The Relationship Between Traffic Volume, Pollutant Emissions and Traffic Accidents: A Case Study of Shiraz City Entrances

Rozita SAKI and Ara TOOMANIAN and Najmeh Neysani SAMANI, Iran
Abbas RAJABIFARD, Australia

Key words: Remote sensing, Spatial planning, Climate Change, Road Accidents

SUMMARY

Air pollution is one of the most important crises that most countries are dealing with today Industry and technology are facing it. The country of Iran and especially the city of Shiraz is not exempt from this phenomenon. Effect Urban air pollution on the environment and human health is a growing concern for researchers and policy makers To reduce the negative effects of air pollution on health, its timely measurement High temporal and spatial resolution is very important. On the other hand, the air pollution measuring stations in the city, despite high accuracy in pollutant measurement, due to time and space limitations and point measurement capability.

They are not generalizable. The complementary and sometimes alternative solution is the use of remote sensing and satellite data paying attention to the optimal cost and wide coverage is considered a suitable method for air pollution monitoring. Furthermore, traffic emissions, coupled with accidents at key urban intersections, significantly contribute to these pollutants. This study uses remote sensing data to extract and analyze NO2 levels at several traffic-heavy points in Shiraz, particularly the city entrances, which are prone to high congestion and road accidents. Advanced programming techniques are applied to process satellite imagery and other geospatial data, enabling precise mapping of pollutant levels. Complex interconnections between traffic volume, pollutant emissions-specifically nitrogen oxides (NO2)-and the frequency of traffic accidents at the city entrances of Shiraz. Through advanced remote sensing techniques, the research aims to quantify nitrogen pollutants and assess their correlation with traffic congestion and accident-prone zones. Programming tools, including deep learning and programming, are applied to analyze traffic data and satellite imagery. The results suggest a direct relationship between increased traffic flow, elevated pollutant levels, and a higher frequency of accidents. This research underscores the importance of integrating real-time traffic and environmental data for smarter urban Spatial planning and traffic management solutions aimed at reducing both accidents and air pollution.

1. Introduction

The rapid urbanization and increasing vehicular traffic in metropolitan areas have led to critical environmental and safety challenges, including elevated air pollution and a surge in road accidents at urban entrances. Among various pollutants, nitrate pollution has emerged as a significant environmental concern due to its adverse effects on human health and ecosystems (Zhang et al., 2022). Addressing these challenges requires integrating advanced analytical techniques with spatiotemporal data to provide actionable insights.

In this study, we leverage the synergy between geostatistical methods and machine learning algorithms to analyze the complex interactions among high traffic volumes, nitrate pollution, and road accidents at urban entrances. Specifically, we employ the Inverse Distance Weighting (IDW) method for spatial interpolation of traffic data, enabling a detailed spatial representation across the study area. Subsequently, the Random Forest algorithm, a robust machine learning approach, is utilized to model and predict nitrate pollution levels, incorporating spatiotemporal data from traffic sensors, accident records, and satellite imagery.

Remote sensing-based methods are one of the most powerful and accurate methods for intelligent monitoring of land changes and spatial mapping of air pollution monitoring. Google Earth Engine, as a powerful open-source processing system in remote sensing, was first launched by Google in December 2010 to enable monitoring, discovery, and measurement of global changes. This system contains a large database of location-based information and images from around the world that can be easily retrieved and analyzed. It also provides the ability to quickly view data with the ability to use zoom, change location at any point on the globe, and time-spatial series data to examine changes over time and depending on location.
In addition, the capabilities of this system include free access, high processing speed, support for time-spatial series processing, as well as spatial analysis and data transfer capabilities, which can be a reason for the use of this valuable system in most reputable global articles.
The interaction between air pollution and climate change is multifaceted. For example, increased global temperatures accelerate the formation of ground-level ozone (O_3), worsening air pollution (Jacob & Winner, 2009). Additionally, aerosols from air pollution can affect climate by either cooling the atmosphere through increased albedo or warming it by absorbing solar radiation, depending on their composition and atmospheric behavior (Andreae et al., 2005).
Mitigating these dual challenges requires integrated approaches. Transitioning to renewable energy, implementing stricter emissions regulations, and using advanced

technologies to monitor and reduce pollutants are essential measures. These actions not only improve air quality but also contribute to reducing greenhouse gas emissions, addressing both air pollution and climate change simultaneously (Smith et al., 2020). This integrated approach aligns with the United Nations Sustainable Development Goals (SDGs), promoting health, environmental protection, and sustainability.

