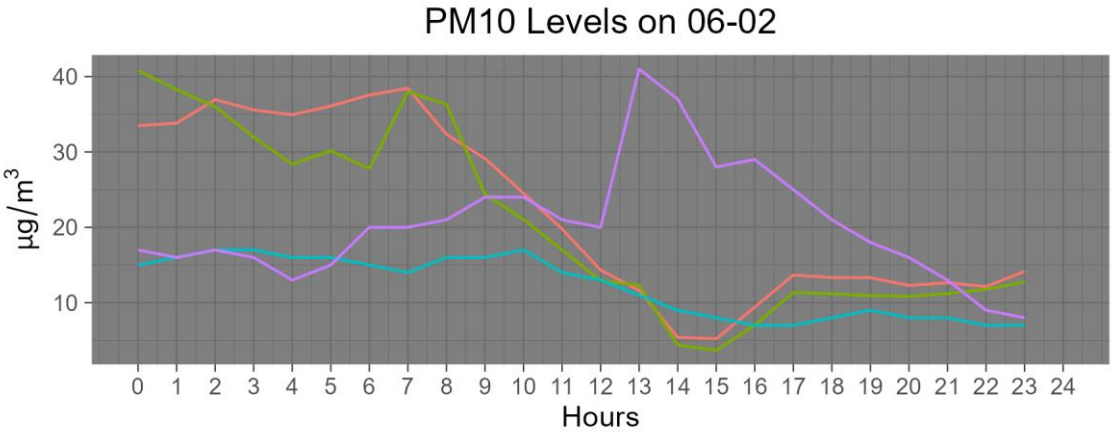


Figure 32: Sensor to reference station performance with calibration

6.3 PM10 Analysis: Static and official Monitors

Figure 33 illustrates the 24-hour PM10 levels recorded from February 5 to February 9, 2023. It includes data from Monitor 2 (depicted in red) and Monitor 3 (in green), aligned with readings from two official monitoring stations in Wuerzburg: Kopfklinik (blue) and Stadtring Süd (purple). Monitor 4 was excluded from the visualization due to an excess of missing values in its dataset.





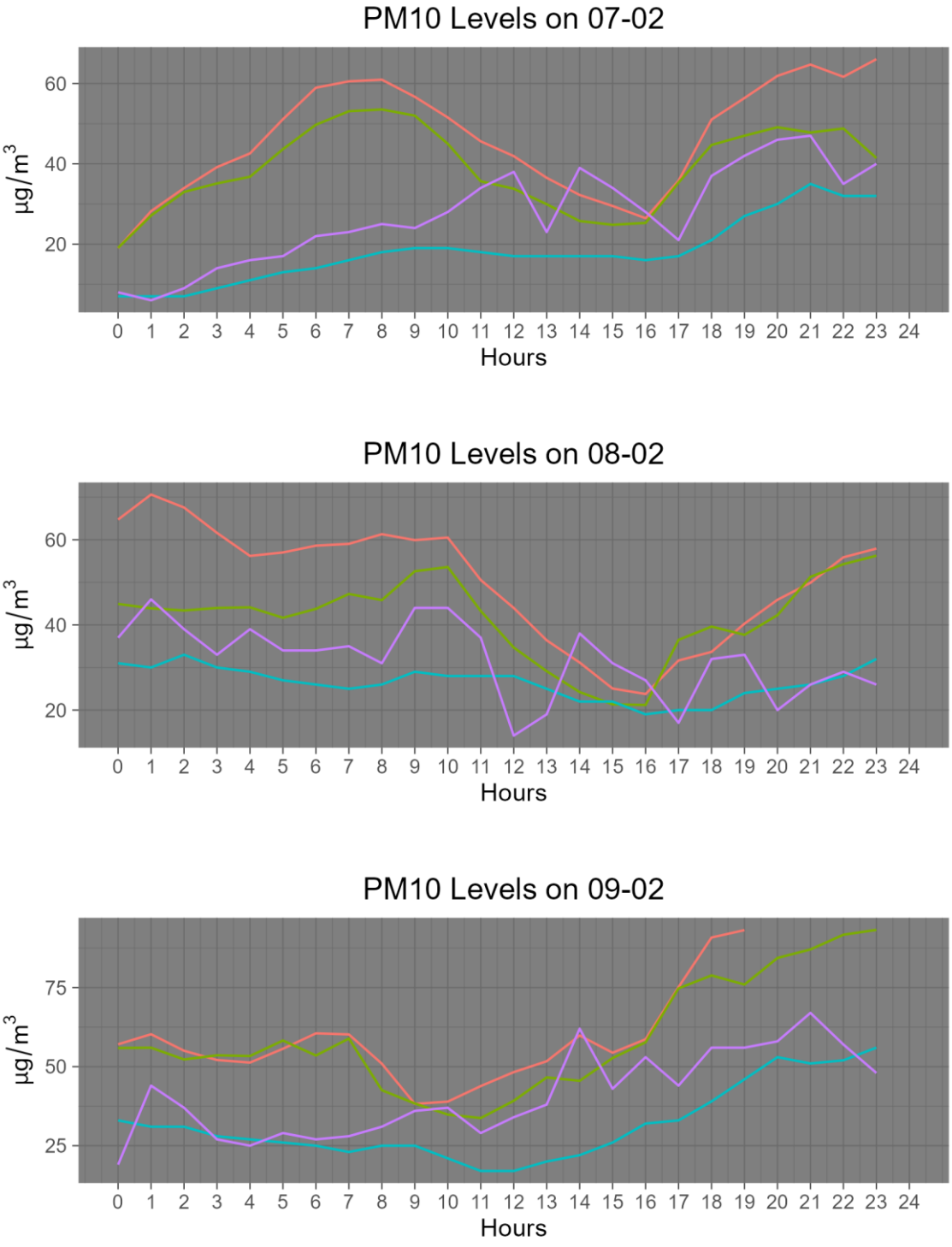
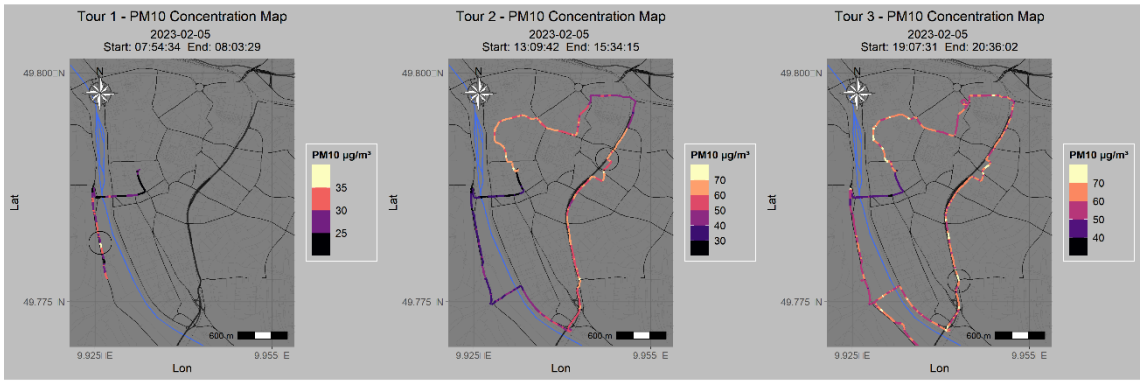


