IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 15 January 2024, accepted 18 January 2024, date of publication 22 January 2024, date of current version 26 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3356731

RESEARCH ARTICLE

Outlier Detection Performance of a Modified Z-Score Method in Time-Series RSS Observation With Hybrid Scale Estimators

ABDULMALIK SHEHU YARO[1,2], FILIP MALY[1], PAVEL PRAZAK[1], AND KAREL MALÝ[1]

[1]Department of Informatics and Quantitative Methods, Faculty of Informatics and Management, University of Hradec Králové, 500 03 Hradec Králové, Czech Republic

[2]Department of Electronics and Telecommunications Engineering, Faculty of Engineering, Ahmadu Bello University, Zaria 810106, Nigeria

Corresponding author: Abdulmalik Shehu Yaro (abdulmalik.yaro@uhk.cz)

This study is supported by the Excellence project run at the Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic.

ABSTRACT The modified Z-score (mZ-score) method has been used to detect outliers in time series received signal strength (RSS) observations. Its performance is dependent on the scale estimator used, and each has advantages and disadvantages over the others. One approach to developing a scale estimator that combines the advantages of two or more scale estimators is through scale estimator hybridization. In this paper, the outlier detection performance of a mZ-score method with different hybridization approaches for Sn and median absolute deviation (MAD) scale estimators is determined and analysed. Three different hybrid scale estimators are identified, namely weighted, maximum, and average hybrid scale estimators. The performance of the mZ-score method using the three different hybrid scale estimators is determined using three experimentally generated and publicly available time-series RSS datasets. Based on the simulation results, the weighted hybrid scale estimator results in the best outlier detection performance amongst the three hybrid scale estimators. When compared to the mean-shift-based outlier detection (MOD) technique, the k-means clustering-based technique, and the density-based spatial clustering (DBSCAN) technique, the mZ-score method with the weighted hybrid scale estimator performs better with little or no false alarm and false negative detections.

INDEX TERMS Average, hybrid scale estimator, MAD, maximum, mZ-score, outlier, Sn, weighted.

I. INTRODUCTION

Fingerprint-based localization, which uses received signal strength (RSS) measurements obtained from spatially deployed wireless access points (APs), is an emerging technology that has found applications in a variety of areas, including indoor localization [1], [2]. Due to the dynamic nature of the indoor environment, such as the presence and absence of crowds and furniture, as well as variations in temporal and ambient conditions, the signals from the wireless APs fluctuate, causing the RSS measurements to fluctuate, resulting in poor localization accuracy [3], [4]. To deal with the RSS fluctuations, several research works have suggested collecting several RSS measurement observations over a time period and finding the mean average of these observations using techniques such as the mean-averaging filter, moving average filter, median filter, or Kalman filter [5], [6], [7]. During the collection of the time-series RSS observations, it is possible to have RSS outliers [5]. RSS outliers are RSS observations that appear to differ significantly from the rest of the observations. The presence of outliers in time-series RSS observations has a significant impact on mean average determination. Techniques such as the Z-score method, interquartile range (IQR), moving average, moving median,

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamed M. A. Moustafa.

and density-based techniques have been used to detect and remove outliers [8], [9], [10], [11], [12], [13]. There is no best outlier detection method, as each method has its own set of advantages and disadvantages over the others. Furthermore, the performance of each method is dependent on the statistical distribution of the time-series RSS observation. This paper considers the use of the Z-score method for outlier detection due to its ease of implementation, reliability, and robustness to outliers. It can be used on a wide range of datasets for outlier detection and can detect outliers in large datasets quickly and efficiently [8], [10], [14], [15].

The outlier detection performance of the Z-score method depends on the scale estimator used [10], [14]. The scale estimator is used to measure the deviation of each RSS observation from the mean or median of the entire time-series RSS observation. The Z-score method, by default, uses the standard deviation (SD) as its scale estimator. To improve the robustness and efficiency of the Z-score method for outlier detection, several methods have been proposed. One of these methods is replacing the default SD scale estimator with more robust estimators like the Sn and median absolute deviation (MAD) [14], [16]. The Z-score method that is based on either a Sn or MAD scale estimator is referred to as the modified Z-score (mZ-score) method. Another method to improve the performance of the Z-score method is by using the winsorization method [17], [18]. The winsorization method involves removing extreme RSS observations to reduce their influence on the SD scale estimator. The bootstrap method is another method used to improve the outlier detection performance of the Z-score method [19]. It involves resampling the time-series observation to estimate the variation in the scale estimator. Another method known as the trimmed mean method is used to improve the performance of the Z-score method [11], [20]. The trimmed mean method involves removing a certain percentage of extreme RSS values from both sides of the time-series observation and then calculating the mean of the remaining RSS observations.

Another approach to improving the outlier detection performance of the Z-score method that has yet to be investigated is the use of scale estimator hybridization. Scale estimator hybridization is a technique that combines the advantages of two or more different scale estimators to create a single, more robust, and more flexible scale estimator. A scale estimator based on scale estimator hybridization is proposed in this paper, and the outlier detection performance of the mZ-score method with the proposed hybrid scale estimator is determined and evaluated. For hybridization, the Sn and MAD scale estimators are considered. This paper makes three important contributions: (a) it identifies the various hybridization methods for Sn and MAD scale estimators; (b) it evaluates the hybrid scale estimators for outlier detection using the mZ-score method; and (c) it determines the best hybrid scale estimator for use with the mZ-score method, taking into account different time-series RSS datasets.

The remainder of the paper is organized as follows: Section II provides an overview of outlier detection using the Z-score method. Section III presents the mZ-score method based on the hybridization of the Sn and MAD scale estimators, and Section IV presents the simulation results and discussion. Section V presents the conclusion.

II. OVERVIEW OF OUTLIER DETECTION USING Z-SCORE METHOD

This section describes in detail the Z-score method for RSS outlier detection and its modifications. Detecting RSS outliers in any given time-series RSS dataset using the Z-score method requires determining the z-score value of each RSS observation. The z-score value of an RSS observation is determined by first finding the difference between the RSS observation and the mean of the entire time-series RSS dataset. The result is then divided by the SD of the entire time-series dataset to obtain the z-score value. An RSS observation that has a z-score value greater than a pre-defined threshold is considered an outlier [8], [10], [16].

Mathematically, let $\mathbf{rss}_i$ represent a time-series RSS dataset of N RSS observations collected from the i-th wireless AP.

$$\mathbf{rss}_i = [rss\,(1)\,, \cdots rss(N)] \quad \text{for } 1 \leq n \leq N \tag{1}$$

where $rss(n)$ is the n-th RSS observation.

The z-score value of the n-th RSS observation using the SD as the scale estimator is calculated as follows [10], [12], [21]:

$$Z_{SD}\,(rss\,(n)) = \frac{rss\,(n) - \mu_{\mathbf{rss}_i}}{\sigma_{SD}} \tag{2}$$

where "$\mu_{\mathbf{rss}_{i,}}$" and "σ_{SD}" are the mean and SD of the time-series RSS dataset in (1), respectively.

A z-score value of 0 in (2) indicates that the RSS observation has the same value as the mean of the entire time-series RSS dataset in (1). A positive z-score value indicates that the RSS observation has a value that is above the mean value of the entire time-series RSS dataset. A negative z-score value indicates that it is below the mean. If the z-score value of an RSS observation is above a predefined threshold, it is considered an outlier. The commonly used z-score threshold value is ±3.0 [15].

The Z-score method based on (2), which uses SD as a scale estimator, is very sensitive to RSS outliers. Furthermore, the use of SD as a scale estimator is under the assumption that the time-series RSS dataset has a statistically normal distribution. This is not always the case for experimentally generated datasets. To overcome these limitations, scale estimators that are more robust and less sensitive to outliers are used, resulting in the mZ-score method. These scale estimators are the MAD and Sn scale estimators [10], [12]. A detailed description of the mZ-score method with the MAD and Sn scale estimators is presented in the following subsections.