2. Case study and Data

2.1 Study area

Shiraz is the capital of Fars Province, with a length of 40 km and a width of between 15 and 30 km, with an area of 1,268 square kilometers in a rectangular shape and geographically located in southwestern Iran and in the central part of Fars. Shiraz is surrounded by relatively high mountain ranges in the form of a solid fence. The city is limited to the west by Mount Drak, and to the north by the Bemo, Sabzeposhan, Chehelmoqam and Babakouhi mountains (of the Zagros Mountain range).The geographical coordinates of Shiraz are 29 degrees and 36 minutes north and 52 degrees and 32 minutes east, and its altitude above sea level varies between 1,480 and 1,670 meters in different parts of the city. The Shiraz Dry River is a seasonal river that, after passing through the city of Shiraz, bends to the southeast of its basin and flows into Maharlu Lake.
The climate of Shiraz city has the highest precipitation in winter, 26.2 mg, the lowest precipitation in summer, which was zero precipitation. Also, the highest temperature (in degrees Celsius) in summer is 28.5. The lowest temperature in winter, which was 9.29. And finally, the highest humidity in winter is 53.33. The lowest humidity in summer, which was 30.44. In general, Shiraz city has a hot and semi-arid climate. Shiraz city is one of the densely populated cities that suffers from air pollution. Due to the city's location between two mountain ranges at the southern end of the Zagros and the relatively high population growth, as well as the increasing number and variety of pollutants, it seems that the pollution problem will take an upward trend over time and become less preventable [1]. According to statistics from the Director General of Fars Environmental Protection, there are currently more than 500,000 vehicles in the city of Shiraz daily, and more than 75 percent of air pollution is caused by fuel combustion in motor vehicles. Also, worn-out vehicles over 20 years old, which emit a large amount of polluting gases into the air due to high gasoline consumption, cause 22 percent of air pollution in Shiraz by 25,000 worn-out vehicles. According to the statistics provided, 93 percent of city minibuses, 47 percent of city buses, and 37 percent of taxis in Shiraz are worn-out. A low-end car produces an average of 30 grams of carbon monoxide per

kilometer, polluting parameters that include carbon oxides, nitrogen oxides, sulfur oxides, soot, particulate matter, and hydrocarbons.

2.2 Data

--Spatiotemporal Data from Daily Traffic Count Sensors

For this study, data from 8 traffic sensors on the Shiraz-Kazeroun Road were analyzed during the period 2019 to 2022.The corresponding columns of the data are shown in Table 1. The data include traffic flow measurements rate (vehicle count) and average speed at each sensor location, per hour, per traffic lane and per vehicle category. This means that we have very granular data. To simplify the data, we aggregate all vehicle categories and traffic lanes. This results in data at each location, per hour. Here, we summarized the traffic flow rate (vehicle count) and calculated the average speeds. Finally, we removed all data from sensors that were not recorded correctly.

