POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)
URL: http://dergipark.org.tr/politeknik

Detection of anomalous Nitrogen Dioxide (NO_2) concentration of a district in Ankara: a reconstruction-based approach

Ankara'nın bir ilçesindeki Azot Dioksit (NO_2) konsantrasyonunun anomali tespiti: rekonstrüksiyona-dayalı bir yaklaşım

Yazar(lar) (Author(s)): *Mustafa Murat ARAT[1]*

ORCID[1]: 0000-0003-3740-5135

To cite to this article: Arat M. M., "Detection of anomalous Nitrogen Dioxide concentration of Ankara: a Reconstruction-based approach", ***Journal of Polytechnic***, 28(1): 101-114, (2025).

Bu makaleye şu şekilde atıfta bulunabilirsiniz: Arat M. M., "Detection of anomalous Nitrogen Dioxide concentration of Ankara: a Reconstruction-based approach", ***Politeknik Dergisi***, 28(1): 101-114, (2025).

Erişim linki (To link to this article): http://dergipark.org.tr/politeknik/archive

DOI: 10.2339/politeknik.1419512

Detection of Anomalous Nitrogen Dioxide (NO_2) Concentration of A District in Ankara: A Reconstruction-Based Approach

Highlights

- *A first study in the whole literature of environmental science about air quality for anomaly detection of NO2 levels using SOTA.*
- *Instead of utilizing raw univariate time series data, this study involves incorporating engineered features based on past and current values as well as the integration of periodic variables.*
- *Anomalies were detected using Variational Autoencoder, a deep learning approach in a supervised manner.*
- *A model was developed with a high true positive rate and minimized false alarms.*

Graphical Abstract

In this study, an anomaly detection algorithm using Variational Autoencoder was developed for NO2 concentration levels of Bahçelievler district of Ankara.

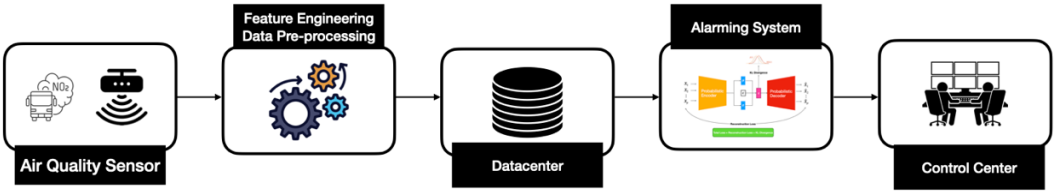

Figure. A system design for detection of anomalies of ground-level NO_2

Aim

Detection of anomalies of ground-level NO2 concentration levels with the intention of just-in-time intervention and/or detecting a potential malfunction in error-prone sensors.

Design & Methodology

After high quality, relevant, and appropriately transformed data was obtained, classical statistical approaches as well as machine learning- and deep learning-based anomaly detection algorithms were benchmarked.

Originality

This study addresses a critical gap in environmental science by developing an algorithm to diagnose anomalies in ground-level NO2 concentration level, with focus on specifically a densely trafficked district with old households, introducing novel feature engineering incorporating past and current values and periodic variables for enhanced anomaly detection, surpassing traditional reliance on meteorological predictors.

Findings

Variational Autoencoder emerges as noteworthy, exhibiting superior performance with %98 recall, %82 precision and %0.12 false alarm rate.

Conclusion

As a first study in the whole literature for anomaly detection of NO2 levels using SOTA, the model was successfully developed and it can be serialized to be put behind a mobile or web application in order to detect anomalies automatically so to give alerts when needed.

Declaration of Ethical Standards

The author(s) of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

Detection of Anomalous Nitrogen Dioxide (NO_2) Concentration of A District in Ankara: A Reconstruction-Based Approach

Araştırma Makalesi / Research Article

Mustafa Murat ARAT[1*]

[1]Department of Statistics, Hacettepe University, Ankara Türkiye

(Geliş/Received : 14.01.2024 ; Kabul/Accepted : 06.03.2024 ; Erken Görünüm/Early View : 03.06.2024)

ABSTRACT

Air quality significantly impacts human health, particularly in urban areas, leading to global morbidity and mortality. Elevated air pollutant levels pose health risks, emphasizing the need for timely monitoring and detection. This study adopts an innovative approach to identify anomalies of hourly NO2 concentration levels in a district of Ankara, Turkey. Leveraging both traditional statistical approaches and state-of-the-art techniques, the research aims to provide real-time alerts. Employing a multivariate strategy, the study generates new features based on historical data, and incorporates periodic variables, as well. Among the methods explored, Variational Autoencoder emerges as noteworthy, exhibiting superior performance with %98 recall, %82 precision and %0.12 false alarm rate. This approach not only demonstrates a high true positive rate, enhancing its efficacy in anomaly detection but also effectively mitigates false alarms, preventing alert fatigue. By using advanced methodologies with a focus on NO2 levels, the study contributes to proactive measures for public health, enabling prompt responses to potential air quality issues.

Keywords: Air quality, anomaly detection, variational autoencoders, NO_2 concentration, machine learning.

Ankara'nın Bir İlçesindeki Azot Dioksit (NO_2) Konsantrasyonunun Anomali Tespiti: Rekonstrüksiyona-Dayalı Bir Yaklaşım

ÖZ

Hava kalitesi, özellikle kentsel bölgelerde önemli ölçüde insan sağlığını etkilemekte ve tüm dünyada morbiditeye ve mortaliteye yol açmaktadır. Bu nedenle kirleticilerin zamanında izlenmesi ve potansiyel problemlerin erken bir şekilde tespit edilmesi gerekmektedir. Bu çalışma, Ankara'nın bir ilçesindeki saatlik NO2 konsantrasyon seviyesindeki anomalileri belirlemek için yenilikçi bir yaklaşım benimsemektedir. Geleneksel istatistiksel yöntemler ile birlikte son teknoloji teknikleri kullanan bu araştırma ile gerçek-zamanlı uyarıların elde edilmesi amaçlanmaktadır. Uygulanan çok değişkenli strateji, geçmiş verilere dayalı yeni özniteliklere ek olarak ayrıca periyodik değişkenleri de kullanmaktadır. İncelenen yöntemler arasında, %98 duyarlılık, %82 kesinlik and %0.12 yanlış alarm oranı ile Varyasyonel Otokodlayıcı yöntemi üstün performans sergileyerek dikkat çekmektedir. Elde edilen model, sadece yüksek bir gerçek pozitif oranı elde etmekle kalmayıp, aynı zamanda yanlış alarmları etkili bir şekilde azaltarak alarm yorgunluğunu önlemektedir. Özellikle NO2 hava kirleticisi üzerinde gelişmiş metodolojiler kullanan bu çalışma, böylelikle halk sağlığına yönelik proaktif önlemlere katkıda bulunarak potansiyel hava kalitesi sorunlarına hızlı bir şekilde yanıt verilmesini desteklemektedir.

Anahtar Kelimeler: Hava kalitesi, anomali tespiti, varyasyonel otokodlayıcılar, NO_2 konsantrasyonu, makine öğrenmesi.

1. INTRODUCTION

Air quality profoundly impacts human health and well-being, particularly of those in urban areas causing global morbidity and mortality. Exposure to elevated levels of air pollutants can lead to respiratory problems [1] [2], cardiovascular issues [3], and other adverse health effects [4] [5]. Thereby, nations and international organizations are actively engaged in concerted efforts to monitor and alleviate the pervasive issue of air pollution [6]. Implementing comprehensive monitoring systems, fostering collaborative research initiatives, and formulating stringent environmental policies are some of the strategies being employed to curb the adverse effects of air pollution, which is measured by particulate matter (fine particles PM2.5 and coarse particles PM10), sulfur dioxide (SO2), carbon monoxide (CO), nitrogen dioxide (NO2), nitric oxide (NO), Nitrogen Oxides (NOX), and Ozone (O3) [7].

