INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

REVIEW ARTICLE

Computing Robust Measure of Location on Multivariate Statistical Data Using Euclidean Depth Procedures

R Muthukrishnan [1], **Surabhi S Nair**[2]*

1 Professor, Department of Statistics, Bharathiar University, Coimbatore, 641046, Tamil Nadu, India

2 Research Scholar, Department of Statistics, Bharathiar University, Coimbatore, 641046, Tamil Nadu, India

OPEN ACCESS

Received: 11-04-2023

Accepted: 05-06-2023

Published: 27-06-2023

Citation: RM, Nair SS (2023) Computing Robust Measure of Location on Multivariate Statistical Data Using Euclidean Depth Procedures. Indian Journal of Science and Technology 16(26): 1927-1934. https://doi.org/10.17485/IJST/v16i26.847

* **Corresponding author**.

surabhinair93@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To suggest reliable location parameters (central value) in multivariate datasets using data depth procedures in order to reduce the presence of outliers. **Methods**: Applying depth techniques in both outlier-free and outlier-containing scenarios, the data sets starsCYG and delivery time data are utilized to determine the measure of location. Various classical and robust data depth procedures are used to find the location parameters, namely Mahalanobis depth, Tukey's half space depth, Projection depth, Zonoid depth, Spatial depth, and L2 Depth (Euclidean Depth). Distance-Distance plot is used for identifying the outliers. Further, it has been researched how well the data depth processes work by computing the parameters under actual and simulation environments, with and without outliers by considering different levels of contaminations (0%, 1%, 2%, 3%, 5%, 10%, 20%, 30%, 40%). **Findings:** From the two data sets studied, Halfspace depth and Euclidean (using MCD estimator) give the same location parameters if the anomalies are present. These two procedures work equally well and more effectively than the others. The robust depth procedures work well if the outliers are present in the datasets and from the simulation study it can handle a certain level of contamination present in the data set. **Novelty:** Any dataset that contains outliers makes analysis results risky. Robust statistical techniques can tolerate some getting contaminated. According to the study, even if the data contains outliers, the Depth processes employing robust estimators can withstand a certain amount of contamination and still produce accurate findings.

Keywords: Location; Data Depth; Outliers; Mahalanobis Distance; Robust

1 Introduction

Identifying the center of a data cloud and calculating how close a particular data point is to the center are common statistical tasks that include measures of centrality. We are interested in the subject of how central an observation is in a probability distribution in a probabilistic context. Outlyingness is the antithesis of centrality. The Mahalanobis distance is a commonly used measure of outlyingness in multivariate statistics. This is the standard Euclidean distance used to transform "whitened" data using a scatter matrix and a center point. Since the early 1990s, more general statistics for calculating the centrality and outlyingness of data in R^d as well as for locating the center portions of a data cloud made up of points with at least a certain degree of centrality have been created. Data depth is one of the main concepts used to measure the deepest point[1]. This concept is emerging nowadays and is used in the field of non-parametric statistics. Data depth measures how deep a given data is located in the entire data cloud. It is also defined as the position of the sample point to the entire data or the position of the sample point concerning the probability distribution. Various notions of depth preliminaries have been established in the literature by several researchers. Relying upon this depth functions, procedures of signs and ranks, median, order statistics, and outlyingness indicators might be simply expanded from their basic versions[2]. Data exploration, asymptotic distributions, ordering, and robust estimation are key statistics areas where data depth is essential[3,4]. Data depth procedures in both univariate and multivariate are based on the idea of ranks. It results in an ordering of observations from the center outward as opposed to the usual ranking from smaller to largest[5,6].

The depth of a point $X \in R^p$ should not depend on the fundamental coordinate system or in particular the scale of the fundamental measurement. For any distribution with a uniquely distinct center, the depth function should attain a maximum value at its center[7]. As a point moves away from the deepest point, along a fixed line through the center, the depth at X should decrease monotonically for any x having the deepest point[8]. The depth of point x should be moved toward zero when the observation is far away from the data cloud.

