INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

Air Quality Predication System Using SPATIO Temporal Constrained Neuro –Fuzzy Classifier

Duganapalle Suresh Reddy[1], Research Scholar, Department of CSE,
J.S. University, Shikohabad,Firozabad (Dist.) , U.P.-283135
Email: sureshdugana@gmail.com

Dr.Vikram Singh Rathore[2] ,Professor, Department of CSE,
J.S. University, Shikohabad, Firozabad (Dist.), U.P.-283135
Email: rathorevsmail74@gmail.com

Dr. M. V.Rathnamma[3], Associate. Professor, Department of CSE,
K.S.R.M College of Engineering, Kadapa –A.P - 516003
Email: rathnamma@ksrmce.ac.in

--ABSTRACT--

Abstract - Concerns about the environment, such as pollution in the air and water, have a profound influence on people. Many people are losing their lives as a direct result of the filthy air because of the pollution. It is well known that the lungs and other respiratory organs are susceptible to the effects of air pollution. There is a constant increase in the levels of air pollution due to the fact that businesses are developing at a dizzying pace and an increasing number of people are utilising vehicles. In this circumstance, reducing the amount of time spent driving helps to maintain lower levels of air pollution. By increasing public awareness, giving statistics on the scope of the issue, and encouraging people to walk or use public transit more often rather than driving, the government is also doing its bit to reduce the amount of pollution that is released into the atmosphere. In addition, the specialists are collecting and evaluating data from a variety of areas throughout the city in order to keep an eye on the quality of the air. The purpose of this study is to offer a unique and effective technique for predicting air quality that is based on spatiotemporal limited fuzzy neural networks. Through the use of a spatiotemporal limited fuzzy neural network, this study introduces a unique technique to the forecasting of air quality. The newly developed algorithm allows for more accurate forecasting of the air quality. These letters, NFTCA-S, are where it all began. In addition, a trustworthy Butterfly Optimisation Algorithm (BOA) is used in order to extract the qualities that are the most beneficial from the natural data. Through this method, both the reduction of processing time and the improvement of efficiency are accomplished.

Keywords- NFTCA, BOA,FUZZY, air pollution.

I. INTRODUCTION

Concerns about the environment and technology improvements are contributing to the high population increase that is occurring in a number of locations around the world. The heavily populated nations of India and China are without a doubt facing a significant challenge in the form of air pollution. Many people, regardless of their age, are battling with a variety of ailments, including lung cancer and asthma, all of which are prevalent. Methods to detect and predict the levels of air pollution are being developed by experts from all over the world in order to fight the issue of air pollution. As an additional precautionary action in response to the projections, people all around the world are decreasing the amount of time they spend driving their vehicles and increasing the amount of time they spend planting trees. At the moment, the technologies that we use to anticipate air quality are not sufficient enough to reduce the amount of pollution that is present in the air. When calculating a nation's overall health score, it is necessary to take into account the quality of the air. For the goal of defining the health rate, which is also referred to as the air quality index, it is necessary to evaluate the effect that air pollution has on human health in relation to the present state of the air. Increasing public awareness of the significance of air quality and giving a daily estimate of the levels of air pollution are two current techniques to reduce air pollution that are in compliance with rules imposed by the government. In order to lessen the amount of pollution that is released into the environment, it is necessary to raise the air quality index rate. An innovative method for predicting air quality is presented in this study, and it is anticipated that it will be successful in the environment for which it was designed.

Monitoring and analysis of the PM 2.5 level has been carried out in New Delhi, India, from the year 2013. Since that time, a number of other major cities in India, including Chennai, Hyderabad, and Mumbai, have followed these same steps. The purpose of this approach to monitoring air pollution is to offer information both about the present circumstances and the trends that are occurring. It is possible for researchers and members of the general public to have a better understanding of the different metropolitan areas by using the real-time

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

PM2.5 data that is available to them. It is also documented and made accessible for research reasons that the PM2.5 levels in each municipality are measured and reported. In addition to this, a statistical analysis is performed on each place from season to season. It is our intention to apply machine learning approaches for the purpose of data pre-processing and categorisation in order to investigate regional trends in PM2.5 measurements.