A. MODIFIED Z-SCORE METHOD WITH MAD SCALE ESTIMATOR

As previously stated, the SD scale estimator is extremely sensitive to RSS outliers. The MAD scale estimator, on the other hand, is less sensitive and more robust to outliers when compared to the SD scale estimator. The MAD scale estimator calculates the scaling value of the mZ-score method by taking the median of the absolute differences between each RSS observation and the median of all RSS observations in the time-series dataset. This makes it more resilient to RSS outliers. Some statistical properties of the MAD scale estimator include 37% Gaussian efficiency and a 50% breakdown point [14].

Given the time-series RSS dataset in (1), the scale value of an RSS observation using the MAD scale estimator is calculated mathematically as [14] and [16]:

$$\sigma_{MAD}\left(rss(n)\right) = B_n \times Mdn\left\{\left|rss(n) - Mdn\left\{\mathbf{rss_i}\right\}\right|\right\} \tag{3}$$

where B_n is a constant known as the scaling factor. A common choice for $B_n = 1.4826$. This makes the MAD scale estimator's performance consistent with the SD scale estimator for a normally distributed dataset.

The z-score value of an RSS observation based on the MAD scale estimator in (3) is mathematically obtained as [14] and [16]:

$$Z_{MAD}\left(rss\left(n\right)\right) = \frac{\left|rss\left(n\right) - Mdn\left\{\mathbf{rss_i}\right\}\right|}{\sigma_{MAD}(rss\left(n\right))} \tag{4}$$

The z-score value obtained using (4) is a measure of how many MADs away an RSS observation is from the median of the entire time-series RSS dataset in (1).

B. MODIFIED Z-SCORE METHOD WITH Sn SCALE ESTIMATOR

Another robust scale estimator for determining the variability of a time-series RSS dataset is the Sn scale estimator. The scaling value obtained using the Sn scale estimator is based on the median of pairwise absolute differences between RSS observations. That is, it considers the difference between each pair of RSS observations and takes the median of those differences. The Sn scale estimator has a Gaussian efficiency of 58%, which is higher than that of the MAD scale estimator but has the same breakdown point of 50% [14].

Given the time-series RSS dataset in (1), the scaling value of the entire RSS observation obtained using the Sn scale estimator can be calculated mathematically as [10] and [14]:

$$\sigma_{Sn}\left(\mathbf{rss_i}\right) = C_n \times Mdn_m\left\{Mdn_k\left\{\left|rss(m) - rss(k)\right|\right\}\right\} \tag{5}$$

where $m \in [1, 2, .., N]$, $n \in [1, 2, .., N]$, $m \neq n$ and C_n is a scale constant that is used to make the Sn scale estimator perform as an unbiased estimator. According to [22], the value of C_n is a function of the RSS observation size and is given as:

$$C_n = \begin{cases} \dfrac{N}{N-0.9} & \textit{for odd } N \\ 1 & \textit{for even } N \end{cases} \tag{6}$$

TABLE 1. Performance comparison between the Sn and MAD scale estimators.

Performance metric	Scale estimator MAD	Sn
Accuracy	more accurate for normally distributed data	greater accuracy for non-normally distributed data
Robustness	Less robust	More robust
Efficiency	Less efficient	More efficient
Sensitivity	More sensitive	Less sensitive
False positive rate	Higher	Lower
Computational complexity	Lower	Higher

The z-score value of an RSS observation based on the Sn scale estimator in (5) is mathematically obtained as [10]:

$$Z_{Sn}(rss(n)) = \frac{\left|rss(n) - Mdn\left\{\mathbf{rss_i}\right\}\right|}{\sigma_{Sn}\left(\mathbf{rss_i}\right)} \tag{7}$$

The z-score value obtained from (7) is a measure of how far an RSS observation is from the median when the pairwise absolute differences between each RSS observation are considered.

Both the Sn and MAD scale estimators have their advantages and disadvantages. Based on the influence function analysis presented in [14], the Sn scale estimator is more efficient and robust to outliers than the MAD scale estimator. However, the MAD scale estimator is more sensitive (in terms of gross error sensitivity) to outliers than the Sn scale estimator. This means that under the same conditions, the MAD scale estimator will detect more outliers [10], [14]. In terms of computational complexity, the MAD scale estimator is lower than the Sn scale estimator. A summary of the performance comparison between the Sn and MAD scale estimators for outlier detection is shown in Table 1 [14].

From Table 1, it can be seen that both scale estimators have one or two advantages and disadvantages over each other. However, despite their individual advantages and disadvantages, it is possible to have a scale estimator that combines the advantages of both scale estimators and can be used with the mZ-score method for improved outlier detection. This can be accomplished through scale estimator hybridization. By hybridizing the two scale estimators and using the resulting scale estimator with the mZ-score method, it is possible to have an outlier detection technique based on the mZ-score method that is as efficient and robust as the Sn scale estimator. Furthermore, it can be as sensitive to outliers as the MAD scale estimator and has a low false alarm and false negative rate as that of the Sn scale estimator. Thus, in the next section, the mZ-score method based on the different hybrid scale estimators is presented.

III. MODIFIED Z-SCORE METHOD WITH HYBRID SCALE ESTIMATOR

Scale estimator hybridization, as previously stated, can improve the performance of the mZ-score method for detecting outliers. Three different scale estimator hybridization approaches have been identified: weighted, maximum, and average scale estimator hybridization approaches. The detailed description of each hybridization approach in the generation of the hybrid scale estimator for use with the mZ-score method is presented as follows:

A. MODIFIED Z-SCORE METHOD BASED ON WEIGHTED HYBRID SCALE ESTIMATOR

The first hybrid scale estimator presented in this paper is based on weight assignment. Weights are assigned to the z-score values obtained with the MAD and Sn scale estimators.

Let $Z_{MAD}(rss(n))$ and $Z_{Sn}(rss(n))$ be the z-score value obtained for any given RSS observation using (4) and (7) respectively. The weighted hybrid z-score value is mathematically obtained as:

$$Z_{wgt}(rss(n)) = w \times Z_{MAD}(rss(n)) + (1-w) \times Z_{Sn}(rss(n)) \tag{8}$$

The "w" in (8) denotes the weight assigned to the z-score value obtained using the MAD scale estimator, and "$(1-w)$" denotes the weight assigned to the z-score value obtained using the Sn scale estimator. The assignment of weights to each scale estimator is based on its importance or performance. The weighted hybridization method allows for greater flexibility in deciding whether to prioritize the robustness of the Sn scale estimator or the sensitivity of the MAD scale estimator. At $w = 0$, the weighted hybrid scale estimator functions fully as an Sn scale estimator, while at $w = 1$, it functions fully as an MAD scale estimator. The hyrbidization is valid only when w is set between 0 and 1.

B. MODIFIED Z-SCORE METHOD BASED ON MAXIMUM HYBRID SCALE ESTIMATOR

The second hybrid scale estimator presented in this paper is based on maximum value scale estimator hybridization. In this approach, the maximum value between the z-score values obtained using the MAD and the Sn scale estimators is taken as the hybrid z-score value. The hybrid z-score value based on the maximum hybridization method is mathematically obtained as:

$$Z_{max}(rss(n)) = max\{Z_{MAD}(rss(n)), Z_{Sn}(rss(n))\} \tag{9}$$

By choosing the maximum value between the two estimates, this approach considers outliers based on either the robustness of the Sn scale estimator or the sensitivity of the MAD scale estimator, whichever has a higher z-score value.

C. MODIFIED Z-SCORE METHOD BASED ON AVERAGE HYBRID SCALE ESTIMATOR

The last hybrid scale estimator presented in this paper is based on the average z-score value. The hybrid z-score value is the average of the z-score values obtained with the MAD scale estimator and the Sn scale estimator. The hybrid z-score value is calculated as follows:

$$Z_{avg}(rss(n)) = 0.5 \times (Z_{MAD}(rss(n)) + Z_{Sn}(rss(n))) \tag{10}$$

By taking the average, the robustness of the Sn scale estimator and the sensitivity of the MAD scale estimator are both incorporated into this hybridization approach, resulting in a well-balanced hybrid z-score value. It is worth noting that the weighted hybrid scale estimator performs the same as the average hybrid scale estimator at w = 0.5.