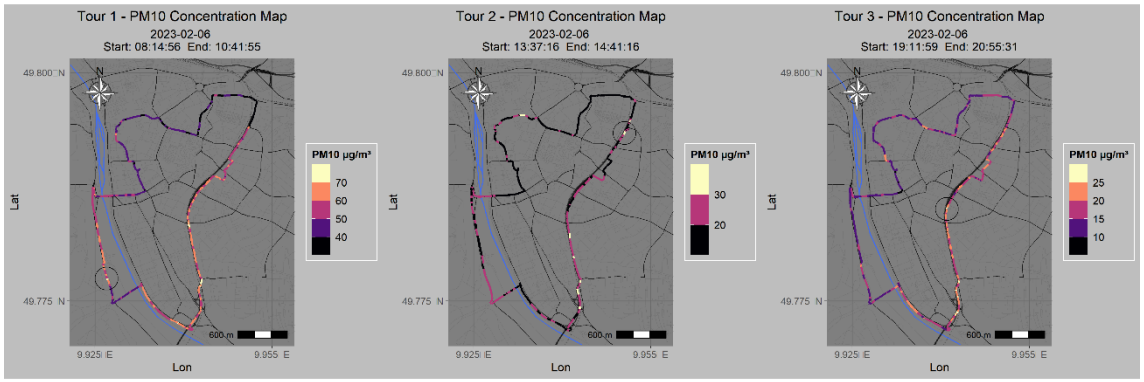
Figure 33: LCS aligned with official monitors

6.4 Analysis of mobile Air Quality Monitoring

6.4.1 Variation in PM10 Levels: A Spatio-Temporal Examination

During the first day of recording on February 5, 2023, unforeseen issues with the SD-Card Module led to incomplete data for the first tour. Figure 34 presents the three individual tours conducted each day, with starting times around 8 am, 1 pm, and 7 pm. The PM10 levels ranged from a minimum of approximately 10 µg/m³ to a maximum of around 150 µg/m³ across all days.