An example of a job that machine learning systems often do is the categorisation of air pollution levels in different locations, such as business and study. Enhancing the performance of segmentation is accomplished by the use of machine learning techniques such as DT, RF, NB, SVM, and NN. Combining a large number of machine learning models is made possible via the use of ensemble methods. In this particular setting, stacking is an essential part of the ensemble method to learning that is being used. You will be able to see the operation of the algorithms as they are applied to the dataset during the training process by looking at the result column. With regard to performance, it is probable that alternative models will fall short of the layered machine learning model [7]. Just-in-time manufacturing and ski injury prediction are two examples of the many applications that may be found for this technology, which is also often used to win competitions [8-9]. Nevertheless, machine learning strategies are not accurate enough on their own, therefore they need to be supplemented with deep learning methods in order to get even better outcomes.

The following is a list of the most important contributions that this effort has included:

in order to come up with a fresh method for predicting air quality that is based on the neuro-fuzzy temporal classification algorithm (NFTCA).

You may make accurate assumptions about the degree of air pollution in different areas by using the information that is provided by the assessment of the air quality in the respective region.

When integrating the PM2.5 datasets with the neural system that has been presented, it is difficult to arrive at reasonable findings owing to the geographical limits of the system.

In order to improve the process of feature optimisation, the Butterfly optimisation Algorithm (BOA) that is already in use is being used.

so as to get forecasts that are very accurate while having a root mean square error rate that is as low as possible.

This chapter makes available a detailed explanation of the approach that has been suggested for predicting the quality of the air you breathe. In it, you will discover the formula, the operations that are necessary, and an explanation of the architecture of the system. In addition, comparative analysis is included in the part that discusses the results and the discourse.

2. SOFTWARE DESIGN

As can be seen in Figure 4.1, the newly developed prediction system is functioning in a clear and concise manner. A number of critical components of the process include the classification of data, the optimisation of features, the user interface module, the spatial agent, the temporal agent, the forecast manager, the rule base, and the rule manager. The relevance of each component is also examined in detail, and this is accomplished by evaluating the approach or plan that was used in doing this study. Following the completion of the PM2.5 collection, the user engagement module is provided with the essential raw data, which is referred to as PM2.5 data. The forecast manager receives all of the necessary data in order to guarantee that the air quality measurement is carried out in an appropriate manner.

This is the point at which the information that was gathered from the user interface module is sent to the probability manager. Utilising the BOA optimisation strategy, the prediction manager continues to optimise the features in order to achieve optimal performance. After the dataset has been preprocessed, it is then submitted to the data classification step so that the newly generated NFTCA-S may be classified using the feature selection approach. Following the classification of data by the application of spatial and temporal restrictions, as well as fuzziness through the utilisation of facts, fuzzy temporal rules, and expert opinion that is recorded in the rule base, the decision manager ends up making the final choice about the forecasting process. In this section, the decision manager provides the rule manager with instructions on how to handle the facts and regulations that have been stored in the rule base. Following that, the rule manager is the one who handles it. The decision manager, who is in charge of the overall design of the system, provided the fuzzy rule manager with suggestions, and the fuzzy rule manager used

