

Article

Machine Learning Techniques for Spatio-Temporal Air Pollution Prediction to Drive Sustainable Urban Development in the Era of Energy and Data Transformation

Mateusz Zareba ^{*,†}, Szymon Cogiel [†], Tomasz Danek [†] and Elzbieta Weglinska [†]

Department of Geoinformatics and Applied Computer Science, Faculty of Geology, Geophysics and Environmental Protection, AGH University of Krakow, 30-059 Krakow, Poland; cogielszymon1@gmail.com (S.C.); tdanek@agh.edu.pl (T.D.); weglinska@agh.edu.pl (E.W.)

* Correspondence: zareba@agh.edu.pl

† These authors contributed equally to this work.

Abstract: Sustainable urban development in the era of energy and digital transformation is crucial from a societal perspective. Utilizing modern techniques for analyzing large datasets, including machine learning and artificial intelligence, enables a deeper understanding of historical data and the efficient prediction of future events based on data from IoT sensors. This study conducted a multidimensional historical analysis of air pollution to investigate the impacts of energy transformation and environmental policy and to determine the long-term environmental implications of certain actions. Additionally, machine learning (ML) techniques were employed for air pollution prediction, taking spatial factors into account. By utilizing multiple low-cost air sensors categorized as IoT devices, this study incorporated data from various locations and assessed the influence of neighboring sensors on predictions. Different ML approaches were analyzed, including regression models, deep neural networks, and ensemble learning. The possibility of implementing such predictions in publicly accessible IT mobile systems was explored. The research was conducted in Krakow, Poland, a UNESCO-listed city that has had long struggle with air pollution. Krakow is also at the forefront of implementing policies to prohibit the use of solid fuels for heating and establishing clean transport zones. The research showed that population growth within the city does not have a negative impact on PM_x concentrations, and transitioning from coal-based to sustainable energy sources emerges as the primary factor in improving air quality, especially for PM_x, while the impact of transportation remains less relevant. The best results for predicting rare smog events can be achieved using linear ML models. Implementing actions based on this research can significantly contribute to building a smart city that takes into account the impact of air pollution on quality of life.

Keywords: big data; energy transition; smart cities; machine learning; air pollution; urban development



Citation: Zareba, M.; Cogiel, S.; Danek, T.; Weglinska, E. Machine Learning Techniques for Spatio-Temporal Air Pollution Prediction to Drive Sustainable Urban Development in the Era of Energy and Data Transformation. *Energies* **2024**, *17*, 2738. <https://doi.org/10.3390/en17112738>

Academic Editor: Jun Liu

Received: 13 May 2024

Revised: 27 May 2024

Accepted: 29 May 2024

Published: 4 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Smart urban development planning requires multidisciplinary collaboration, including political science, environmental engineering, and spatial planning. Life satisfaction depends on policymakers' effectiveness, living standards, education access, and a clean environment [1]. Research shows that environmental comfort, particularly air pollution, is crucial for human quality of life [2]. With increasing urbanization and environmental issues, there is a need for a holistic global approach to urban development [3]. Liang et al. emphasized the importance of particulate matter analysis in balanced urban planning [4]. Jonek-Kowalska pointed out the research gap in long-term air pollution studies in cities aspiring to be smart [5]. Urban development may increase energy demand. Relying on certain energy mixes can strengthen air pollution. In 2021, Poland's per capita household energy consumption from hard coal was almost 22%, compared to the

EU average of 2.5% [6]. Therefore, advanced technologies and IoT sensors are essential for assessing the impact of energy transition on air pollution.

Krakow has struggled with air pollution for years. Initially, the metallurgical industry was the main source of pollution, but fossil fuel heating has become an increasingly significant source [7]. Despite a total ban on solid fuel use for heating, it remains the main source of pollution in winter [8,9]. PM10 consists of over 40% carbon, with coal responsible for 50% of it in winter and 20% in summer. Car transportation is the second highest contributor. Natural sources account for 30% of the carbon fraction across all seasons [10]. Air pollution generated outside the city is migrating to the city—especially during the cold season [11]. The city's location, surrounded by small hills locally and the Carpathian Mountains to the south [12], contributes to this phenomenon. Krakow, a UNESCO World Heritage site, faces challenges affecting both residents and tourists. The COVID-19 pandemic led to a decline in tourism, prompting efforts to revitalize the sector sustainably [13]. Episodic smog events make Krakow one of the most polluted cities globally and may deter tourists, especially amid COVID-19 respiratory concerns. Improving air quality in the city benefits both the economy and public health. Krakow aims to become a smart city with modern, pleasant living conditions [14], and it plans to address the challenges of air pollution, thereby improving its attractiveness to tourists. In addition, Poland implemented the mObywatel system [15], a digital platform accessible to every citizen, offering real-time data on air pollution levels. This innovation opens the possibility of demonstrating the tangible impact of zero-emission policies on the environment. It can improve societal attitudes and influence the belief that sustainable urban development is attainable. By providing such data, the gap identified by Jonek-Kowalska [5] can be addressed.

In this study, we utilize big data from various sources to assess factors possibly affecting air pollution and to determine the best techniques for predicting rare peaks in pollution influx. The main focus is on how to drive policy-making in the era of energy and digital transformation by using these technologies in the most optimal and environmentally friendly way. Krakow serves as a good benchmark city for other regions worldwide, as it needs to implement strict laws according to European Union regulations while still being located in a country with a predominantly coal-based energy system.

This work is divided into two parts aiming to close part of a mental gap. The first part consists of a historical, descriptive, and diagnostic analysis in which we examine data on population changes, types and quantities of transport means, and heating types to answer the question of what happened and what impact it had on pollution levels over the past 10 years. This is crucial for future informed city planning, particularly in the context of energy policy. The second part consists of a predictive and prescriptive analysis, where we analyze the predictive potential of various methods to optimally forecast relatively rare but significant smog episodes from a public health perspective. Such episodes are outliers in data analysis, requiring the selection of appropriate modern techniques because traditional methods often fail in such applications. In this case, different neural network architectures were analyzed in comparison to the autoregressive moving average.