This paper mainly focused on computing location parameters based on depth procedures. Calculates a depth value for each observation and the deepest point—the one with the highest depth value—is subsequently taken into account as a location. In this article, various depth procedures, including classical and robust methods are discussed and examined by computing location parameters in both outlier-free and outlier-containing scenarios for real and simulated environments, taking into account varying levels of contamination. The conventional measure of location (sample mean vector) is incredibly data-sensitive when the data deviate from the model assumptions. Many robust alternatives are set up to assess the measure of location. These are mainly based on the concept of robust estimation and have been created in recent years. The effectiveness of the technique is examined in real and simulations (level of contaminations: 0%, 1%, 2%, 3%, 5%, 10%, 20%, 30%, and 40%), and the results are given in Section 3 and the appendix.

2 Methodology

To compare inference methods based on depth and evaluate measures such as location and scale many graphical and quantitative methods are established. Throughout the past few decades, various depth theories have been put forth. The well- well-known techniques, including the Mahalanobis Depth[9], Half-Space Depth[10], Projection Depth[11,12], Zonoid Depth[13], Spatial Depth[14], L2 depth (Euclidean Depth)[15,16], and Euclidean Depth using MCD estimator[17] that are simply described in this section.

2.1 Mahalanobis Depth

Mahalanobis depth was first described by Liu et al. (1993) from Mahalanobis distance. Mahalanobis (1936) established the statistical idea of generalized distance which is calculated by using a classical mean vector and covariance matrix. For determining the Mahalanobis depth of observation, the Mahalanobis distance is used. The positive inverse of Mahalanobis distance is termed as Mahalanobis depth. For an observation y $\in S_n \subset R^d about a - dimensional data$,, Mahalanobis depth is specified as

$$MD(Y,\ Sn) = [1 + (Y - \bar{Y})S^{-1}(Y - \bar{Y})]^{-1} \tag{1}$$

where $\bar{Y}$ and S are the mean vector and dispersion matrix of S_n . Since it is reliant on non-robust parameters like the mean and dispersion matrix, this algorithm lacks to be reliable.

2.2 Half-space Depth

The concept of half-space depth was introduced by Tukey (1975). Suppose dividing a certain number y into two parts: each point equal to or less than y is considered as a closed half-space, and each value less than y is an open half-space. Similarly, all values equal to or greater than y as closed halfspace. For $d > 1$, the lowest depth of every one-dimensional projection of the dataset can be used to define the halfspace depth. The procedure for computing halfspace depth along with numerical illustration has been given by Xiongtao (2022).

That is the smallest number of observations in a closed half-space with a border through Y, which is defined as the halfspace depth of an observation regarding a d-dimensional data set Sn. Halfspace depth of a point $\mathrm{y} = (\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_n) \in S_n \subset R^d relativetod - dimensionaldatasetS_n$ is, given by

$$\mathrm{P}(z \mid Y) = \inf\{P[Y \in H], closedhalfspace, z \in H\} \tag{2}$$

Here the median is the deepest point. As the point that has the highest depth value in the multivariate scenario, the concept of median can be used widely. The Tukey median is referred to as the multivariate median. Tukey depth is another name for Half Space depth.

2.3 Projection Depth

Projection depth is initiated by Liu et.al (1993). It is based on the outlyingness and further, it was explored by Zuo and Serfling (2000). This procedure reflects the projection pursuit methodology and involves supremum over infinitely numerous direction vectors hence the computation of projection depth appears intractable. Initially, classical location and scale can be used to compute projection depth. Later it was replaced by a robust measure such as median and Median Absolute Deviation (MAD).

Let $\mu(.)$ be the location and $\sigma($. be a scale for univariate observations. For a distribution function G, the outlyigness of observation y is given as

$$\mathrm{o}(\mathrm{y}, G) = \sup |(u, \mathrm{y}, G)| \|u \mid -1 \tag{3}$$

where $Q(u, \mathrm{y}, G) = \left(u^T \mathrm{y} - \mu\left(G_u\right)\right) / \sigma\left(G_u\right) andG_{u\,\mathrm{is}}\, thedistributionofu^T \mathrm{y}$.