*Sorumlu Yazar (Corresponding Author)
e-posta : muratarat@hacettepe.edu.tr

Air quality in cities in Turkey is constantly monitored thanks to the National Air Quality Monitoring Network operated by the Ministry of Environment, Urbanisation and Climate Change (MoEUCC), and the data collected at Clean Air Centers are published via its website and/or mobile application [8] and can be easily visualized on a country map [9]. While numerous cities across the country are equipped with multiple air quality measurement stations, our study concentrates on Ankara, second-largest city, which has seen rapid population growth and irregular urban expansion, leading to more vehicles and industry. As a key trade hub, it's also a central point in the nation's road network.

Surveillance for regulatory and scientific purposes, along with enhancing air quality, are vital initiatives to protect public health and address the potential enduring effects of air pollution, particularly in densely populated metropolitan areas. Developing a timely anomaly detection-based early warning system is among the most effective strategies to minimize health risks and economic losses associated with air pollution. However, while PM2.5, PM10 and O3 remain extensively studied air pollutants, there is a notable lack of research on NO2. It is well-established that NO2 significantly contributes to the increase in concentrations of other detrimental outdoor air pollutants, such as ground-level PM2.5, PM10, and O3. [10]. NO2 predominantly stems from excessive coal combustion and motor vehicle emissions. Notably, in Turkish cities, road transport and residential coal burning are major contributors to pollution, with vehicle density emerging as the predominant factor influencing air pollution levels.

Anomaly detection (AD) is the process of identifying patterns in data that deviate significantly from the expected behavior. This technique is particularly valuable in applications such as fraud detection and intrusion monitoring. Over the past few years, numerous AD techniques on univariate time series have been developed, which can be categorized as (1) statistical methods (2) classical machine learning (ML) methods, and (3) methods using neural networks, based on the type of the problem such as supervised, semi-supervised, unsupervised [11]. In this context, for this study, we opt for the adoption of widely utilized methods in all these three categories, namely, Z-score method, robust Z-score method, interquartile range method Winsorization method and Hampel filter for statistical approach, isolation forest and one-class support vector machines for ML approach, autoencoders and variational autoencoders for deep learning (DL) approach.

Previous studies predominantly focus on forecasting for NO2 levels, leaving the realm of AD largely unexplored [12] [13] [14] [15] [16]. For AD literature for NO2 concentration, to this date, there are a few studies. Firstly, Aggarwal and Toshniwal (2019) [17] employ a hybrid approach that integrates proximity- and clustering-based AD methods. Their primary objective is to identify anomalies in air quality data across diverse locations in India. Secondly, van Zoest et al. (2018) [18] employs the mean and standard deviation (SD) of the normal distribution underlying the truncated normal distribution of hourly NO2 observations to detect outliers. Lastly, Torres et al. (2011) [19] employ functional outlier detection, treating gas emissions over time as curves and then they identify outliers by comparing these curves rather than using vector-based methods.

In response to this critical scenario, our study addresses an existing gap by diagnosing anomalies (outliers from the norm) of ground-level NO2 concentration levels with the intention of just-in-time intervention and/or detecting a potential malfunction in error-prone sensors [20]. This effort represents a pioneering initiative, marking the first attempt to create such an algorithm specifically tailored for Ankara, with a particular focus on a densely trafficked district with full of old households in the city. Furthermore, this study marks the inaugural step in feature engineering for such a system in environmental science literature. Instead of utilizing raw univariate data, this study involves incorporating features based on past and current values as well as the integration of periodic variables, as opposed to relying on meteorological variables such as wind speed and temperature This is because studies have demonstrated that models incorporating periodic parameters exhibit superior performance compared to models relying solely on meteorological predictors [21]. Finally, the majority of AD methods prioritize achieving high true positives. Through employing an effective methodology, we were able to develop an approach characterized by both a high true positive rate and minimized false alarms.

2. DATA

Study area

This study deviates from the common practice of aggregating measurements across various stations, focusing exclusively on a specific monitoring station in the vibrant Bahçelievler district of Ankara, renowned as a pivotal hub for social activities. This decision stems from the disparities in observations collected from different spatial and temporal contexts, where factors like terrain, emission sources, and monitoring device biases contribute to data inequities.

Named after its Turkish translation, meaning "houses with gardens", the district serves a diverse community, primarily housing college students and in close proximity to numerous business/shopping centers, workplaces and official buildings. The buildings surrounding the station rely on coal/natural gas for heating, contributing to the dynamic atmosphere of the district. The roads adjacent to the station witness consistent and vibrant traffic, accentuating the significance of monitoring air quality fluctuations in this densely populated area. Ankara, as per Turkish Statistical Institute data for 2023, boasts the highest number of vehicles among provinces [22], with one car for every three people in its population of 5,397,000. Given the multifaceted nature of this location, the station located at this area, which can be seen in

Figure 1, holds strategic importance in understanding the broader environmental implications within this urban setting.

Data preparation

The detection of anomalies in this context allows researchers to focus on only one air pollutant, without mixing the effects of others, especially, pollutants with huge fluctuation are very critical to catch. Therefore, this study concentrates on hourly N02 concentration level measured in micrograms per cubic meter air or $\mu g/m^3$.

Figure 1. (a) Satellite image showing the location of Bahçelievler station (annotated by big red dot), (b) Bahçelievler air quality measurement station [23]

The raw data comprises NO2 concentration levels obtained from MoEUCC, spanning from May 1st, 2008, to December 17th, 2023. However, due to the significant number of missing values between May 1st, 2008 and May 23rd, 2011, we opt for an analysis focusing on univariate series of 99,101 readings starting from May 23rd, 2011, as shown in **Hata! Başvuru kaynağı bulunamadı.**. Data pre-processing procedure involves engineering new features based on current and past quantities, scaling the numeric variables, and encoding categorical variables.

Feature Engineering

To enhance the dataset, distinct columns for day, month, and year are created by parsing the existing date column. This feature engineering strategy aims to enable a detailed analysis of temporal patterns. Concurrently, a new column indicating the day of the week is introduced to unveil daily patterns in NO2 levels. Additionally, a Boolean variable categorizing each day as a weekend or weekday is introduced. This variable serves to distinguish traffic patterns, aiding in the exploration of variations in NO2 levels influenced by distinct traffic dynamics during weekdays (characterized by rush hours) and weekends. Furthermore, the incorporation of the week of the year variable adds another layer of temporal refinement. This variable, alongside other temporal features, helps identify potential periodicity in NO2 levels associated with specific weeks throughout the year. An additional variable is introduced to indicate official holidays, which contributes a valuable dimension to the analysis by considering factors like reduced vehicular traffic, office closures, altered travel patterns and changes in residential combustion patterns during holidays. This perspective is crucial for understanding how external factors may contribute to variations in NO2. Also, we incorporate rolling-window features, which calculations are depicted in Figure 1, to capture short-term variability in NO2 levels. The rationale for employing a rolling window is in the recognition that such series can exhibit temporal fluctuations influenced by daily or weekly patterns. The first introduced variable is the rolling median of the last 7 days, aimed at smoothing temporal fluctuations and capturing underlying trends in the dataset. The choice of a rolling median, rather than mean, is motivated by its robustness to outliers, mitigating their impact on data interpretation. This stability in representation is crucial for understanding the central tendency of NO2 levels over time.