D. RSS OUTLIER DETERMINATION

After calculating the hybrid z-score values using any of the approaches described in (8), (9), and (10), the next step is to determine whether or not an RSS observation is an outlier. To achieve that, a z-score threshold value is set. If the z-score value of an RSS observation is above this set threshold, it is regarded as an outlier. An RSS observation is considered to be an outlier if its z-score value agrees with (11) below.

$$\begin{aligned} &Z_n(rss(n)) > \gamma \\ &\text{for } Z_n(rss(n)) \epsilon [Z_{wgt}(rss(n)), \\ &\qquad Z_{max}(rss(n)), Z_{avg}(rss(n))] \end{aligned} \tag{11}$$

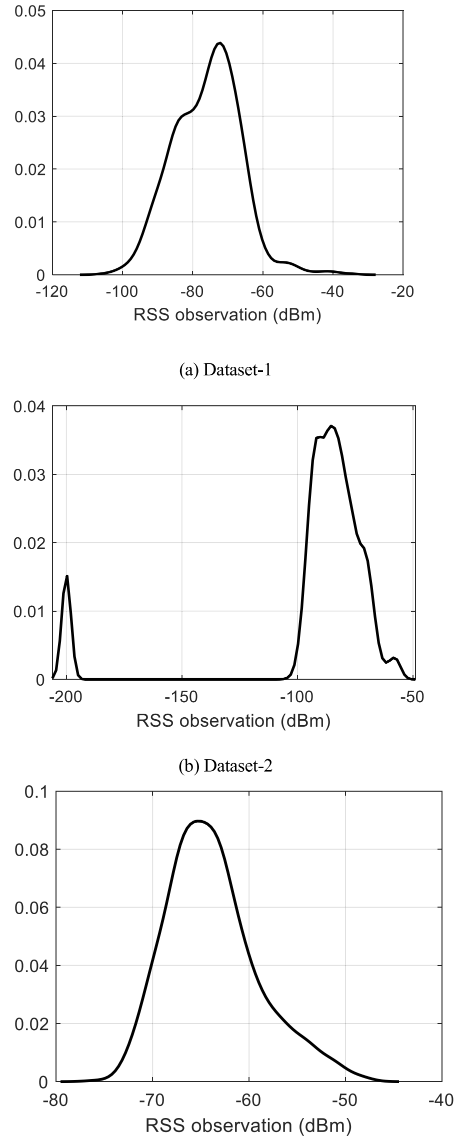
where γ is the z-score threshold value.

In the following section, the performance of each hybrid scale estimator with the mZ-score method is determined and compared using three publicly available time-series RSS datasets. Furthermore, the performance of the best hybrid scale estimator among the three is used with the mZ-score method and compared to other univariant time-series RSS dataset-based outlier detection techniques.

IV. SIMULATION RESULT AND DISCUSSION

This section of the paper uses three experimentally generated time-series RSS datasets from [23], [24], and [25] to determine and compare the outlier detection performances of the mZ-score method with the three hybrid scale estimators described in Section III. The RSS dataset obtained from [23], which will be referred to as Dataset-1, contains time-series LTE RSS observations obtained from Covenant University, Nigeria. The time-series RSS observation was taken between 7:30 a.m. and 11:00 p.m. at 30-min intervals for 30 days, resulting in approximately 960 RSS observations. The second RSS dataset obtained from [25], which will be referred to as Dataset-2, contained a total of 4,920 BLE-based time-series RSS observations obtained from the Physics and Mathematics building of the University of Extremadura, Spain. The last time-series RSS dataset, which will be referred to as Dataset 3, contains a total of 1,000 RSS observations obtained in a large laboratory at the University of Minho, Spain [24].

The distribution of RSS observations plays a crucial role in outlier detection. As earlier mentioned, RSS outliers are RSS observations that deviate significantly from the majority of the RSS observations, and their presence can distort statistical



(a) Dataset-1

(b) Dataset-2

(c) Dataset-3

FIGURE 1. Time-series RSS observation statistical distribution.

analyses. Figure 1 shows the statistical distribution of the RSS observations in each of the datasets considered.

As shown in Figure 1(a), the statistical distribution of the RSS observations in Dataset-1 can be approximately compared to a normal distribution. This indicates that the RSS observations in this dataset are fairly evenly distributed around the mean, with a SD that is relatively consistent. In contrast, the RSS observations in Dataset-2 exhibit two distinct distributions, as depicted in Figure 1(b). The majority of RSS observations follow a non-normal positive-skewed distribution, which suggests that the dataset contains outliers that are skewing the distribution towards the right. Similarly, all the RSS observations in Dataset-3 also display a non-normal right-skewed distribution, as shown in Figure 1(c). This also suggests that the dataset is skewed towards higher

TABLE 2. Outlier detection sensitive for Dataset-1.

Weight value (w)	z-score threshold ±1.0	±2.0	±3.0
0.1	331	46	8
0.2	331	46	8
0.3	331	46	8
0.3	331	46	8
0.4	331	46	8
0.6	331	46	8
0.7	331	46	8
0.8	331	46	8
0.9	331	65	8

TABLE 3. Outlier detection sensitive for Dataset-2.

Weight value (w)	z-score threshold ±1.0	±2.0	±3.0
0.1	1535	498	381
0.2	1535	498	381
0.3	1535	498	381
0.3	1535	498	381
0.4	1535	498	381
0.6	1535	498	381
0.7	1535	507	381
0.8	1535	507	381
0.9	1535	507	381

RSS values, and there might be a small number of extreme RSS outliers.

A. WEIGHT VALUE DETERMINATION FOR THE WEIGHT HYBRID SCALE ESTIMATOR

To determine and compare the outlier detection performance of the mZ-score method that is based on the weighted hybrid scale estimator to the other two hybrid scale estimators, there is a need to determine the optimum weight value to be assigned to each scale estimator hybridised in (8). The weight assignment depends on the statistical distribution of the time-series RSS observation and the required sensitivity and robustness to outliers that are desired.

As previously stated, the weighted hybrid scale estimator functions as a Sn scale estimator at $w = 0$, and as a MAD scale estimator at $w = 1$. Only for $0 > w > 1$ does the scale estimator function as a hybrid. By varying the weight value from 0 to 1 at intervals of 0.1, the number of outliers detected for each weight value is determined using z-score threshold values of ±1, ±2, and ±3 for all three datasets considered. Tables 2, 3, and 4 show the sensitivity of the mZ-score method using the weighted hybrid scale estimator for Datasets 1, 2, and 3, respectively. The green highlights indicate weight values with the least number of detected outliers for each z-score threshold value.

From Tables 2, 3, and 4, irrespective of the weight values, as the z-score threshold value increases from ±1.0 to ±3.0, the number of outliers detected decreases. This is because an increase in the z-score threshold value reduces the total number of RSS observations in each dataset taken into consideration for outlier detection. For z-score threshold values of ±1.0 and ±3.0, the number of outliers detected

TABLE 4. Outlier detection sensitive for Dataset-3.

Weight value (w)	z-score threshold ±1.0	±2.0	±3.0
0.1	313	55	5
0.2	313	55	5
0.3	313	55	5
0.3	313	55	5
0.4	313	55	7
0.6	313	55	7
0.7	313	55	7
0.8	313	55	7
0.9	313	72	7

is approximately the same, irrespective of the weight value. This is because above or below these z-score threshold values, the performances of any scale estimator, irrespective of its sensitivity or robustness to outliers, will be the same, as there are either too many or too few observations to consider for outlier detection. As such, it is difficult to determine whether a scale estimator is robust or too sensitive to outliers.

