IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 14 September 2024, accepted 12 November 2024, date of publication 15 November 2024,
date of current version 26 November 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3499316

RESEARCH ARTICLE

Assertion of Soil Data Consistency by Detecting and Removing Spatial Outliers Using Iterative Techniques for Precision Agriculture

ARUN KUMAR HIREMATH[1], K. R. NAVEEN KUMAR[1], MANJUNATHA HIREKERI MALLESHAPPA[2], BHASKAR AWADHIYA[3], AND YASHWANTH NANJAPPA[3], (Senior Member, IEEE)

[1]Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology, Davanagere, Karnataka 577004, India
[2]Department of Electrical and Electronics Engineering, Bapuji Institute of Engineering and Technology, Davanagere, Karnataka 577004, India
[3]Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India

Corresponding author: Yashwanth Nanjappa (yashwanth.n@manipal.edu)

This work was supported by the Manipal Academy of Higher Education, Manipal.

ABSTRACT In Precision Agriculture, a Decision Support System (DSS) is the ultimate stage that incorporates the basic findings derived from earlier procedures. In many cases, a DSS that is exclusively focused on data-driven methodologies might be highly influenced by data sources if these data resources do not provide any sense of intended conclusion. As a result, on-farm experiment inputs may result in incorrect site-specific crop management in the end. Since data is an essential element of DSS, irregular patterns such as outliers in spatial data that can change the nature of expected outcomes must be avoided during data-driven manipulations. Many of the approaches developed to detect outliers were not designed to deal with masking and swamping effects. With this consideration, the work presented here uses two iterative techniques to locate and remove spatial outliers based on their neighbourhood relationship. As a result of this technique, the masking and swamping effects are reduced. The methods we use are iterative, implying that each iteration discovers and eliminates the expected number of outlying observations. R-Studio is used to demonstrate the use of iterative approaches. The efficacy of both iterative procedures was analysed and compared using one of the current graphical approaches, such as the semivariogram. The research specifically looks at how well these strategies perform on a real-world dataset incorporating spatial observations. The statistical iterative techniques outperformed the graphical approach, according to the findings.

INDEX TERMS Decision support system (DSS), masking and swamping effects, precision agriculture (PA), site-specific crop management (SSCM), spatial autocorrelation, spatial neighborhood, spatial outliers, Iterative-R, Iterative-Z, variogram.

I. INTRODUCTION

In order to cope with Soil, Zones, and Plant qualities in relation to natural and economic situations, economically sustainable farming requires the adaption of precise data-driven methodologies and Principles developed during decision support systems in on-farm experiments [1]. Precision agriculture is information-intensive, Site-specific crop management (SSCM) deals with precise application of Chemical-agro inputs to the subfields as per the inference indicated by data-driven techniques in a Decision support system (DSS). Data-driven methods [2] are regarded as the heart of DSS, requiring investigation of obtained data to assure data-feature regularity. As a result, irregular observations, also known as Outliers, are more fascinating patterns than normal data since they provide the most crucial guidance to aid in decision-making. Spatial outliers are a type of contextual outlier in which the contextual attribute is the spatial attributes, such as spatial relationships like distance or adjacency. Such outliers must be eliminated because they may misread an expert's affirmation of decision-making processes. In most cases [3], detecting an outlier is challenging, since identifying the pattern that outliers might follow is difficult. The straightforward attempt is to utilize

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

the knowledge of the underlying situation. A few innovative spatial-outlier detection techniques offered to supplant existing statistical approaches in the context of Precision agriculture aided data mining. The study of statistical techniques for dealing with outliers, as well as comparisons with existing methodologies, has yet to be undertaken. Although iterative techniques [4], [5]were proposed many years ago with a rich and wide background, their employment in precision agriculture data-related activities such as site-specific precise nutrient management has remained a strikingly new concern.

For optimal soil management and the planning of precise agro-inputs, accurate estimation of soil physicochemical characteristics is necessary. Soil characteristics are complex as the associated heterogeneity makes interpretations uneasy and assumptions made over the fields are not always reliable. But the knowledge of variability in soil parameters helps to understand the production variability. Optimal soil management is always concerned with soil variability estimation which affects crop variability. The spatial data analysis helps in developing useful information for the proper soil estimation. Soil samples from subfields are usually analyzed to measure the concentration of nutrients and to indicate fertility and deficiency. Based on this analysis, the precise application of fertilizers for crop management is then recommended at various stages. Hence, at this point, it can be concluded that accurate estimation of soil physic-chemical characteristics is necessary for optimal soil management as well as precise agro-inputs planning, where the quality of collected data is evaluated in association with conduction of outlier detection mechanism, as any sort of spatial outliers can alter the actual predictions resulting in unrealistic inferences. Thus, in such early management phases, any spatial variability, if associated with the soil characteristics, needs to be carefully assessed to isolate it from the remainder of the data.

The spatial outlier detection mechanism provides a strong base for recommending the precise application of fertilizers. A spatial variability associated with the soil characteristics needs to be assessed to expedite the presence of outliers to isolate them from the remainder of the data. Spatial outliers in many cases disclose significant speculations in which they may reveal true information about their outlyingness. The process of detecting outliers not only helps in isolating the irregular data observations from the remaining data but also attempts to focus on masking and swamping effects which are critical challenges in outlier detection. Masking occurs when true outliers are hidden by nearby normal data points, leading to their incorrect classification as non-outliers. Conversely, swamping occurs when non-outliers are falsely identified as outliers due to the influence of nearby true outliers. These effects are particularly problematic in spatial data analysis, where neighborhood relationships play a significant role.

Several statistical techniques consider the relationships between the observations in the neighbourhood to expel the outlyingness of data observation, and most of these conventional technologies concentrate on detecting single outliers of attributes, so when their neighbourhoods include specific spatial outliers with extreme attribute values, they may misclassify normal cases as outliers. Hence, these associated techniques provide no clear idea to overcome this adverse effect. The research work presented here strongly attempts to serve the process of detecting outliers, which not only helps in isolating the irregular data observations from the remaining data but also focuses on masking and swamping effects, on which significant attention has been made. So, the work mainly contributes towards understanding the spatial data observations and the neighbourhood relationships between them. Using statistical iterative techniques, an effort has been made to provide the solution to address the issues involved in finding the outliers in spatial data. With the statistical iterative techniques, the outliers in the collected data are stabilized with their neighbour's cooperation and this methodology finally makes the normal data values follow a common generating mechanism.

The objective of the current research work is to conduct statistical iterative techniques [6] as a binary classification exercise to form an outlier detection methodology which works on observations neighbourhood relationship. The iterative techniques considered for the work demonstration rely on the astute principle, which observes a data-point's attribute values at a location against an aggregate-function outlining the neighbourhood attribute-values. The comparison is then normalized across the entire dataset. The observations with the highest-outlier score are considered as spatial-outliers against the observations with low score. Further, one of the graphical approaches such as semivariogram has been considered to work on visualization of data.

The work is initially concerned with employing the iterative techniques for the outlier detection and removal from the spatial data collected and then the computational efficiency of both the techniques is analysed. And, Performance evaluation of both the iterative techniques has been explored using relevant performance metrics. The iterative and graphical approaches are compared to show their performance effectiveness. The major concern of this research work is to reduce the swamping and masking effects using iterative techniques and to analyze the computational efficiency in this regard. Table 1 shows the real-time Spatial Data related to Soil characteristics and agriculture is collected from the Davanagere jurisdiction (District in Karnataka State, India) in association with ''Taralabalu Krishi Sanshodhana Kendra'' and the Agricultural department, Davangere.

As shown in Table 1, an agricultural dataset for the davangere region has been considered with many important factors not only related to the soil but also containing environmental aspects. Consideration of too many parameters with complex data behaviour for a model may be difficult to perform. And also, the involvement of many irrelevant parameters may not give satisfactory results. Hence, for the outlier detection techniques, the current work considers only those parameters which are significant and highly relevant to the work.

The current research work, contributes to the farming society through the deployment of statistical and graphical

techniques altogether to detect, analyze and remove outliers and any anomalous data which leads to the misinterpretation of the data during decision making for the optimal soil management. In the end, in the context of outlier detection, the work carried out claims the most suitable statistical technique for the optimal application of agro-inputs.

TABLE 1. Spatial data containing 683 observations made at various villages from all six talukas of davangere district, Karnataka,India.

Location	y	x	N	P	K	pH	EC
Khudapura	14.421	76.554	103.21	22.31	105.32	6.98	0.29
Manamainahatti	14.425	76.522	124.02	27.08	124.63	6.77	0.11
Turuvanur	14.400	76.430	102.31	12.05	58.8	5.98	0.12
Kolahal	14.334	76.187	160.16	23.01	58.22	5.8	0.09
Ganjigatte	14.234	76.115	62.01	8.15	58.32	8.14	0.36
Gyarahalli	14.228	76.107	53.02	9.75	69.73	8.16	0.29
Muthugaduru	14.219	76.115	102.32	19.75	68.11	7.54	0.18

The rest of this paper is organized as follows; the related work is presented in Section-II with most relevant work done in connection with the current research focus. In Section III, major steps involved in the spatial iterative outlier detection techniques and the details of associated study area under usage, have been explored. Performance evaluation of iterative techniques is assessed in Section IV. A detailed discussion of the results has been made in Section V.