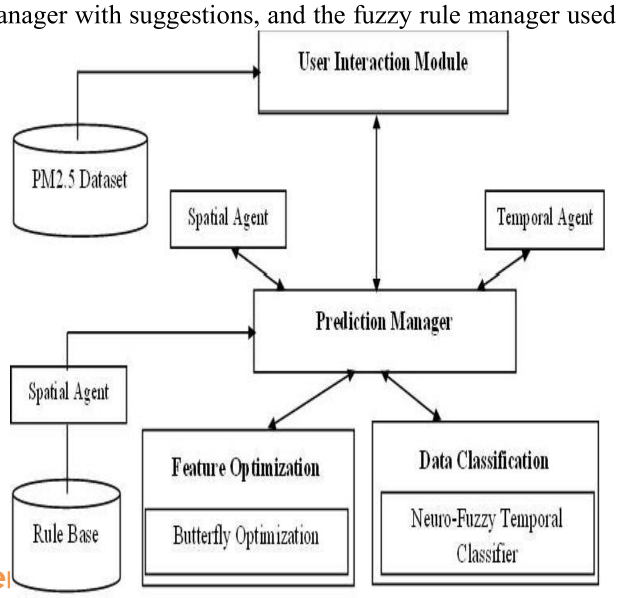


Figure1 System

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

those recommendations to develop the fuzzy temporal rules and add them to the rule base. The purpose of this article is to describe the new method of predicting air quality, which involves making intelligent guesses about the degree of air pollution in a number of different cities and countries. In the course of this study, a unique NFTCA-S model that takes into account geographical restrictions was constructed for the goal of predicting smog and air quality. According to Arora and Singh (2018), the use of a capable BOA has the potential to improve the accuracy of the predictions produced. For the purpose of demonstrating the new system, two components, such as "feature optimisation" and "classification," are used. The first thing that is presented is an explanation of the crucial formulas and techniques that are required for feature enhancement.

3. OPTIMISING FEATURES

The process of simplifying features is an important factor to consider when determining which features should be prioritised for donation. The feature reduction technique is used by the forecasting system in order to cut down on the quantity of information that may be utilised. As a consequence of this, it cuts down on the amount of time needed for testing and training while still delivering high-quality categorisation results. This research makes use of BOA (Arora and Singh 2018) to improve the quality of the input data in order to get a more effective and efficient categorisation. Within this part, the user interface module supplies the input dataset PM2.5 with the essential features and values. These include PM2.5, PM10, NO2, CO, O3, and SO2. There are also PM2.5 and PM10 included. Following the completion of the project, it will proceed to the phase of feature optimisation, when it will undergo further improvement.Organising Data

By offering some background information and explaining the fundamental techniques for brain classification utilising equations, this section provides an explanation of the working of the newly developed spatially limited NFTCA.

Intelligent Classifier

Neural networks (NNs) have a number of practical applications, one of which is the quick categorisation of incoming data according to supervisory criteria that have been set. Over the course of the last several years, NN-aware classifiers have been used in a number of different data categorisation endeavours. There are a variety of professions that make use of NNs, including as the medical field, social services, and the military. Since neural networks are capable of self-adaptation, they do not need a training technique to begin with. According to the second point, neural networks are regarded as a general method that may be used to any methodology in order to get a prediction result. Adaptable models are developed by the use of nonlinear models, often known as NNs. Issues that occur in the actual world may then be addressed using these models. Next, neural networks are used in order to provide predictions about the posterior probability. These are the fundamental pieces of information that are used to construct the rules and statistics. The input, the processing, and the output levels are the three fundamental levels. The input layer is in charge of collecting responses from users and then transmitting those responses to the process layer so that they may be processed further. The data must first be received by the processing layer before it can be processed. After that, the output layer is responsible for producing the final result. In the next step, the output layer will communicate the result to the client. There has been a significant amount of writing done on neural predictors such as ANN, BP, and FN by authors.

Non-Linear Analysis

According to Zadeh (1998), the fuzzy set is a valuable tool that may be used when dealing with or showing data that is confusing. There are many different forms of database uncertainty, some of which include inconsistency, haziness, imprecision, and a lack of clarity for the data. By using ambiguity, fuzzy sets are able to produce systems that are unintelligible. In order to provide a workable method for dealing with data that is both incomplete and confusing, fuzzy sets are presently being developed. In order to increase performance, there is a common practice of using all of these to muddy the data. Utilising the fuzzification and de-fuzzification processes, you may transform fuzzy values into crisp values or crisp values into fuzzy values according on your preference. It is important to use fuzzy membership functions in order to determine the fuzzy ranges. Some examples of these functions are the trapezoidal, triangular, Mamdani, and Gaussian functions. The rules are formulated with the help of the fuzzy bands that are produced by the fuzzy membership function, and the results are also acquired for each individual occurrence.