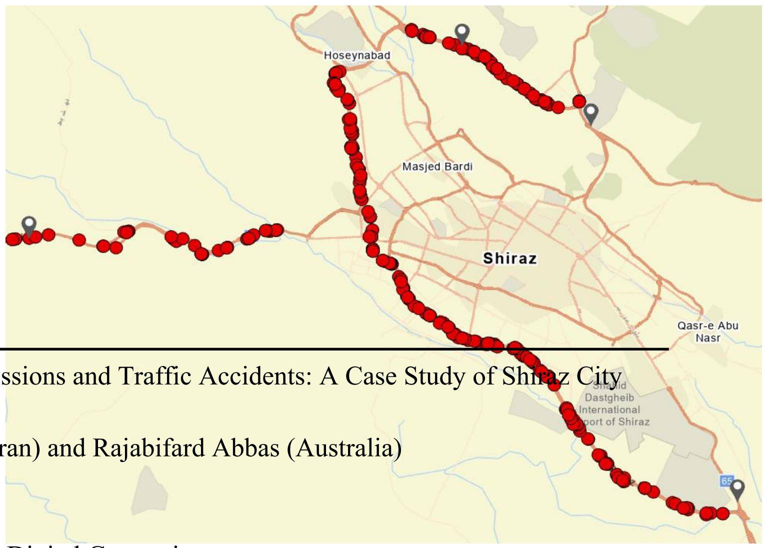
This resulted in a final dataset that included 4 GeoSensor, whose data were linked to the locations of these sensors. The are shown in Figure 1.

Fig. 1. GeoSensor location and City Entrances of Shiraz

Location of sensor X	Location of sensor y	Day of the week	Working time	Number of vehicles	Average hourly traffic	Number of heavy vehicles	Percentage of heavy vehicles	Passenger vans	Trucks, and minibuses	Two-axle truck s	Buses	Three-axle trucks	Medium speed	Illegal speed	Illegal distance	Number of vehicles Counted after correcting
52.297789	29.618526	Sunday	24	7376	307	330	4.47	7046	117	57	61	95	85.7	496	918	7376
52.297789	29.618526	Monday	24	8442	367	332	3.93	8110	125	50	64	93	84.9	417	1087	8645
52.297789	29.618526	Tuesday	24	10459	439	518	4.95	9941	183	77	83	175	81.8	373	1651	10575
52.297789	29.618526	Wednesday	24	11644	485	533	4.57	11111	174	97	65	197	80.9	393	1937	11644
52.297789	29.618526	Thursday	24	14465	609	663	4.58	13802	193	107	80	283	80.8	442	3022	14703
52.297789	29.618526	Friday	24	13473	563	701	5.2	12772	234	132	76	259	83.2	515	2373	13542
52.297789	29.618526	Saturday	24	12281	513	1036	8.43	11245	299	197	74	466	82.4	485	2087	12342
52.297789	29.618526	Sunday	24	11193	483	818	7.3	10375	229	174	64	351	82.1	474	1970	11802

Table 1. Column description of sample of sensor data

Fig. 2. Location of Road accidents near by the City Entrances of Shiraz



-- Sentinel 5P Satellite Data

The use of remote sensing and satellite data is a relatively new method for monitoring Earth's surface phenomena .The Sentinel-5P satellite was introduced by the European Space Agency to monitor the environment and air pollutants around the Earth, and for this purpose, the sensor of this satellite, called TROPOMI, was launched into space on October 23, 2008, to monitor the troposphere. This sensor is currently the latest satellite for monitoring air pollution, which, using the Pushbroom imaging geometry, images the Earth's atmosphere in three ranges: ultraviolet and visible (UV) (816 to 260 nm), near-infrared (NIR) (1 to 11 nm), and short-wavelength-infrared (SWIR) (836 to 832 nm). This satellite also uses other similar sensors such as SCIAMACHY and OMI with more advanced technologies for recording and recording information and provides users with up-to-date information with better spectral, spatial, and temporal resolution. Also, complete information on the Sentinel 5P satellite bands is given in Table2. One of the main objectives of this satellite is to measure the concentration of NO pollutants, which are measured using bands 3 and 0 and in the ultraviolet and visible regions. In order to extract satellite data from the Sentinel 5P sensor, we must extract images taken by the Sentinel 5P satellite by coding in the Google Earth Engine space and specifying information such as the type of band of each pollutant, the spatial range and the time period of the study. The data of all sets of this satellite are presented in two forms: NRTI and OFFL. NRTI data covers a smaller area than OFFL, but is made available to users faster. The 25P OFFL NO-Sentinel band and column_number_density2NO were used to extract the 2NO pollutant data. This dataset provides high-resolution offline images of the 2NO pollutant concentration.