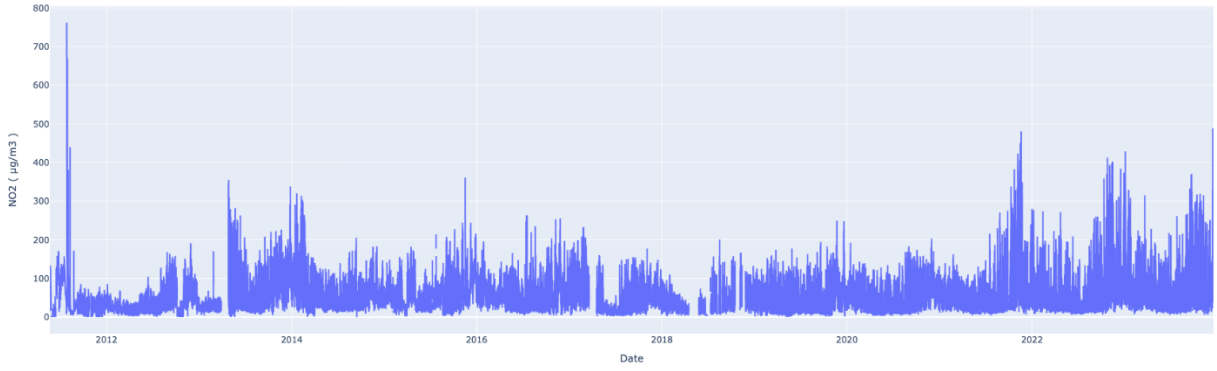

Figure 2. Graphical representation of N02 levels of Bahçelievler district of Ankara from May 23rd, 2011, to December 17th, 2023.

Further variables include the rolling SD of last 7 days, rolling skewness of last 7 days, rolling kurtosis of the last 7 days, coefficient of variation, and percent change from the previous day. The rolling SD provides insight into the variability in NO2 levels within the specified window, assessing deviations from the rolling mean. Rolling skewness aids in understanding the distribution's shape. Also, monitoring rolling kurtosis helps identify periods of increased or decreased tail risk in the NO2 level distribution. The coefficient of variation, calculated as the ratio of rolling SD to rolling mean, provides a normalized measure of dispersion, facilitating comparisons relative to the mean. Calculating the percentage change from the previous day quantifies the daily rate of change in NO2 levels, offering insights into abrupt shifts or trends.

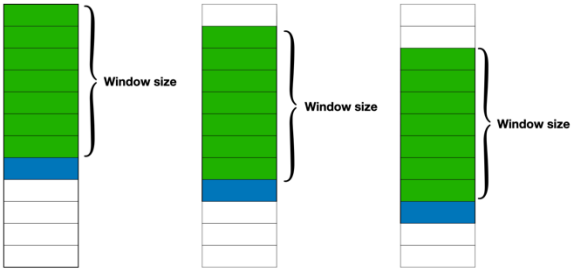

Figure 1. Rolling-window features use values from previous rows in a given fixed window. For any given row (blue rows) we are only using the past n values (n being the window size, shown green rows) to calculate the rolling window features

Training and Testing Split

Unlike traditional tasks, time series data has a sequential structure with temporal order. Random shuffling poses a risk of data leakage, introducing bias and unrealistic performance measures. Preserving temporal order is crucial; earlier periods constitute the training set, and later periods the testing set. This prevents look-ahead bias, ensuring the model relies solely on historical data. Training set spans May 23rd, 2011, to December 31st, 2022; testing uses data from the entire year of 2023. No splitting has been done for classical statistical approaches.

Data Pre-Processing

Following feature engineering, a critical preprocessing step involves scaling numerical features and encoding categorical variables. Scaling standardizes the range of numeric variables, ensuring a common scale. All numeric variables are scaled and centered by subtracting the median and dividing by the interquartile range. This choice minimizes outlier influence, as the median and interquartile range are less susceptible to outliers than the mean and SD. This approach enhances stability, especially in the potential presence of anomalies [24].

Handling categorical variables involves encoding them into numerical form for machine understanding. One-hot encoding transforms each category into a binary vector, assigning one to the category's index and zero to others [25]. However, this method overlooks cyclical patterns present in variables like day of the month or day of the week. Therefore, circular encoding is utilized, resulting in two new variables for each categorical variable X, capturing sine and cosine transformations, as expressed in Equation (1) and Equation (2).

$$\text{var\_sin}(x_i) = \sin\left(\frac{x_i \times 2\pi}{\max(X)}\right) \quad (1)$$

$$\text{var\_cos}(x_i) = \cos\left(\frac{x_i \times 2\pi}{\max(X)}\right) \quad (2)$$

Creating label variable

Despite the inherently unsupervised nature of AD, acquiring a label variable without human intervention is facilitated in this study. Türkiye, a candidate country for the European Union, is in the process of transposing national legislation on the environment, including air quality. Turkish Environment Law [26] emphasizes air quality protection and pollution prevention, allowing for the definition of rules and principles by MoEUCC through regulations. Key legislation regarding Air Quality Management includes the By-law on Air Quality Assessment and Management (O.G. 06.06.2008, no: 26898) [27] and Circular on Air Quality Assessment and Management (Circular No 2013/37) [28]. This study aligns with the strategy outlined by the Law. Since January 1, 2008, the NO2 limits of ambient air for human health and issued by the Ministry of Environment, Urbanization, Culture, and Climate (MoEUCC), currently in effect, are outlined in Table 1.

Table 1. NO_2 air pollutant limit values for Türkiye between 2008 and 2024.

Years	***2008***	***2009***	***2010***	***2011***	***2012***	***2013***	***2014***	***2015***	***2016***
Hourly (for human health)	300	300	300	300	300	300	300	290	280
Annual (for human health)	100	100	90	80	70	60	60	56	52
Years	***2017***	***2018***	***2019***	***2020***	***2021***	***2022***	***2023***	***2024***	
Hourly (for human health)	270	260	250	240	230	220	210	200	
Annual (for human health)	48	44	40	40	40	40	40	40	

In accordance with the specified limits in this table, any hourly measurement throughout the year exceeding the corresponding limit is labeled as an anomaly, while values below this threshold are classified as normal. This newly-established variable serves as label, where it can be used to assess the evaluation of model performance as well as performance of hyperparameter tuning for the selected ML models. Applying this rule to NO2 concentration levels measured hourly between May 23rd 2011 and December 17th 2023, results in 98,510 normal values and 597 anomaly values.

3. METHODOLOGY

In this study, we prioritize point-wise detection, pinpointing isolated time points with anomalous NO2 levels. This choice aligns with our objective of swiftly addressing specific air quality concerns for effective environmental monitoring and public health management in the capital.

Z-score Method

Detecting anomalies in one-dimensional data can be achieved by computing Z-scores for each point. The calculation is straightforward and is done using Equation (3).

$$z_i = \frac{x_i - \mu}{\sigma}, \quad i = 1, 2, \dots, n \tag{3}$$

where μ and σ are the mean and SD of the series, respectively. A Z-score of 0 signifies that an observation aligns with the time-series mean, while positive and negative Z-scores indicate observations above and below the mean, respectively. To identify anomalies, one can set a predefined threshold for Z-score values. The widely adopted threshold for outlier detection is ± 3. When the Z-score value of an observation surpasses ± 3, it is considered an anomaly.

Robust Z-Score Method

The conventional Z-score method is sensitive to outliers, distorting the mean and compromising its accuracy in representing the data's central tendency. To mitigate this, the modified Z-score method given in Equation (4) offers a robust alternative for AD, replacing the mean with the more resilient median and SD with the Median Absolute Deviation (MAD) [29]. This modification ensures less sensitivity to extreme values, making it a more resilient method for identifying anomalies in time-series data. Incorporating the median and MAD provides a more accurate reflection of central tendency, enhancing AD's reliability in the presence of anomalies.

$$zMAD_i = \frac{x_i - MED(x_i)}{\alpha MAD(x_i)}, \quad i = 1, 2, \dots, n \tag{4}$$

ere, $\alpha = 1.4285$ serves as the scale factor for MAD in non-normal distributions, ensuring consistency with the data's distribution. $MED(x_i)$ represents the median of the observations. Additionally, $MAD(x_i)$ is the MAD scale estimator mathematically obtained using Equation (5).

$$MAD(x_i) = MED(|x_i - MED(x_i)|), i = 1, 2, \dots, n \tag{5}$$

Similarly, the threshold is a user-defined cut-off, usually 3.0. An observation is considered an anomaly if its computed score value exceeds this threshold.