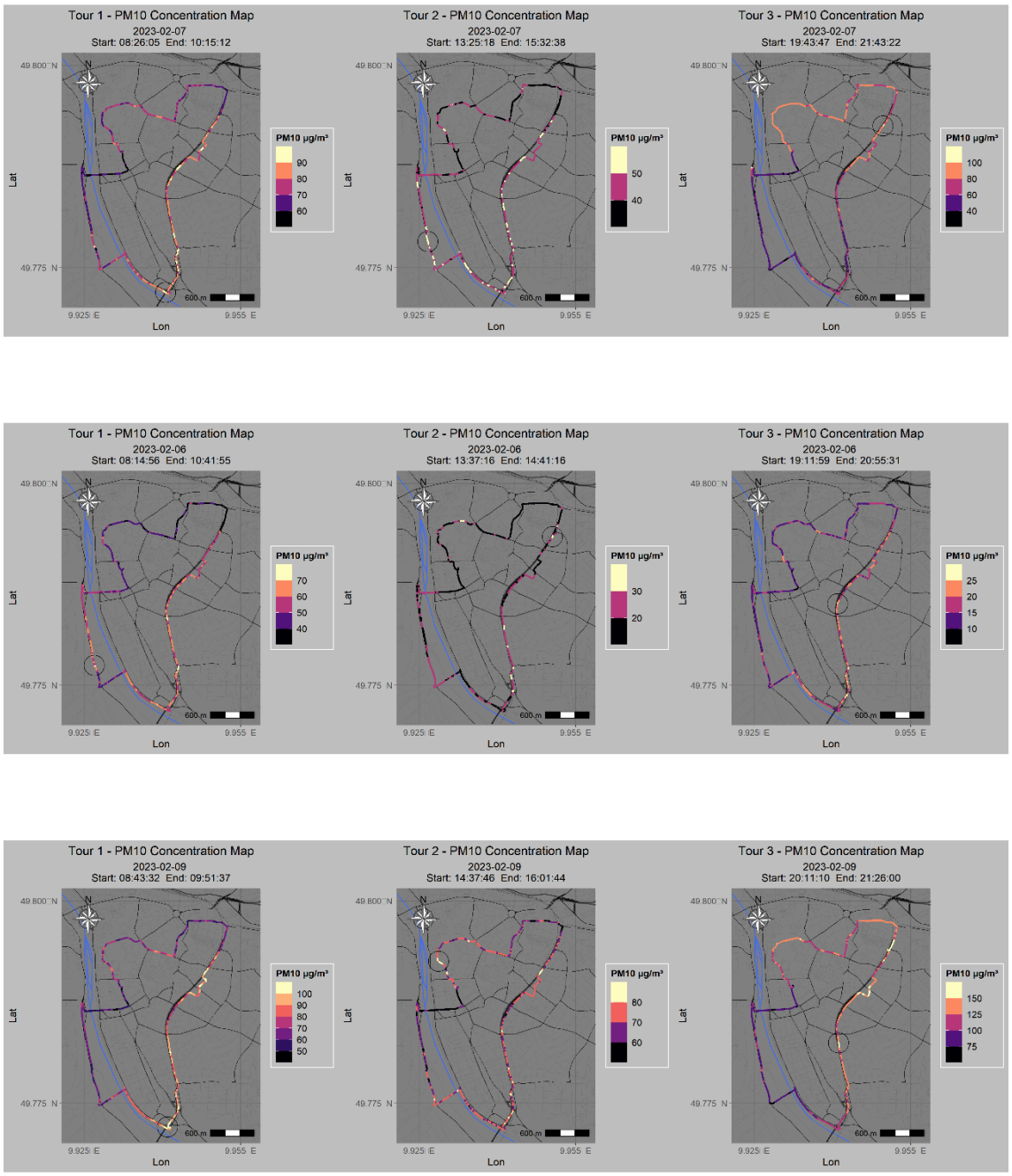
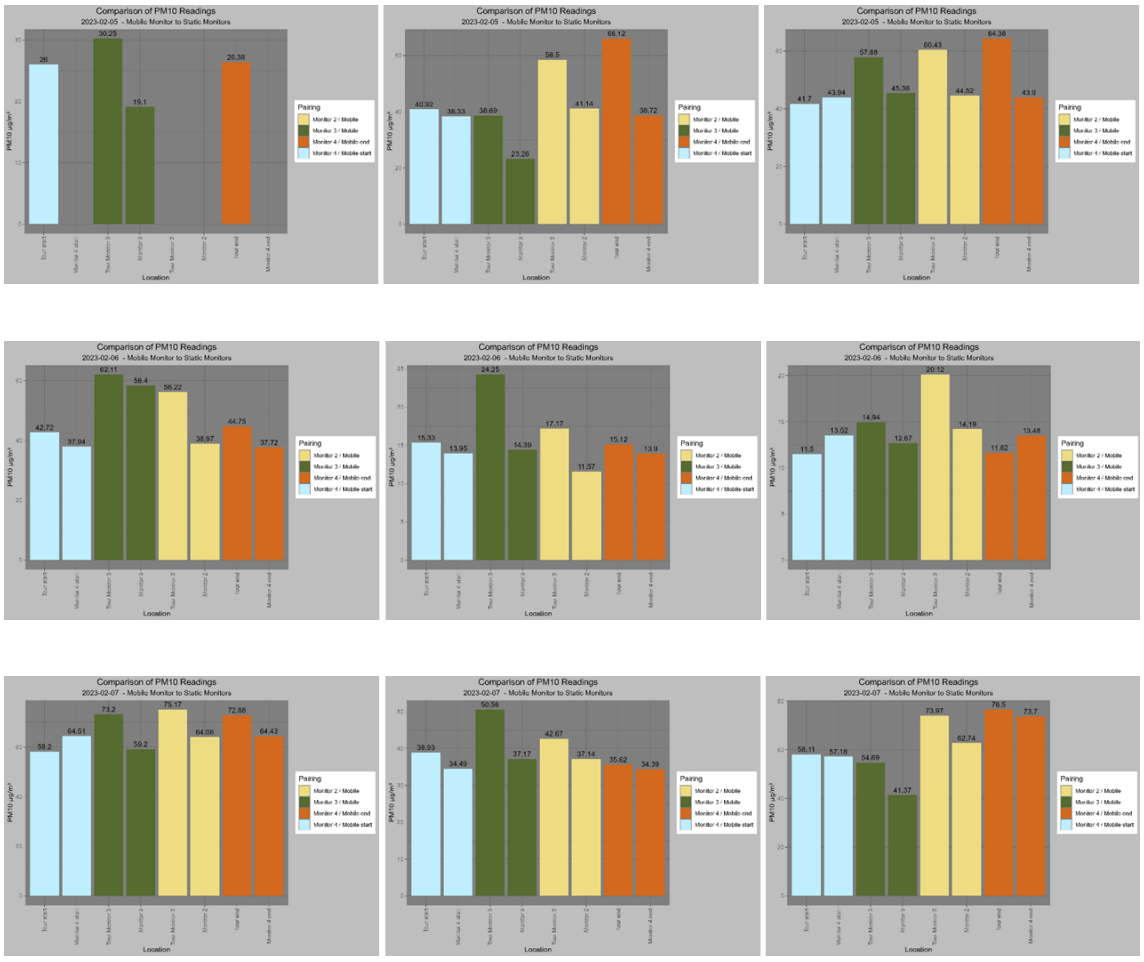


Figure 34: Mobile field collection of PM10 concentration

6.4.2 Mobile Sensor 2 fixed Sensor Quality Check

Figure 35 evaluates the performance discrepancy between the mobile monitoring system and the static Monitors 2, 3, and 4 during the three daily field trips. The graph commences with blue bars representing the initial PM10 measurements taken by the mobile system at the start of each tour in Altstadt, compared to the static Monitor 3. Subsequently, green bars depict the mobile system's readings when in proximity to Monitor 3 located in Steinbachtal. Yellow bars correspond to the mobile system near Monitor 2 in Frauenland, and orange bars signify the end of each tour, aligned with Monitor 4's data. Some bars are absent due to missing data, as indicated in Figure 35.