Looking at the number of outliers detected as the weight value varies from 0.1 to 0.9, considering a z-score threshold of ±2.0, it can be seen that there are weight values above which there is a sudden increase in the number of outliers detected for all three datasets considered. From Table 2, as w increased from 0.8 to 0.9, the number of detected outliers increased significantly from 46 to 65. Also, in Table 3, as w increased from 0.6 to 0.7, the number of outliers detected increased from 498 to 507. So also in Table 4, as w increased from 0.8 to 0.9, the number of outliers increased from 55 to 72. This sudden increase in the number of outliers detected shows that the weighted hybrid scale estimator has become more sensitive to outliers. That is, prior to that, there was a balance between the influence of the MAD and Sn scale estimators; however, at that instance, the MAD had more influence on the outlier detection performance of the weighted hybrid scale estimator. Detecting a large number of outliers does not imply good performance. There is a high possibility of false alarm detection. The goal is to reduce false alarm detection while at the same time detecting as many valid outliers as possible. Thus, the largest weight value that results in the least number of detected outliers across all z-score threshold values is considered the optimum weight value. With this weight value, there will be a reduction in false alarm rates while at the same time increasing both the accuracy and robustness of the weighted hybrid scale estimator for outlier detection. As such, the optimum weight values for Datasets 1, 2, and 3 are $w = 0.8$, $w = 0.6$, and $w = 0.3$, respectively.

B. OUTLIER DETECTION PERFORMANCE COMPARISON OF MZ-SCORE METHOD WITH DIFFERENT HYBRID SCALE ESTIMATORS

In this subsection, the outlier detection performances of the mZ-score method with the three hybrid scale estimators are determined and compared for each dataset. The analysis will be conducted at a z-score threshold value of ±2, as this is considered to be the optimal value [26], [27]. What this means is that any RSS observation whose z-score value is more than ±2 SD is considered an outlier. Also, the ±2 SD z-score threshold indicates that about 98% of the entire RSS observation should be considered valid, which is non-outlier. With this, for each dataset, the number of RSS observations identified as outliers by the mZ-score method with each of the hybrid scale estimators is determined.

TABLE 5. Number of outliers detected for Dataset-1.

Scale Estimator	Number of Outliers detected
Maximum hybrid	65
Average hybrid	46
Weighted hybrid	46
MAD	65
Sn	45

1) OUTLIER DETECTION PERFORMANCE FOR DATASET-1

This subsection presents the outlier detection performance of the mZ-score method with the hybrid scale estimators using Dataset-1. Table 5 shows the total number of outliers detected by the mZ-score method with each of the three hybrid scale estimators—weighted hybrid, maximum hybrid, and average hybrid—as well as with parent scale estimators hybridised, that is, MAD and Sn scale estimators. A graphical comparison of the number of outliers detected for Dataset-1 is shown in Figure 2.

The MAD scale estimator is known for its high sensitivity to outliers, and from Table 5, it can be seen to detect a large number of outliers of about 65. The Sn scale estimator with a lower sensitivity to outliers detected a small number of outliers, about 45. Extending the analysis to the hybrid scale estimators, the maximum hybrid scale estimator performed equally to the MAD scale estimator, detecting the same number of outliers of about 65. The average and weighted hybrid scale estimators, when compared to the maximum hybrid scale estimator, detected a smaller number of outliers of about 46, approximately the same as the Sn scale estimator.

Overall, based on Dataset-1, the average and the weighted hybrid scale estimators are considered to be the best scale estimators to be used with the mZ-score method for outlier detection. This is because they detected a moderate number of outliers that are slightly higher than those of the Sn scale estimator, indicating slightly higher sensitivity. The two hybrid scale estimators detected outliers that are significantly lower than the number detected by the MAD scale estimator, indicating a higher robustness to outliers. In summary, the average and weighted hybrid scale estimators considering Dataset-1 are as robust to outliers as the Sn scale estimator and as sensitive to outliers as the MAD scale estimator.

2) OUTLIER DETECTION PERFORMANCE FOR DATASET-2

This subsection presents the performance of the mZ-score method with the three hybrid scale estimators using Dataset-2. Table 6 shows the number of outliers detected with each of

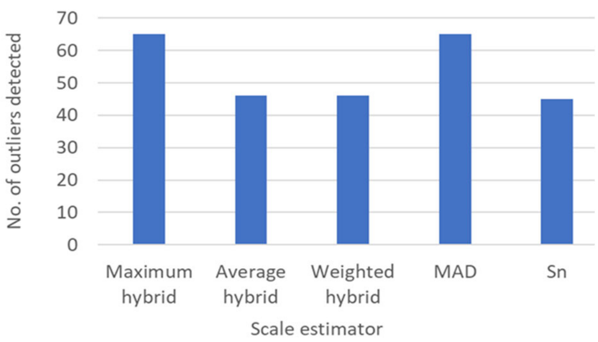


FIGURE 2. A graphic illustration of the number of outliers detected by the mZ-score method considering different scale estimators for Dataset 1.

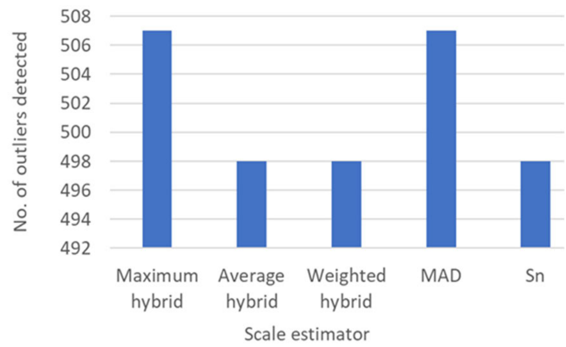


FIGURE 3. A graphic illustration of the number of outliers detected by the mZ-score method considering different scale estimators for Dataset 2.

TABLE 6. Number of outliers detected for Dataset-1.

Scale Estimator	Number of Outliers detected
Maximum hybrid	507
Average hybrid	498
Weighted hybrid	498
MAD	507
Sn	498

TABLE 7. Number of outliers detected for Dataset-1.

Scale Estimator	Number of Outliers detected
Maximum hybrid	74
Average hybrid	55
Weighted hybrid	55
MAD	74
Sn	55

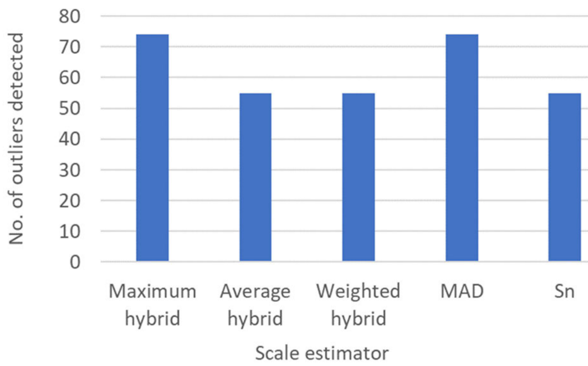


FIGURE 4. A graphic illustration of the number of outliers detected by the mZ-score method considering different scale estimators for Dataset 3.

the scale estimators with the graphical comparison shown in Figure 3.

Both the Sn and MAD scale estimators detected a large number of outliers; however, there are no significant differences in the number of outliers detected. The Sn scale estimator detected about 498 outliers, while the MAD scale estimator detected about 507, which is 9 outliers higher. The maximum hybrid scale estimator, as usual, performed equally to the MAD scale estimator, while the weighted and average hybrid scale estimators performed equally to the Sn scale estimator. Amongst the hybrid scale estimators, the weighted and average hybrid scale estimators performed the best, as they detected fewer outliers, even though the difference is not that significant.