Remote sensing and satellite data are an innovative method for monitoring Earth surface phenomena (Sohrabinia and Khorshiddoust, 2007). The Sentinel 5P satellite was launched by the European Space Agency to monitor the environment and air pollutants around the world, and for this purpose, the TROPOMI1 sensor on this satellite was launched into space on October 13, 2017 to monitor the troposphere. This Pushbroom imaging geometry probe, It images the Earth in three bands: ultraviolet and visible (UV) (270 to 490 nanometers), near infrared (NIR) (675 to 775 nanometers), and shortwave infrared (SWIR) (2305 to 2385 nanometers).

The advantages of this sensor include high spatial resolution, daily global coverage, and high signal-to-noise ratio (Veefkind et al., 2012). This satellite is currently the most advanced satellite for air pollution monitoring, and with the TROPOMI sensor, it can use similar sensors with more advanced technologies such as SCIAMACHY and OMI to record and record information, and also Provide users with emerging information with better spectral, spatial, and temporal resolution (Veefkind et al., 2012).

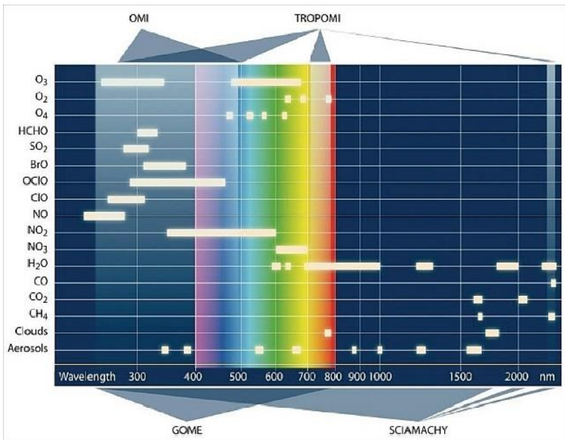


Fig 3: TROPOMI spectral window compared with GOME, SCIAMACHY and OMI (image credit: Dutch Space, TNO)

Product	Spectrometer	Application
Ozone	UV, UVIS	Ozone layer monitoring, UV-index forecast, Climate monitoring
NO_2	UVIS	Air quality forecast and monitoring
CO	SWIR	Air quality forecast and monitoring
CH_2O	UVIS	Air quality forecast and monitoring
CH_4	SWIR	Climate monitoring
SO_2	UVIS	Air quality forecast and monitoring, Climate monitoring, Volcanic plume detection
Aerosol	UVIS, NIR	Air quality forecast and monitoring, Climate monitoring, Volcanic plume detection
Clouds	UVIS, NIR	Climate monitoring
UV-Index	UVIS	UV index forecast

Table 2. List of Sentinel-5P level 2 products are show in the table (credit: KNMI):

3. Methods

The integration of spatial artificial intelligence (GeoAI) with geostatistical techniques has proven to be an effective approach for solving multifaceted urban problems (Li & Goodchild, 2020). By combining spatial interpolation with machine learning, this research not only identifies the spatial correlations between traffic, pollution, and accidents but also forecasts pollution levels with high precision. This innovative approach highlights the importance of spatiotemporal datasets in developing predictive models and informs targeted urban management strategies to mitigate pollution and enhance road safety.