For a multivariate d- dimensional data set with location $\mu(.)andscale\sigma(.)$, the description of the projection depth is

$$PD(y, G) = \frac{1}{1 + o(y, G)} \tag{4}$$

2.4 Zonoid Depth

The Zonoid depth was demonstrated by Koshevoy and Mosler (1996) and is equivalent to two times Tukey's data depth of a suitably modified distribution. Zonoid depth differs from the remaining ideas. This perception has many properties in which these concepts have some limitations most effective, Liu et al. (1993) for Tukey's and projection depth. The zonoid depth, $D_\mu(\mathrm{y}) ofanobservation \in \mathbb{R}^d isdefinedby,$

$$D\mu(\mathrm{y}) = \begin{cases} \sup\{\alpha : \mathrm{y} \in \mathrm{D}_\alpha(\mu)\}, & \text{if y} \in \mathrm{D}_\alpha(\mu) \text{ for some } \alpha \\ 0, & \text{elsewhere.} \end{cases} \tag{5}$$

Depth of an observation y is the maximum height $\alpha andwhere \alpha \in \mathrm{proj}_\alpha \hat{Z}(\mu)$.

Now $NowD_\alpha(\mu) = \frac{1}{\alpha} \mathrm{proj}_\alpha(\hat{Z}(\mu))$

where $0 < \alpha \leq 1$.

Moreover, the depth of y is equal to one if y is the expectation and zero if y sits exterior $(\mu) forall \alpha. Intheeventthat \alpha > 0, (\mu)$ is the collection of each observation with data depths that are either greater or equal to α .

2.5 Spatial Depth

Vardi and Zhang (2000) formulated Spatial depth from the idea of spatial quantiles which was introduced by Chaudhuri (1996). It was extended by Zuo and Serfling (2000) and is given by, Supposes G be a cumulative distribution function of a d- dimensional random vector X. Formerly the multidimensional spatial depth of $y \in \mathbb{R}^d$ is therefore outlined I n the following manner.

$$SD(x, G) = I - \|\mathrm{E}(\mathrm{y} - \mathrm{x})\|_{\mathrm{e}} \tag{6}$$

where $\|.\|_e$ is the Euclidean model in d-dimensional space.

The concept of spatial (sometimes referred to as geometric) quantiles for multidimensional data acts as a basis for the depth function known as spatial depth. L1-depth is another name for this spatial depth.

2.6 Euclidean Depth (L_2Depth)

Euclidean Depth was introduced by Zuo and Serfling (2000). According to a point's mean outlyingness, the Euclidean Depth, $D_{L2} ismeasuredbytheL_2$distance and is given by

$$D_z(v \mid Y) = (1 + \mathrm{E}\|v - Y\|)^{-1} \tag{7}$$

holds $\alpha_{\max} = 1. Thedepthliftis \hat{D}^{L_2}(Y) = \{(\alpha, v) : \mathrm{E}\|v - \alpha Y\| \leq 1 - \alpha\}$ and curved.

It is provided for the distribution of n observations,

$$D_{L_2}(v \mid y_1, \ldots, y_n) = \left(1 + \frac{1}{n}\sum_{i=1}^{n} \|v - y_i\|\right)^{-1} \tag{8}$$

L2 depth is extreme at its spatial median of Y and disappears towards infinity. The extreme depth value is obtained at the center whereas if the distribution is completely symmetrical, in which case the spatial median serves as the center. The triangle inequality is the sole source of the deepest point's monotonicity, as well as the compactness and convexity of the center areas. Further, Euclidean Depth depends continuously on v. Additionally, depth merges in the probability distribution when considering a weakly convergent along with an identically integrable sequence. $P_n \to P itholds \lim_n D_{L_2}(v \mid P_n) = D_{L_2}(v \mid P)$.

Yet, the ordering of scatter produced by the Euclidean Depth is not an appropriate ordering of scatter because it violates its dilation order. Euclidean Depth is also known L2 depth.