3) OUTLIER DETECTION PERFORMANCE FOR DATASET-2

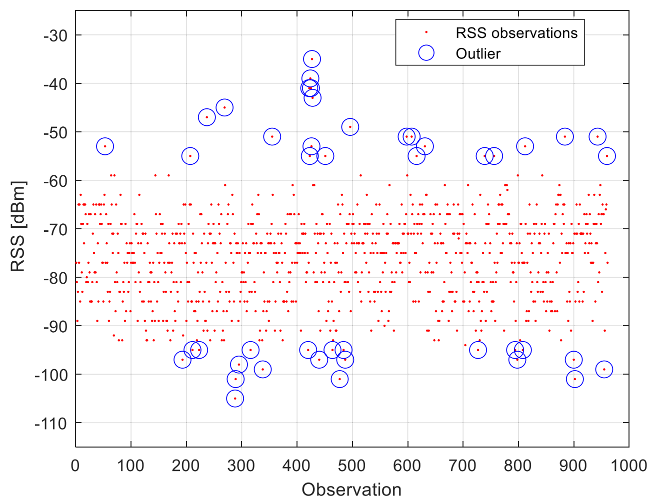
The performance of the mZ-score method with the hybrid scale estimators is determined and presented in this subsection using Datase-3. Table 7 shows the number of outliers detected by each scale estimator, with a graphical illustration in Figure 4.

Based on the data presented in Table 7, the same conclusion is reached as in Datasets 1 and 2 with regards to the best hybrid scale estimator. The weighted and average hybrid scale estimators both detected the least number of outliers. Their performance is comparable to the performance of the Sn scale estimator, which has about the same robustness to outliers while still having the same sensitivity to outliers as the MAD scale estimator.

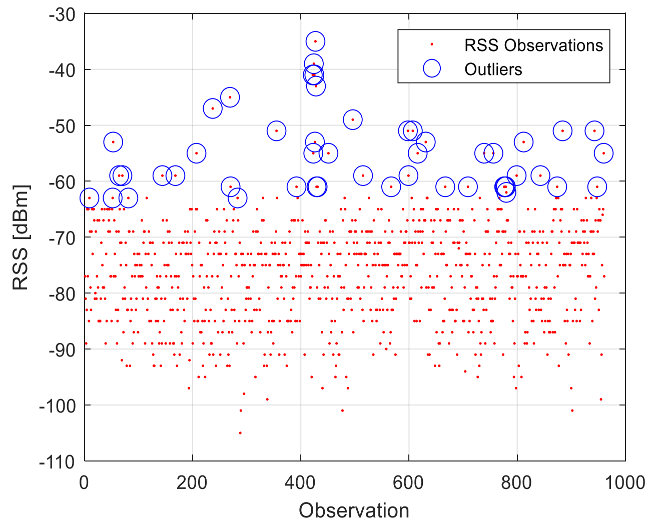
Based on the performances analysis of the hybrid scale estimator considering the three RSS dataset, the weighted and average hybrid scale estimators have shown better performance than the maximum hybrid scale estimators. As earlier mentioned, at $w = 0.5$, the weighted hybrid scale estimator functions as an average hybrid scale estimator. As such, it can be said that of all the hybrid scale estimators considered, the weighted hybrid scale estimator performs the best. The weighted hybrid scale estimator tries to strike a balance between sensitivity and robustness to outliers, as determined by the weight value. Thus, the selection of the weight value for the weighted hybrid scale estimator determined how well the mZ-score method detected outliers, and this is dependent on the distribution of the RSS observation.

C. OUTLIER PERFORMANCE COMPARISON OF THE MZ-SCORE + WEIGHT HYBRID SCALE ESTIMATOR WITH OTHER UNIVARIANT-BASED OUTLIER DETECTION TECHNIQUES

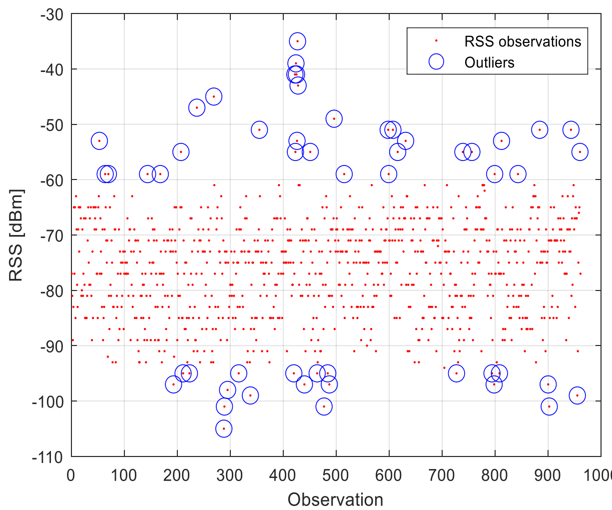
Based on the results analysis and conclusions from the preceding subsections, the weighted hybrid scale estimator



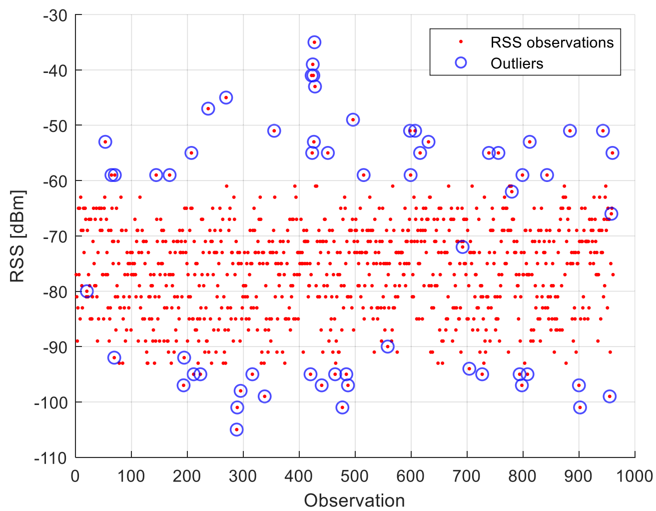
(a) mZ-score method with weighted hybrid scale estimator



(b) MOD technique



(c) K-means clustering-based technique



(d) DBSCAN based technique

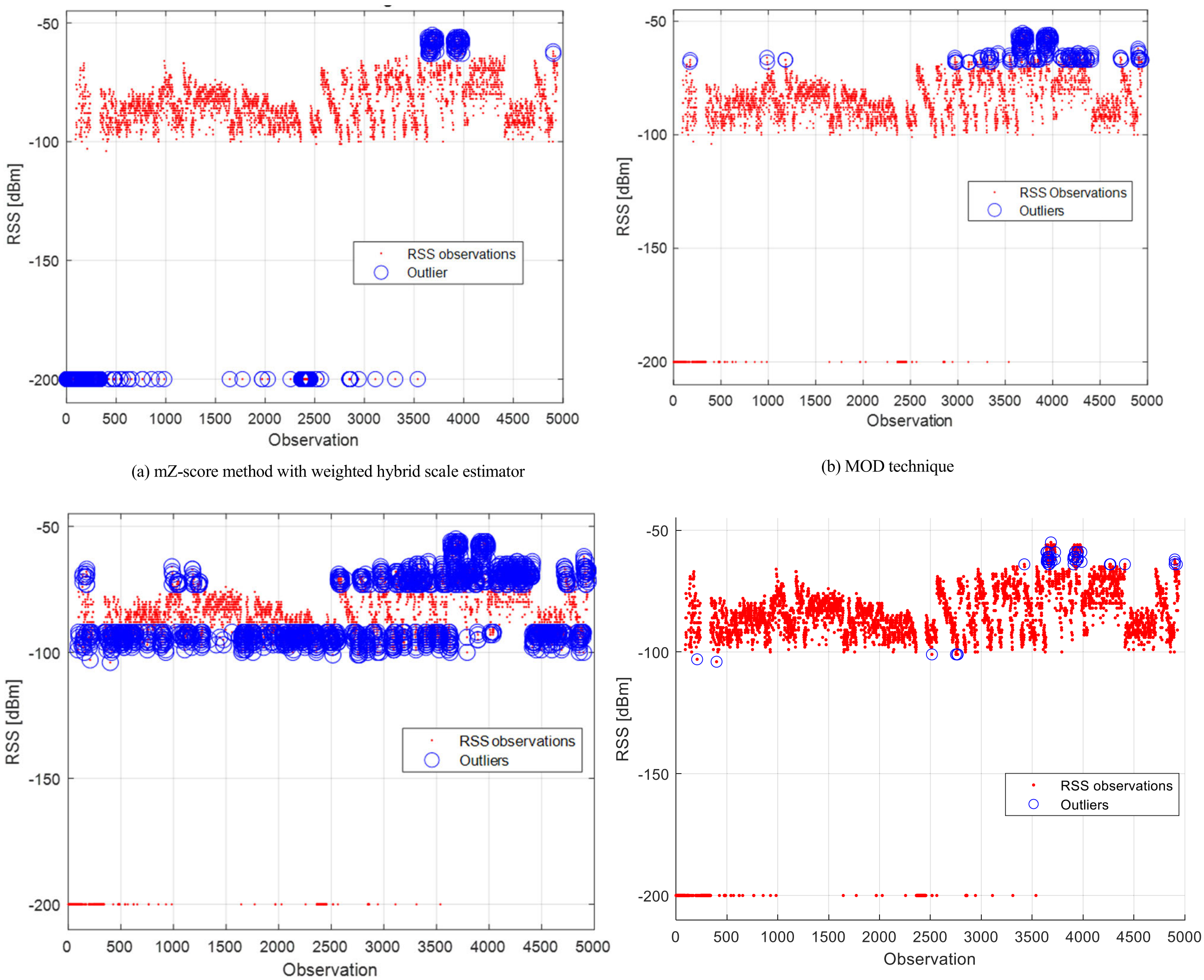
FIGURE 5. Graphical illustration of outliers detected in Dataset-1.