For the prediction part, various supervised ML methods were used. In general, ML techniques can be divided into two different sub-domains: supervised and unsupervised learning. In supervised learning, data is labeled, and the dataset is typically randomly split into training and testing sets. This allows for the calculation of model performance and estimation of how the model will perform when new data is introduced [16]. However, supervised learning for time-series analysis may require a special kind of data splitting because the dataset cannot be randomly split due to the time dimension. In this study, the spatial component was additionally accounted for by including information from all 52 receivers around the city. The reason that research on ML potential for air pollution prediction is rare is that many researchers have found it useful for environmental monitoring. Rana et al. [17] demonstrated that AI technology can be successfully applied to water pollution analysis. Zaresefat and Derakhshani [18] showcased how AI has revolutionized groundwater management. Uriarte-Gallastegi et al. [19] highlighted the potential of AI

for sustainable energy management. Zareba et al. [20] illustrated the use of unsupervised machine learning for spatial analysis of PM time-series. ML technologies play a key role in various forecasting-related problems; they also enable more effective urban environmental management focused on air pollution. ML is crucial due to its ability to predict and adapt to dynamic conditions. It allows for the analysis of large volumes of data, pattern identification, and the delivery of real-time predictions, which are invaluable in urban planning and environmental protection.

The novelty of this research lies in bridging the gap between advanced AI-assisted spatial analysis of air pollution big data and sustainable urban planning, with a particular emphasis on energy transition. We selected Krakow as our European benchmark city for two main reasons: 1. Krakow stands out as a unique anti-smog city, being isolated from neighboring areas where there is no anti-smog law in force. 2. The city is also located in a country where the transition from coal-based energy production holds significant importance and will continue to do so. It is essential to address how to effectively predict, communicate, and create society-friendly solutions, as well as to establish systems to accurately assess the positive effects of energy transformation on the development of smart cities. This research specifically focuses on analyzing how air pollution can be predicted effectively, which is beneficial not only for warning residents but also for planning transformation and determining the consequences of planned changes. We also analyzed historical data, considering various factors that could influence PM_x air pollution, such as population, the number of vehicles, or the state of public transportation. Additionally, we considered the impact of the program aimed at changing the energy mix used for heating homes in the city area. This research aims to explore the effectiveness of using big data and ML/AI methods to shape smarter cities by learning from the past to improve the future. To achieve this goal, we will conduct statistical analysis on spatio-temporal data, focusing on efficient, reliable ML techniques. Our research will determine the impact of population changes, types of transportation, and energy policy on air quality and answer whether it is possible to accurately predict rare smog events by using dense spatio-temporal time series.

2. Materials and Methods

2.1. Urban Development and Energy Transition

For quantitative analysis of actions taken by the city of Krakow to mitigate atmospheric air pollution, an analysis of data provided by the Krakow City Council and Chief Inspectorate for Environmental Protection (<https://powietrze.gios.gov.pl/>, accessed on 10 May 2024) was conducted. The analysis took into account long-term PM_{2.5} trends, information such as the city's population in respective years, and data on road transport—both private and public. Regarding data on the number of registered passenger cars, the total number of vehicles was considered, while for public transport, metrics such as the number of bus and tram lines and their lengths in kilometers were utilized. Information on bicycle traffic was also considered, focusing solely on linear infrastructure, specifically the absolute length of bicycle paths. Point infrastructure, such as bicycle rental stations, was entirely omitted from the analysis, as there is currently no municipal bike rental system in operation in the city.

Population and transport data were analyzed over the years from 2010 to 2019. Data provided by the city council in the form of an informational folder [21] required extraction into a numerical format for processing. Each year's data is provided in a separate folder, and the document structure varies slightly from year to year, posing difficulties in automatic data extraction. Additionally, an analysis of data for the PONE program—the Low-Emission Reduction Program in Krakow—was conducted [22]. The program aimed to reduce the use of coal furnaces and replace them with more environmentally friendly heating systems, such as gas-fired furnaces. Data available for the PONE program covers recent years starting from 2014. Each year, the report includes information on the number of coal furnaces and boilers removed, the installation of renewable energy sources, and the amount

spent on the low-emission reduction program (in PLN). Furthermore, general aggregate statistics provided by the city council related to the PONE program were also utilized.

2.2. Machine-Learning Data Pipeline

In the world of ML today, designing reproducible, maintainable, and modular data processing pipelines is becoming increasingly crucial. This approach not only enhances the efficiency and scalability of projects but also allows for quick adaptation to changing requirements and easier project management [23]. In this project, a comprehensive pipeline was developed, as illustrated in Figure 1. This pipeline is divided into four distinct sub-pipelines: preprocessing, feature engineering, modeling, and Explainable AI (XAI) and evaluation.

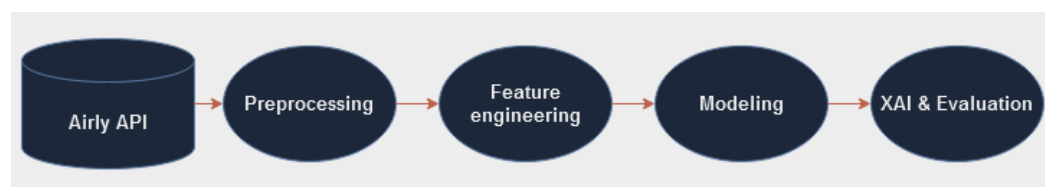


Figure 1. General visualization of ML data pipeline used in this study.

In the preprocessing sub-pipeline, we interpolated missing values and scaled the data using a robust scaler [24] to ensure robustness against outliers, with a particular emphasis on preserving rare picks, which are crucial for our analysis. During the feature engineering phase, cyclic features such as the time of day and wind direction were created to capture the inherent periodicity in these features. Additionally, features such as cardinal wind direction and sunrise and sunset times were developed, along with the inclusion of holidays and social events in Krakow. PM2.5 components such as seasonality and trend, derived from STL decomposition [25], were also integrated. Lag features based on autocorrelation analysis and exploratory data analysis (EDA) were created to enhance the predictive model. In the modeling phase, a model factory was established that generated various models available in the Darts library [26], with specific configurations and hyperparameters, enabling systematic experimentation and optimization. For model evaluation, backtesting with expanding window optimization was used, focusing on regression metrics to assess performance and reliability. Additionally, reports were generated on the analysis of model residuals and XAI analyses, including techniques like SHAP (SHapley Additive exPlanations) [27].