II. LITERATURE REVIEW

Data mining, as an established- area of research, has its major success in deploying applications and serving the results to the end user in various-fields ranging from medical research to Precision agriculture [3].To be specific, Precision agriculture, to uplift Site-Specific-Crop Management [7],with the involvement of Software-Based-wireless-sensor-network monitoring system [8] and yield-optimization strategies [2] at a smaller scale [9],is turning exponentially into data-driven approach. Existing traditional outlier-detection-methods can be categorized into: Spatial Neighbourhood Information based, Extreme value based, Information theoretic Based, Distribution-based, Proximity-based, Density-based, Probabilistic or Statistical methods and Depth-based Outlier-detection methods etc.

Many techniques are developed to handle outliers specifically in the context of spatial data. Their aim was to showcase the temporal correlation and spatial autocorrelation in expressing the data-inference for the prediction/forecasting models. An unsupervised filtering method for the labelling of faulty spatial neighbourhood observations has been suggested [7] in which a non-parametric and unsupervised approach is detected by Outliers. Using multi-level restricted Delaunay triangulation [1], Neighbourhood information fusion [10], Robust Metric Learning for Contextual Neighbourhood Exploration [11], Likelihood Displacement Statistic method (LD) and Likelihood Ratio Statistic for a Mean Shift method (LR) [12] were concerned on dealing with contextual attributes. The Clustering-based techniques [13], [14] identifies exceptional observations as anomalies whose interest does not belong to any of the existing clusters. The efficacy of techniques with this idea is not optimized since these are not exactly developed for outlier-detection. Proximity-based statistical algorithms [15] compare the value of an attribute of an instance with its neighbourhood attribute-values aggregate.

Distance-based-methods such as Solving Set-based approach and Fast Solving Set approach [16] Partial least squares (PLS) for detecting multivariate outliers [17] declare exceptionally far distant data points as outliers. In [18], kriging and inverse distance weighting (IDW) are evaluated for spatial analysis of soil bulk density to reflect true variation of bulk density. In [13], Minimum volume ellipsoid (MVE) with principal component analysis (PCA) extension has been applied to detect multivariate outliers. Cook's Distance identifies the outliers if the data-points exceed their cut-off value.

Density-based-algorithms [19], [20] such as Local-Outlier-Factor (LOF) evaluate the outlier ness of an object in terms of its local-reachability-densities. Quantitative methods [7], [21], [22] investigates the data to differentiate spatial outliers from the normal observations. From the literature survey, it has been examined that, Outlier detection techniques in spatial context [5] and prediction models describing the relationship between soil properties and the yield are developed with an inadequate knowledge in which spatial outlier detection techniques are superior. Traditional-outlier-detection-approaches developed for the detection of outliers are ineffective due to the sparsity of the data objects with many attributes. As a major concern, in the context of spatial autocorrelation [23], it is examined that existing-spatial-outlier-detection-techniques primarily concentrate only on how to identify an outlier with a single attribute. To analyze the robustness of swamping and masking [24], several authors gave a fundamental and conceptual framework [25,26,4] with basic foundations to figure out the robustness of outlier detection procedures [27].

As a major concern, in the context of spatial autocorrelation, it is examined that existing-spatial-outlier-detection-techniques primarily concentrate only on how to identify an outlier with a single attribute. Where the frequency of anomalies in the test is uncertain, the masking and swamping effect may occur [8], [12]. To analyze the robustness of swamping and masking, several authors gave a fundamental and conceptual framework [28], [29] with basic foundations to figure out the robustness of outlier detection procedures [30], [31]. In addition, general descriptions of the masking swamping breakdown points were formulatedwithin a coherent framework and lemmas were created to test robustness measures in practical applications.

The literature survey has provided several background works done in the past in connection with outlier detection, with which, it can be understood that the outlier detection mechanism is more important than any other stage in data processing as it poses crucial information to be manipulated. The Methodologies followed by most of the above-mentioned

techniques have come up with the intention of completely removing the outliers from the dataset. This is acceptable only if, outliers cause the data to lose its quality, but in many situations, investigation of the outliers is necessary as they not only pose extreme values but also portray neighbourhood points as outliers. Hence, it can be concluded that, there is an inadequate idea to handle the effects of masking and swamping effects in addition to outlier detection, as the available spatial Outlier detection techniques provide little or inadequate knowledge in deciding which outlier detection technique is most suitable for the spatial data that works on spatial neighbourhood information. The Problem of minimizing the masking and swamping effects has not been properly addressed by any of the techniques. It is also noted that consideration of any visualization technique or any other traditional method is also not a better choice in this regard. So, there is a need for a proper investigation to address this paradox of determining irregular patterns in spatial-data which is apparently non-trivial.

III. SPATIAL OUTLIER ANALYSIS METHODOLOGY

Figure.1 shows a systematic workflow for the detection of spatial outliers with various stages involved in the research work. Iterative techniques [5] are performed as a binary classification exercise in parallel with semivariogram. The two techniques, the one which works on spatial neighbourhood information and the other based on visualizing the data, are compared to analyze their outlier detection abilities.

A Real and Spatial dataset containing 683 soil sample observations made at various locations within the Davangere district has been collected. The dataset is checked to see all the essential soil physicochemical properties which are helpful in optimal soil management. The statistical techniques and semivariogram are considered to detect the spatial outliers from the soil sample dataset containing soil physicochemical properties. The spatial autocorrelation among the measured data points can be illustrated using Semivariogram models. Figure 1 illustrates the use of semivariogram analysis to assess spatial autocorrelation in the data. While it does not directly estimate outliers, it provides a visual representation of spatial continuity, which is crucial for understanding the underlying spatial relationships in the data.

Iterative techniques are examined with respect to swamping and masking effects. The influence of outliers on their neighbours is examined. Iterative-R and Iterative-Z are compared in order to determine their performance. Several analytical measures are estimated as performance measure to find the classification accuracy of Iterative techniques and the ROC has been used to see the classification ability of both of these techniques. The complexity of each technique is calculated, which generally comprises computing nearest neighbours for each individual observation and then determining the aggregate neighbourhood function as well as the value for a comparison function.

A. DEFINITION AND NOTATIONS FOR THE PROBLEM

Let the set S be explicitly defined as a collection of spatial data observations. Each element of S represents a spatial data point, which is used in the subsequent analysis. So,the set-of-points $S = \{S_1, S_2, S_3 \ldots S_n\}$ be spatial-data-observation, with $f_{attr}()$ as the attribute-values of every spatial-data-point with $\delta \geq 1$, so that $f_{attr}(S_i)$ indicates attribute-value of spatial-data-point 'Si'. For each 'S_i', $NN_k(S_i)$ are k-nearest-neighbours. A Neighbourhood function$f_{aggr}(S_i)$, gives a summary-statistics of attribute-values of the spatial-data-points inside$NN_k(S_i)$. With these notations, Iterative-R and Iterative-Z are defined to detect outliers. We consider Iterative- R to deal with it, first. For each S_i, we calculate $NNk(S_i)$ and Summary-statistics of the neighbours can be determined using

$$f_{aggr}(S_i) = \frac{1}{k} \sum S \in NN_k(S_i) f_{attr}(S_i) \quad (1)$$

Equation (1) defines the summary statistics as the aggregate function of the neighborhood attribute values. This is calculated for each data point based on the attribute values of its k-nearest neighbors, providing a basis for comparison in the iterative techniques.

The comparison- function $f_{ratio}()$ is the ratio of $f_{attr}()$/ $f_{aggr}()$.

i.e, for each S_i,

$$f_{ratio}() = f_{attr}(S_i) / f_{aggr}(S_i) = y_i() \quad (2)$$

where $y_i = f^i_{ratio}(S)$ for $i = 1, 2, \ldots n$.

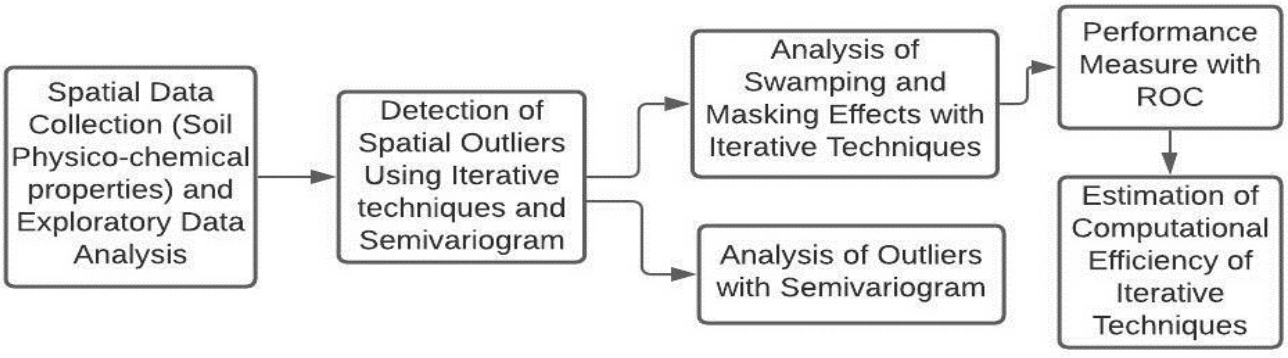

FIGURE 1. A systematic workflow for spatial outlier detection and removal.