2.7 Euclidean Depth (using MCD estimator)

The Euclidean Depth is not affine invariant, although it is invariant under rigid Euclidean motions. The following is an affine invariant version when faced with a positive definite $d \times dmatrix$H,, and H-norm is given in the form

$$\|v\|_H = \sqrt{v' H^{-1} v}, v \in \mathbb{R}^d \tag{9}$$

Let S_Y be a positive definite $d \times d$ matrix that continuously reflects on the allocation and assesses the dispersion of Y in an affine equivariant manner. Later indicates that

$S_{Yb+d} = BS_Y B$ holds for some matrix B of full rank and some d.

The affine invariant form of L_2depth is taken for more robust choices for S_Y, the covariance matrix, and modified L_2 depth is

$$D_{L_2}(v \mid Y) = (1 + E\|v - Y\|s_x)^{-1}$$

It shares the same characteristics as the L$_2$depth aside from invariance. The covariance matrix $\Sigma_X ofXisastraightforwardoptionforS_X$,Zuo, and Serfling (2000). The minimum covariance determinant (MCD) in addition to minimum volume ellipsoid (MVE) estimators are more reliable options for S_X , Rousseeuw et al. (1987), Lopuha et al. (1991).

3 Results and Discussion

A comparison of various depth procedures to find location parameters of two real data sets namely starsCYG and delivery time is enlisted in Tables 1 and 2 respectively. Also, a simulation study was done for the same. The depth procedures discussed above are conducted in real and simulation environments. Various notions of depth procedures are used to locate the deepest point, both with and without outlier conditions. Comparing various depth functions, robust estimators can withstand some contamination and still provide accurate findings, even when the data contains outliers. These estimators are employed in the depth procedures to identify location parameters.

3.1 Real Data

Real data from starsCYG, Rousseeuw, and Leroy (1987, pp.27) is taken into consideration to examine the effectiveness of the aforementioned depth techniques. Two variables, the logarithm of the effective surface temperature of each star and the logarithm of its light intensity are included in the dataset, which contains 47 stars in the direction of Cygnus. Found that there are 4 extreme outliers and 2 potential outliers using a distance-distance plot. Table 1 provides an overview of the highest depth value and corresponding observation under various depth processes (with and without outliers).

Table 1. Maximum depth value and observation number under various depth procedures of stars CYG data

Methods	MD	HSD	PD	ZD	SD	L2d(T)	L2d(mcd)
Depth value	0.941 (0.924)	0.382 (0.372)	0.659 (0.605)	0.891 (0.821)	0.960 (0.850)	0.465 (0.449)	0.425 (0.328)
Observation no.	25 (28)	28(28)	25(42)	25(28)	25(42)	25 (25)	28(28)

(.) - Without outliers

From Table 1 it is noticed that Halfspace depth and L2 depth (using the MCD estimator) give the same measure of location in both cases, the 28th observation. These two procedures work equally well and more effectively than the others. L2 depth using the moment estimator gives the 25th observation as a measure of location in both cases. Other depths namely Mahalanobis, Projection, Zonoid, and Spatial give different locations corresponding maximum depth values.

Delivery Time Data from Montgomery and Peck (1982), which is multivariate real data, is taken into consideration to study the effectiveness of the mentioned depth methods. The data set contains 25 observations and 3 variables namely the number of products, the distance walked by the route driver, and Delivery time. Found that there are 6 outliers using distance-distance plot, the two are extreme outliers and the other 4 are considered as potential outliers. Table 2 displays the greatest depth value and observation number for several depth techniques with and without outliers.

Table 2. Maximum depth value and observation number under various depth procedures of Delivery time data

Methods	MD	HSD	PD	ZD	SD	L2d(T)	L2d(mc d)
Depth value	0.934 (0.932)	0.4 (0.3)	0.753 (0.517)	0.771 (0.683)	0.858 (0.78)	0.36 (0.319)	0.408 (0.395)
Observation no.	15 (6)	6 (6)	6(6)	15 (17)	15 (6)	17 (6)	6 (6)

(.) - Without outliers

Table 2 shows that Halfspace depth, projection depth, and L2 depth (using MCD estimator) give the same measure of location in both cases, 6th observation with and without outliers. These techniques perform similarly well as well as being more efficient than the others. Other depths namely Mahalanobis, Zonoid, Spatial, and L2 depth using moment estimator give different locations corresponding maximum depth values.