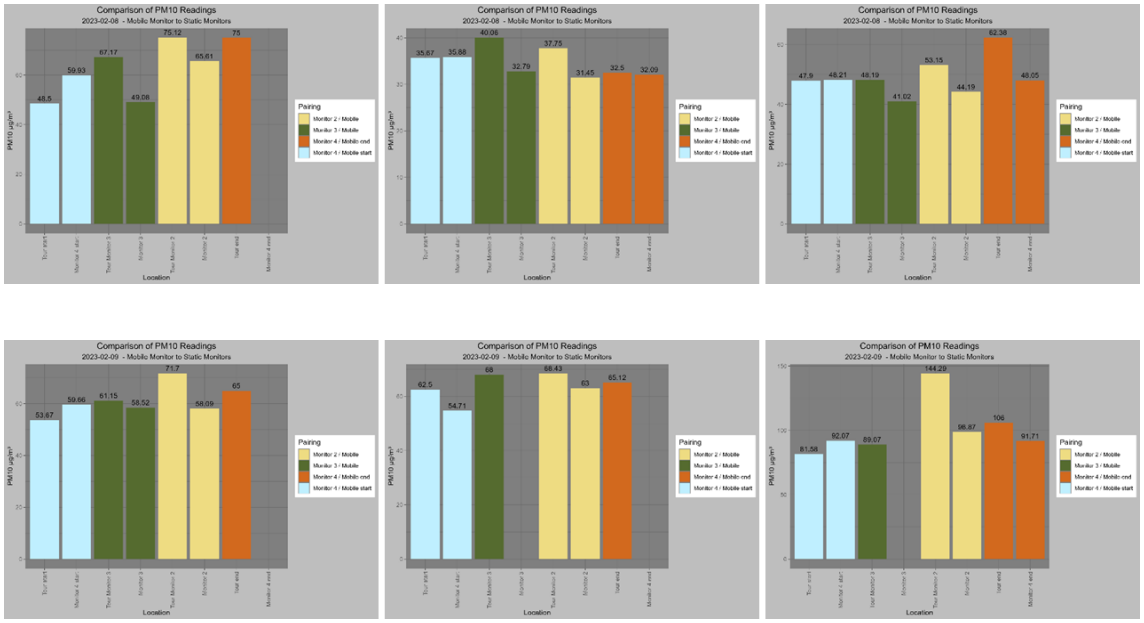


Figure 35: Mobile compared to static sensors (05.02.2023 to 09.02.2023)

VII Discussion

Key Findings

The central focus addressed in this study is the enduring poor air quality within the European Union, particularly in cities. Although official measuring stations provide localized and reliable data, their spatial limitation restricts the comprehensiveness of their findings over entire urban areas. The question aimed to answer was whether combining air quality related datasets from official sources, satellite-based remote sensing data from the Sentinel 5P system, and self-developed stationary and mobile low-cost sensors can provide a more detailed understanding of air pollution distribution in urban spaces. The analysis of satellite-based remote sensing data revealed that the selected Bavarian cities Munich, Regensburg, Nuremberg, Augsburg, and Wuerzburg vary in their geographical locations and areas but exhibit minimal fluctuations in their tropospheric NO2 average levels over the study period. Munich was the weakest performer with an average of 70 µmol/m², whereas Regensburg had the least polluted air with an average of 52.4 µmol/m² from 2019 to 2022. Upon a detailed temporal analysis, a clear seasonal component was identified, indicating that tropospheric NO2 levels rise in colder months and decline in warmer periods, corroborating the significant role meteorology plays in air quality [68]. Spatially, elevated levels of tropospheric NO2 were confined to urban areas and dissipated towards rural regions, as confirmed by a Focal Correlation analysis. The urban core was particularly marked by spatiotemporal contrasts. While individual district analyses revealed minimal discrepancies in average tropospheric NO2 levels, a consistent ranking pattern emerged over the study period. This pattern affirms that districts closer to the urban centre generally experience higher tropospheric NO2 concentrations. Delving into the seasonal differences, Table 5 presents an extremely low p-value ($<2^{-16}$), which is significantly below the standard significance level of 0.05. This strongly supports the rejection of the null hypothesis, suggesting no significant differences in mean NO2 values across seasons. This assertion is further strengthened by a high F-statistic of 230.4, emphasizing that the variance in mean NO2 values between seasons substantially exceeds the variance within seasons.

Turning our attention to district-level seasonal differences, the data from Table 6 reveals a p-value of 0.173, which is above the threshold of 0.05. Thus, we cannot dismiss the null hypothesis, suggesting that significant variations in NO2 concentrations among districts on a seasonal basis may not be present. When examining monthly differences at the district scale, Table 7 shows a p-value of $< 2^{-16}$, far below the accepted significance level. This leads to the rejection of the null hypothesis, which posits no substantial differences in tropospheric NO2 mean values among months. This underscores the existence of significant fluctuations in tropospheric NO2 concentrations on a monthly basis. The results from Figure 25, which illustrates the Urban Atlas land-use classes overlaid with mean tropospheric NO2 levels, show elevated tropospheric NO2 concentrations in the city's central areas, potentially a result of higher traffic or industrial activities. However, the observations are derived from averaged data, and transient factors may have influenced these readings. In Figure 26, the violin plot reveals the distribution of tropospheric NO2 concentrations for the four examined clusters. The mean values for Industrial Areas, Transport, Urban/Built-up, and Vegetation clusters are closely grouped, all within the range of 54.3 to 54.9 µmol/m². Notably, the Urban/Built-up cluster has a significant number of datapoints in the _max concentration areas. The close grouping of NO2 values across diverse land-use types indicates a spatial coarseness in the dataset. This lack of granularity means that the dataset may not effectively differentiate the nuanced effects of land-use on tropospheric NO2 dispersion. The overlapping mean values between the Urban/Built-up and Vegetation clusters, for example, further stress the necessity for a dataset with finer spatial resolution. Linear regression analysis, comparing air quality with topographical data from a Digital Elevation Model (DEM), indicated that despite marginal differences in NO2 levels from an averaged dataset of 2019 to 2022, a correlation exists. Pollution levels declined with increasing elevation; for each 5-meter elevation zone, the NO2 content decreased by 0.083 µmol/m². The results from the elevation-NO2 correlation analysis provide insights into the topographical factors that might be affecting air quality in Wuerzburg. The data suggests an association between higher tropospheric NO2 levels and lower elevations, which could be seen in Figure 27. One possible explanation for this observation is the accumulation of emissions in valleys or lower-lying areas due to certain atmospheric conditions. This relationship is further quantified by the scatterplot in Figure 28, indicating that with an increase in elevation, there is a decrease in mean tropospheric NO2 levels. The multiple R-squared