2.3. Machine Learning Forecasting

The study made use of Global Forecasting Models. It is an approach that enables the construction of a single predictive model for multiple, geographically dispersed time series simultaneously. Its aim is to capture the core patterns governing the series, thereby minimizing the potential noise that each series may introduce. This approach is computationally efficient, easy to maintain, and stable in generalizations across various time series. However, it comes at the cost of a shallower understanding of the individual characteristics of each series separately [28].

2.3.1. Models

This study was initially conducted on 12 distinct models, selecting 5 of them for parameter tuning based on their performance. The models were trained globally using data from a total of 455,520 measurements collected from 52 sensors. Each sensor recorded 8760 measurements, covering the entire year from 1 January 2022, at 00:00:00, to 1 January 2023, at 00:00:00.

The first group includes linear models such as ridge regression, which extends traditional linear regression by incorporating L2 norm loss functions [24]. In addition to basic linear models, specially adapted linear models for time series such as DLinear and NLinear were examined. The DLinear model decomposes input data into trend and seasonal

components, processes these through single-layer linear transformations, and combines the results for final prediction, showing particular efficacy with trend-heavy data. The NLinear model enhances data adaptability by subtracting the last value of the sequence prior to processing through a linear layer, and reintegrating it post-transformation, which improves adaptation to data changes and enhances overall model performance [29].

Traditional time series models, such as ARIMA (AutoRegressive Integrated Moving Average), were also evaluated. ARIMA effectively integrates autoregressive and moving average components to model time series data, adeptly capturing underlying trends and seasonality [30].

Models based on decision trees and gradient boosting also demonstrated robust performance. Decision tree models utilize trees to make predictions and have proven effective across various tasks. Gradient boosting models enhance this approach by sequentially correcting the errors of previous models in the ensemble and continuously refining predictions.

The evaluation also covered advanced deep learning models, including variants of recurrent neural networks (RNNs) such as GRU, LSTM, and NBEATS, as well as transformers like the Temporal Fusion Transformer (TFT). The GRU, a streamlined version of the LSTM, uses update and reset gates to manage information flow and is characterized by having fewer trainable parameters, which can reduce training time but may impact performance in complex scenarios. The TFT employs the transformer architecture to precisely model temporal dependencies using an attention mechanism, adept at incorporating both global trends and local fluctuations and introduces a probabilistic element that is especially useful in forecasting time series data with inherent uncertainties [31].

2.3.2. Evaluation

In order to evaluate the models, backtesting with expanding window optimization was utilized, as shown in Figure 2. Given that the training is based on a historical dataset, the amount of data will increase over time. Backtesting allows for a more precise assessment of the model's efficiency after incorporating new data [32]

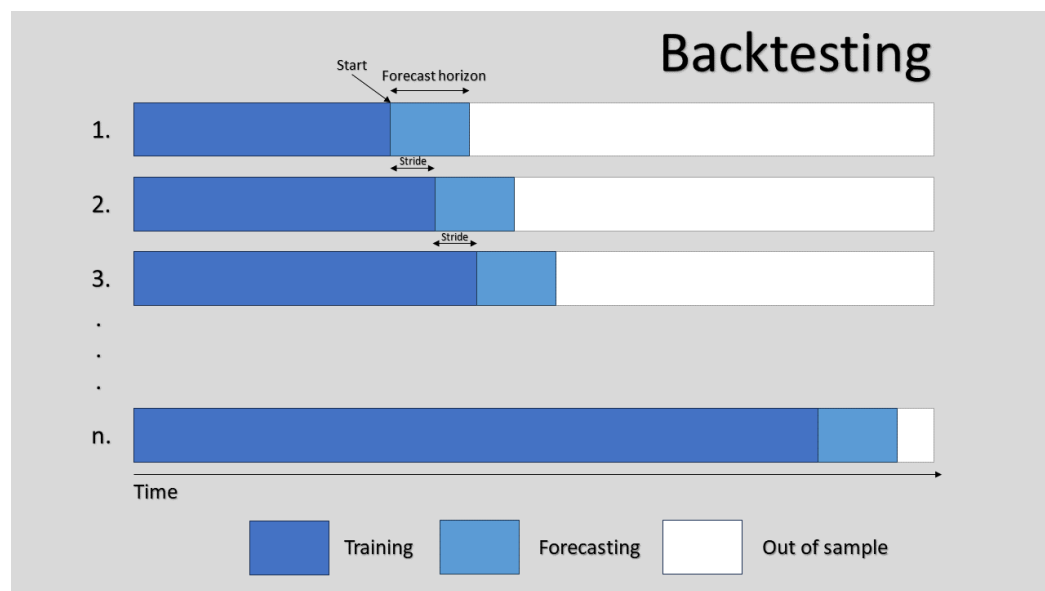


Figure 2. General conception of backtesting with expanding window optimization.

Predicted data after backtesting were validated using metrics such as MAE, RMSE, R2, and MAPE. The Mean Absolute Error (MAE) offers a straightforward interpretation of the error in the units of the predicted variable, allowing for easy understanding of prediction accuracy. In contrast, the Root Mean Square Error (RMSE) is more sensitive to outliers compared to MAE, which makes it valuable for highlighting issues in prediction models

where outliers are significant. Relative Risk Error (MARRE) provides an additional layer of understanding by measuring the relative error in predictions compared to a benchmark or base model, helping to rate model performance in a comparative context. Additionally, the Mean Absolute Percentage Error (MAPE) evaluates the model by examining its capacity to explain the variability of the data, providing insights into the accuracy of predictions relative to the actual values [24,26].

$$\text{MAE}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (1)$$

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (2)$$

$$\text{R2}(y, \hat{y}) = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (3)$$

$$\text{MAPE}(y, \hat{y}) = \frac{100\%}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i} \quad (4)$$

$$\text{MARRE}(y, \hat{y}) = 100 \cdot \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{\max(y_i) - \min(y_i)} \right| \quad (5)$$

where N is the number of observations, y_i is the actual value of the i -th observation, \hat{y}_i is the predicted value of the i -th observation, \bar{y} is the mean of dependent variables, $\max(y_i)$ is the maximum value in the series, and $\min(y_i)$ is the minimum value in the series.