Thus, with the given functions, i.e., $f_{attr}()$, k, $f_{aggr}()$ and $f_{ratio}()$ a spatial-observation'S_i' is an outlier, if y_i is a very small or large value of the set $\{y_1, y_2, \ldots y_n\}$. Such outlying $'y_i'$ is normalized by replacing its $f_{attr}(S_i)$ with $f_{aggr}(S_i)$.

In case of Iterative-Z, the comparison-function $f_{diff}()$ is the function of $f_{attr}()$ and $f_{aggr}()$. i.e., we consider $f_{diff}()$ as the $f_{attr}()-f_{aggr}()$.

i.e., for each S_i,

$$y_i = f_{diff}(S_i) = f_{attr}() - f_{aggr} \tag{3}$$

where, the neighbourhood function is as given in (1)

Let the 'μ' be the sample mean and 'σ' sample standard deviation of the dataset i.e., $\{f_{diff1}, f_{diff2}, f_{diff3}, \ldots\}$.

$$\mu = \frac{\sum_{i=1}^{n} S_i}{n} \tag{4}$$

$$\sigma = \sqrt{\frac{\sum (S_i - \mu)^2}{n-1}} \tag{5}$$

then, we Standardize the data set and compute the absolute value y_i for i = 1, 2, 3, ……n.

$$y_i = |f_{diff} - \frac{\text{sample mean}}{\text{standard deviation}}| \tag{6}$$

We choose $'y_i'$ i.e., max in the $\{f_{diff1}, f_{diff2}, f_{diff3}, \ldots\}$ and it will be termed as an outlier. i.e., Comparison function $f_{diff}(S_i)$ is taken to be the difference of $f_{attr}(S_i)$ and $f_{aggr}(S_i)$. For $f_{diff}(S_i)$ if the value is very large or very small then it is an indication that 'S_i' might turn out to be a Spatial outlier.

B. SPATIAL DATA OF STUDY-AREA UNDER USAGE

A Spatial data analysis eventually associates non-spatial attributes, however, it is essential to know the fertility status of soil in a specific field to apply the required intake of nutrients, and fertilizers during the management of small-scale heterogeneous sub fields for variable rate technology (VRT). For this, the soil samples are collected from various grid points all over the district of Davangere. Figure 2 shows the Pre-processed Soil data samples collected from various villages of Davangere district.

Each data observation is checked against the expected standard ranges of N, P, and K specified by the government survey to label it as ''Normal'' or ''Outlier''. A Rule based classifier indicated 577 instances as Normal and 106 instances tend to show outlier ness.

Figure 3 shows that the possible values for N are ranging from 8 to 230 kg/ha, for P, the possible values are in between 4 and 78 kg/ha and for K, the possible values ranges from 4 to 324 kg/ha.From the survey made by the state government, it is observed that for a regional crop the permissible and susceptible standard values of N, P and K ranges between 48 to 236 kg ha-1, 8 to 78 kg ha-1 9 to 312 kg ha-1 respectively. Table 2 summaries the Dataset with 683 observations containing N, P, and K values with class labels.If we check the conformity of each of the N, P, K values in our pre-processed dataset with those declared standard ranges forN,P,K provided by the district soil health care centre survey, the following observation can be noted.

- Out of 652 observations, we will get 23 nonconforming values for 'n', 77 nonconforming values for 'p', and 29 nonconforming values for 'k'.
- 629 Complying values of 'N', 575 Complying values of 'P', 623 Complying values of 'K' from the Prepared Dataset are found.

Here, outliers considered are global, since they are based on global comparisons and these observations are not pertaining to the standard ranges. However, these outlier's score cannot be accepted and this situation necessitates the detection of spatial outliers based on their local comparisons, derived from neighborhood comparison.

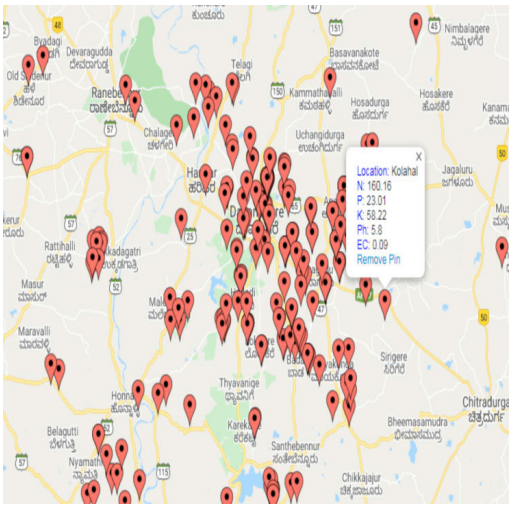

FIGURE 2. Pre-processed soil data samples from various villages of davangere district.

C. INVESTIGATION OF THE SPATIAL DATA USING ITERATIVE TECHNIQUES

The work attempts to find the outliers using the Iterative-R (ratio) technique that works in iterations to identify the expected number of outliers, so that detected outliers will not in turn impact the subsequent iteration negatively. At this point, it is noted that the expected number of outliers can be set to any smaller integer value (say 10) which will be used by the technique while determining the total number of outliers. The scoring of each data point is estimated, and this must be used to decide its class. The decision for converting the scoring of a data point into a class label is achieved with the decision threshold and the default value for this parameter is 0.5 or any scores between 0 and 1. We are performing a binary classification exercise where any score exceeding the threshold would be treated as an outlier and if not so, the score is retained within the set of the normal score. With a threshold value 'Θ' set to 1 and the total number of expected outliers 'm' set to 10, the Iterative techniques consider all those estimated scores as outliers which exceeds the threshold value, where the number of outliers to be detected using

TABLE 2. Dataset with 683 observations containing N, P, and K values with class labels.

Location	y	x	N	P	K	pH	EC	Is Outlier?
Khudapura	14.42	76.55	103.21	22.31	105.32	6.98	0.29	No
Manamainahatti	14.42	76.52	124.02	27.08	124.63	6.77	0.11	No
Mallenahalli	14.20	76.07	43.02	6.75	65.01	8.02	0.85	Yes
Saasalu	14.19	76.11	102.32	14.01	78.02	6.43	0.15	No

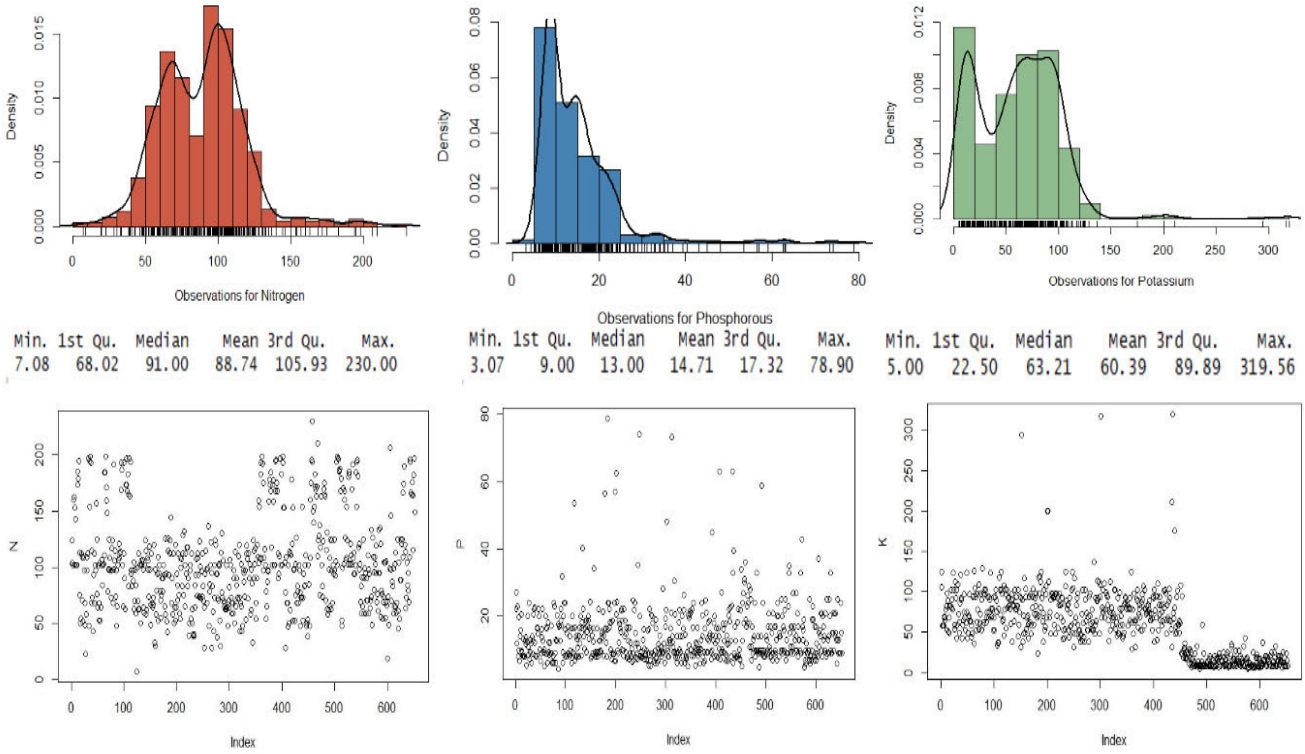

FIGURE 3. Density plots for collected N, P, K, pH and EC values.