value shows that elevation can account for approximately 22.62% of the variance in NO2 levels. However, this also indicates that a large portion of the variance is still unexplained. While elevation appears to be a significant factor in determining NO2 levels, it is not the only determinant. The limited scatter of residuals around the regression line indicates that the model fits the observed data well. However, the remaining unexplained variance suggests the presence of other influencing factors. Potential variables might include factors such as wind speed, temperature, or localized sources of NO2 emission that might not have a direct correlation with elevation. Seasonal variations and elevation, or topography, were confirmed as valuable parameters for assessing urban spaces using Sentinel 5P data. Analysis of TROPOMI data for Bavaria, with a particular focus on Wuerzburg, as depicted in Table 4, Table 5, and Figure 22, Figure 23, reveals a marked reduction in tropospheric NO2 levels. It is plausible that this decrease is attributed to the impacts of the Covid-19 pandemic and the subsequent lockdown measures implemented [69]. One of the key findings of the study is the high correlation between the self-developed low-cost sensors, evidenced by R^2 values, which revolve around 0.9109 to 0.9251, suggest a high degree of consistency between the different low-cost sensors. Specifically, the pairings like PM10_1 and PM10_3 showing the highest R-squared value of 0.9251 suggests that these two sensors were the most aligned in their readings during the study duration. Nevertheless, these sensors exhibited high RMSE values between 16 to 19 when compared to the reference station, with an R^2 of 0.441. This suggests that the sensors could only deliver "moderate" quality data and are not readily usable in their current state. An approach to calibrate the measurements of the low-cost sensors against a reference station using a Random Forest model yielded promising results. This calibration improved the R^2 value to 0.83 (Figure 31and Figure 32), bringing the data closer to real-world conditions. The necessity of calibration is highlighted by the identified unreliability of the raw data logs during winter. External factors such as temperature and humidity have a pronounced effect on the readings of low-cost particulate matter sensors. This observation is supported by existing literature [70]. The focus on the stationary low-cost sensors showed no discernible, recurring temporal pattern in PM10 levels during the study period. Additionally, the variance between the low-cost systems themselves and among the official systems was less than that between the two types of systems. However, caution is advised when interpreting these data without appropriate prior calibration. The low-cost sensor data provided reasonable

hourly-averaged PM readings even without calibration. Location-based calibration over an extended period could significantly improve the data quality of these sensors, thereby contributing to the cost-effective expansion of the existing monitoring network. The analysis of data collected through the bicycle-mounted sensors also did not reveal clear trends in the temporal distribution of PM10 levels. The only exception was around midday, during the second tour of each data collection day, where the mobile data showed comparatively lower pollution levels. In terms of spatial distribution, elevated concentrations of PM10 were noted near the city Stadtring, a high-traffic area, as well as in the district Altstadt. The performance analysis of the bicycle-mounted mobile system compared to the static low-cost sensors indicated that the mobile sensor readings are consistently higher than the stationary ones during the same time periods. Data collected at the start of each mobile session were generally close to those from Monitor 4, unlike the readings taken at the end of each tour. The discrepancy in the readings at the end of each tour may be attributed to possible issues with the sensor housing. For instance, if the measuring chamber is not adequately ventilated, it might not exchange air with high PM concentrations effectively, leading to skewed readings. The study yielded mixed results, aligning with some predictions while revealing new insights. The Sentinel-5P Offline NO2 satellite data were adept at identifying meteorological and topographical factors impacting air quality. Nonetheless, their coarse grid resolution meant that they couldn't provide intricate spatial insights. As anticipated, urban zones, given their dense traffic and industrial activities, showcased elevated tropospheric NO2 profiles. Moreover, topographical nuances did influence concentration metrics. Interestingly, certain low-elevated regions that were also deficient in ventilation exhibited heightened tropospheric NO2 levels. This concurred with preliminary assumptions and underscores the multifaceted nature of urban air pollution dynamics. Turning to the sensor evaluations, the low-cost sensors' performance was found to range from moderate to subpar. This echoed expectations, underscoring the need for further calibration. Intriguingly, bicycle-mounted sensors consistently recorded higher PM10 values than their stationary counterparts during equivalent timespans, even accounting for adaptation periods. Such discrepancies could potentially stem from issues related to sensor housing.

Learned Limitations

Throughout the study's progression, valuable lessons learned illuminated additional limitations that had not been initially considered. One of the key findings was the significant improvement in the accuracy of the low-cost sensors after calibration, highlighting their potential for more reliable air quality monitoring. This progress notably altered my initial assessment of these sensors' limitations and underscores their utility when properly calibrated. However, even the calibrated mobile sensors failed to capture the expected daily variations in PM10 concentrations, particularly during high-traffic commuting hours. This unexpected result complicates the interpretation of the data and its application for public policy decisions. Furthermore, the influence of topography on tropospheric NO2 levels, initially not a primary focus of the study, emerged as a factor warranting further scrutiny. Another unexpected observation was the lack of distinct PM10 concentration trends in high-traffic areas compared to residential zones, calling into question the robustness of the methodologies applied. These newfound limitations not only serve as cautionary notes for the interpretation of this study's results but also pinpoint specific areas that merit deeper investigation in future research. The absence of multiple linear regression analysis with relevant parameters, for instance, may limit the depth of interpretation of the relationships among different variables. This leaves room for questions related to the true impact of variables like topography, time of day, and specific geographic locations on air quality levels. Budgetary and methodological constraints, particularly concerning the spatial resolution of the Sentinel 5P data and the limitations of low-cost sensors, shaped the scope of the study.