3. Results

3.1. Urban Development

Figure 3 illustrates the trend of PM2.5 concentration in Krakow over 10 years, along with a population chart of the city. A steady increase in the population residing in Krakow is evident. The growth rate during the studied period can be divided into two phases—a low growth rate of the population from 2010 to 2014, followed by a rapid increase in the years 2015 to 2019. The opposite behavior is observed for the PM2.5 trend line—a steady decrease is visible over the past few years, with disruption of the declining trend in 2015 and 2016. The most rapid decline in PM2.5 concentration is observed in the periods 2013–2014, in the year 2017, and 2019.

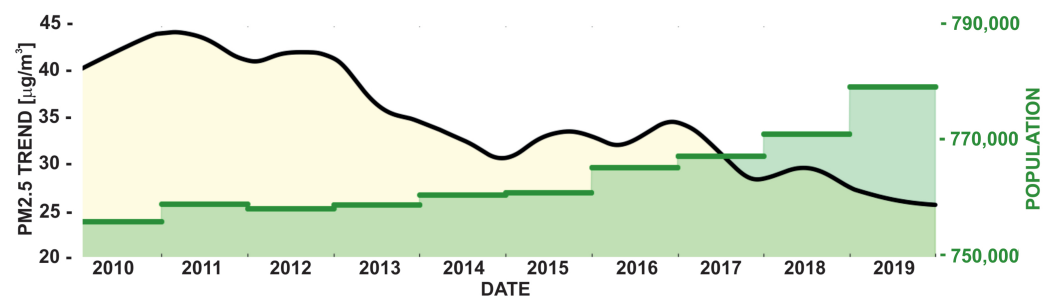


Figure 3. Krakow's population (green) and PM2.5 trend (black/yellow) in the period 2010–2019.

Figure 4 illustrates the development of urban infrastructure over the years 2010–2019. A clear, linear, and steady increase in the number of registered passenger cars is evident. It is noteworthy that out of the total number of registered cars in the city of Krakow in 2019, exceeding 650,000 vehicles, the total share of electric cars in the entire country is just over 6000, of which almost 3000 are hybrid cars [33]. There is also a noticeable increase in the number of bicycle lanes in the city, with their total length increasing by over 50 km over 10 years. The tram lines measured in kilometers remain relatively stable during the analyzed

period, except for a significant decline in the year 2012. As for the total bus lines length, also expressed in kilometers, a decreasing trend is visible until 2014, followed by a clear rebound towards an increasing trend in the subsequent years.

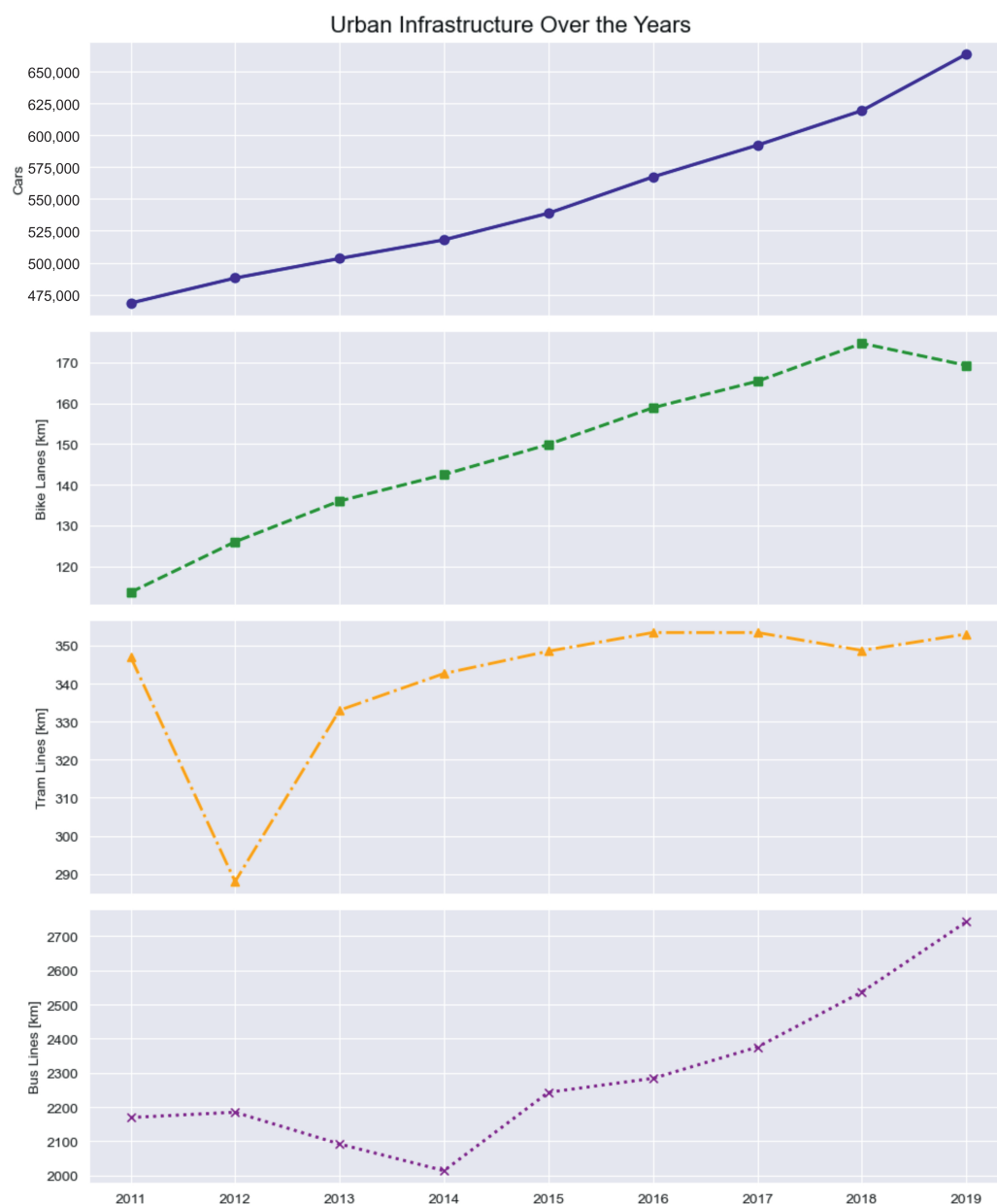


Figure 4. Krakow’s urban infrastructure over the years 2010–2019—number of registered cars (blue), bike lines in kilometers (green), tram lines in kilometers (orange), and bus lines in kilometers (purple).