Iterative techniques is restricted by the value specified for the ''expected number of outliers'' i.e., 'm'. In each iteration, for every single spatial data point, the ratio of a point's attribute value to the average of the aggregated attribute value of its neighbours, (r-value) as per the equation (2) is computed. The neighborhood-function $f_{aggr}(S_i)$ of the data point is calculated using equation (1), whereas value of $f_{attr}(S_i)$ is known. For every point 'Si', we compute the Comparison function $f_{ratio}(S)$ which is taken to be the ratio of $f_{attr}(S_i)$ to $f_{aggr}(S_i)$ and the point with the extreme R-value has been declared as an outlier. Then $f_{attr}(S_i)$ of such outlier is substituted with$f_{aggr}(S_i)$.

With Iterative-R 113 data points are identified as outliers as shown in Figure 4.After having both, the summarized attribute value $f_{attr}(S_i)$ and the average attribute value of the k-nearest neighbours of $'S_i'$ i.e, $f_{aggr}(S_i)$, for every point, We compute the difference between a point $'S_i$'s attribute value $f_{attr}(S_i)$ and its neighbours average attribute values $f_{aggr}(S_i)$. i.e., Comparison function $f_{diff}(S)$ is equals to $f_{attr}(S_i)$ - $f_{aggr}(S_i)$ as per the eqn (3).With Sample mean and sample standard deviation of the data set $\{f_{diff1}, f_{diff2}, f_{diff3}, \ldots\}$, we Standardize the data set and compute the absolute value using (6) for $i = 1, 2, 3, \ldots n$. Any value exhibiting extremity in this set is considered as outlier and replaced with the average attribute value of the k- nearest neighbours. As a result, we have obtained 134 data observations as extreme, i.e., outlying observations as shown in Figure 5.

D. ANALYSIS OF OUTLIERS WITH SEMIVARIOGRAM BASED ON SPATIAL AUTOCORRELATION

A semivariogram is a crucial tool in geostatistics, providing insights into the degree of spatial autocorrelation within a dataset. It measures the variance of the difference between data values as a function of the distance between data points. In the context of spatial outlier detection, the semivariogram can be used to identify points where the spatial correlation

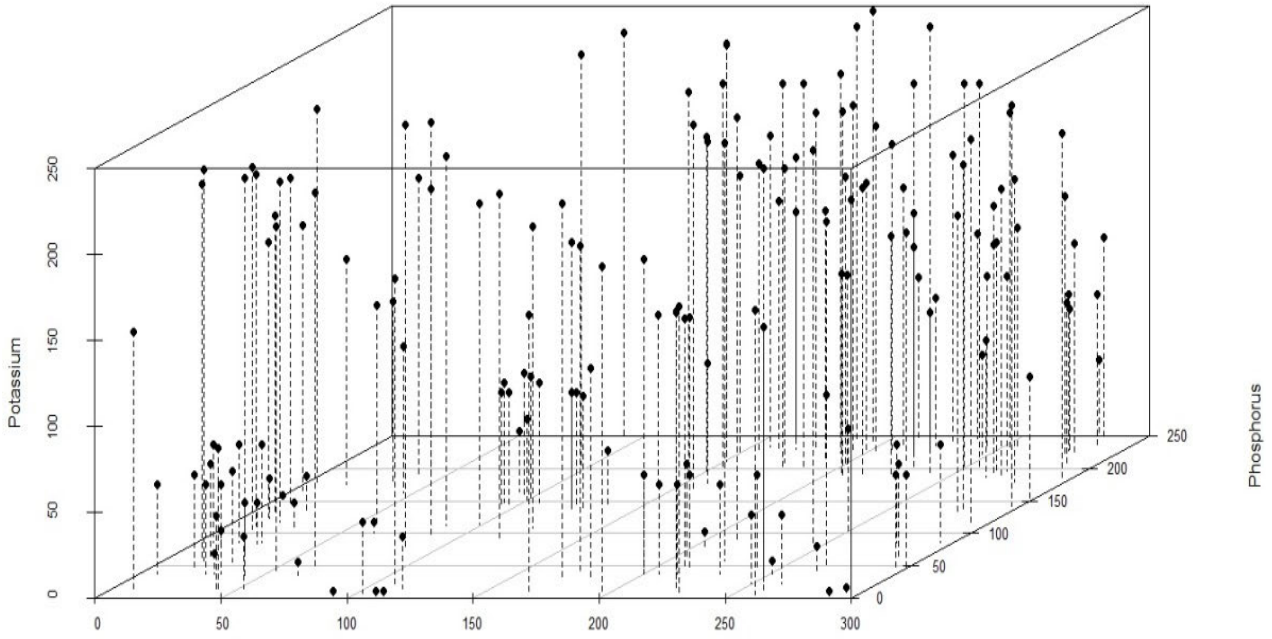

FIGURE 4. Outliers removed by iterative–R technique.

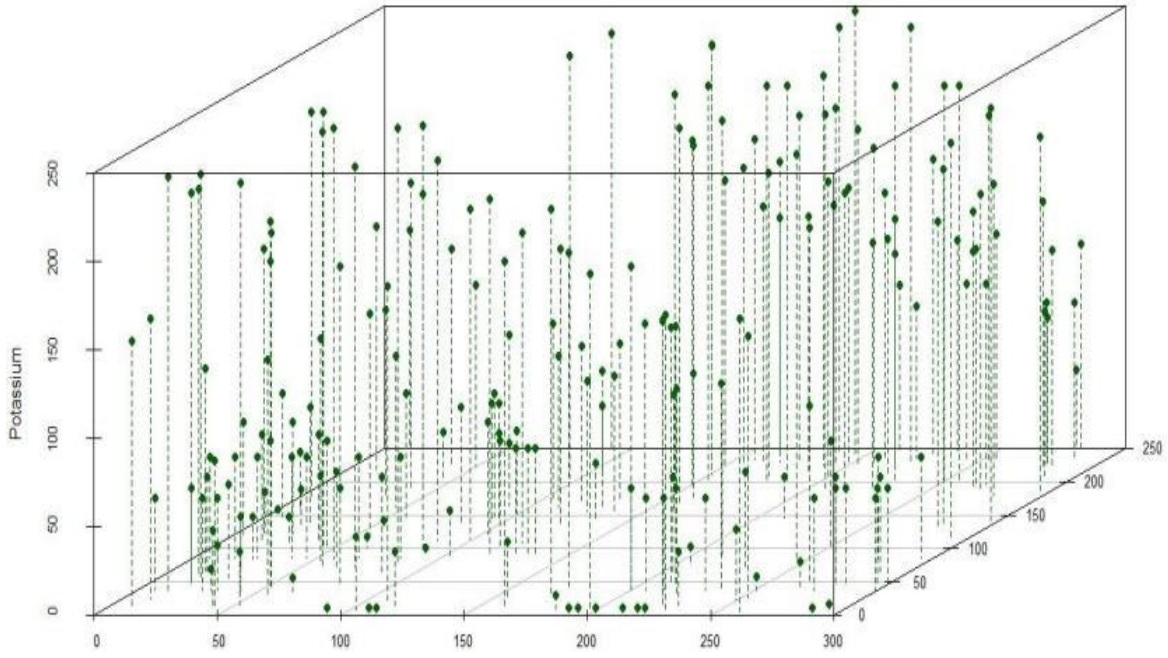

FIGURE 5. Outliers removed by iterative-Z technique.

deviates significantly from the expected pattern. Outliers may manifest as points where the semivariogram exhibits a sharp increase, indicating that these points have a relationship with their neighbors that is markedly different from the general trend. By analyzing these deviations, the semivariogram helps to visually and quantitatively assess the presence of spatial outliers. A Variogram showing the spatial continuity of the data can be used to fit a model of the spatial correlation of the observations. While semivariogram does not directly estimate outliers, it provides a visual representation of spatial continuity, which is crucial for understanding the underlying spatial relationships in the data.