is the best scale estimator proposed for the efficient and accurate detection of outliers using the mZ-score method. In this section, the performance of the mZ-score method with the weighted hybrid scale estimator is compared to that of other univariant-based outlier detection techniques. The univariant-based outlier detection techniques considered are the mean-shift-based outlier detection (MOD) technique described in [28], the k-means clustering-based outlier detection technique described in [29], [30], and [31], and the density-based spatial clustering (DBSCAN)-based technique described in [32] and [33]. Figures 5–7 show a graphical comparison of the outlier detection performances of the mZ-score method with the weighted hybrid scale estimator with the MOD, k-means clustering-based, and DBSCAN-based techniques for Datasets 1, 2, and 3, respectively. Table 8 displays a summary of the number of outliers detected by each technique for each dataset.

TABLE 8. Outlier detection performance comparison across three datasets: mZ-score method with weighted hybrid scale estimator vs. related techniques.

Outlier detection technique	Number and percentage of outliers detected		
	Dataset-1	Dataset-2	Dataset-3
mZ-score with weighted hybrid scale estimator	46	498	55
MOD technique	49	248	51
k-means clustering-based technique	53	1727	3
DBSCAN-based technique	61	55	42

Looking at the outlier detection performance result comparison for Dataset-1, it can be seen that all four outlier



(a) mZ-score method with weighted hybrid scale estimator

(b) MOD technique

(c) K-means clustering-based technique

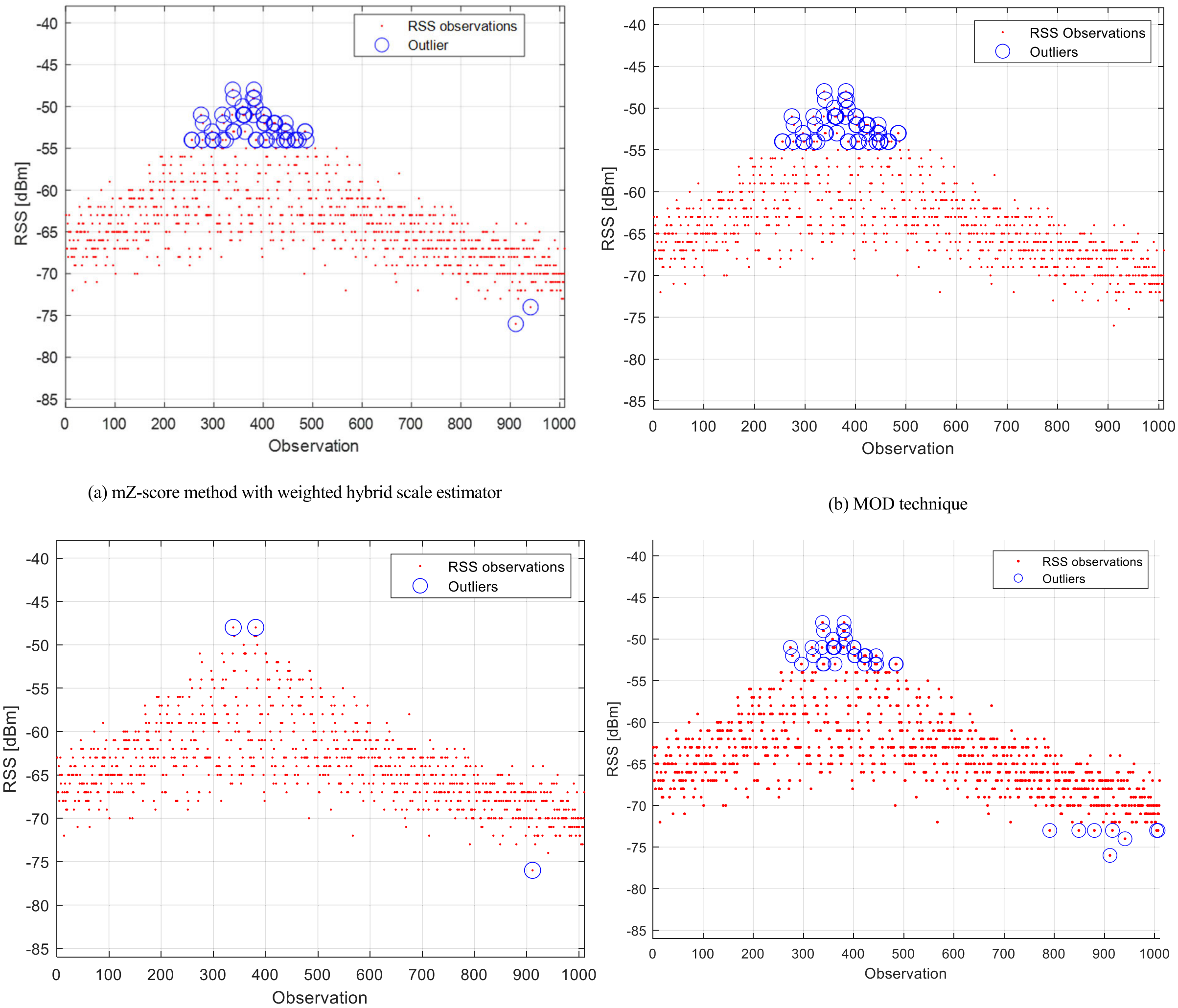
(d) DBSCAN-based technique

FIGURE 6. Graphical illustration of outliers detected in Dataset-2.

detection techniques detected a moderate number of outliers, about 6% of the overall RSS observations. However, from Figure 5, it can be seen that the location of the outliers detected by each technique is different. The distribution of the RSS observations for the dataset is considered to be approximately normal, as shown in Figure 1(a), as outliers are expected to be at both extreme ends of the distributions. The mZ-score method with the proposed scale estimator was able to detect outliers at both ends of the distribution, as shown in Figure 5(a), while the MOD technique only detected outliers at the upper end of the RSS distribution, as shown in Figure 5(b). The k-means clustering-based technique detected outliers at both ends of the distribution, as shown in Figure 5(c), just like that of the mZ-score method with the proposed scale estimator. However, it detected higher numbers of outliers than the MOD technique and the mZ-score method, indicating higher sensitivity to outliers. The DBSCAN-based technique detected the highest number of outliers in Dataset-1, indicating higher sensitivity to outliers, and was also able to detect outliers at both ends of the RSS observation distributions. However, from Figure 5(d), it can be seen that there are RSS observations that are falsely detected as outliers that are present in the middle of the RSS observation, which suggests a false detection.