3.1Spatial Interpolation

The Inverse Distance Weighting (IDW) method was applied to interpolate traffic sensor data collected at the entrances of Shiraz's Road network. IDW is a deterministic spatial interpolation technique that estimates unknown values based on the proximity of known data points. It assumes that points closer to the location of interest have a higher influence on the interpolated values. To implement IDW, Traffic Sensor Data Preparation: Traffic volume data from sensors at key road entry points were spatially referenced. Interpolation Process: Each point's traffic value was weighted inversely proportional to its distance from the unknown location. The formula used was:

$$Z(x,y) = \frac{\sum_{i=1}^N \frac{Z_i}{d_i^p}}{\sum_{i=1}^N \frac{1}{d_i^p}}$$

where $Z(x,y)$ is the interpolated value, Z_i represents the traffic value at point i, d_i is the distance to the unknown point, and p is the power parameter controlling the weight's influence. Validation: The interpolated results were validated against known traffic data from nearby points to ensure accuracy. This interpolation expanded the spatial scope of traffic data, enabling analysis along the entire road network.

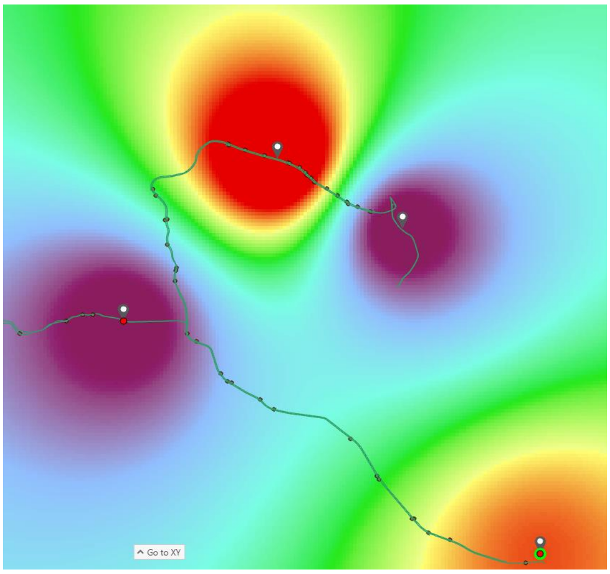


Fig 4: Traffic Sensor Data IDW map of entrances of Shiraz's Road network

3.2Correlation Analysis

Spatiotemporal datasets for traffic volume, road accidents, and nitrate pollution were aligned both spatially and temporally for analysis. Nitrate pollution levels were extracted from Sentinel-5P satellite imagery, which provides high-resolution measurements of atmospheric constituents. The analysis revealed a strong correlation between high traffic volumes and elevated nitrate pollution and areas with frequent road accidents and increased nitrate pollution.This correlation underscored the interconnectedness of traffic, accidents, and environmental pollution in urban areas.

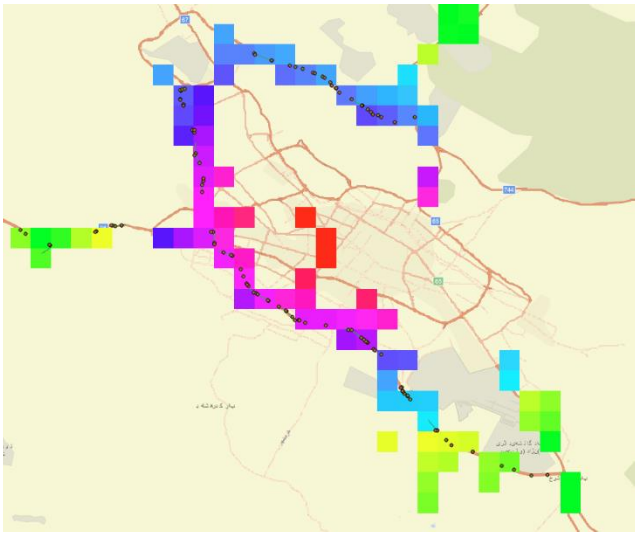


Fig 5: Traffic Sensor Data No2 pollution map for 2020 entrances of Shiraz's Road network

3.3Predictive Modeling with Machine Learning

To predict nitrate pollution levels for 2023, the Random Forest Machine Learning Algorithm was employed. Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mean prediction of the individual trees. It is particularly effective in handling complex, nonlinear relationships and integrating diverse datasets.