3.2. Energy Transition

In 2015, almost 24,000 furnaces, boilers, and combustion chambers fueled by coal-based products were registered in Krakow [34]. Figure 5 describes the number of decommissioned furnaces of this type—both those associated with building insulation and those intended for water heating. The graph also includes the number of installed renewable energy installations. A noticeable increase in the number of dismantled furnaces and boilers was observed in 2017 and subsequent years, with a moderate number of decommissions from 2014 to 2016. Unfortunately, in the case of installations based on renewable energy sources, there is no upward trend observed during the analyzed period, and their number remains low and relatively stable each year except for 2018, when the number almost reached 500.

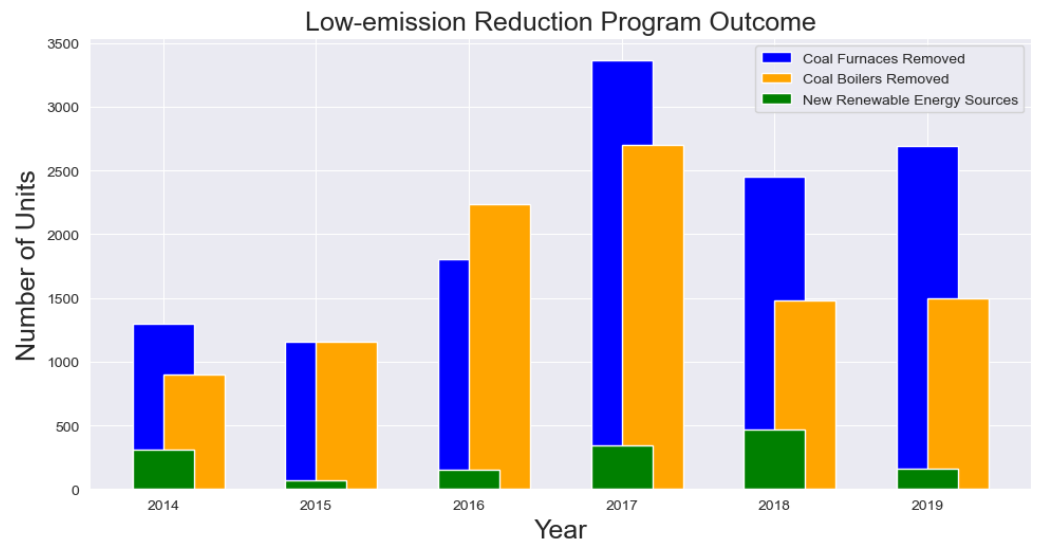


Figure 5. Krakow’s low-emission reduction program (PONE) in years 2014-2019. Number of removed coal furnaces (blue), coal boilers (orange), and number of new renewable energy sources (green).

3.3. Machine Learning Forecasts

All ML models were trained and evaluated using the backtesting technique with expanding window optimization (Figure 2). The starting point for backtesting was set at 80% of the available data, corresponding to the date 20 October 2022 at 01:00:00. The forecast horizon was systematically shifted in increments of one time step, while maintaining a constant forecast horizon of 24 h.

Analyzing the results of various models from Table 1, it can be observed that linear models such as DLinear, NLinear, and Ridge showed the best PM2.5 forecasting potential. During training, the loss function, specifically MSELoss, was monitored to prevent overfitting, and for the DLinear model, training was concluded after five epochs (Figure 6).

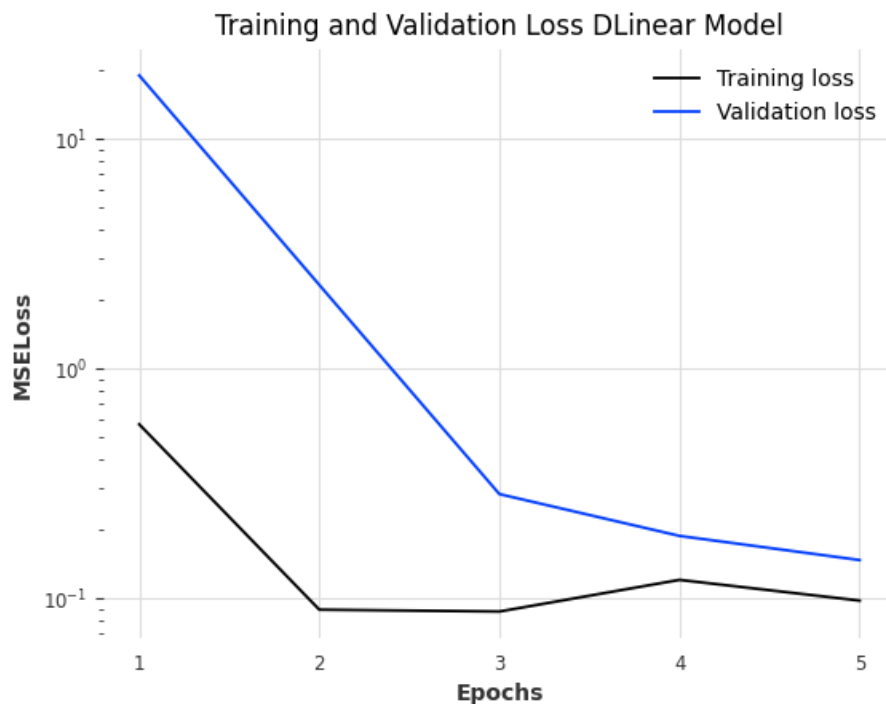


Figure 6. Loss function DLinear.

They are characterized by low values of average errors, such as MAE and RMSE, and a high coefficient of determination R2. Particularly, the DLinear and Ridge Regression models stand out, achieving very high R2 values of 0.947 and 0.956, respectively. The deep learning and linear models adjusted for time series, achieved the best results with a time window of 168 samples (7 days).

Table 1. Model performance table.