3.2 Simulation Data

The different depth procedures have all been put through simulation testing. The experiments were carried out by computing the maximum depth values that correspond to location measurements and the observation numbers that belong to maximum depth values under simulation studies while considering various levels of contamination. First data is generated with mean vector μ= (0, 0), covariance matrix Σ=I2 for sample size, n=100. Further same experiments were performed under various levels of contaminations, such as ε= 0%, 1%, 2%, 3%, 5%, 10%, 20%, 30%, and 40%. (For Location μ= (4, 4), Σ=I2, Scale μ= (0, 0), Σ=1.5I2, Location, and Scale, μ= (4, 4), Σ=1.5I2 are taken into account, and the outcomes are compiled in the tables titled Tables 3, 4 and 5. It is concluded that half space depth and L2 depth using the MCD estimator can tolerate certain levels of contamination and gives the same deepest point. Other depth procedures fail to provide identical location measurements even if the data contamination is very low.

Table 3. Measure of location and deepest points under Location contamination

Error	MD	HSD	PD	ZD	SD	L2D	L2D(mcd)
0.00	0.968 (-.177,.169) 41	0.42 (.177,.169) 41	0.783 (.177,.169) 41	0.876 (-.177,.169) 41	0.917 (-.177,.169) 41	0.445 (-.177,.169) 41	0.442 (-.177,.169) 41
0.01	0.952 (-.177,.169) 41	0.42 (.177,.169) 41	0.762 (.177,.169) 41	0.813 (-.294, -.465) 20	0..915 (-.177,.169) 41	0.426 (-.177,.169) 41	0.442 (-.177,.169) 41

Continued on next page

Table 3 continued

0.02	0.947 (-.177,.169) 41	0.42 (.177,.169) 41	0.778 (.177,.169) 41	0.782 (-.294, -.465) 20	0..925 (-.294, .465) 20	0.426 (-.177,.169) 41	0.442 (-.177,.169) 41
0.03	0.939 (-.177,.169) 41	0.42 (.177,.169) 41	0.777 (.177,.169) 41	0.784 (-.294, -.465) 20	0..904 (-.177,.169) 41	0.417 (-.177,.169) 41	0.442 (-.177,.169) 41
0.05	0.905 (.235,.33) 90	0.42 (.177,.169) 41	0.738 (.177,.169) 41	0.718 (0.414,0.183) 82	0.845 (0.235,0.33) 90	0.408 (0.235,0.33) 90	0.442 (-.177,.169) 41
0.1	0.849 (.235,.33) 90	0.42 (.177,.169) 41	0.670 (.177,.169) 41	0.698 (-.294, -.465) 20	0.817 (0.235,0.33) 90	0.398 (-.006,.485) 36	0.442 (-.177,.169) 41
0.2	0.741 (.414,0.183) 82	0.41 (.177,.169) 41	0.590 (.414,0.183) 82	0.690 (.414,0.183) 82	0.768 (.414,0.183) 82	0.388 (.414,0.183) 82	0.442 (-.177,.169) 41
0.3	0.664 (.235,.33) 90	0.39 (.414,.183) 82	0.53 (.235, -0.33) 90	0.677 (-0.006,0.485) 36	0.699 (.235, -0.33) 90	0.377 (-.006,.485) 36	0.442 (-.177,.169) 41
0.4	0.579 (.235,.33) 90	0.38 (-.294,.465) 20	0.534 (.235, -0.33) 90	0.661 (.235,0-.33) 90	0.647 (-.294, .465) 20	0.365 (.235, -0.33) 90	0.405 (.235,.33) 90

(First value indicates the highest depth value, (.) indicates measure of location and the last value indicates the observation number).