A. Time Limitations

A significant amount of time is required in order to accurately forecast future occurrences, such as illnesses, earthquakes, and air pollution. A certain amount of time is required in order to make intelligent judgements about day-to-day activities, which is particularly important in the fast-paced digital environment of current times. Even more time-consuming is the process of gathering information from a broad range of different sources. As a consequence of the fact that it takes very little time to grasp and produce the outcome, every approach is considered to be effective in every circumstance. As a

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE
ISSN: 2057-5688

consequence of this, time is considered to be a significant barrier in this investigation when it comes to picking relevant input records. In the same way, the amount of time that is required to collect data in each country and area is constantly regulated. Data sets that are collected over a longer period of time are likely to provide more precise results than data sets that are collected on a daily basis.

Limitations on Space

Geographic data is an essential component in the process of acquiring information on topics such as health, transportation, weather, and air quality. This study takes use of the PM2.5 data in order to accurately classify items and make predictions on the quality of the air. In order to accurately forecast the future air quality of a region, the PM2.5 dataset uses the date, the time, and the location as very important factors. Therefore, the training technique is carried out on each of the data sets in order to do this. The reason for this is because judgements about the methodologies that are used to make forecasts of air quality should be spatially limited.

B. *A New Neural Classifier Being Poised*

In this part of the article, we will discuss the NFTCA-S approach, which was established not too long ago, and how it contributes to more accurate predictions of air quality. Within the scope of this section, some background material on neural networks (NNs), fuzzy reasoning, and spatial and temporal restrictions is discussed. This part also discusses the newly developed NFTCA-S as well as the important following actions that are to be implemented. When it comes to this particular scenario, we train NN to swiftly infer conclusions from input on air quality by using spatiotemporal fuzzy rules. Additionally, a sigmoid function is used as an activation function in order for this neural network to be able to make judgements. Furthermore, the trapezoidal fuzzy membership function is used in order to construct the fuzzy gaps that are necessary for the completion of the decision-making process itself.

4. Conclusion and Analysis

Throughout the several subsections that make up this part, the evaluation metrics, the description of the dataset, as well as the experimental and comparative results, are all shown. The first thing that is presented is an inventory of the evaluation instruments, along with short explanations of each method.

4.4.1 Assessment Criteria

TheimportantandnecessaryevaluationmetricssuchasRoot MeanSquaredErrorperformance in terms of prediction accuracy. The RMSE value is calculated by usingtheequation(4.1).

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

What happens when N is the record length? The variables represent the real numbers that help in the process of making a good prediction. The root-mean-squared error (RMSE) was used to determine the average number of horizons for each prediction procedure using the horizon numbers. In the course of this investigation, a series of experiments were carried out in order to evaluate the newly created method for forecasting the quality of the air environment.

4.4.2 Description of the Data Set

In order to build the open-source PM2.5 dataset, the air quality index measurements from a number of monitoring points in Beijing, China were combined. If you are looking for it, you may locate it at the UCI Machine Learning Repository. Particulate matter, dust, ozone, and sulphur dioxide are all examples of substances that fall within this group. Data pertaining to the climate, such as pressure, temperature, humidity, wind speed and direction, and other meteorological factors, are also contained in the file. Furthermore, the Meteorological Data Centre in Chennai is equipped with all of the data that is required for the purposes of training and verification.