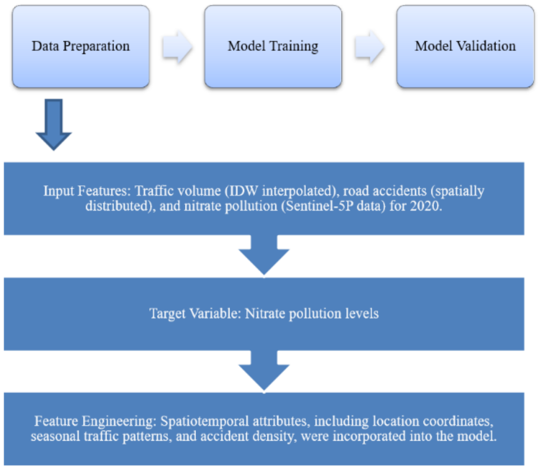


Fig 6: Steps in the Random Forest Implementation

1- Data Preparation:

- Input Features: Traffic volume (IDW interpolated), road accidents (spatially distributed), and nitrate pollution (Sentinel-5P data) for 2020.

- Target Variable: Nitrate pollution levels.

- Feature Engineering: Spatiotemporal attributes, including location coordinates, seasonal traffic patterns, and accident density, were incorporated into the model.

2. Model Training:

- Tree Construction: The Random Forest algorithm built multiple decision trees using bootstrap sampling. Each tree was trained on a subset of the training data to reduce variance. For Split Criteria: At each node, features were split based on their ability to minimize the impurity (e.g., Gini index or mean squared error for regression tasks). The ensemble Output: Predictions from all trees were averaged to produce the final output, reducing overfitting and enhancing generalizability. Finally, using trained Random Forest models, nitrate pollution levels were predicted for 2023 by inputting 2023 traffic and accident data (interpolated using the same IDW method).

4. Model Validation:

- Predicted nitrate pollution levels for 2023 were compared with Sentinel-5P satellite-derived data for the same year.

- Validation metrics, including R^2, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE), confirmed high predictive accuracy.

Advantages of Random Forest in this Study:

- Robustness to Overfitting: By averaging multiple trees, the model avoided overfitting to noise in the training data.

-Handling Nonlinear Relationships: The Random Forest algorithm effectively captured complex relationships between traffic, accidents, and nitrate pollution.

- Feature Importance: The model identified the most influential factors in nitrate pollution predictions, such as peak traffic volumes and accident hotspots.

3.4Validation Results

The predicted nitrate pollution levels for 2023 demonstrated strong alignment with actual pollution levels observed in Sentinel-5P satellite imagery. This validated the model's ability to leverage traffic and accident data to forecast environmental pollution trends accurately. The predicted spatial patterns closely matched the observed nitrate concentration distribution, underscoring the efficacy of integrating spatiotemporal datasets with machine learning methods.

4.Discussion and Conclusion

The spatial analysis of traffic volume, road accidents, and nitrate pollution revealed significant correlations, highlighting the interconnected nature of urban traffic dynamics, safety concerns, and environmental pollution. The spatial maps derived from the interpolated traffic sensor data, Sentinel-5P satellite imagery, and accident records demonstrated a strong spatial alignment between areas of high traffic density, frequent accidents, and elevated nitrate pollution. This spatial coherence underscores the value of integrating geospatial data in urban studies.

Using the Random Forest algorithm, nitrate pollution levels for 2023 were predicted based on the actual traffic and accident data for that year, alongside historical nitrate pollution data from 2020. The machine learning model demonstrated its capability to effectively utilize the complex relationships between the input variables, yielding accurate predictions for nitrate pollution levels. The spatial correlation analysis between the predicted nitrate pollution for 2023 and the actual nitrate pollution observed from Sentinel-5P satellite imagery revealed a 69% spatial correlation, indicating a robust agreement between the model's outputs and the real-world data.