Model Performance for 24 h Prediction					
Model	MAE	RMSE	R2	MAPE	MARRE
Ridge	2.666	3.867	0.956	15.621	1.859
ARIMA	6.688	9.282	0.755	36.417	4.622
XGBoost	4.104	6.763	0.879	19.234	2.747
CatBoost	3.839	6.236	0.897	18.463	2.569
LGBM	4.863	7.445	0.850	27.175	3.340
GRU	5.170	7.790	0.831	25.855	3.582
LSTM	5.258	7.704	0.830	27.206	3.682
NBEATS	12.000	17.915	0.079	76.314	8.547
TCN	13.276	19.651	−0.108	68.585	9.448
TFT	3.915	5.971	0.900	17.675	2.710
NLinear	3.356	4.695	0.932	20.706	2.418
DLinear	2.947	3.888	0.947	20.354	2.210

Non-linear models such as tree-based, gradient boosting, or deep learning models struggle with rapid peaks in PM2.5 concentration. Figure 7 shows the comparison between the DLinear and XGBoost models. Linear models handle these sudden jumps better. However, it is evident that the larger the peak, the greater the error made by the models.

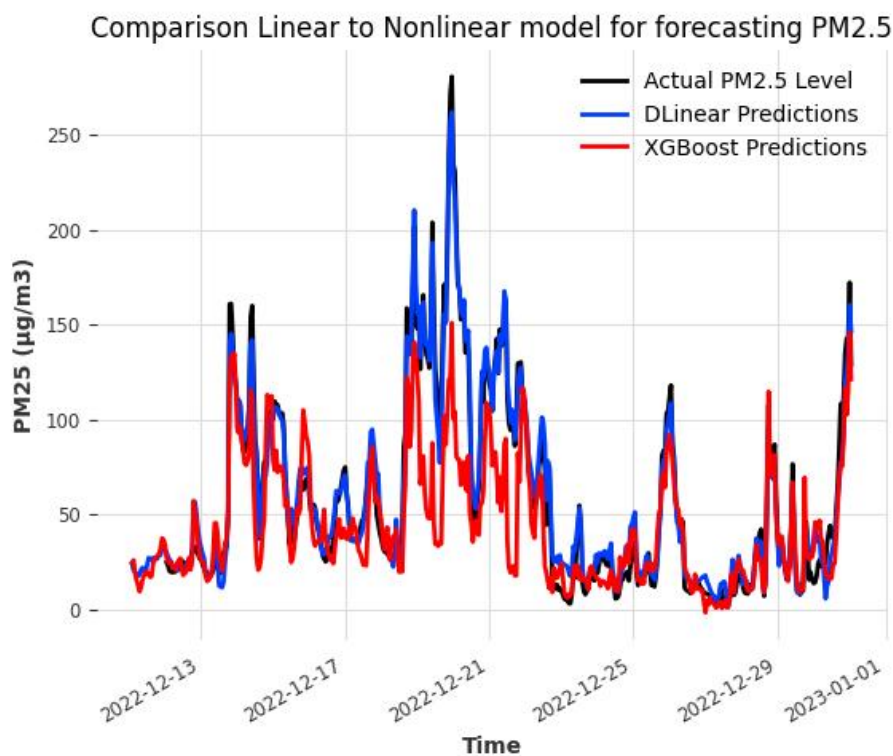


Figure 7. Comparison between DLinear and XGBoost for forecasting PM2.5.

4. Discussion

The rapid population growth in the city, particularly since 2014, does not appear to have a negative impact on air pollution levels in the analyzed urban area. This positive observation suggests the feasibility of constructing densely populated urban centers while maintaining healthy, high-quality air. It is particularly important currently when we can observe people migrating to Krakow. The significant decline in pollution in Krakow during the years 2013–2014 may be directly linked to the adoption of the first anti-smog law in the country by the provincial assembly of Malopolska [35]. This legislation included defining permissible types of fuels for use in the Krakow Municipal District. Additionally, during a comparable period, the PONE program, which subsidizes replacing coal-fired heating installations, began to be implemented. These legal changes were preceded by numerous grass-roots actions and social protests, resulting in the establishment of the Krakow Smog Alarm in 2012 [36].

A clear relationship is evident between the type of fuel used for home heating and air quality, independent of population growth. However, a multidimensional analysis of transportation in the city in the context of population growth and declining PM_{2.5} trends provides particularly intriguing insights. There is a steady increase in the number of registered vehicles, predominantly fueled by combustible engines, which correlates positively and demonstrates a clear cause-and-effect relationship with the city's growing population. However, the PM_{2.5} pollution trend is decreasing. With the increase in individual transportation, the number of available public transportation routes also rises. The rapid growth of individual transportation may negatively impact residents' willingness to use public transport, especially buses, which travel on the same streets as cars. In the concept of a smart city, the relatively stable number of available tram lines should be systematically increased. The expansion of bicycle routes has been positive, as this mode of transportation is truly zero-emission in terms of propulsion and provides additional health benefits from aerobic exercise. Therefore, the observed negative correlation between the increased number of passenger cars and decreased PM_{2.5} concentration may be easily misinterpreted. The number of registered cars does not necessarily involve the highest number of cars in the city. As we observe a general increase in available public transportation, personal cars may be used for weekend trips, not for everyday commuting. Secondly, it was proved that the most important factor was coal-based heating according to isotropic [10] and geostatistical [37] studies.

Based on this data, the relatively small impact of transportation on the overall PM_{2.5} trend in Krakow is evident. The trend in PM_{2.5} concentrations is dominated by the combustion of solid fuels, mainly coal for heating homes and water, and is strongly influenced by annual seasons, as confirmed by various analyses [38] including big data analyses [39]. It should be emphasized that these studies focus solely on the analysis of suspended particulate matter in the air and do not consider other health-detrimental volatile compounds, the occurrence of which may be dominated by transportation.

As demonstrated above, the impact of transportation is significantly smaller compared to household heating on PM_x concentrations. Therefore, it is justified, especially during the autumn–winter–spring period, to utilize data analysis and machine learning techniques for predicting and optimally managing the city in the event of negative episodes concerning public health. These actions must encompass not only issues related to promoting public transportation, which, as reiterated, does not have a dominant influence on this type of pollution but also actions considering limitations on pollution influx from neighboring municipalities, which is a dominant factor [11].