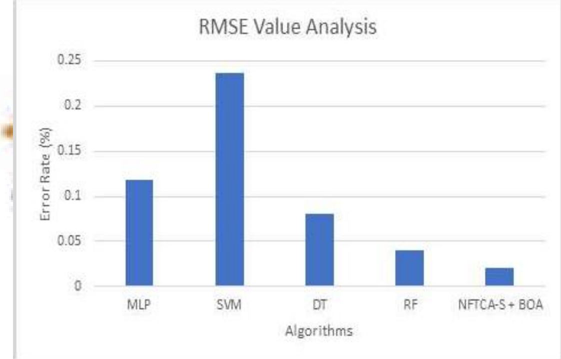
4.4.3 Findings from the Experiment

Thedifferentnumbersofexperimentshavebeendonetoevaluatethenewly

input records are used to perform the training process to identify the patterns and theremaining 20% of the input records in the PM2.5 dataset is considered to perform thetestingprocess.TheinputrecordsarepreprocessedbyusingBOAforobtainingoptimal features that are helpful for enhancing the prediction accuracy. In this datapre-processing stage, the irrelevant data are eliminated and the relevant that are havingvaliddataonlyconsideredforfurtherprocess.

Figure4.2demonstratestheperformanceoftheNFTCA-Swithrespecttotheprediction accuracy of the newly developed air quality prediction system over thetraining and testing dataset. Here, five different experiments have been done with theconsiderationofthevarious records as input.

ISSN: 2057-5688



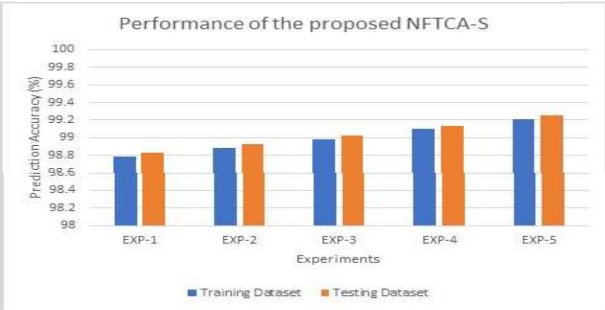


Figure 4.2 The suggested NFTCA-S's performance in the absence of feature selection

The recently built NFTCA-S demonstrates satisfactory results in each of the five tests, as seen in Figure 4.2. With the passage of time and the accumulation of more data, the accuracy of the prediction increases. utilising newly designed spatio-fuzzy temporal rules to generate snap choices utilising input datasets results in improved accuracy. This is brought about by the use of the input datasets. One is able to create judgements via the use of fuzzy logic by using a collection of facts in conjunction with specific geographical and temporal limits. For the purpose of forecasting for the future, it makes use of previous data on the state of the air and the weather.

Figures 4.3 and 4.4, which make use of the most recent BOA, illustrate how well the new air quality prediction system performs in terms of formulating projections throughout the training and testing datasets. For each investigation, they make use of the most effective highlighted dataset that is available. The location in

Figure4.3. Optimising NFTCA-S for the Training Dataset and Its Prediction Accuracy

question has been subjected to five separate examinations, each of which has used a different collection of data as input.

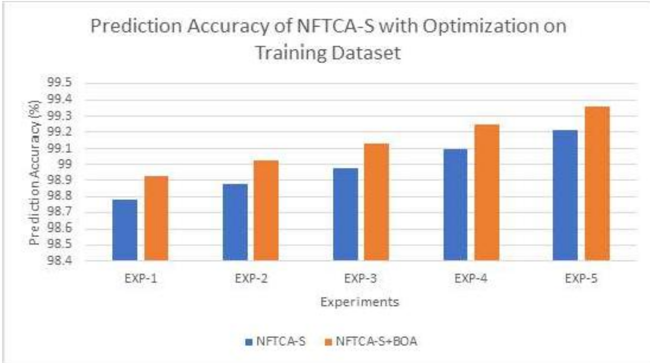


Figure4.4.The Testing Dataset and the Prediction Accuracy of NFTCA-S after Optimisation