Interquartile Range (IQR) Method

Anomalies identified by this method are the observations that fall outside the range defined by the IQR. In this process, we first calculate the first quartile (Q1) and the third quartile (Q3) and then find IQR (IQR = Q3 - Q1), afterwards defining the lower bound as (Q1 - 1.5 × IQR) and the upper bound as (Q3 + 1.5 × IQR), so we identify outliers by considering any observation below the lower bound or above the upper bound as an anomaly. The 1.5 multiplier in bounds corresponds to approximately 3 SD from the median, covering 99.7% of data in a distribution.

Winsorization Method (Percentile Capping)

This method involves the adjustment of data by constraining extreme values to mitigate the impact of potentially misleading outliers [30]. The resulting effect is analogous to clipping in signal processing. This method is very much similar to the IQR method. In this approach, data values that fall below the 1st percentile (or any k-th percentile) or exceed the 99th percentile (or any $(100 - k)$-th upper percentile) of the given values are considered anomalies.

Hampel Filter

The Hampel identifier is recognized as a robust outlier detector in time series data [31]. Utilizing MAD, it incorporates a rolling window for AD. Two key parameters are essential: (1) A symmetric moving window size for anomaly assessment, and (2) threshold, a critical parameter requiring careful selection to avoid inadvertent AD for valuable data. The typical threshold value is 3.

For a given sequence of length of n and a window half-width l and a threshold value t, l samples from a single channel form the population for computing the median and SD. The chosen threshold requires the center of the sample of the rolling window to be at least t SD units away from the window median for anomaly flagging. The filtering involves computing *(1)* the rolling median (Equation (6)), *(2)* rolling MAD (Equation (7)) and *(3)* estimating the rolling SD as rolling MAD multiplied by the scaling constant (Equation (8)), for each timestamp in the data series ($i = 1, 2, \dots, n$).

$$MED_i = median(x_{i-l}, x_{i-l+1}, x_{i-l+2}, \dots, x_i, \dots, x_{i+l-2}, x_{i+l-1}, x_{i+l}) \tag{6}$$

$$MAD_i = median(|x_{i-l} - MED_i|, \dots, |x_{i+l} - MED_i|) \tag{7}$$

$$S_i = k\, MAD_i \tag{8}$$

where $k = 1.4826$. After all these steps done, we can decide whether a point is anomalous based on the following equation: $|x_i - MED_i > tS_i|$. In words, whenever the difference between a data point and its

rolling median is greater than the predefined multiple of the rolling SD, this observation is said to be anomalous.

One-class Support Vector Machines (OCSVM)

OCSVM is tailored for unsupervised learning, specifically targeting anomaly and novelty detection [32]. The primary objective is to find a boundary around the densest part of the data, essentially encapsulating the normal data points. Then, any observations that lie outside this boundary are considered **anomalies**. Given n unlabelled training data $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n\}$, kernel function $\phi(\boldsymbol{x})$ maps vector $\boldsymbol{x}$ from input space into a high dimensional feature space F. Then, we can construct a hyperplane in feature space F as given in Equation (9) to separate as many as possible of the mapped vectors from the origin in feature space F.

$$f(\boldsymbol{x}) = \boldsymbol{\omega}^T \phi(\boldsymbol{x}) - \rho \tag{9}$$

$\boldsymbol{w}$ and ρ in Equation (9) is obtained by solving an optimization problem given in Equation (10).

$$\begin{aligned} &\min_{\boldsymbol{\omega},\xi_i,\rho} \frac{\|\boldsymbol{\omega}\|_2^2}{2} - \rho + \frac{1}{\nu n}\sum_{i=1}^{n}\xi_i \\ &s.t. \quad \boldsymbol{\omega}^T\phi(\boldsymbol{x}_i) \geq \rho - \xi_i, \quad 1 \leq i \leq n \\ &\qquad \xi_i \geq 0, \quad 1 \leq i \leq n \end{aligned} \tag{10}$$

Here, ξ_i's are slack variables allowing some points to be located on the wrong side of the hyperplane and $\nu \in (0, 1]$ indicates the percentage of anomalies. Some kernels have a coefficient, γ, for Polynomial, Gaussian, and Sigmoid kernels. Anomaly score is the signed distance to the separating hyperplane, with positive values for inliers and negative values for outliers.

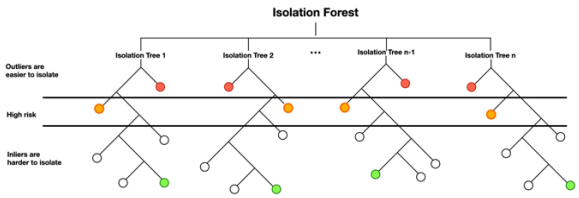

Figure 2. A basic diagram of isolated data points.

Isolation Forest

Isolation Forests (iForest) share similarities with Random Forest as they constitute a machine-learning approach designed for unsupervised AD based on decision trees [33]. AD typically involves profiling normal data, identifying anomalies as deviations from this profile. However, instead of modeling normal instances, iForest explicitly isolates potential anomalies in the dataset. The algorithm forms an ensemble (forest) of isolation trees (iTrees), each generating partitions on the dataset by randomly selecting a feature and then a split value between the maximum and minimum values of that feature. Shorter paths in the trees are expected for anomalies, as distinguishable attribute-values lead to early separation, as can be seen in Figure 2. To flag an anomaly, the iForest computes the average path length (the path length being the number of edges traversed from the root node) across all trees for a data point. A shorter average path length signifies an anomaly.

For training an iForest, there are three main parameters to be tuned: (1) number of trees to grow in the forest, (2) the number of observations to draw from training set to train iTree, and (3) number of features to draw from training set to train each iTree.

Deep Autoencoders

Autoencoders (AEs) are neural networks used for unsupervised learning, aiming to efficiently reconstructs the original input while compressing data in the process [34]. Comprising an encoder and a decoder which is represented in Figure 3, the encoder maps input to a condensed latent representation (the "bottleneck"), and the decoder reconstructs the input. While encoder and decoder architectures can be complex, the fundamental structure involves multi-layer feedforward neural networks. The bottleneck dimension, intentionally smaller than the input, captures meta-variables. Equation (11) presents the encoder transformation, which maps the input vector X to a hidden representation H, and Equation (12) reconstructs the initial input space from hidden representation H.

$$H = \sigma(\boldsymbol{W}_{xh}\boldsymbol{X} + \boldsymbol{b}_{xh}) \tag{11}$$

$$H = \sigma(\boldsymbol{W}_{h\hat{x}}\boldsymbol{H} + \boldsymbol{b}_{h\hat{x}}) \tag{12}$$

Here, σ is an activation function which introduces non-linearity to the network, $\boldsymbol{W}$ is a weight matrix and $\boldsymbol{b}$ is a bias vector.

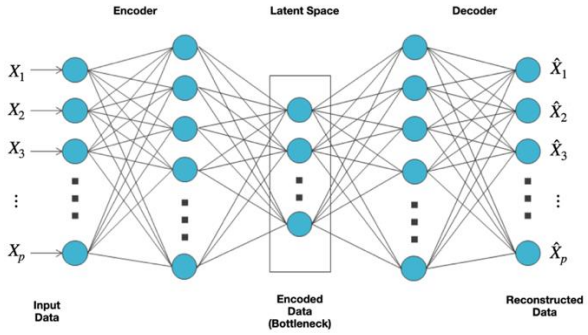

Figure 3. Illustrative description of a shallow AE

AEs serve various purposes, including dimensionality reduction, feature learning, and data denoising. During training, the model minimizes the reconstruction error (RE) between the reconstructed vector $\hat{\boldsymbol{X}}$ and the original input $\boldsymbol{X}$. This process encourages the autoencoder to learn a concise representation, while capturing essential input information. AEs can be employed for AD by utilizing the learned representation of normal data. This involves identifying anomalies through RE, which, in this case, it is mean squared error. When this error surpasses a predefined threshold, AEs are valuable for uncovering irregularities or outliers in datasets [35]. However, this might be challenging due to the lack of human supervision [37].