In summary, for Dataset-1, the mZ-score method with the proposed hybrid scale estimator has better outlier detection performance compared to the other 3 techniques. It is not as sensitive to outliers as the DBSCAN and k-means clustering-based techniques, and there is no false alarm detection. It is also as robust to outliers as that of the MOD technique but is able to detect outliers at both ends of the RSS observation distribution, which is not done by the MOD technique.

Extending the analysis to Dataset-2, there is a significant difference in the number of outliers detected by each technique from Table 8, and the locations of outliers detected by each technique differ according to Figure 6. As shown



(a) mZ-score method with weighted hybrid scale estimator

(b) MOD technique

(c) K-means clustering-based technique

d) DBSCAN-based technique

FIGURE 7. Graphical illustration of outliers detected in Dataset-3.

in Figure 1(b), the RSS observations in Dataset-2 have two distinct distributions. The few RSS observations with mean values around -200 dBm are all considered outliers. As for the majority of the RSS observations that followed the non-normal positive-skewed distribution, it is expected that outliers are present towards the right side of the distribution. The mZ-score method with the proposed hybrid scale estimator identified all RSS observations with a mean of around -200 dBm as outliers. It was also able to identify outliers on the right side of the RSS distribution, as shown in the upper side of the RSS observations in Figure 6(a). The MOD technique fails to identify the RSS observations with a mean of around -200 dBm as outliers, resulting in false negative performance, as shown in Figure 6(b). However, it is able to identify the RSS observations towards the right of the RSS distribution as outliers. The k-means clustering-based technique performed the worst in Dataset-2, as it incorrectly identified a large number of RSS observations as outliers, as shown in Figure 6(c), indicating a high false alarm performance. It also exhibits false negative performances as it fails to detect the RSS observations with a mean of around -200 dBm as outliers. The DBSCAN-based technique also fails to detect outliers with mean values around -200 dBm. However, it was able to detect outliers to the right of the RSS distribution but incorrectly identified a few RSS observations as outliers, as shown in Figure 6(d).

In summary, for Dataset-2, the proposed mZ-score technique with the proposed hybrid scale estimator has the best outlier detection performance, as it was able to detect all outliers with a mean value around -200 dBm and those present on the right side of the RSS distribution.

The majority of the outliers in Dataset-3 are to the positive right of the RSS distribution, which translates to observations with high RSS values. As shown in Figures 7(a), 7(b), and 7(c), the mZ-score method with the proposed scale estimator, the MOD, and the DBSCAN-based techniques, respectively, were able to detect these outliers. The mZ-score method and the DBSCAN-based method have few false alarm detections. The mZ-score technique incorrectly identified 2 RSS observations as outliers, accounting for approximately 3.5% of all outliers detected. The k-means clustering-based technique performs the worst, detecting only three outliers, as shown in Figure 7(a).

In summary, for Dataset-3, the MOD technique has the best outlier detection performance, closely followed by the mZ-score method with the proposed hybrid scale estimator.

Overall, the mZ-score method with the proposed hybrid scale estimator, which is based on the weighted hybrid scale estimator, has the best outlier detection performance across the three datasets. It can detect outliers accurately with no false negative detection, as seen with MOD, k-means clustering-based, and DBSCAN-based techniques. It also has very little or no false alarm detection performance, as observed with the other three techniques. This demonstrates the high robustness as well as the low sensitivity to outliers the weighed hybrid scale estimator made the mZ-score method to be.

V. CONCLUSION

The performance of three different scale estimator hybridization methods, namely the weighted, maximum, and average, was evaluated in this paper for use in conjunction with a mZ-score method for outlier detection. The MAD and Sn scale estimators are the two scale estimators considered for hybridization. The study employs three publicly available time-series RSS datasets, each with a unique statistical distribution. The simulation results based on the time-series RSS datasets considered show that the weighted hybridization method creates a scale estimator that is optimal for detecting outliers when used with the mZ-score method. When compared to the MOD, k-means clustering-based, and DBSCAN-based outlier detection techniques, the mZ-score method with the weighted hybrid scale estimator has superior outlier detection performance with little or no false alarm and false negative detection.

ACKNOWLEDGMENT

This study is supported by the Excellence project run at the Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic.

REFERENCES

[1] A. S. Yaro, F. Maly, and P. Prazak, "A survey of the performance-limiting factors of a 2-dimensional RSS fingerprinting-based indoor wireless localization system," *Sensors*, vol. 23, no. 5, p. 2545, Feb. 2023, doi: 10.3390/s23052545.

[2] A. S. Yaro, F. Malý, and K. Malý, "A two-nearest wireless access point-based fingerprint clustering algorithm for improved indoor wireless localization," *Emerg. Sci. J.*, vol. 7, no. 5, pp. 1762–1770, Oct. 2023, doi: 10.28991/esj-2023-07-05-019.

[3] A. S. Yaro, F. Malý, and K. Malý, "Improved indoor localization performance using a modified affinity propagation clustering algorithm with context similarity coefficient," *IEEE Access*, vol. 11, pp. 57341–57348, 2023, doi: 10.1109/ACCESS.2023.3283592.

[4] S. Shang and L. Wang, "Overview of WiFi fingerprinting-based indoor positioning," *IET Commun.*, vol. 16, no. 7, pp. 725–733, Apr. 2022, doi: 10.1049/cmu2.12386.

[5] L. Flueratoru, V. Shubina, D. Niculescu, and E. S. Lohan, "On the high fluctuations of received signal strength measurements with BLE signals for contact tracing and proximity detection," *IEEE Sensors J.*, vol. 22, no. 6, pp. 5086–5100, Mar. 2022, doi: 10.1109/JSEN.2021.3095710.

[6] R. Zhou, F. Meng, J. Zhou, and J. Teng, "A Wi-Fi indoor positioning method based on an integration of EMDT and WKNN," *Sensors*, vol. 22, no. 14, p. 5411, Jul. 2022, doi: 10.3390/s22145411.

[7] H. Zhu, K.-F. Tsang, Y. Liu, Y. Wei, H. Wang, C. K. Wu, and H. R. Chi, "Extreme RSS based indoor localization for LoRaWAN with boundary autocorrelation," *IEEE Trans. Ind. Informat.*, vol. 17, no. 7, pp. 4458–4468, Jul. 2021, doi: 10.1109/TII.2020.2996636.

[8] I. Bae and U. Ji, "Outlier detection and smoothing process for water level data measured by ultrasonic sensor in stream flows," *Water*, vol. 11, no. 5, p. 951, May 2019, doi: 10.3390/w11050951.

[9] P. Sadhukhan, S. Gain, K. Dahal, S. Chattopadhyay, N. Garain, and X. Wang, "An efficient clustering with robust outlier mitigation for Wi-Fi fingerprint based indoor positioning," *Appl. Soft Comput.*, vol. 109, Sep. 2021, Art. no. 107549, doi: 10.1016/j.asoc.2021.107549.

[10] A. S. Yaro, F. Maly, and P. Prazak, "Outlier detection in time-series receive signal strength observation using Z-Score method with SN scale estimator for indoor localization," *Appl. Sci.*, vol. 13, no. 6, p. 3900, Mar. 2023, doi: 10.3390/app13063900.

[11] P. J. Rousseeuw and M. Hubert, "Robust statistics for outlier detection," *WIREs Data Mining Knowl. Discovery*, vol. 1, no. 1, pp. 73–79, Jan. 2011, doi: 10.1002/widm.2.

[12] R. R. Wilcox, "3—Summarizing data," in *Applying Contemporary Statistical Techniques*, R. R. Wilcox, Ed., Burlington: Academic Press, 2003, pp. 55–91, doi: https://doi.org/10.1016/B978-012751541-0/50024-9.