VIII Conclusion

The findings from this study underscore that satellite-based remote sensing, especially the Sentinel-5P system, is a potent tool for discerning large-scale seasonal fluctuations in urban locales and gaining insights into how topographical attributes influence air quality. While the spatial resolution of the Sentinel-5P data may not be optimal for intricate, district-level analysis or the granularities showcased by platforms like the Urban Atlas land cover classifications, the patterns and trends observed in the data were unambiguously evident and aligned with initial expectations. This suggests that while the raw data might not be perfectly suited for hyper-localized analysis, it still offers meaningful insights into broader air quality dynamics. Earth observation data emerges as a precious asset, offering pertinent parameters like the Normalized Difference Vegetation Index (NDVI) and Land Surface Temperature (LST), which deepen our grasp of micro-scale air quality shifts. Established governmental monitoring frameworks offer consistent and trustworthy data but come with substantial costs, hindering their ubiquitous adoption. Their data presentation, often in hourly averages, poses challenges for instantaneous evaluations. Low-cost sensors emerge as budget-friendly additions, enhancing both the geographical spread and time-bound precision of measurements. Their swift manufacturability, deployability, and easy replacement in the event of malfunctions make them indispensable. Yet, precise calibration remains pivotal, relying on trustworthy data about ambient conditions and reference frameworks. Prolonged engagement might shed light on the enduring accuracy of calibrated data and the recalibration periodicity. The mobile data collection methodology introduced challenges when trying to compare its results with stationary datasets. When sensors were placed at fixed reference points for extended durations, the readings did not consistently match those from uncalibrated measurements. The study faced limitations in collecting comprehensive data via bicycle-mounted systems due to time and personnel constraints. Combining data from Earth observation, in situ governmental measurements, and low-cost sensors (both mobile and static) offers a comprehensive perspective on urban air quality dynamics. By integrating these different layers, a more detailed and holistic insight into the factors influencing air quality can be obtained. This integrated approach is particularly valuable in evaluating

the progression of air quality management strategies, assessing the effectiveness of current measures, and making necessary adjustments when needed. In summary, this research provides insights into air quality dynamics across selected Bavarian urban areas, especially in Wuerzburg, highlighting the significant roles of time, topography, and land use. Furthermore, the potential of calibrating low-cost sensors for enhanced air quality monitoring stands out as a promising avenue for future developments in the field.

Literature

[1] “Air quality in Europe - 2020 report — European Environment Agency.” Accessed: May 19, 2022. [Online]. Available: https://www.eea.europa.eu/publications/air-quality-in-europe-2020-report

[2] W. H. Organization, “Ambient air pollution: a global assessment of exposure and burden of disease.” World Health Organization, 2016.

[3] “EUR-Lex - L:2008:152:TOC - EN - EUR-Lex.” Accessed: Jun. 01, 2022. [Online]. Available: https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32008L0050&from=DE

[4] G. Collaborators, J. Ärnlöv, and others, “Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019,” *The Lancet*, vol. 396, no. 10258, pp. 1223–1249, 2020.

[5] I. Scientific, “Air Pollution and Cancer,” *Lyon IARC Sci.*, 2013.

[6] A. Hajat, C. Hsia, and M. S. O’Neill, “Socioeconomic Disparities and Air Pollution Exposure: a Global Review,” *Curr. Environ. Health Rep.*, vol. 2, no. 4, pp. 440–450, Dec. 2015, doi: 10.1007/s40572-015-0069-5.

[7] G. Bolte, C. Bunge, C. Hornberg, and H. Köckler, “Umweltgerechtigkeit als ansatz zur verringerung sozialer ungleichheiten bei umwelt und gesundheit,” *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz*, vol. 61, no. 6, pp. 674–683, 2018.

[8] C. Nagl, J. Schneider, and P. Thielen, “Implementation of the Ambient Air Quality Directive,” 2016.

[9] World Health Organization, *WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization, 2021. Accessed: Feb. 17, 2023. [Online]. Available: https://apps.who.int/iris/handle/10665/345329

[10] P. Sicard, E. Agathokleous, A. De Marco, E. Paoletti, and V. Calatayud, “Urban population exposure to air pollution in Europe over the last decades,” *Environ. Sci. Eur.*, vol. 33, no. 1, p. 28, Mar. 2021, doi: 10.1186/s12302-020-00450-2.

[11] F. Karagulian *et al.*, “Review of the performance of low-cost sensors for air quality monitoring,” *Atmosphere*, vol. 10, no. 9, p. 506, 2019.

[12] P. Kumar *et al.*, “The rise of low-cost sensing for managing air pollution in cities,” *Environ. Int.*, vol. 75, pp. 199–205, 2015.

[13] M. Badura, P. Batog, A. Drzeniecka-Osiadacz, and P. Modzel, “Regression methods in the calibration of low-cost sensors for ambient particulate matter measurements,” *SN Appl. Sci.*, vol. 1, no. 6, p. 622, May 2019, doi: 10.1007/s42452-019-0630-1.

[14] “Managing air quality in Europe — European Environment Agency.” Accessed: Feb. 17, 2023. [Online]. Available: https://www.eea.europa.eu/publications/managing-air-quality-in-europe

[15] M. Demuzere, B. Bechtel, A. Middel, and G. Mills, “Mapping Europe into local climate zones,” *PloS One*, vol. 14, no. 4, p. e0214474, 2019.

[16] T. Erbertseder, H. Taubenböck, and J. Meyer-Arnek, “Dicke Luft–Stadtregionen als globale Zentren der Luftverschmutzung,” in *Globale Urbanisierung*, Springer, 2015, pp. 191–203.

[17] E. T. Wilkins, “Air pollution and the London fog of December, 1952,” *J. R. Sanit. Inst.*, vol. 74, no. 1, pp. 1–21, 1954.