1) ROLE OF THE SEMIVARIOGRAM IN SPATIAL OUTLIER DETECTION

The semivariogram provides a graphical representation that plots the variance of data point differences against the distance separating them. Key elements of the semivariogram include the nugget (representing micro-scale variations or measurement errors), the sill (the total variance when data

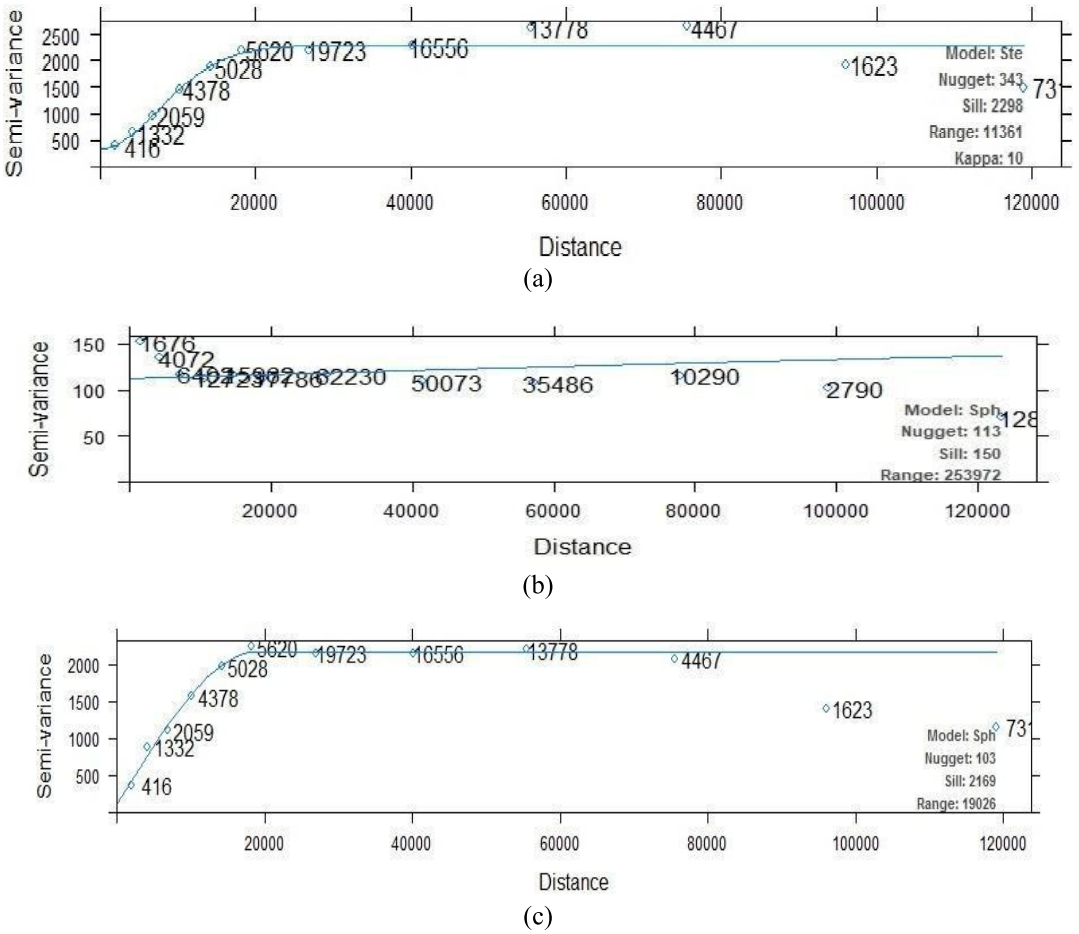

FIGURE 6. Semivariogram on (a) Nitrogen (b) Phosphorous(c) Potassium.

points become uncorrelated), and the range (the distance at which the semivariogram reaches the sill, indicating the limit of spatial correlation). In the context of spatial outlier detection, the semivariogram can highlight areas where the spatial correlation deviates from the norm. These deviations are crucial because they may indicate the presence of outliers—data points that significantly differ from their neighbors. For example, in a dataset of soil nutrient levels, an unexpected spike in the semivariogram could suggest that a particular sample's nutrient content deviates sharply from surrounding samples, flagging it as a potential outlier.

2) STATISTICAL INTERPRETATION AND JUSTIFICATION

The semivariogram analysis of soil properties such as Nitrogen (N), Phosphorus (P), and Potassium (K) is shown in Fig.6. The Figure shows the value of Sill, range and nugget for N, P and K from the dataset collected. Each semivariogram reveals different patterns of spatial correlation:

a: Nitrogen (N)

The semivariogram for Nitrogen shows a sharp increase in variance at a certain distance, followed by a plateau. This suggests that some samples differ significantly in nitrogen content compared to others at specific distances, indicating potential outliers. The presence of a nugget effect indicates minimal measurement error, reinforcing that these deviations are likely genuine outliers rather than noise.

b: Phosphorus (P)

The semivariogram for Phosphorus also indicates potential outliers, though the variance increases more gradually. This suggests that while there are outliers, they are less extreme than those found in the Nitrogen dataset.

c: Potassium (K)

The semivariogram for Potassium shows the smoothest increase in variance, implying strong spatial continuity with fewer significant outliers. The gradual curve suggests a more uniform distribution of potassium levels across the sampled locations.

The Nugget effect can be attributed to measurement errors and indicates there is no spatial continuity of the data. But this cannot be accounted to the spatial outlier detection phenomenon as the true observations may be included in the range above/below the nugget specification in each Semivariogram.

This analysis provides initial clues about the presence of spatial outliers, highlighting areas where the spatial

correlation breaks down. However, while the semivariogram indicates where these deviations occur, it does not definitively identify outliers. This is where the semivariogram must be used in conjunction with robust statistical techniques to confirm and quantify these outliers.

And, it can be stated that false positives and true negatives may not be dealt with such a graphical approach as the influence of an outlier on a particular observation can be analysed only with the help of statistical tests such as iterative techniques where the calculation of correlation among the data observations plays a vital role.

IV. RESULTS AND DISCUSSION

The Iterative techniques are performed as a Binary classification Exercise and the observations are summarized in Table 3. Figure 7 shows relevant performance measures to understand the ability of iterative techniques. The Iterative-Z as a binary classifier with less error rate and more classification accuracy leaves an impression that it performs better than the Iterative-R technique.

TABLE 3. Observation summary.

Normal Instances			Outliers		Total Instances	
Iterative – R	Iterative - Z	Actual	Iterative – R	Iterative – Z	Actual	683
570	549	577	113	134	106	

A. INFERENCES FROM THE ANALYSIS OF SWAMPING AND MASKING EFFECTS

The robustness of the iterative techniques against masking and swamping effects is assessed by calculating the false positive rate (FPR) and false negative rate (FNR). FPR measures the proportion of non-outliers incorrectly classified as outliers (swamping), while FNR measures the proportion of true outliers that are incorrectly classified as non-outliers (masking). Lower values of FPR and FNR indicate higher robustness.

The breakdown point of an outlier detection method is a critical measure of its robustness. It is defined as the point at which the method fails to correctly classify outliers and non-outliers as the number of outliers in the dataset increases. This is evaluated by progressively adding outliers to the dataset and observing the method's performance, specifically through changes in the false positive and false negative rates.

The Iterative techniques are examined from the perspective of masking and swamping effects. We calculate the total percentage of swamping effect incurred by iterative techniques which is the False Positive Rate corresponds to the proportion of negative data points that are mistakenly considered as positive, with respect to all negative data points.

- With respect to Iterative-R $\rightarrow$ FPR = 9.201%
- With respect to Iterative-Z $\rightarrow$ FPR = 9.688%

On the other hand, total percentage of masking effect incurred by Iterative techniques is the False Negative Rate corresponds to the proportion of positive data points that are ignored as negatives, with respect to all positive data points.

- With respect to Iterative-R $\rightarrow$ FNR = 43.925 %
- With respect to Iterative-Z $\rightarrow$ FNR = 22.857 %

It can be concluded that an iterative technique which exhibits minimum masking and swamping effect is a better statistical approach to reduce the swamping and masking effect. The percentage of swamping and masking effects scored by Iterative techniques indicates that 9.2% non-outlying observations with respect to Iterative-R and 9.7% non-outlying observations with respect to Iterative-Z, are considered (mistaken) as outliers. This means that the percentage of swamping effect in case of Iterative-Z has increased up to 0.5% than that of Iterative-R. This exaggeration still can be accepted. Hence, both the techniques perform almost equally well in minimizing the swamping effect. In case of masking effect, 43.92%, true-outliers by Iterative-R and 22.85% true-outliers by Iterative-Z are considered as non-outliers due to the presence of nearby true-outliers. Hence, the masking effect in case of Iterative-R is relatively higher than that of Iterative-Z which leaves the impression that the latter is better in minimizing the masking effect.

B. PERFORMANCE ASSESSMENT OF ITERATIVE TECHNIQUES

A particular observation is probed to check whether it is entangled in masking and swamping effects due to the presence of true outliers. Both the Iterative techniques performed as binary classifiers can form two types of errors: They can label an instance as ''normal'' who defaults to an abnormal category or vice versa. It is an obvious case of interest to know the root cause of this consequence that emerges these types of errors. So, we prefer analytical tools which are convenient ways to display this information. Table 4 and 5 shows the instances classified and labeled using iterative - R Technique.Table 6 shows the Prediction Summary of the Iterative-R and Iterative-Z as a Binary Classifier

The proposed iterative-R and iterative-Z techniques are treated as binary classification exercises which predict whether a given input sample is an outlier, i.e., any sample from the given soil dataset belongs to one of the two classes: i.e., OUTLIER (YES) or NON-OUTLIER (NO). A Confusion matrix for our techniques on 683 samples of given dataset has been formed from the investigation of both the techniques. Table 7 and 8 show the Values for the Confusion Matrix with Iterative-R and Iterative-Z.