This finding not only validates the reliability of the proposed method but also emphasizes the potential of combining geospatial data with advanced machine learning algorithms for environmental monitoring and urban planning. The integration of traffic and accident data with pollution analysis provides a comprehensive framework for understanding urban challenges and developing targeted mitigation strategies. By leveraging such innovative approaches, urban policymakers can better address the dual issues of traffic safety and environmental sustainability.

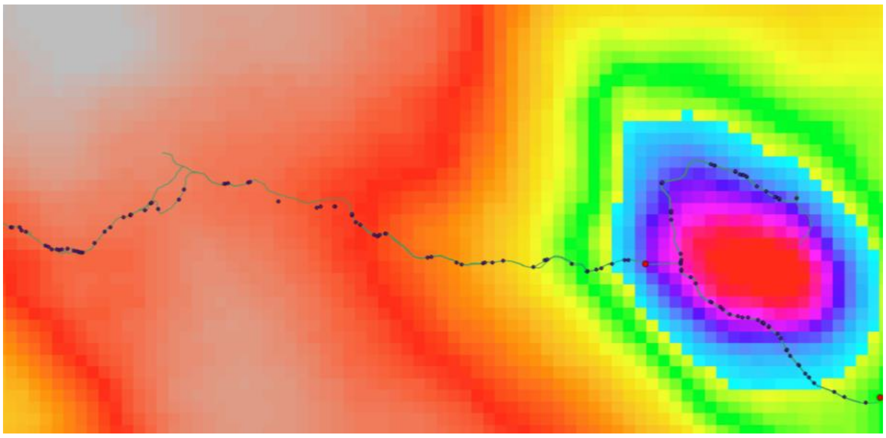


Fig 6: Traffic Sensor Data No2 pollution predictive map for 2023 entrances of Shiraz's Road network

5.REFERENCES

1.Breiman, L. (2001). "Random Forests." Machine Learning, 45(1), 5–32.

https://doi.org/10.1023/A:1010933404324

2. Shepard, D. (1968). "A two-dimensional interpolation function for irregularly spaced data." Proceedings of the 1968 ACM National Conference.

https://doi.org/10.1145/800186.810616

3. Zhang, Y., Liu, H., & Sun, J. (2022). “Application of machine learning in air pollution prediction.” Atmospheric Pollution Research, 13(2), 101–116.

https://doi.org/10.1016/j.apr.2021.101256

4. Li, S., & Goodchild, M. F. (2020). “Spatial artificial intelligence.” Annals of GIS, 26(1), 1–13.

https://doi.org/10.1080/19475683.2020.1736994

5. Donnelly, A., Misstear, B., & Broderick, B. (2010). “Real-time air pollution monitoring using GIS and machine learning methods.” Environmental Monitoring and Assessment, 170(1–4), 407–424.

https://doi.org/10.1007/s10661-009-1247-1

6. Shafizadeh-Moghadam, H., et al. (2017). “Integrating GIS and machine learning algorithms for urban land-use change modeling.” Journal of Environmental Management, 192, 252–261.

https://doi.org/10.1016/j.jenvman.2017.01.054

7. ESA Sentinel-5P Mission Overview. (2018). European Space Agency.

https://sentinel.esa.int/web/sentinel/missions/sentinel-5p

8. Chakraborty, A., & Sarkar, S. (2021). “Spatial analysis of road traffic accidents using GIS and statistical methods.” Accident Analysis & Prevention, 152, 105975.

https://doi.org/10.1016/j.aap.2021.105975

9. Hasan, M., et al. (2020). “Spatiotemporal analysis of air pollution and its relationship with road traffic.” Environmental Research, 185, 109400.

https://doi.org/10.1016/j.envres.2020.109400

10. Goovaerts, P. (1997). Geostatistics for Natural Resources Evaluation. Oxford University Press.

https://doi.org/10.1093/oso/9780195115383.001.0001

CONTACTS

Rozita SAKI
Tehran University
Address
City Tehran
COUNTRY Iran
Tel. 989171009162
Email: rozita.saki@ut.ac.ir