It can be shown in Figures 4.3 and 4.4 that the newly built NFTCA-S with BOA works very well in each of the five tested conditions. Given this situation, it can be concluded that the accuracy of forecasts progressively improves as the amount of data points in each graph rises. Making use of BOA was the impetus for the move. A prognostication

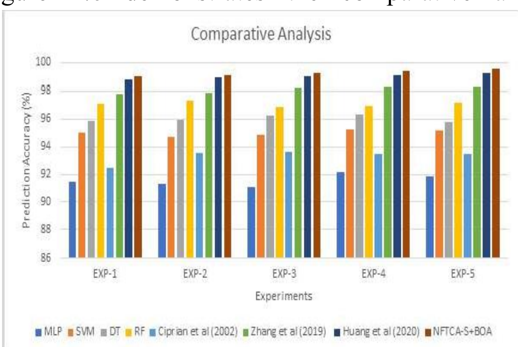
The training dataset was able to reach an accuracy of 99.36%, while the testing dataset was able to obtain an accuracy of 99.56%.

Figure4.5. Comparison of the suggested and current methods' root-mean-squared error values

Figure 4.5 presents a comparison between the new air quality prediction system, which make use of the NFTCA-S and BOA, and the prior classification approaches, which include SVM, RF, DT, and MLP. The comparison is made by examining the RMSE values of each of these techniques.

As shown in Figure 4.5, the performance of the new air quality prediction system, which integrates the NFTCA-S and BOA, is inferior to that of the earlier models, which included MLP, SVM, DT, and RF. The better prediction outcome is a direct result of the spatio-fuzzy temporal rules that were created not too long ago. These rules make it easier to make sensible decisions on the datasets. For the purpose of making accurate selections based on the data that is presented, fuzzy logic is used. Additionally, spatial boundaries are utilised in order to develop informed estimates about the previous weather and air quality in space environment.

Figure 4.6 demonstrates the comparative analysis



INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

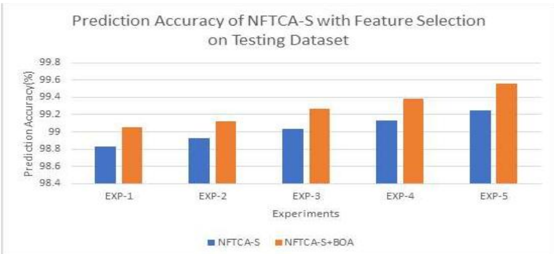
between the newly developed airquality prediction system and the existing classifiers namely SVM, DT, RF, MLP andthe prediction systems developed by Ciprian et al (2002), Zhang et al (2019) andHuangetal(2020).Thenewlydevelopedpredictionsystemthatcombinesthe NFTCA-SandBOAobtainedhighprediction accuracy with low error rate. Thisexperimental result demonstrates the effectiveness of the newly developed air qualityprediction system by conducting five experiments over the different set of air qualityrecords.

Figure4.6. Evaluation in Comparison

As can be shown in Figure 4.6, the innovative air quality prediction system (-S+BOA) performs better than the current classifiers (SVM, DT, RF, and MLP) as well as the prediction systems created by Ciprian et al. (2002), Zhang et al. (2019), and Huang et al. (2020) in each of the five trials. The enhanced performance may be attributed to the recently created spatio-fuzzy temporal rules, which are used for the purpose of making choices in a timely manner across many datasets. The trapezoidal fuzzy membership function makes it possible to make decisions based on the input data as well as past location predictions depending on the conditions of the air and the weather.

Conclusion

During the course of this research endeavour, new techniques for making precise forecasts about the quality of the air have been established. For the purpose of forecasting, a new neural network (NN) with spatial and temporal fuzzy rules is used, and the input data is used to



conduct the procedure. The process of feature tuning, which increases the accuracy of predictions and minimises the root mean square error (RMSE), is also simplified by a contemporary BOA. In order to arrive at judgements in a short amount of time, this method employs a single-neural network (NN) that is bound by time, space, and fuzzy logic. The incorporation of PM2.5 data into this prediction technique provides further evidence that the precision of the forecasted air quality is maintained.

to make a decision on the utilisation of nitrogen and carbon monoxide. There is a possibility that this endeavour might be taken to the next level by using contemporary deep learning methods to enhance the accuracy of the forecasts.