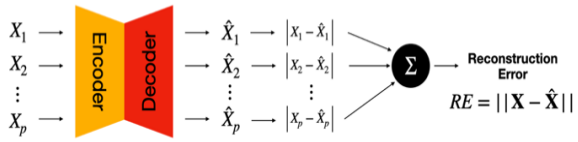

Figure 4. The process of computing RE in an Autoencoder-based AD method

Variational Autoencoders

A limitation of classical AEs lies in the potential for overfitting and learning sparse representations, compromising the model's ability to generalize to new inputs. Variational Autoencoders (VAEs) address this by introducing probabilistic latent representations, enhancing regularization and representation capabilities by combining variational bayes inference and DL [37]. Unlike AEs, VAEs generate a distribution over the latent space, typically a Gaussian distribution, capturing the mean (μ) and variance (σ) with an encoder. The decoder reconstructs the input using vectors sampled from this distribution, in other words, the input of the decoder is stochastic. VAEs' probabilistic nature makes them more robust in handling anomalies, as the learned distribution captures the variability of the data, highlighting anomalies as deviations. VAEs also incorporate the Kullback-Leibler (KL) divergence in the loss function, optimizing a combination of the reconstruction error and KL divergence to guide the latent space towards the desired distribution. This improvement enhances interpretability and adaptability to varying input patterns in AD applications. A general layout of VAEs and its loss function can be seen in Figure 5. Similar to AE, once we train the VAE and define a threshold, we can evaluate anomalies.

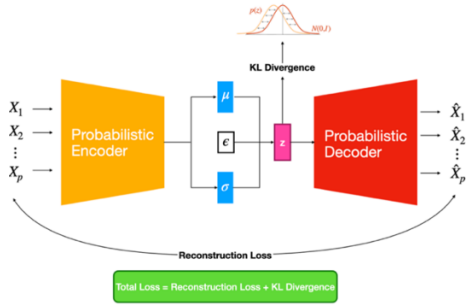

Figure 5. General layout of VAEs and its loss function

Algorithmic Evaluation

In the context of AD, we transformed the problem into a supervised two-class classification by creating a target variable, considering "anomaly" as the positive class. The confusion matrix given in Table 2 serves as a valuable tool for evaluating classification model performance [38].

Table 2. Confusion matrix

	Predicted Class	
Actual Class	*Anomaly*	*Not Anomaly*
Anomaly	TP	FN
Not Anomaly	FP	TN

Key metrics derived from the confusion matrix which is given in Table 2 include True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), used to calculate balanced accuracy, precision, recall, specificity, and F1 score, which are also given in Table 3.

.

Table 3. Evaluation metrics

Metric's Name	**Formula**	**Explanation**
Precision	$\frac{TP}{TP + FP}$	Ratio of correctly classified anomalous observations to the total number predicted as anomalous
Recall (Sensitivity/True Positive Rate)	$\frac{TP}{TP + FN}$	Ratio of correctly classified anomalous observations to the total number of actual anomalous observations
Specificity (True Negative Rate)	$\frac{TN}{TN + FP}$	Ratio of correctly classified non-anomalous observations to the total number of actual non-anomalous observations
Balanced Accuracy	$\frac{Sensitivity + Specificity}{2}$	**Average of sensitivity and specificity**
F1 Score	$\frac{2 \times Precision \times Recall}{Precision + Recall}$	Harmonic mean of precision and recall
False Alarm Rate	$\frac{FP}{FP + TN}$	Ratio of incorrectly classified non-anomalous observations to the total number of non-anomalous observations

While all metrics are reported, the study places particular emphasis on selecting the model with the highest recall considering the significant consequences of overlooking anomalous cases (false negatives) as well as the lowest

false alarms (false alarm rate) to prevent alert fatigue and unnecessary interventions, emphasizing the importance of minimizing these occurrences to maintain the effectiveness of the overall monitoring system.

4. RESULTS AND DISCUSSION

We firstly investigate the mean disparities within distinct groups. To compare the means between these groups, unbalanced One-Way ANOVA test was performed by pingouin library of Python programming language. In Table 4, summary statistics of the aggregated data can be seen across different years and seasons.

As per Table 4, The annual average concentration of NO2 in this station decreasing year by year from 2013 to 2018, with NO2 level decreasing from 63.26 to 38.03 $\mu g/m^3$, and from 2018 to 2023 the annual average NO2 concentration shows an increasing trend from 38.03 to 66.39 $\mu g/m^3$. Even in 2020, a year marked by the global challenges of the COVID-19 pandemic, there appears to be no notable reduction in NO2 levels. The results indicate that at 95% confidence level there is statistically significant difference between the mean concentrations of NO2 measured in different years ($F = 377.30, p - value < 0.05$). Additionally, the average NO2 concentrations tend to be higher during the fall (69.94 $\mu g/m^3$) and winter (61.12 $\mu g/m^3$) months, while the lowest concentrations of NO2 were observed during the summer and spring months. This is because NO2 has a notably shorter chemical life expectancy in summer compared to other seasons. [39]. Cold months see elevated concentrations due to weather conditions—cold temperatures, cloudiness, and reduced solar radiation—that impede photochemical reactions, causing NO2 to linger longer than in summer and spring. [40] [41]. Additionally, out of all other nitrogen oxides emitters, increased usage of internal combustion engines (largely diesel) and higher residential coal burning may contribute to the increasing trend [42]. The results suggest that, at a 95% confidence level, there exists a statistically significant difference between the mean concentrations of NO2 measured in different seasons ($F = 2{,}805.46, p - value < 0.05$).

Table 4. Comparison of years and seasons on NO_2 concentration levels

	NO2 ($\mu g/m^3$)				
Years	**Mean**	**Standard Deviation**	**Median**	**Minimum**	**Maximum**
2011	48.30	51.85	32.47	0.01	761.9
2012	35.63	26.45	28.05	0.01	191.17
2013	63.26	48.30	44.71	0.1	355.45
2014	59.60	40.32	48.41	0.01	320.05
2015	56.76	37.18	46.98	1.57	368.3
2016	54.77	33.75	47.65	4.47	263.56
2017	51.48	38.99	43.1	1.71	232.81
2018	38.03	25.64	31.6	2.32	200.76
2019	48.74	33.57	41.885	0.01	249.36
2020	50.46	33.18	43.35	3.07	183.04
2021	58.29	50.35	42.83	4.93	480.57
2022	63.34	55.36	44.48	6.58	412.62
2023	66.39	54.38	49.06	6.25	488.3
Autumn	69.94	48.71	60.38	0.02	488.3
Spring	43.22	34.05	33.11	0.01	355.45
Summer	42.00	38.05	29.58	0.01	761.9
Winter	61.12	41.77	49.79	0.74	428.96
Overall	**54.04**	**42.76**	**41.8**	**0.01**	**761.9**

Table 5. Comparison of weekdays and weekends on NO2 concentration levels

	NO2 ($\mu g/m^3$)				
	Mean	**Standard Deviation**	**Median**	**Minimum**	**Maximum**
Weekdays	55.00	42.96	42.90	0.01	761.9
Weekends	51.49	42.20	38.77	0.01	667.97

Table 6. Comparison of days of the week on NO2 concentration levels

	NO2 ($\mu g/m^3$)				
Day of the Week	**Mean**	**Standard Deviation**	**Median**	**Minimum**	**Maximum**
Monday	54.65	42.51	42.44	0.01	450.34
Tuesday	54.48	41.55	42.56	0.01	428.96
Wednesday	54.3	40.92	43.17	0.01	439.36
Thursday	55.78	45.74	42.60	0.06	761.9
Friday	55.76	43.89	43.90	0.01	488.3
Saturday	53.23	42.84	41.07	0.02	667.97
Sunday	49.73	41.47	36.53	0.01	369.2

Table 7. Effects of holidays on NO2 concentration levels

	NO2 ($\mu g/m^3$)				
	Mean	**Standard Deviation**	**Median**	**Minimum**	**Maximum**
Holiday	54.66	42.96	42.49	0.02	761.90
Not Holidays	36.84	33.30	25.67	0.02	300.91

In the Bahçelievler district of Ankara, Turkey, a comparison of NO2 levels between weekdays and weekends reveals interesting patterns, given in Table 5. On average, NO2 concentrations are slightly lower during weekends (51.49 $\mu g/m^3$) compared to weekdays (55 $\mu g/m^3$). The SD values indicate similar variability in NO2 levels for both periods. Median concentrations further support the observation of lower NO2 levels on weekends, with values of 38.77 $\mu g/m^3$ compared to 42.9 $\mu g/m^3$ on weekdays. These findings suggest a potential reduction in NO2 pollution during weekends, possibly influenced by decreased vehicular and industrial activities. The results indicate that at 95% confidence level there is statistically significant difference between the mean concentrations of NO2 in weekdays and weekends ($F = 135.97, p - value < 0.05$).