[18] “Clean Air Act 1956 (c. 52).” Accessed: Sep. 12, 2023. [Online]. Available: https://www.legislation.gov.uk/ukpga/Eliz2/4-5/52/enacted/data.xht?view=snippet&wrap=true

[19] “EUR-Lex - L:1996:296:TOC - EN - EUR-Lex.” Accessed: Sep. 12, 2023. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AL%3A1996%3A296%3ATOC

[20] European Commission. Directorate General for Environment., VVA., Toegepast natuurwetenschappelijk onderzoek., Tecnalia., ANOTEC., and Universitat Autònoma de Barcelona., *Assessment of potential health benefits of noise abatement measures in the EU: Phenomena project*. LU: Publications Office, 2021. Accessed: Sep. 12, 2023. [Online]. Available: https://data.europa.eu/doi/10.2779/24566

[21] “New WHO Global Air Quality Guidelines aim to save millions of lives from air pollution.” Accessed: Jun. 01, 2022. [Online]. Available: https://www.who.int/news/item/22-09-2021-new-who-global-air-quality-guidelines-aim-to-save-millions-of-lives-from-air-pollution

[22] “ETC HE Report 2022/3: Status report of air quality in Europe for year 2021, using validated and up-to-date data,” Eionet Portal. Accessed: Oct. 21, 2023. [Online]. Available: https://www.eionet.europa.eu/etcs/etc-he/products/etc-he-products/etc-he-reports/etc-he-report-2022-3-status-report-of-air-quality-in-europe-for-year-2021-using-validated-and-up-to-date-data

[23] “39. BImSchV. Bundesgesetzblatt: Teil I,” Aug. 2010.

[24] “Würzburg Stadtprofil: Statistische Daten 2022.” Accessed: Sep. 05, 2023. [Online]. Available: https://www.wuerzburg.de/media/www.wuerzburg.de/org/med_5478/582744_00_wuerzburg_2022_stadtprofil.pdf

[25] “classification-of-monitoring-stations-and.pdf.” Accessed: Oct. 23, 2023. [Online]. Available: https://www.eea.europa.eu/downloads/cb32af951deb4e40aef444bdd37d9306/1680621730/classification-of-monitoring-stations-and.pdf

[26] I. Müller, T. Erbertseder, and H. Taubenböck, “Tropospheric NO2: Explorative analyses of spatial variability and impact factors,” *Remote Sens. Environ.*, vol. 270, p. 112839, 2022.

[27] D. Oxoli, J. R. Cedeno Jimenez, and M. A. Brovelli, “ASSESSMENT OF SENTINEL-5P PERFORMANCE FOR GROUND-LEVEL AIR QUALITY MONITORING: PREPARATORY EXPERIMENTS OVER THE COVID-19 LOCKDOWN PERIOD,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLIV-3/W1-2020, pp. 111–116, Nov. 2020, doi: 10.5194/isprs-archives-XLIV-3-W1-2020-111-2020.

[28] P. Schneider, P. D. Hamer, A. Kylling, S. Shetty, and K. Stebel, “Spatiotemporal Patterns in Data Availability of the Sentinel-5P NO2 Product over Urban Areas in Norway,” *Remote Sens.*, vol. 13, no. 11, Art. no. 11, Jan. 2021, doi: 10.3390/rs13112095.

[29] European Commission. Joint Research Centre., *Guidance on low-cost sensors deployment for air quality monitoring experts based on the AirSensEUR experience*. LU: Publications Office, 2022. Accessed: Oct. 21, 2023. [Online]. Available: https://data.europa.eu/doi/10.2760/14893

[30] A. Cavaliere *et al.*, “Development of low-cost air quality stations for next-generation monitoring networks: calibration and validation of NO_2 and O_3 sensors,” *Atmospheric Meas. Tech.*, vol. 16, no. 20, pp. 4723–4740, Oct. 2023, doi: 10.5194/amt-16-4723-2023.

[31] P. Nowack, L. Konstantinovskiy, H. Gardiner, and J. Cant, “Machine learning calibration of low-cost NO_2 and PM_{10} sensors: non-linear algorithms and their impact on site transferability,” *Atmospheric Meas. Tech.*, vol. 14, no. 8, pp. 5637–5655, Aug. 2021, doi: 10.5194/amt-14-5637-2021.

[32] M. Zusman *et al.*, “Calibration of low-cost particulate matter sensors: Model development for a multi-city epidemiological study,” *Environ. Int.*, vol. 134, p. 105329, Jan. 2020, doi: 10.1016/j.envint.2019.105329.

[33] A. Anjomshoaa, P. Santi, F. Duarte, and C. Ratti, “Quantifying the Spatio-Temporal Potential of Drive-by Sensing in Smart Cities,” *J. Urban Technol.*, vol. 28, no. 1–2, pp. 199–216, Apr. 2021, doi: 10.1080/10630732.2020.1791679.

[34] “DWD - Temperatur 1991 - 2020.” Accessed: Sep. 05, 2023. [Online]. Available: https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/multi_annual/mean_91-20/Temperatur_1991-2020.txt

[35] “BayernAtlas,” bayernatlas.de. Accessed: Oct. 20, 2023. [Online]. Available: https://www.bayernatlas.de

[36] “Urban Atlas 2018 — Copernicus Land Monitoring Service.” Accessed: Sep. 05, 2023. [Online]. Available: https://land.copernicus.eu/local/urban-atlas/urban-atlas-2018

[37] J. Van Geffen, H. J. Eskes, K. F. Boersma, J. D. Maasakkers, and J. P. Veefkind, “TROPOMI ATBD of the total and tropospheric NO2 data products,” *DLR Doc.*, 2019.

[38] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, “Google Earth Engine: Planetary-scale geospatial analysis for everyone,” *Remote Sens. Environ.*, 2017, doi: 10.1016/j.rse.2017.06.031.

[39] “Sentinel-5P OFFL NO2: Offline Nitrogen Dioxide | Earth Engine Data Catalog | Google Developers.” Accessed: Feb. 13, 2023. [Online]. Available:

https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S5P_OFFL_L3_NO2

[40] M. Sneep, “Sentinel 5 precursor/TROPOMI KNMI and SRON level 2 Input Output Data Definition,” no. 16.