C. OUTLIER ANALYSIS OF MASKING AND SWAMPING EFFECTS

The observations which are misclassified as outliers (False Positives) due to the influence of surrounding true outliers, have been shown in Figure8, which portrays 53 non-outlying

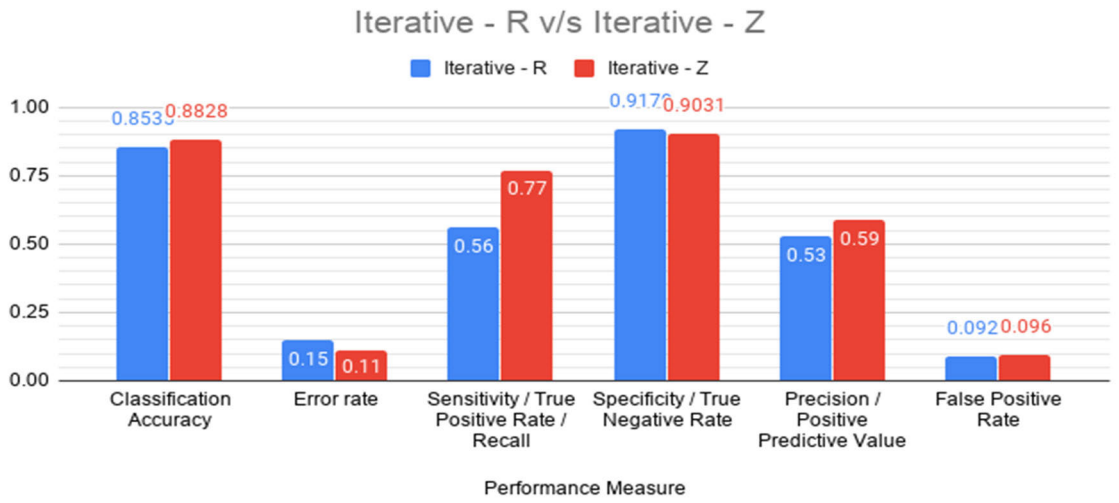

FIGURE 7. Summary of performance measures for iterative techniques.

TABLE 4. Instances classified and labeled using iterative - R and iterative - Z technique.

Location	y	x	N	P	K	pH	EC	Is Outlier ?	Predicted by Iterative-R	Match
Khudapura	14.4212	76.55	103.21	22.31	105.32	6.98	0.29	No	No	✓
Kolahal	14.33	76.18	160.16	23.01	58.22	5.8	0.09	No	Yes	X
Mallenahalli	14.2063	76.07	43.02	6.75	65.01	8.02	0.85	Yes	Yes	✓

TABLE 5. Instances classified and labeled using iterative-Z technique.

Location	y	x	N	P	K	pH	EC	Is Outlier?	Predicted by Iterative-Z	Match
Khudapura	14.4212	76.55	103.21	22.31	105.32	6.98	0.29	No	Yes	X
Kolahal	14.23	76.11	62.01	8.15	58.32	8.14	0.36	No	No	✓
Mallenahalli	14.20	76.07	43.02	6.75	65.01	8.02	0.85	Yes	Yes	✓

TABLE 6. Prediction summary of the iterative-R and iterative-Z as a binary classifier.

Instances	Iterative-R		Iterative-Z	
Correctly Classified	630	92.24 %	639	93.55 %
Incorrectly Classified	053	07.75 %	044	06.44%
Total	683			

observations are considered as outliers due to the presence of 60 true Outliers in their surroundings.

TABLE 7. Values for the confusion matrix with iterative-R.

N = 683	Predicted: NO	Predicted: YES
Actual: NO	TN = 523	FP = 53
Actual: YES	FN = 47	TP = 60

TABLE 8. Values for the confusion matrix with iterative-Z.

N = 683	Predicted: NO	Predicted: YES
Actual: NO	TN = 522	FP = 56
Actual: YES	FN = 24	TP = 81

The Figure 9 shows 56 non-outlying observations are considered as Outliers due to the presence of 81 actual Outliers in their surroundings.Figure 10 shows the actual outliers which are classified as Non-Outliers (False Negatives) due to the influence of surrounding true outliers (True Positives). Here, 47 True-Outlying observations are considered as Non-Outliers due to the influence of surrounding 60 True Outliers.Similarly, Figure 11 Shows 24 True-Outlying observations are considered as non-outliers due to the influence of surrounding 81 True Outliers.

D. PERFORMANCE ASSESSMENT USING ROC

The Area Under the Curve (AUC) is a critical metric used in ROC (Receiver Operating Characteristic) analysis to evaluate the performance of binary classification models, including outlier detection methods. In this context, the semivariogram approach, while primarily a tool for assessing spatial autocorrelation, can also be used as a comparative graphical method for identifying potential outliers, but, as stated earlier, since the semivariogram is not directly designed to classify outliers, creating a ROC curve for it is more conceptual.

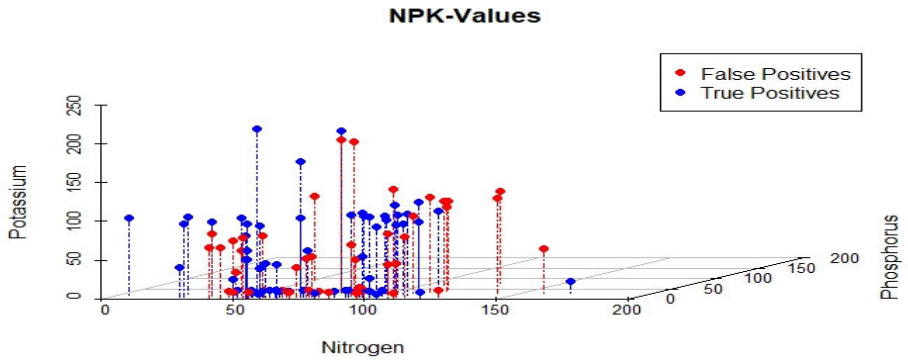

FIGURE 8. Swamping effect with iterative-R.

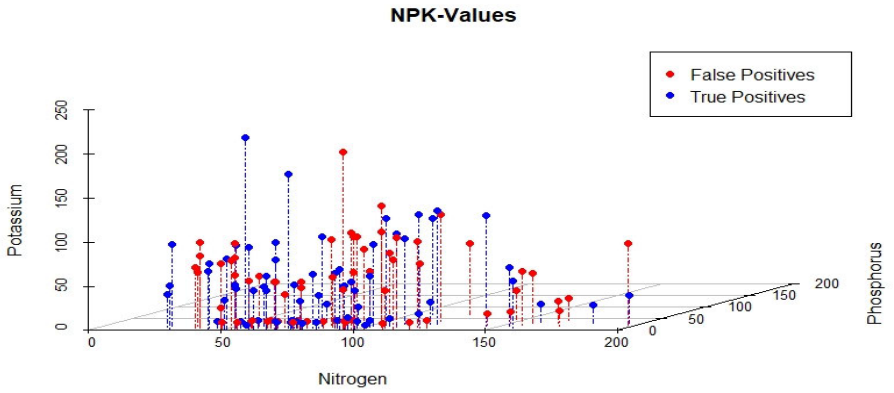

FIGURE 9. Swamping effect with iterative-Z.

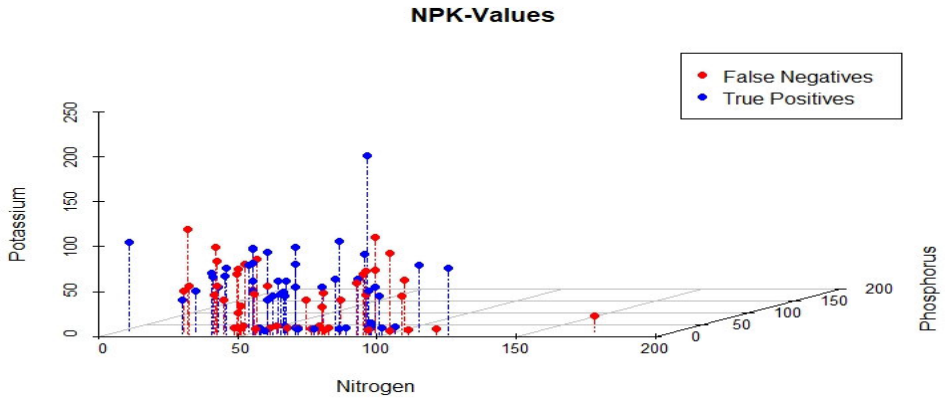

FIGURE 10. Masking effect with iterative-R.

In the semivariogram approach, the AUC value represents the method's ability to distinguish between true spatial outliers and non-outliers based on deviations in spatial autocorrelation. Since the semivariogram is not a direct outlier detection technique but rather a diagnostic tool, its AUC value is typically lower than those of specialized statistical methods like Iterative-R and Iterative-Z.