Based on the concentration levels for NO2 for each day of the week, we also provide summary statistics for the days of the week in Table 6.It appears that Sunday has the lowest median NO2 concentration (36.53 μg/m³), followed by Saturday (41.07 μg/m³). The weekdays show slightly higher median concentrations, with Wednesday having the highest median NO2 level (43.17 μg/m³) among the weekdays. However, observed variations seems to be very similar throughout the week. The findings suggest that, at a 95% confidence level, there exists a statistically significant difference between the mean concentrations of NO2 measured on different days of the week. ($F = 33.25, \ p - value < 0.05$). In addition to the comparison between weekdays and weekends, the analysis was extended to include holidays.

In Table 7, it appears that NO2 levels experience a significant decrease during holidays. The mean, median, and maximum concentrations are notably lower during holidays, indicating a potential reduction in pollution. The SD values suggest a lower degree of variability in NO2 levels on holidays compared to non-holidays. The statistics underscore the influence of holidays in NO2, possibly attributed to decreased anthropogenic activities during these periods. The results indicate that at 95% confidence level there is statistically significant difference between the mean concentrations of NO2 measured in holidays and not-holidays ($F = 619.69, p - value < 0.05$).

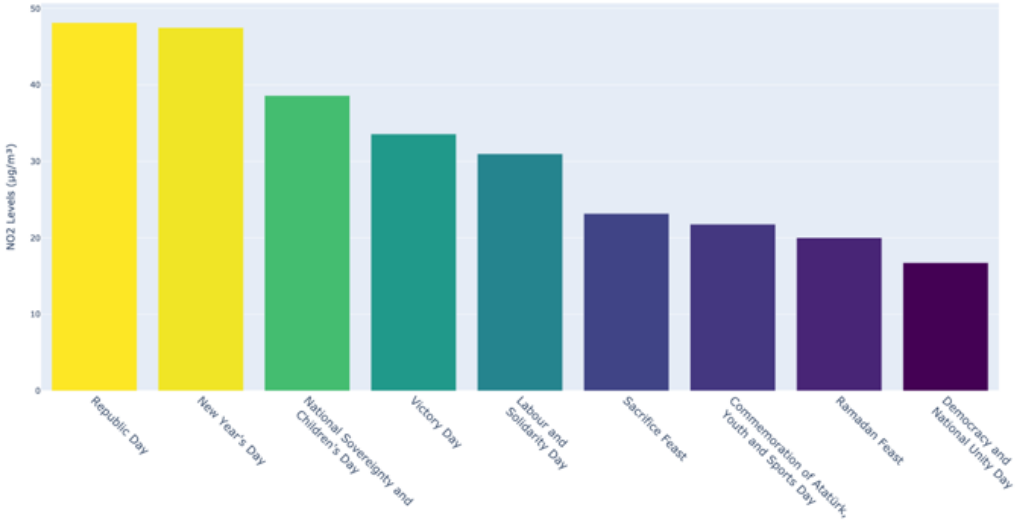

Figure 6. Median concentration of NO2 levels on different holidays

Republic of Türkiye has total 9 holidays, which 2 of them (Ramadan Feast and Sacrifice Feast) are Islamic and the rest are general/national holidays, i.e., 1 January New Year's Day, 23 April National Sovereignty and Children's Day, 1 May Labour and Solidarity Day, 19 May Commemoration of Atatürk, Youth and Sports Day, 15 July Democracy and National Unity Day, 30 August Victory Day, and 29 October Republic Day. Out of these holidays, only for "15 July Democracy and National Unity Day" the data is collected starting from July 15th 2016, since it is commemoration of the national unity against the coup d'état attempt for democracy in 2016. Median concentration of NO2 levels aggregated across all years can be seen in Figure 6.

Figure 8 clearly shows that the lowest median NO2 concentration occurred on "Democracy and National Unity Day", reflecting the limited data available for this relatively new holiday. Conversely, the Bahçelievler district experiences higher NO2 levels on "Republic Day", "National Sovereignty and Children's Day", and "Victory Day", attributed to increased visitor influx due to its close proximity to Anıtkabir (the mausoleum of Mustafa Kemal Atatürk). The elevated NO2 concentrations (48.13 $\mu g/m^3$, 38.58 $\mu g/m^3$, 33.57 $\mu g/m^3$ respectively) are likely due to higher vehicular traffic during these events, as vehicular emissions are a common source of NO2. The district's vibrant nature contributes to elevated NO2 levels on "New Year's Day" and "Labour and Solidarity Day" due to increased societal activities and events attracting larger crowds. The lower NO2 median concentration (i.e., 21.77 $\mu g/m^3$) on "Commemoration of Atatürk, Youth and Sports Day" is explained by the holiday's primary location around the "Grand National Assembly of Türkiye", situated in a different district. Islamic holidays, such as Sacrifice Feast (23.17 $\mu g/m^3$) and Ramadan Feast (20.00 $\mu g/m^3$) display decreasing NO2 concentrations as residents, particularly students, go back to their hometowns out of Ankara for celebrations. This phenomenon aligns with a common pattern seen in urban areas during holiday seasons when residents travel to be with their families and friends. Statistical analysis at a 95% confidence level reveals a significant difference in mean NO2 concentrations across various official holidays ($F = 38.14, p - value < 0.05$).

After delineating the variations in median NO2 levels across specific individual variables, the subsequent focus will pivot towards the domain of AD. Initially, we adopt statistical methodologies renowned for their simplicity, focusing exclusively on a singular column denoting the concentration levels of NO2. Then, we train ML and DL methods in multivariate approach.

The Scikit-Learn library (version 1.2.0) was employed for the implementation of iForest and OCSVM algorithms. Regarding iForest algorithm settings, a coarse grid-search was conducted on the hyperparameters: contamination as $\{'auto', 0.001, 0.01, 0.03, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5\}$, bootstrap as $\{True, False\}$, maximum samples to train each iTree as $\{'auto', 0.05, 0.1, 0.3, 0.5, 0.6, 0.7\}$, and the number of trees as $\{50, 100, 150, 200, 250, 300, 500\}$. It is noteworthy that all features were utilized in the analysis. Number of trees has been found to be 250 for the forest with contamination set to '$auto$'. No bootstrapping has been applied. The number of samples to draw to train each iTree is found to be 0.5. For OCSVM, a similar strategy was applied. With setting kernel to radial basis function, gamma to 0.052, tolerance to 1e-3 and nu to 0.6, the best test score has been obtained without overfitting.