Table 4. Measure of location and deepest points under Scale contamination

Error	MD	HSD	PD	ZD	SD	L2D	L2D(mcd)
0.00	0.944 (.053,.351) 27	0.36 (0.053,0.351) 27	0.657 (0.053,0.351) 27	0.856 (0.053,0.351) 27	0.863 (.053,.351) 27	0.435 (.053,.351) 27	0.442 (.053,.351) 27
0.01	0.933 (.053,.351) 27	0.36 (0.053,0.351) 27	0.651 (0.053,0.351) 27	0.848 (0.053,0.351) 27	0.846 (.053,.351) 27	0.418 (.053,.351) 27	0.442 (0.053,.351) 27
0.02	0.924 (.053,.351) 27	0.36 (0.053,0.351) 27	0.629 (0.053,0.351) 27	0.863 (0.053,0.351) 27	0.858 (0.053,.351) 27	0.412 (.053,.351) 27	0.442 (.053,.351) 27
0.03	0.920 (.053,.351) 27	0.36 (0.053,0.351) 27	0.662 (0.053,0.351) 27	0.808 (0.053,0.351) 27	0.877 (.053,0.351) 27	0.411 (0 .053,.351) 27	0.442 (.053,.351) 27
0.05	0.902 (.053,.351) 27	0.36 (0.053,0.351) 27	0.626 (0.053,0.351) 27	0.808 (0.053,0.351) 27	0.863 (0.053,.351) 27	0.405 (.053,.351) 27	0.442 (0.053,.351) 27
0.1	0.851 (.053,.351) 27	0.34 (0.053,0.351) 27	0.594 (0.053,0.351) 27	0.795 (0.053,0.351) 27	0.863 (0.053,.351) 27	0.400 (.053,.351) 27	0.442 (0.053,.351) 27
0.2	0.747 (.053,.351) 27	0.33 (0.053,0.351) 27	0.544 (0.053,0.351) 27	0.746 (0.053,0.351) 27	0.754 (.053,0.351) 27	0.383 (.053,.351) 27	0.442 (0.053,.351) 27
0.3	0.648 (.053,.351) 27	0.36 (0.053,0.351) 27	0.501 (0.053,0.351) 27	0.713 (0.053,0.351) 27	0.719 (.053,.351) 27	0.373 (.129, - .162) 70	0.442 (.053,.351) 27
0.4	0.582 (.612, -.162) 15	0.34 (0.053,0.351) 27	0.500 (0.573,0.607) 27	0.603 (-0.108, -.250) 78	0.641 (.129, -.162) 70	0.363 (.181, .138) 70	0.418 (.181, .138) 70

Table 5. Measure of location and deepest points under Location-Scale contamination

Error	MD	HSD	PD	ZD	SD	L2D	L2D(mcd)
0.00	0.984 (.184,.023) 34	0.44 (-.184,.023) 34	0.771 (-.184,.023) 34	0.893 (-.184,.023) 34	0.945 (-.184,.023) 34	0.446 (-.184,.023) 34	0.442 (-.184,.023) 34
0.01	0.980 (-.184,.023) 34	.44 (.184,.023) 34	0.713 (-.184,.023) 34	0.832 (-.103,.327) 64	0.94 (-.184,.023) 34	0.405 (-.184,.023) 34	0.441 (-.184,.023) 34
0.02	0.974 (-.184,.023) 34	0.46 (-.184,.023) 34	0.767 (-.184,.023) 34	0.870 (.038,.189) 12	0.94 (-.184,.023) 34	0.421 (-.184,.023) 34	0.441 (-.184,.023) 34
0.03	0.966 (-.184,.023) 34	0.42 (-.184,.023) 34	0.767 (-.184,.023) 34	0.870 (.038,.189) 12	0.939 (-.184,.023) 34	0.414 (-.184,.023) 34	0.441 (-.184,.023) 34
0.05	0.947 (-.184,.023) 34	0.43 (-.184,.023) 34	0.742 (-.184,.023) 34	0.893 (-.355,.276) 34	0.932 (-.184,.023) 34	0.418 (-.184,.023) 34	0.441(- .184,.023) 34
0.1	0.870 (.038,.189) 12	0.42 (-.184,.023) 34	0.714 (-.184,.023) 34	0862 (-.362,.327) 98	0.893 (-.184,.023) 34	0.402 (-.184,.023) 34	0.441 (-.184,.023) 34
0.2	0.791 (-.184,.023) 34	0.42 (-.184,.023) 34	0.642 (-.184,.023) 34	0.800 (-.184,.023) 34	0.835 (-.184,.023) 34	0.387 (-.184,.023) 34	0.441 (-.184,.023) 34
0.3	0.695 (.060,.128) 50	.46 (-.184,.023) 34	0.588 (-.103,.327) 64	0.838 (-.103,.327) 64	0.808 (-.103,.327) 64	0.378 (-.103,.327) 64	0.432 (.038,.189) 12
0.4	0.587 (-.342,.100) 81	0.44 (.038,.189) 12	0.870 (.038,.189) 12	0.812 (-.103,.327) 64	0.798 (-.103,.327) 64	0.378 (-.184,.023) 34	0.427 (-.342,.100) 81