In general, linear models, such as DLinear, NLinear and Ridge Regression, perform better for predicting PM_{2.5} pollution levels, especially in the face of rapid concentration peaks. Thanks to the L2 regularization mechanism, which limits the size of the coefficients, these models are more resistant to extreme values, which is crucial when pollution data may exhibit sudden changes due to unusual pollution influx—especially in the winter season. Tree-based models, gradient-boosted, and deep neural networks may struggle

with such data because their nonlinear nature leads them to overfit local details in the training data, including atypical pollution patterns. This overfitting can result in modeling overly complex patterns that may not be effective in predicting future changes in PM_{2.5} concentrations, especially under dynamically changing external conditions. Consequently, linear models can offer greater stability and predictability in the analysis of time series data concerning PM_{2.5} pollution, which is important for environmental planning and industrial regulations.

5. Conclusions

Our investigation into the impacts of population growth and other factors on PM_x concentrations reveals significant insights. Over the past decade, the area we studied experienced a 3% increase in population. However, our data indicate that this growth did not negatively affect air quality, as PM_x concentrations actually decreased by 40%. This suggests that the area has successfully managed air quality despite increased urban density since 2014.

Moreover, our findings highlight that the predominant factor for enhancing air quality, specifically concerning PM_x, has been the transition from coal-based energy sources to more environmentally sustainable options such as natural gas or renewable energy sources. Interestingly, while transportation was initially considered a potential major factor, it does not constitute the primary influence on fluctuations in PM_x pollution levels within the city.

Additionally, our study indicates that the energy transition for PM_x mitigation appears to occur at two speeds: there has been a fast change in heating energy sources, which has been a main factor, while the transition in transportation has been slower. Furthermore, the potential of big data and automated prediction tools has been shown to be crucial for better crisis response planning within smart cities.

Regarding the performance of linear models (specifically DLinear with an MAE of 2.95, NLinear with an MAE of 3.36, and Ridge Regression with an MAE of 2.67) in predicting PM_{2.5} levels often outperform nonlinear models such as TCN with an MAE of 13.27 and classic ARIMA with an MAE of 6.69. This performance is especially notable when rapid changes occur within the data. This is due to their stability and resistance to overfitting, which is attributed to mechanisms like L2 regularization that constrain coefficient sizes, making them less sensitive to extreme values in pollution data.

Author Contributions: Conceptualization, M.Z. and S.C.; methodology, M.Z., S.C., E.W. and T.D.; validation, S.C., M.Z., E.W. and T.D.; formal analysis, M.Z., S.C., E.W. and T.D.; investigation, M.Z., S.C., E.W. and T.D.; data curation, T.D.; writing—original draft preparation, M.Z., S.C. and E.W.; writing—review and editing, M.Z., S.C., E.W. and T.D.; visualization, S.C., M.Z., E.W. and T.D.; supervision, M.Z. and T.D.; project administration, T.D. and M.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported as a part of the statutory project by AGH University of Krakow, Faculty of Geology, Geophysics and Environmental Protection.

Data Availability Statement: Availability of data and materials. Publicly available datasets from Airly sensors were analyzed in this study and can be found here: (<https://map.airly.org/>, accessed on 29 April 2024). API documentation from Airly is available here: (<https://developer.airly.org/en/docs>, accessed on 29 April 2024). Publicly available datasets from the Chief Inspectorate for Environmental Protection database were analyzed in this study. This data can be found here: (<http://powietrze.gios.gov.pl/pjp/home>, accessed on 29 April 2024). API documentation is available here: (<http://powietrze.gios.gov.pl/pjp/content/api>, accessed on 29 April 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ARIMA	AutoRegressive Integrated Moving Average
EDA	Exploratory Data Analysis
GRU	Gated Recurrent Unit
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MARRE	Relative Risk Error
ML	Machine Learning
NBEATS	Neural Basis Expansion Analysis Time Series
PLN	Polish Zloty
PONE	Low-Emission Reduction Program in Krakow
R ²	R-squared (Coefficient of Determination)
RMSE	Root Mean Square Error
SHAP	SHapley Additive exPlanations
TFT	Temporal Fusion Transformer
XAI	Explainable AI

References

- Zhang, L.; Su, D.; Guo, W.; Li, S. Empirical study on urban sustainable development model based on identification of advantages and disadvantages. *Front. Sustain. Cities* **2022**, *4*, 894658. [CrossRef]
- Wesz, J.G.B.; Miron, L.I.G.; Delsante, I.; Tzortzopoulos, P. Urban Quality of Life: A Systematic Literature Review. *Urban Sci.* **2023**, *7*, 56. [CrossRef]
- Keith, M.; Birch, E.; Buchoud, N.J.A.; Cardama, M.; Cobbett, W.; Cohen, M.; Elmqvist, T.; Espey, J.; Hajer, M.; Hartmann, G.; et al. A new urban narrative for sustainable development. *Nat. Sustain.* **2023**, *6*, 115–117. [CrossRef]
- Liang, L.; Gong, P. Urban and air pollution: A multi-city study of long-term effects of urban landscape patterns on air quality trends. *Sci. Rep.* **2020**, *10*, 18618. [CrossRef] [PubMed]
- Jonek-Kowalska, I. Assessing the Effectiveness of Air Quality Improvements in Polish Cities Aspiring to Be Sustainably Smart. *Smart Cities* **2023**, *6*, 510–530. [CrossRef]
- Central Statistical Office of Poland. *Energy Statistics—Energy Report 2023*; Central Statistical Office of Poland: Warsaw, Poland, 2023.
- Bokwa, A. Environmental impacts of long-term air pollution changes in Krakow, Poland. *Pol. J. Environ. Stud.* **2008**, *17*, 673–686.
- Oleniacz, R.; Gorzelnik, T. Assessment of the variability of air pollutant concentrations at industrial, traffic and urban background stations in Krakow (Poland) using statistical methods. *Sustainability* **2021**, *13*, 5623. [CrossRef]
- Danek, T.; Zareba, M. The Use of Public Data from Low-Cost Sensors for the Geospatial Analysis of Air Pollution from Solid Fuel Heating during the COVID-19 Pandemic Spring Period in Krakow, Poland. *Sensors* **2021**, *21*, 5208. [CrossRef] [PubMed]
- Wojewodzki Inspektorat Ochrony Środowiska w Krakowie. Jakość powietrza w Krakowie. Podsumowanie wyników badań. 2020. Available online: <https://krakow.wios.gov.pl/2020/09/jakosc-powietrza-w-krakowie-podsumowanie-wynikow-badan/> (accessed on 12 April 2024).
- Danek, T.; Weglinska, E.; Zareba, M. The influence of meteorological factors and terrain on air pollution concentration and migration: A geostatistical case study from Krakow, Poland. *Sci. Rep.* **2022**, *12*, 11050. [CrossRef]
- Zareba, M.; Danek, T.; Zając, J. On Including Near-surface Zone Anisotropy for Static Corrections Computation—Polish Carpathians 3D Seismic Processing Case Study. *Geosciences* **2020**, *10*, 66. [CrossRef]
- Government of Poland. *A Sustainable Tourism Policy for Kraków in the Years 2021–2028*; Government of Poland: Warsaw, Poland, 2022.
- European Commission, Directorate-General for Research and Innovation. *EU Missions—100 Climate-Neutral and Smart Cities*; European Commission: Brussels, Belgium, 2024.
- Government of Poland. *mObywatel 2.0*; Government of Poland: Warsaw, Poland, 2024.
- Zareba, M.; Danek, T.; Stefaniuk, M. Unsupervised Machine Learning Techniques for Improving Reservoir Interpretation Using Walkaway VSP and Sonic Log Data. *Energies* **2023**, *16*, 493. [CrossRef]
- Rana, R.; Kalia, A.; Boora, A.; Alfaisal, F.M.; Alharbi, R.S.; Berwal, P.; Alam, S.; Khan, M.A.; Qamar, O. Artificial Intelligence for Surface Water Quality Evaluation, Monitoring and Assessment. *Water* **2023**, *15*, 3919. [CrossRef]
- Zaresefat, M.; Derakhshani, R. Revolutionizing Groundwater Management with Hybrid AI Models: A Practical Review. *Water* **2023**, *15*, 1750. [CrossRef]
- Uriarte-Gallastegi, N.; Arana-Landín, G.; Landeta-Manzano, B.; Laskurain-Iturbe, I. The Role of AI in Improving Environmental Sustainability: A Focus on Energy Management. *Energies* **2024**, *17*, 649. [CrossRef]