Table 8. Confusion matrices and calculated evaluation metrics for all the methods

Metrics	**Z-score**	**Robust Z-Score**	**IQR**	**Winsorization**	**Hampel Filter**	**iForest**	**OCSVM**	**AE**	**VAE**
TP	471	487	505	530	224	543	550	576	590
TN	94,734	97,584	95,550	97,123	95,799	98,106	98,051	98,228	98,380
FP	3770	920	2954	1381	2705	398	453	276	**124**
FN	126	110	92	67	373	54	47	21	7
Balanced Accuracy	0.87	0.90	0.90	0.93	0.67	0.95	0.95	0.98	0.99
Precision	0.11	0.34	0.14	0.27	0.071	0.57	0.54	0.67	0.82
Recall (Sensitivity)	0.78	0.81	0.84	0.88	0.37	0.90	0.92	0.96	**0.98**
F1-Score	0.19	0.48	0.24	0.42	0.12	0.70	0.68	0.79	0.90
Specificity	0.96	0.99	0.97	0.98	0.97	0.99	0.99	0.99	0.99
False Alarm Rate	0.0382	0.0093	0.0299	0.0140	0.0274	0.0040	0.0045	0.0028	**0.0012**

PyTorch is used to implement AE and VAE models, leveraging the Optuna library for hyperparameter tuning [43]. Early stopping was also utilized. In each experiment, we employed a batch size of 32 and used Adam as an optimizer with default parameter values, and the learning rate is taken as 1e-4. For AE, the input layer consists of 18 neurons, the encoder component consists of 2 layers each containing respectively 32 and 16 with the dimension of latent space set to 5 neurons. Between the layers a batch normalization was deployed in order to always obtain the same scale among the layers. Decoder has 2 layers and each layer has 16 and 32 neurons, respectively. The last layer has the same number of neurons with the input layer, exactly 18 neurons. The activation function for all the layers is Scaled Exponential Linear Units (SELU). The training has been completed in 150 epochs. For VAE, the best validation score has been obtained for the architecture with 3 layers of encoders, each with 64, 32, 16 neurons, respectively, and 3 layers of decoder, each with 16, 32, 64 neurons respectively, and the latent variable consisting of 3 neurons, where each z represents different latent variable and each of them have a different mean μ and SD σ. Between the layers a batch normalization was deployed as well. The activation function for all the neuron layers is Rectified Linear Units (ReLU). The training has been done in 200 epochs. The elements of the confusion matrices and evaluation metrics obtained from the resulting models have been reported in Table 8.

iForest, OCSVM, AE, and VAE emerge as methods that not only effectively detect anomalies but also minimize the occurrence of triggered false alarms. While iForest and OCSVM are known for their ability to handle complex data distributions, reconstruction-based approaches show the best overall performance with high recall and balanced accuracy, as well as low false alarm rate. In the end, though, VAE stands out for the superior performance, making it a promising choice for AD in this context.

Anomalies in NO2 concentration of the Bahçelievler district of Ankara for the year 2023 based on the anomaly scores of VAE model are displayed in Figure 7. It is evident that the model has effectively captured the data patterns based on the created features. For the year 2023, there seems to be 231 anomalies really existing and the algorithm can correctly pinpoint all the out-of-ordinary readings with only one error, which is the NO2 level for the day of January 26th (The other 6 false negatives belong to previous years). The NO2 concentration level for this particular day was 208.11 $\mu g/m^3$, which is not an anomaly according to the law but algorithm says it is. After the best model is obtained, it can be serialized to be put behind a mobile or web application in order to detect anomalies automatically so to give alerts when needed.

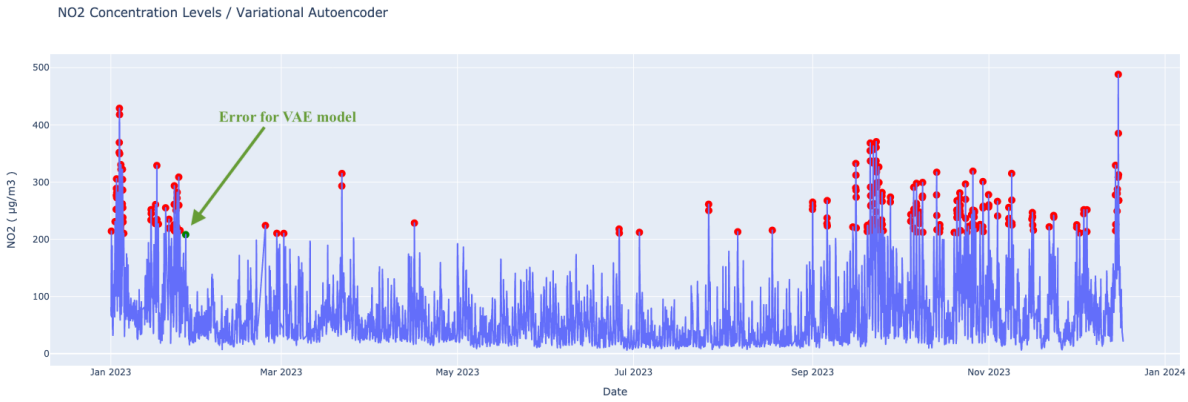

Figure 7. Anomaly points for NO2 concentration levels of Bahçelievler district of Ankara, for the year 2023, detected by VAE model

5. CONCLUSION

As a summary we obtained a fully-working algorithm to pinpoint any potential anomalies and/or malfunctioning sensor automatically for ground-level NO2 concentration level in Bahçelievler district of Ankara, the capital of Republic of Türkiye using one of the deep learning approaches, namely Variational Autoencoder, which gives the best true positive rate and false alarm rate. Developed model can be serialized and be used easily behind any application for air quality surveillance by creating a prediction function with an API endpoint.

To the best of our knowledge, this study represents the first research utilizing state-of-the-art (SOTA) method for such a system in environmental science literature. Besides, instead of using only raw univariate time series data, we also show that if one follows the right feature engineering steps in order to incorporate additional features, one can create such real-time AD system. This paper utilizes engineered rolling-window features to capture the short-term temporal patterns, variability, volatility, and statistical characteristics of data and periodic variables to identify potential periodicity in NO2 levels as well as additional categorical attributes to explore the factors like reduced vehicular traffic, office closures, altered travel patterns and changes in residential combustion. Even though such systems are unsupervised in nature, a label variable using Turkish Environmental Law was created to be a teacher for the algorithm itself.

It is essential to point out once again that most of the environmental science literature about air quality revolves around the prediction and/or forecasting of any air pollutant. However, this study aims for AD, especially for NO2 which seems to be devoid of studies on it.

The main advantage of VAEs is that they can learn to encode data in an unsupervised manner, which means they can detect anomalies without being explicitly trained on what constitutes an anomaly. They are generally robust to noise in the data, which is an important feature for monitoring environmental pollutants like NO2 where readings can be affected by various factors such as sensor malfunctions or external environmental conditions. They can also be adapted to various data scales, making them suitable for different monitoring setups, from small-scale local environments to large-scale urban or regional networks. However, even though ML and DL approaches are easy to implement with ready-to-use programming frameworks, they require a delicate hyperparameter tuning.

Additionally, the scope of the current work is limited to one pollutant in one district of one city. Therefore, for future studies, the replication of the research for different pollutants in different cities and their individual districts could provide additional insights. Additionally, new features can be added or created, for example, meteorological predictors.

DECLARATION OF ETHICAL STANDARDS

The authors of this article declare that the materials and methods they use in their studies do not require ethics committee permission and / or legal-specific permission.

AUTHORS' CONTRIBUTIONS

Mustafa Murat ARAT: Collected data from relevant institutions, performed the experiments, analyze the results, wrote the manuscript.

CONFLICT OF INTEREST

There is no conflict of interest in this study.

REFERENCES

[1] Brauer M., Hoek G., Smit H. A., d. Jongste J. C., Gerritsen J., Postma D. S., Kerkhof M. and Brunekreef B., "Air pollution and development of asthma, allergy and infections in a birth cohort", ***Eur. Respir. J.***, 29: 879-888, (2007).

[2] Liu Y., Pan J., Zhang H., Shi C., Li G., Peng Z., Ma J., Zhou Y. and Zhang L., "Short-Term Exposure to

Ambient Air Pollution and Asthma Mortality", ***Am. J. Respir. Crit. Care Med.***, 200(1): 24-32, (2019).

[3] Rajagopalan S., Al-Kindi S. G. and Brook R. D., "Air Pollution and Cardiovascular Disease: JACC State-of-the-Art Review", ***J. Am. Coll. Cardiol.***, 72(17): 2054-2070, (2018).

[4] Tuśnio N., Fichna J., Nowakowski P. and Tofiło P., "Air Pollution Associates with Cancer Incidences in Poland", ***Appl. Sci.***, 10(21), (2020).