Based on the analysis, the AUC for the semivariogram approach is approximately 0.65. This value indicates moderate performance in distinguishing between outliers and non-outliers, reflecting the semivariogram's role as a preliminary tool for identifying areas that require more in-depth statistical analysis.

The Iterative-R and Iterative-Z techniques, which are designed specifically for outlier detection, achieve higher AUC values of 0.743 and 0.821, respectively. These values demonstrate the superior performance of these methods in accurately identifying spatial outliers.

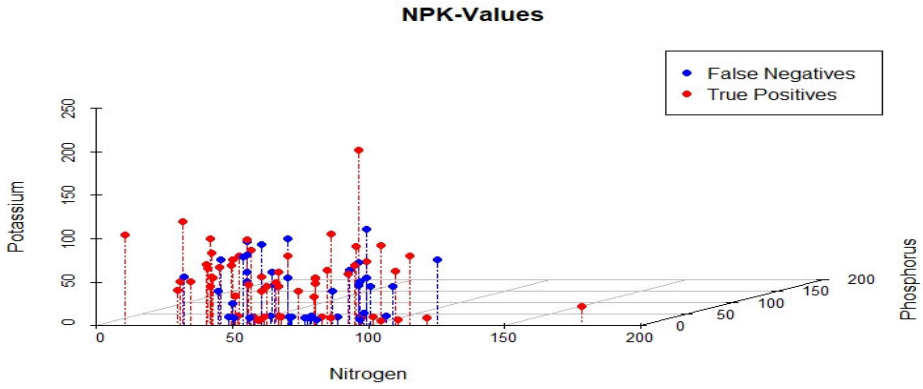

FIGURE 11. Masking effect with iterative-Z.

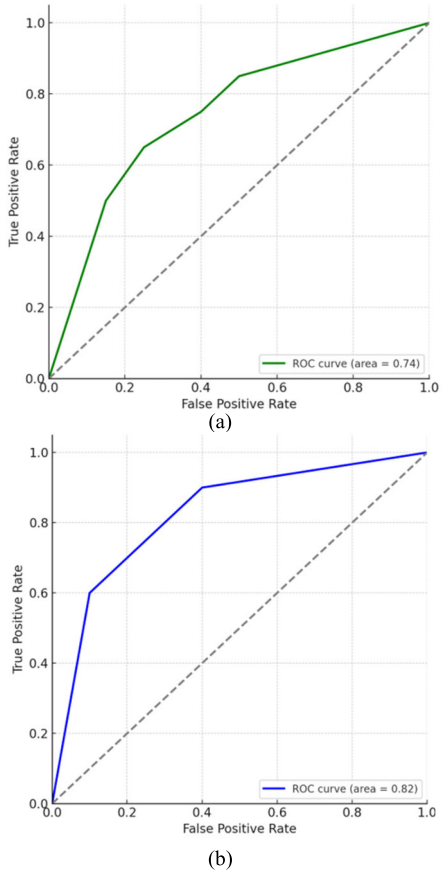

FIGURE 12. ROC curve for (a) iterative-R and (b) iterative-Z.

The diagnostic potential of binary classifiers is illustrated using the Receiver Operator characteristic (ROC) curve. With respect to the ROC curve for the Iterative techniques shown in Figure 12, the Area under Curve for the Iterative-R and Iterative-Z is 0.743 and 0.821 respectively.

The Area Under Curve (AUC) is proportional to the probability that a classifier would rank a positive instance selected randomly higher than a negative one randomly selected (claiming that 'positive' ranks higher than 'negative'). AUC with value 0.7, indicates that there is a 70% chance that the model will be able to identify between positive class and negative class. With this, it can be concluded that Iterative-Z as a classifier can distinguish positive class observations from the negatives better than Iterative-R. The value for the threshold equal to 1 or 1.1 can be considered as an optimal or decision threshold for both the Iterative techniques.

E. ESTIMATION OF COMPUTATIONAL EFFICIENCY OF ITERATIVE TECHNIQUES

The computational complexity of iterative techniques is largely determined by the k-Nearest Neighbors (kNN) query, which is a core operation in both the Iterative-R and Iterative-Z techniques. The kNN query involves finding the k closest neighbors for each data point, which is crucial for calculating the aggregate neighborhood function.

Step 1 (k-Nearest Neighbors Calculation):

- For each data point, the algorithm computes the distance to all other data points to identify the k-nearest neighbors. This involves looping through all n data points, with each loop requiring a comparison across k neighbors.
- Given that the data points have d dimensions, the time complexity for calculating the k-nearest neighbors for a single data point is O(n * d). Since this must be done for each of the n data points, the overall complexity for this step is O(n^2 * d).

Step 2 (Aggregate Neighborhood Function Calculation):

- Once the k-nearest neighbors are identified, the next step is to calculate the aggregate neighborhood function for each data point. This step has a complexity of O(n * k * d), as it involves aggregating the k-nearest neighbors' attributes for each data point.

Step 3 (Iterative Updates):

- The iterative process, which refines the outlier detection through repeated calculations, typically involves m iterations. During each iteration, the entire process (kNN calculation and neighborhood function aggregation) is repeated.

- The complexity of each iteration involves re-evaluating k-nearest neighbors and updating the aggregate function. Thus, for m iterations, this adds a complexity of O(m * (n^2 * d + n * k * d)).

1) OVERALL COMPUTATIONAL COMPLEXITY

- Combining these steps, the overall computational complexity of the iterative technique is estimated as: O(m*(n^2*d+n*k*d)).
- For large datasets, where n is significantly larger than k or m, the term O(n^2*d) dominates, leading to a complexity close to O(m*n^2*d). This indicates that the computational cost grows quadratically with the number of data points, linearly with the number of dimensions, and linearly with the number of iterations.
- The computational complexity of the Iterative-R and Iterative-Z techniques has been carefully calculated, considering the key operations of k-nearest neighbors determination, aggregate neighborhood function calculation, and iterative updates. Specifically, the overall complexity is dominated by the kNN calculations and is expressed as O(m*n^2*d), where n is the number of data points, k is the number of neighbors, d is the dimensionality, and m is the number of iterations.

2) KEY POINTS CONSIDERED

- Quadratic Complexity: The complexity grows quadratically with the number of data points (n), which is typical in kNN-based algorithms.
- Dimensionality (d): The computational cost also increases linearly with the number of dimensions (d).
- Iterations (m): The number of iterations (m) directly impacts the overall computational time, which is important for the iterative techniques.

V. CONCLUSION AND FUTURE WORK

The research presented in this paper underscores the critical importance of identifying and mitigating masking and swamping effects in spatial outlier detection, particularly within the highly variable context of Precision Agriculture. Accurate soil data analysis is essential for informed decision-making, and the detection of spatial outliers plays a pivotal role in refining this process. By employing advanced statistical techniques that leverage neighborhood relationships, this study has successfully developed methods to detect and eliminate outliers, thereby enhancing the quality of the spatial data used in agricultural applications. While the semivariogram provides valuable insights into spatial autocorrelation, particularly highlighting the presence of measurement errors through the nugget effect, it falls short in effectively identifying outliers. The nugget indicates areas where spatial continuity is absent, yet this method does not fully account for the presence of outliers, as non-outlier instances may still fall within acceptable ranges above the nugget threshold. Moreover, the semivariogram approach is limited in its capacity to address swamping and masking effects, which are best analyzed through the correlation calculations inherent in statistical iterative techniques.

This study has rigorously evaluated the performance of the Iterative-R and Iterative-Z techniques through binary classification, focusing on their effectiveness in reducing swamping and masking. Our findings reveal that the Iterative-Z technique significantly outperforms Iterative-R in minimizing the masking effect, as evidenced by a False Negative Rate (FNR) of 22.85% compared to 43.92% for Iterative-R. Additionally, both methods demonstrated similar efficacy in reducing the swamping effect, with False Positive Rates (FPR) of 9.20% for Iterative-R and 9.69% for Iterative-Z. The robustness of the Iterative-Z technique is further validated by its superior Area Under the Curve (AUC) value of 0.821, compared to 0.743 for Iterative-R, confirming its enhanced capability in accurately detecting spatial outliers. These results not only affirm the effectiveness of the Iterative-Z method as a reliable tool for optimizing agricultural practices but also highlight the broader implications for Precision Agriculture. The ability to accurately identify and manage spatial outliers ensures that resources are allocated efficiently, leading to better crop management and higher yields. As a promising avenue for future research, the Iterative techniques could be further refined to explore the dependency between sequence values generated by the comparison functions, potentially leading to even more precise outlier detection methods. This ongoing enhancement of spatial data analysis tools will continue to drive advancements in agricultural decision-making and sustainability.