[13] Y. C. Chen, W. C. Sun, and J. C. Juang, "Outlier detection technique for RSS-based localization problems in wireless sensor networks," in *Proc. SICE Annu. Conf.*, Aug. 2010, pp. 657–662.

[14] P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," *J. Amer. Stat. Assoc.*, vol. 88, no. 424, pp. 1273–1283, Dec. 1993, doi: 10.1080/01621459.1993.10476408.

[15] L. Kulanuwat, C. Chantrapornchai, M. Maleewong, P. Wongchaisuwat, S. Wimala, K. Sarinnapakorn, and S. Boonya-Aroonnet, "Anomaly detection using a sliding window technique and data imputation with machine learning for hydrological time series," *Water*, vol. 13, no. 13, p. 1862, Jul. 2021, doi: 10.3390/w13131862.

[16] Y.-C. Chen and J.-C. Juang, "Outlier-detection-based indoor localization system for wireless sensor networks," *Int. J. Navigat. Observ.*, vol. 2012, pp. 1–11, Apr. 2012, doi: 10.1155/2012/961785.

[17] L. Chiaramonte, F. H. Liu, F. Poli, and M. Zhou, "How accurately can Z-score predict bank failure?" *Financial Markets, Institutions Instrum.*, vol. 25, no. 5, pp. 333–360, Dec. 2016, doi: 10.1111/fmii.12077.

[18] O. A. Vsevolozhskaya, K. C. Manz, P. M. Zephyr, and T. M. Waters, "Measurement matters: Changing penalty calculations under the hospital acquired condition reduction program (HACRP) cost hospitals millions," *BMC Health Services Res.*, vol. 21, no. 1, p. 131, Dec. 2021, doi: 10.1186/s12913-021-06108-w.

[19] J. M. Torres, P. J. G. Nieto, L. Alejano, and A. N. Reyes, "Detection of outliers in gas emissions from urban areas using functional data analysis," *J. Hazardous Mater.*, vol. 186, no. 1, pp. 144–149, Feb. 2011, doi: 10.1016/j.jhazmat.2010.10.091.

[20] T. Crosby, "How to detect and handle outliers," *Technometrics*, vol. 36, no. 3, pp. 315–316, Aug. 1994, doi: 10.1080/00401706.1994.10485810.

[21] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection," *ACM Comput. Surveys*, vol. 41, no. 3, pp. 1–58, Jul. 2009, doi: 10.1145/1541880.1541882.

[22] C. Croux and P. J. Rousseeuw, "Time-efficient algorithms for two highly robust estimators of scale," in *Computational Statistics*. Heidelberg, Germany: Physica-Verlag HD, 1992, pp. 411–428, doi: 10.1007/978-3-662-26811-7_58.

[23] H. S. Karanja and A. Atayero, "Cellular received signal strength indicator dataset," *IEEE Dataport*, 2020. Accessed: Jan. 22, 2023. [Online]. Available: https://ieee-dataport.org/open-access/cellular-received-signal-strength-indicator-dataset

[24] J. Torres-Sospedra and A. Moreira, "Analysis of sources of large positioning errors in deterministic fingerprinting," *Sensors*, vol. 17, no. 12, p. 2736, Nov. 2017, doi: 10.3390/s17122736.

[25] F. J. Aranda, F. Parralejo, F. J. Álvarez, and J. Torres-Sospedra, "Multi-slot BLE raw database for accurate positioning in mixed indoor/outdoor environments," *Data*, vol. 5, no. 3, p. 67, Jul. 2020, doi: 10.3390/data5030067.

[26] T. Sakai, H. Onishi, and K. Miyazoe, "Evaluation of optimal threshold of Z score map for statistical brain image analysis," *Japanese J. Radiological Technol.*, vol. 71, no. 4, pp. 325–331, 2015, doi: 10.6009/jjrt.2015_jsrt_71.4.325.

[27] H. Beyer and J. W. Tukey, *Exploratory Data Analysis*, vol. 23. USA: Addison-Wesley, 1977, pp. 413–414, doi: 10.1002/bimj.4710230408.

[28] J. Yang, S. Rahardja, and P. Fränti, "Mean-shift outlier detection and filtering," *Pattern Recognit.*, vol. 115, Jul. 2021, Art. no. 107874, doi: 10.1016/j.patcog.2021.107874.

[29] B. Angelin and A. Geetha, "Outlier detection using clustering techniques–K-means and K-median," in *Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2020, pp. 373–378, doi: 10.1109/ICICCS48265.2020.9120990.

[30] W. Liu, N. Wang, and Y. Huang, "Outlier detection method based on improved K-means clustering algorithm," in *Proc. 5th Int. Conf. Electron. Inf. Technol. Comput. Eng.* New York, NY, USA: ACM, Oct. 2021, pp. 1350–1355, doi: 10.1145/3501409.3501648.

[31] A. Sarvani, B. Venugopal, and N. Devarakonda, "Anomaly detection using K-means approach and outliers detection technique," in *Soft Computing: Theories and Applications*, vol. 724. Singapore: Springer, 2019, pp. 375–385, doi: 10.1007/978-981-13-0589-4_35.

[32] A. Smiti, "A critical overview of outlier detection methods," *Comput. Sci. Rev.*, vol. 38, Nov. 2020, Art. no. 100306, doi: 10.1016/j.cosrev.2020.100306.

[33] M. Chen, K. Zhou, and D. Liu, "Machine learning based technique for outlier detection and result prediction in combustion diagnostics," *Energy*, vol. 290, Mar. 2024, Art. no. 130218, doi: 10.1016/j.energy.2023.130218.

ABDULMALIK SHEHU YARO was born in Kaduna, Nigeria, in 1990. He received the B.Eng. degree in electrical engineering from Ahmadu Bello University (ABU), Zaria, in 2012, and the M.Eng. degree in telecommunications engineering and the Ph.D. degree in electrical engineering from the University of Technology, Malaysia, in 2014 and 2018, respectively. He is currently a Scientific Researcher with the Department of Informatics and Quantitative Methods, University of Hradec Králové, Czech Republic, and an Assistant Professor with the Department of Electronics and Telecommunications Engineering, ABU. His papers have appeared in more than 20 international peer-reviewed journals and conference proceedings. His research interests include indoor wireless localization and sensor networks.

FILIP MALY was born in Czech Republic, in 1981. He received the master's degree in information management from the University of Hradec Králové, Czech Republic, in 2005, and the Ph.D. degree in information and knowledge management, in 2009. Since 2014, he has been an Associate Professor of systems engineering and informatics with the University of Hradec Králové. He is currently a Senior Researcher and a Teacher with the Department of Informatics and Quantitative Methods, University of Hradec Králové. His research interests include indoor localization, wireless sensor networks, programming languages, and operating systems.

PAVEL PRAZAK was born in Czech Republic, in 1968. He received the master's degree in teaching of mathematics and physics and the Ph.D. degree in general problems of mathematics and informatics from the Faculty of Mathematics and Physics, Charles University, Prague, in 1996 and 2006, respectively. Since 2014, he has been an Associate Professor of systems engineering and informatics with the University of Hradec Králové, Czech Republic. He is currently the Head of the Department of Informatics and Quantitative Methods, Faculty of Informatics and Management, University of Hradec Králové. His research interests include indoor localization, dynamical systems and optimization in economics, and applications of statistics.

KAREL MALÝ was born in Czech Republic, in 1973. He received the master's degree in cybernetics and the Ph.D. degree in artificial intelligence and biocybernetics from the Faculty of Electrical Engineering, Czech Technical University in Prague (CTU), Czech Republic, in 1998 and 2005, respectively. Since 2000, he has been an Assistant Professor and a Researcher with the Department of Informatics and Quantitative Methods, University of Hradec Králové, Czech Republic. His research interests include artificial intelligence, robotics, indoor localization, programming languages, and Unix operating systems.