[5] Balogun H., Rantala A., Antikainen H., Siddika N., Amegah A., Ryti N., Kukkonen J., Sofiev M., Jaakkola M. and Jaakkola J., "Effects of Air Pollution on the Risk of Low Birth Weight in a Cold Climate", ***Appl. Sci.***, 10(18), (2020).

[6] **http://tinyurl.com/EUTZPforAWS**, "EU Action Plan: 'Towards Zero Pollution for Air, Water and Soil'",(2021).

[7] **http://tinyurl.com/WHOAQG**, "WHO global air quality guidelines", (2021).

[8] http://tinyurl.com/APDataDownload

[9] https://www.havaizleme.gov.tr

[10] Anenberg S. C., Mohegh A., Goldberg D. L., Kerr G. H., Brauer M., Burkart K., Hystad P., Larkin A., Wozniak S. and Lamsal L., "Long-term trends in urban NO2 concentrations and associated paediatric asthma incidence: estimates from global datasets", ***Lancet Planet Health***, 6(1): 49–58, (2022).

[11] Braei M. and Wagner S., "Anomaly Detection in Univariate Time-series: A Survey on the State-of-the-Art", **arXiv preprint**, (2020).

[12] Shams S. R., Jahani A., Kalantary S., Moeinaddini M. and Khorasani N., "Artificial intelligence accuracy assessment in NO2 concentration forecasting of metropolises air", ***Sci. Rep.***, 11(1): 1805, (2021).

[13] Yammahia A. A. and Aunga Z., "Forecasting the concentration of NO2 using statistical and machine learning methods: A case study in the UAE", ***Heliyon***, 9(2), (2023).

[14] Gastaldo P., Liu B., Zhang L., Wang Q. and Chen J., "A Novel Method for Regional NO2 Concentration Prediction Using Discrete Wavelet Transform and an LSTM Network", ***Comput. Intell. Neurosci.***, 2021, (2021).

[15] Jesemann A.-S., Matthias V., Böhner J. and Bechtel B., "Using Neural Network NO2-Predictions to Understand Air Quality Changes in Urban Areas—A Case Study in Hamburg", ***Atmosphere***, 13(11): 1929, (2022).

[16] Cabaneros S. M. S., Calautit J. K. S. and Hughes B. R., "Hybrid Artificial Neural Network Models for Effective Prediction and Mitigation of Urban Roadside NO2 Pollution", ***Energy Procedia***, 142: 3524-3530, (2017).

[17] Aggarwal A. and Toshniwal D., "Detection of anomalous nitrogen dioxide (NO2) concentration in urban air of India using proximity and clustering methods", ***J&AWMA***, 69(7): 805-822, (2019).

[18] van Zoest V., Stein A., and Hoek G., "Outlier Detection in Urban Air Quality Sensor Networks", ***Water Air Soil Pollut.***, 229(111), (2018).

[19] Torres J. M., Nieto P. G., Alejano L. and Reyes A., "Detection of outliers in gas emissions from urban areas using functional data analysis", ***J. Hazard. Mater.***, 186(1): 144-149, (2011).

[20] Can A. and Özsoy H., "A Different Perspective on Air Pollution Measurements", ***Journal of Polytechnic***, *26(1):* 329 - 344, (2023).

[21] Radojević D., Antanasijević D., Perić-Grujić A., Ristić M., Pocajt V., "The significance of periodic parameters for ANN modeling of daily SO2 and NOx concentrations: A case study of Belgrade, Serbia", ***Atmos. Pollut. Res.***, 10(2): 621-628, (2019).

[22] **http://tinyurl.com/TSIVehicle2023**, "Turkish Statistical Institute - Turkish Statistics for Road Motor Vehicles" (2023).

[23] **http://tinyurl.com/MoEUCCAPCAAP,** "Republic of Türkiye - Ministry of Environment, Urbanisation and Climate Change - Ankara Province Clean Air Action Plan 2020-2024", (2020).

[24] Rousseeuw P. J. and Croux C., "Alternatives to the Median Absolute Deviation," ***JASA***, 88(424): 1273-1283, (1993).

[25] Géron A., "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow", 2nd Edition, **O'Reilly Media, Inc.**, (2019).

[26] **http://tinyurl.com/ISRTELL**, "Information System of Regulations - Turkish Environmental Law", (1983).

[27] **http://tinyurl.com/ISRByLaw,** "Information System of Regulations By-law on Air Quality Assessment and Management", (2008).

[28] **http://tinyurl.com/ISRCircular,** "Information System of Regulations - Circular on Air Quality Assessment and Management", (2013).

[29] Iglewicz B. and Hoaglin D. C., "How to Detect and Handle Outliers", **ASQC Quality Press**, Milwaukee, Wisconsin, (2004).

[30] Tukey J.W. and McLaughlin D.H., "Less Vulnerable Confidence and Significance Procedures for Location Based on a Single Sample: Trimming/Winsorization 1", ***Sankhya: Indian J. Stat.***, 25(3): 331-352, (1963).

[31] Pearson R., "Outliers in process modeling and identification", ***IEEE Trans. Control Syst. Technol.***, 10(1): 55-63, (2002).

[32] Schölkopf B., Platt J. C., Shawe-Taylor J., Smola A.J. and Williamson R.C., "Estimating the support of a high-dimensional distribution", ***Neural Computation***, 13(7): 1443–1471, (2001).

[33] Liu F.T., Ting K.M. and Zhou Z.-H., "Isolation-Based Anomaly Detection", ***ACM Transactions on Knowledge Discovery from Data***, 6(1): 1–39, (2012).

[34] Bengio Y., Lamblin P., Popovici D. and Larochelle H., "Greedy Layer-Wise Training of Deep Networks", ***Advances in Neural Information Processing Systems 19 (NIPS 2006)***, Vancouver, B.C., Canada, (2006).

[35] Zong B., Song Q., Min M.R., Cheng W., Lumezanu C., Cho D.-K. and Chen H., "Deep Autoencoding Gaussian Mixture Model for Unsupervised Anomaly Detection", ***International Conference on Learning Representations***, Vancouver, B.C., Canada, (2018).

[36] Gong D., Liu L., Le V., Saha B., Mansour M.R., Venkatesh S. and Hengel A., "Memorizing Normality to Detect Anomaly: Memory-augmented Deep

Autoencoder for Unsupervised Anomaly Detection", ***arXiv preprint***, (2019).

[37] Kingma D.P. and Welling M., "Auto-Encoding Variational Bayes", ***International Conference on Learning Representations***, Scottsdale, Arizona, USA, (2013).

[38] Awadh K. and Akbaş A., "Intrusion Detection Model Based on TF.IDF and C4.5 Algorithms", ***Journal of Polytechnic***, 24(4): 1691-1698, (2021).

[39] Duncan B.N., Yoshida Y., de Foy B., Lamsal L.N., Streets D.G., Lu Z., Pickering K.E. and Krotkov N.A., "The observed response of Ozone Monitoring Instrument (OMI) NO2 columns to NOx emission controls on power plants in the United States: 2005-2011", ***Atmos. Environ.***, 81: 102-111, (2013).

[40] Matandirotya N. and Burger R., "An assessment of NO2 atmospheric air pollution over three cities in South Africa during 2020 COVID-19 pandemic", ***Air Qual Atmos Health***, 16: 263-276, (2023).

[41] U.S. Environmental Protection Agency, "Overview of Nitrogen Dioxide (NO2) Air Quality in the United States," 29 June 2023. [Online]. Available: http://tinyurl.com/epaNO2. [Accessed 28 December 2023].

[42] Dėdelė A. and Miškinytė A., "A. Seasonal variation of indoor and outdoor air quality of nitrogen dioxide in homes with gas and electric stoves", ***Environ. Sci. Pollut. Res.***, 23: 17784-17792, (2016).

[43] Kavalcı Yılmaz E. and Bakır H., "Hyperparameter Tunning and Feature Selection Methods for Malware Detection", ***Journal of Polytechnic*** **preprint**, (2024).