[41] “(EU) 2021/696.” Accessed: Sep. 12, 2023. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A32021R0696

[42] “Stationäre NOx - Messungen Einzelfragen zu Messmethoden.” Accessed: Sep. 07, 2023. [Online]. Available: https://www.bundestag.de/resource/blob/543806/8704e25eb7690dff66ef16a0949481fc/wd-8-047-17-pdf-data.pdf

[43] “05_wuerzburg_stadtring_sued.pdf.” Accessed: Feb. 11, 2023. [Online]. Available: https://www.lfu.bayern.de/luft/immissionsmessungen/doc/lueb_dokumentation/aktiv/06_Unterfranken/05_wuerzburg_stadtring_sued.pdf

[44] “Welcome.” Accessed: Oct. 20, 2023. [Online]. Available: https://docs.carto.com/

[45] QGIS Development Team, *QGIS Geographic Information System*. QGIS Association. [Online]. Available: https://www.qgis.org

[46] “FreeCAD/FreeCAD: This is the official source code of FreeCAD, a free and opensource multiplatform 3D parametric modeler.” Accessed: Feb. 13, 2023. [Online]. Available: https://github.com/FreeCAD/FreeCAD

[47] “PrusaSlicer | Original Prusa 3D printers directly from Josef Prusa,” Prusa3D by Josef Prusa. Accessed: Feb. 13, 2023. [Online]. Available: https://www.prusa3d.com/page/prusaslicer_424/

[48] K. Arnold, J. Gosling, and D. Holmes, *The Java programming language*. Addison Wesley Professional, 2005.

[49] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2021. [Online]. Available: https://www.R-project.org/

[50] “MicroPython language and implementation — MicroPython latest documentation.” Accessed: Feb. 13, 2023. [Online]. Available: https://docs.micropython.org/en/latest/reference/index.html

[51] A. Amelia *et al.*, “MQTT Protocol Implementation for Monitoring of Environmental Based on IoT,” in *2020 International Conference on Applied Science and Technology (iCAST)*, 2020, pp. 700–703. doi: 10.1109/iCAST51016.2020.9557694.

[52] “AWS IoT Core Documentation.” Accessed: Feb. 13, 2023. [Online]. Available: https://docs.aws.amazon.com/iot/index.html

[53] R. P. Ltd, “Buy a Raspberry Pi Pico,” Raspberry Pi. Accessed: Oct. 20, 2023. [Online]. Available: https://www.raspberrypi.com/products/raspberry-pi-pico/

[54] A. Kumar Pal, A. K. Mohanty, and M. Misra, “Additive manufacturing technology of polymeric materials for customized products: recent developments and

future prospective," *RSC Adv.*, vol. 11, no. 58, pp. 36398–36438, 2021, doi: 10.1039/D1RA04060J.

[55] "4505_PMSA003I_series_data_manual_English_V2.6.pdf." Accessed: Feb. 13, 2023. [Online]. Available: https://cdn-shop.adafruit.com/product-files/4632/4505_PMSA003I_series_data_manual_English_V2.6.pdf

[56] "MicroSD Card Adapter v0.9b Schaltplan.pdf."

[57] "tca9548a.pdf." Accessed: Feb. 13, 2023. [Online]. Available: https://www.ti.com/lit/ds/symlink/tca9548a.pdf

[58] "Tidy Messy Data." Accessed: Sep. 07, 2023. [Online]. Available: https://tidyr.tidyverse.org/

[59] "A Grammar of Data Manipulation." Accessed: Sep. 07, 2023. [Online]. Available: https://dplyr.tidyverse.org/

[60] "Statistical Methods for Research Workers," *Nature*, vol. 131, no. 3307, Art. no. 3307, Mar. 1933, doi: 10.1038/131383b0.

[61] "OpenStreetMap," OpenStreetMap. Accessed: Sep. 10, 2023. [Online]. Available: https://www.openstreetmap.org/copyright

[62] "Ausführliche Messwerte der bayerischen Luftmessstationen - LfU Bayern." Accessed: Feb. 14, 2023. [Online]. Available: https://www.lfu.bayern.de/luft/immissionsmessungen/messwerte/stationen/detail/1405/132

[63] "Applied linear statistical models (Vol. 4). Chicago: Irwin.".

[64] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[65] B. Ilya, "Ilya Bogdanovich. (2021). Geo Tracker - GPS tracker (Version 5.0.2) [Mobile app]. Google Play Store. https://play.google.com/store/apps/details?id=com.ilyabogdanovich.geotracker&gl=US." [Online]. Available: https://play.google.com/store/apps/details?id=com.ilyabogdanovich.geotracker&gl=US

[66] "stadt würzburg shape file." Accessed: Sep. 06, 2023. [Online]. Available: https://opendata.wuerzburg.de/explore/dataset/altstadt/

[67] "Google Maps," Google Maps. Accessed: Oct. 20, 2023. [Online]. Available: https://www.google.com/maps

[68] H. Yin, Y. Sun, J. Notholt, M. Palm, and C. Liu, "Spaceborne tropospheric nitrogen dioxide (NO_2) observations from 2005–2020 over the Yangtze River Delta (YRD), China: variabilities, implications, and drivers," *Atmospheric Chem. Phys.*, vol. 22, no. 6, pp. 4167–4185, Mar. 2022, doi: 10.5194/acp-22-4167-2022.

[69] P. F. Levelt *et al.*, "Air quality impacts of COVID-19 lockdown measures detected from space using high spatial resolution observations of multiple trace gases from Sentinel-5P/TROPOMI," *Atmospheric Chem. Phys.*, vol. 22, no. 15, pp. 10319–10351, Aug. 2022, doi: 10.5194/acp-22-10319-2022.

[70] K. Chan *et al.*, “Low-cost electronic sensors for environmental research: Pitfalls and opportunities,” *Prog. Phys. Geogr. Earth Environ.*, vol. 45, no. 3, pp. 305–338, Jun. 2021, doi: 10.1177/0309133320956567.