REFERENCES

[1] A. Brenning, H. Piotraschke, and P. Leithold, "Geostatistical analysis of on-farm trials in precision agriculture," in *Proc. 8th Int. Geostatistics Congr.*, vol. 2, J. M. Ortiz and X. Emery, Eds., Santiago, Chile, Dec. 2008, pp. 1131–1136.
[2] S. Norng, "Statistical decisions in optimising grain yield," Ph.D. dissertation, Dept. Faculty of Science, Queensland Univ. Technol., Brisbane, QLD, Australia, 2004.
[3] B. Basso, D. Cammarano, and E. Carfagna, "Review of crop yield forecasting methods and early warning systems," in *Proc. First Meeting Sci. Advisory Committee Global Strategy Improve Agricult. Rural Statist.*, vol. 241. Rome, Italy: FAO, 2013, pp. 1–56.
[4] D. Chen, C.-T. Lu, Y. Kou, and F. Chen, "On detecting spatial outliers," *GeoInformatica*, vol. 12, no. 4, pp. 455–475, Dec. 2008.
[5] C.-T. Lu, D. Chen, and Y. Kou, "Algorithms for spatial outlier detection," in *Proc. 3rd IEEE Int. Conf. Data Mining*, Nov. 2003, pp. 597–600.
[6] R. Serfling and S. Wang, "General foundations for studying masking and swamping robustness of outlier identifiers," *Stat. Methodol.*, vol. 20, pp. 79–90, Sep. 2014.
[7] P. C. Su, "Statistical geocomputing: Spatial outlier detection in precision agriculture," Master thesis, Master Environ. Stud. Geography, Univ. Waterloo, Waterloo, ON, Canada, 2011.
[8] S. Babu, "A software model for precision agriculture for small and marginal farmers," in *Proc. IEEE Global Humanitarian Technol. Conf., South Asia Satell. (GHTC-SAS)*, Aug. 2013, pp. 352–355.
[9] R. Gebbers and S. De Bruin, "Application of geostatistical simulation in precision agriculture," in *Geostatistical Applications for Precision Agriculture*. New York, NY, USA: Springer, 2010, pp. 269–303.
[10] H. H. Bosman, G. Iacca, A. Tejada, H. J. Wörtche, and A. Liotta, "Spatial anomaly detection in sensor networks using neighborhood information," *Inf. Fusion*, vol. 33, pp. 41–56, Jan. 2017.
[11] G. Zheng, S. L. Brantley, T. Lauvaux, and Z. Li, "Contextual spatial outlier detection with metric learning," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2017, pp. 2161–2170.

[12] A. T. A. La Ode Saidi, "Outlier detection in multivariate linear models using Lagrange multipliers," *Global J. Pure Appl. Math.*, vol. 13, no. 6, pp. 2563–2578, 2017.
[13] J. Majumdar, S. Naraseeyappa, and S. Ankalaki, "Analysis of agriculture data using data mining techniques: Application of big data," *J. Big Data*, vol. 4, no. 1, p. 20, Dec. 2017.
[14] A. Smiti, "A critical overview of outlier detection methods," *Comput. Sci. Rev.*, vol. 38, Nov. 2020, Art. no. 100306.
[15] A. Dik, K. Jebari, A. Bouroumi, and A. Ettouhami, "Similarity- based approach for outlier detection," 2014, *arXiv:1411.6850*.
[16] F. Angiulli, S. Basta, S. Lodi, and C. Sartori, "Reducing distance computations for distance-based outliers," *Expert Syst. Appl.*, vol. 147, Jun. 2020, Art. no. 113215.
[17] S. N. Thennadil, M. Dewar, C. Herdsman, A. Nordon, and E. Becker, "Automated weighted outlier detection technique for multivariate data," *Control Eng. Pract.*, vol. 70, pp. 40–49, Jan. 2018.
[18] A. Sajid, R. Rudra, and G. Parkin, "Systematic evaluation of Kriging and inverse distance weighting methods for spatial analysis of soil bulk density," *Can. Biosyst. Eng.*, vol. 55, no. 1, pp. 1.1–1.13, Dec. 2013.
[19] Z. Li, Z. Li, N. Yu, and S. Wen, "Locality-based visual outlier detection algorithm for time series," *Secur. Commun. Netw.*, vol. 2017, no. 1, 2017, Art. no. 1869787.
[20] N. Kudnar and M. Rajashekhar, "Using geo-spatial technologies for land and water resource development planning: A case study of Tirora tehsil, India," in *Emerging Technologies for Water Supply, Conservation and Management*. Cham, Switzerland: Springer, 2023, pp. 315–331.
[21] S. Zandi, A. Ghobakhlou, and P. Sallis, "Evaluation of spatial interpolation techniques for mapping soil pH," in *Proc. 19th Int. Congr. Model. Simul.* F. Chan, D. Marinova, and R. S. Anderssen, Eds. Perth, WA, Australia, Dec. 2011, doi: 10.36334/modsim.2011.c2.zandi.
[22] C. G. Karydas, I. Z. Gitas, E. Koutsogiannaki, N. Lydakis-Simantiris, and G. N. Silleos, "Evaluation of spatial interpolation techniques for mapping agricultural topsoil properties in Crete," *EARSeL eProc.*, vol. 8, no. 1, pp. 26–39, 2009.
[23] S. Shekhar, R. R. Vatsavai, and M. Celik, "Spatial and spatiotemporal data mining: Recent advances," in *Next Generation of Data Mining*. New York, NY, USA: CRC Press, 2008, pp. 573–608.
[24] X. Liu, F. Chen, and C.-T. Lu, "Robust prediction and outlier detection for spatial datasets," in *Proc. IEEE 12th Int. Conf. Data Mining*, Dec. 2012, pp. 469–478.
[25] H. A. Issad, R. Aoudjit, and J. J. P. C. Rodrigues, "A comprehensive review of data mining techniques in smart agriculture," *Eng. Agricult., Environ. Food*, vol. 12, no. 4, pp. 511–525, Oct. 2019.
[26] D. Gozdowski, S. Samborski, and E. S. Dobers, "Evaluation of methods for the detection of spatial outliers in the yield data of winter wheat," *Colloq. Biometricum*, vol. 40, pp. 41–51, Jan. 2010.
[27] G. Ruß, "From spatial data mining in precision agriculture to environmental data mining," in *Computational Intelligence in Intelligent Data Analysis*. Berlin, Germany: Springer, 2013, pp. 263–273.
[28] J. John, "Outlier detection and spatial analysis algorithms," 2021, *arXiv:2106.10669*.
[29] M. A. Samara, I. Bennis, A. Abouaissa, and P. Lorenz, "A survey of outlier detection techniques in IoT: Review and classification," *J. Sensor Actuator Netw.*, vol. 11, no. 1, p. 4, Jan. 2022.
[30] J. T. C. Wu and J. Tian, "Spatio-temporal outlier detection: A survey of methods," *Int. J. Frontiers Eng. Technol.*, vol. 2, no. 1, pp. 1–15, Apr. 2020.
[31] H. T. Nguyen and N. H. Thai, "Temporal and spatial outlier detection in wireless sensor networks," *ETRI J.*, vol. 41, no. 4, pp. 437–451, Aug. 2019.

ARUN KUMAR HIREMATH received the Ph.D. degree from VTU, Belagavi, Karnataka, in 2022. He is currently an Associate Professor with the Department of CSE, BIET, Davanagere, Karnataka. He has 12 years of teaching experience with eight years of research exposure in the fields of data mining, machine learning, and precision agriculture. He published many journals and book chapters in reputed publications.

K. R. NAVEEN KUMAR is currently an Associate Professor with the Department of Computer Science and Engineering, Bapuji Institute of Engineering and Technology (BIET), Davanagere. He has 12 years of teaching experience and one year of industrial experience as a Software Developer with CloudByte Tech. Pvt. Ltd., Bengaluru. He is passionate about R, Python, and JavaScript languages. He has published his research articles in many journals and international conferences. His prime areas of research interests include data mining, precision agriculture, natural language processing, and data analytics.

MANJUNATHA HIREKERI MALLESHAPPA received the B.E. degree in electrical and electronics engineering, the M.Tech. degree in computer application in industrial drives, and the Ph.D. degree from Visvesvaraya Technological University, Belagavi, Karnataka, India, in 2009, 2012, and 2022, respectively. He is currently an Assistant Professor with the Bapuji Institute of Engineering and Technology, Davanagere, affiliated with Visvesvaraya Technological University. His main research interests include power system operation and control, distribution automation, artificial intelligence tools applied to smart grids, and energy trading.

BHASKAR AWADHIYA received the B.E. degree in electronics and communication engineering from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, India, in 2011, the M.Tech. degree in microelectronics from Manipal Institute of Technology, Manipal, Udupi, Karnataka, India, in 2014, and the Ph.D. degree from the Pandit Dwarka Prasad Mishra Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, India, in 2021. He is currently an Assistant Professor with the Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education. His current research interests include steep subthreshold slope (SS) devices and low-power devices.

YASHWANTH NANJAPPA (Senior Member, IEEE) received the B.E. degree in electronics and communication engineering from Visvesvaraya Technological University, Belagavi, in 2010, the M.Tech. degree in signal processing and VLSI from Jain University, Bengaluru, in 2012, and the Ph.D. degree from Visvesvaraya Technological University, in 2020. He is currently the Assistant Director (FDW) and an Associate Professor with the Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, where he is actively involved in numerous administrative and academic roles. His primary research interests include wireless sensor networks, communication systems, antennas, VLSI technology, and cyber security.