References

[1] J.Zidan,E.I.Adegoke,E.Kampert,S.A.Birrell,C.R.Ford, and M.Higgins,“GNSSVulnerabilities andExistingSolutions:A ReviewoftheLiterature,” IEEEAccess, earlyaccess, vol. 4,2020.

[2] ICAO,“Performance-basedNavigation(PBN)Manual,” 2013.

[3] Civil Aviation Authority, “CAP1915, Unmanned Aircraft SystemsBVLOSOperationsinSupportoftheCOVID-19Response–Requirements,Guidance&Policy,”2020.

[4] R. Sabatini, T. Moore, and C. Hill, “A new avionics-based GNSSintegrity augmentation system: Part 1 - Fundamentals,” J. Navig., vol.66,no.3,pp.363–384,2013.

[5] N. Zhu, J. Marais, D. Betaille, and M. Berbineau, “GNSS PositionIntegrityinUrbanEnvironments:AReviewofLiterature,”IEEETrans.Intell. Transp. Syst., vol.19,no. 9,pp.2762–2778,2018.

[6] M. Tahsin, S. Sultana, T. Reza, and M. Hossam-E-Haider, “Analysisof DOP and its preciseness in GNSS position estimation,” in 2ndInternational Conference on Electrical Engineering and InformationandCommunicationTechnology,ICEEICT,2015.

[7] J.Januszewski,“SourcesofErrorinSatelliteNavigationPositioning,” TransNav, Int. J. Mar. Navig. Saf. Sea Transp., vol. 11,no.3,pp.419–423,2017.

[8] O. K. Isik, J. Hong, I. Petrunin, and A. Tsourdos, “Integrity analysisfor GPS-based navigation of UAVs in urban environment,” Robotics,vol. 9,no.3,pp.1–20,2020.

[9] Y. Wang et al., “Urban environment recognition based on the GNSSsignal characteristics,” Navig. J. Inst. Navig., vol. 66, no. 1, pp. 211–225,2019.

[10] J. Tan, J. Wang, and D. Lu, “GNSS data driven clustering method forrailway environment scenarios classification,” Proc. 14th IEEE Conf.Ind. Electron.Appl.

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

ICIEA2019,pp.2026–2031,2019.

[11] R.U.R.Lighari,M.Berg,E.T.Salonen,andA.Parssinen, "Classification of GNSS SNR data for different environments andsatelliteorbitalinformation,"in11thEuropeanConferenceonAntennasandPropagation, EUCAP2017,2017,pp.2088–2092.

[12] P.Misra,P.Enge,*GlobalPositioningSystem:Signals,Measurements,andPerformance*(Ganga-JamunaPress,Lincoln,MA,2006)D.Gebre-Egziabher,S.Gleason,*GNSS applications and methods*(ArtechHouse,2009)

[13] M.Karaim,M.Elsheikh,A.Noureldin,R.Rustamov,GNSSerrorsources.MultifunctionalOperationandApplicationofGPSpp.69–85(2018). Doi:10.5772/intechopen.71221

[14] P.Groves,L.Wang,M.Adjrad,C.Ellul, in *Proceedingsofthe28thInternationalTechnicalMeetingofTheSatelliteDivisionoftheInstituteofNavigation* (ION,Tampa,Florida,2015),pp.2421–2443

[15] H. Kuusniemi, G. Lachapelle, in *Proceedings of the 2004 National TechnicalMeetingofTheInstituteofNavigation*(ION,SanDiego,CA,2004),pp.210–224

[16] R.E.Kalman,Anewapproachtolinearfilteringandpredictionproblems.Journalofbasicengineering**82**(1),35–45(1960)