20. Zareba, M.; Dlugosz, H.; Danek, T.; Weglinska, E. Big-Data-Driven Machine Learning for Enhancing Spatiotemporal Air Pollution Pattern Analysis. *Atmosphere* **2023**, *14*, 760. [CrossRef]
21. Biuletyn Informacji Publicznej. Kraków w Liczbach. 2011–2019. Available online: <https://www.bip.krakow.pl/?mmi=6353> (accessed on 13 May 2024).
22. Kraków, M.P.K.M. Program Ograniczania Niskiej Emisji. 2019. Available online: https://www.krakow.pl/aktualnosci/209034,29,komunikat,mobi_short,program_ograniczania_niskiej_emisji_w_pigulce.html (accessed on 13 May 2024).
23. QuantumBlack, M. Kedro Documentation. Available online: <https://docs.kedro.org/en/stable/> (accessed on 10 May 2024).
24. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
25. Cleveland, R.B. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *J. Off. Stat.* **1990**, *6*, 3–73.
26. Herzen, J.; LÃssig, F.; Piazzetta, S.G.; Neuer, T.; Tafti, L.; Raille, G.; Pottelbergh, T.V.; Pasiaka, M.; Skrodzki, A.; Huguenin, N.; et al. Darts: User-Friendly Modern Machine Learning for Time Series. *J. Mach. Learn. Res.* **2022**, *23*, 1–6.
27. Lundberg, S.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
28. Amat Rodrigo, J.; Escobar Ortiz, J. *Skforecast*; Zenodo: Geneva, Switzerland, 2023. [CrossRef]
29. Zeng, A.; Chen, M.; Zhang, L.; Xu, Q. Are Transformers Effective for Time Series Forecasting? *Proc. AAAI Conf. Artif. Intell.* **2022**, *37*, 11121–11128. [CrossRef]
30. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*, 2nd ed.; OTexts: Melbourne, Australia, 2018.
31. Lim, B.; Arik, S.O.; Loeff, N.; Pfister, T. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting. *arXiv* **2020**, arXiv:1912.09363.
32. Yang, R. Omphalos, Uber’s Parallel and Language-Extensible Time Series Backtesting Tool. 2018. Available online: <https://www.uber.com/en-PL/blog/omphalos/> (accessed on 10 May 2024).
33. Elektrycznych, P.S. Licznik Elektromobilności: Wzrost Liczby Samochodów Elektrycznych na Polskich Drogach o Prawie 90% r/r (sierpień 2019). Available online: <https://psnm.org/2019/informacja/licznik-elektromobilnosci-wzrost-liczby-samochodow-elektrycznych-na-polskich-drogach-o-prawie-90-r-r-sierpien-2019> (accessed on 27 April 2024).
34. City of Krakow. Program Ograniczania Niskiej Emisji w Pigulce. City of Krakow Website. 2024. Available online: https://www.krakow.pl/aktualnosci/209034,29,komunikat,mobi_short,program_ograniczania_niskiej_emisji_w_pigulce.html (accessed on 27 April 2024).
35. Marshal’s Office of the Małopolska Region. *Uchwała Nr XVIII/243/16 Sejmiku Województwa Małopolskiego z Dnia 15 Stycznia 2016 r. w Sprawie Wprowadzenia Na Obszarze Gminy Miejskiej Kraków Ograniczeń w Zakresie Eksploatacji Instalacji, w Których Następuje Spalanie Paliw*; Marshal’s Office of the Małopolska Region: Kraków, Poland, 2016.
36. Krakowski Alarm Smogowy. Krakowski Alarm Smogowy z Tytułem Człowieka Roku Polskiej Ekologii. 2016. Available online: <https://www.gramzielone.pl/walka-ze-smogiem/21385/krakowski-alarm-smogowy-z-tytulem-czlowieka-roku-polskiej-ekologii> (accessed on 13 May 2024).
37. Zareba, M.; Danek, T. Analysis of Air Pollution Migration during COVID-19 Lockdown in Krakow, Poland. *Aerosol Air Qual. Res.* **2022**, *22*, 210275. [CrossRef]
38. Wielgościński, G.; Czerwińska, J. Smog Episodes in Poland. *Atmosphere* **2020**, *11*, 277. [CrossRef]
39. Zareba, M.; Weglinska, E.; Danek, T. Air pollution seasons in urban moderate climate areas through big data analytics. *Sci. Rep.* **2024**, *14*, 3058. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.