4 Conclusion

Robust estimators used in the depth procedures to find location parameters may resist a certain level of contamination and still give correct results, even if the data contains outliers. The depth procedures such as halfspace depth and Euclidean Depth using MCD estimator give better results compared with other depth procedures. By finding the deepest point in a dataset instead of relying on a more conventional method of determining location, the research groups can find the best location parameter with greater precision when using these methods.

References

1) Geenens G, Nieto-Reyes A, Francisci G. Statistical depth in abstract metric spaces. *Statistics and Computing*. 2023;33(2). Available from: https://doi.org/10.1007/s11222-023-10216-4.
2) Mosler K, Mozharovskyi P. Choosing Among Notions of Multivariate Depth Statistics. *arXiv*. 2022;37(3):37–37. Available from: https://doi.org/10.48550/arXiv.2004.01927.
3) Dai X, Lopez-Pintado S. Tukey's Depth for Object Data. *Journal of the American Statistical Association*. 2021;2022. Available from: https://doi.org/10.48550/arXiv.2109.00493.
4) Liu X, Li Y. General Notions of Regression Depth Functions. *Statistics Methodology*;2020. Available from: https://arxiv.org/abs/1710.03904.
5) Priya M, Karthikeyan M. Outlier Detection Technique for Multivariate Data: Using Classical and Robust Mahalanobis Distance Method. *Journal of Emerging Technologies and InnovativeResearch*;2019(6):766–769. Available from: https://www.jetir.org/papers/JETIR1906J50.pdf.
6) Muthukrishnan R, Gowri D, Ramkumar N. Measure of Location using Data Depth Procedures. *International Journal of Scientific Research in Mathematical and Statistical Sciences*;2018(5):273–277. Available from: https://doi.org/10.26438/ijsrmss/v5i6.273277.
7) Pawar SD, Digambar T, Shirke. Data Depth-Based Nonparametric Test for Multivariate Scales. *Journal of Statistical Theory and Practice*. 2022;p. 16–16. Available from: https://doi.org/10.1007/s42519-021-00236-6.
8) Vardi Y, Zhang CH. The multivariate L 1-median and associated data depth. *Proceedings of the National Academy of Sciences*. 2000;97(4):1423–1426. Available from: http://dx.doi.org/10.1073/pnas.97.4.1423.
9) Mahalanobis P. On the generalized distance in statistics. *Proceedings of the National Academy India*. 1936;p. 49–55. Available from: https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=1649365.
10) Tukey J. Mathematics and the picturing of data. In: Proceeding of the International Congress of Mathematicians. 1975. Available from: https://www.scirp.org/(S(351jmbntvnsjt1aadkozje))/reference/referencespapers.aspx?referenceid=2502659.

11) Liu X, Zuo Y. Computing projection depth and its associated estimators. *Statistics and Computing*. 2014;24(1):51–63. Available from: https://doi.org/10.1007/s11222-012-9352-6.
12) Muthukrishnan R, Ramkumar N. Computing Adjusted Projection depth using GSO algorithm. *AIP Conference Proceedings*. 2020. Available from: https://doi.org/10.1063/5.0016972.
13) Koshevoy G, Mosler K. Zonoid trimming for multivariate distributions. *The Annals of Statistics*. 1997;25(5). Available from: https://www.jstor.org/stable/2959013.
14) Chaudhuri P. On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association*. 1996;(434):862–872. Available from: https://doi.org/10.2307/2291681.
15) Rousseeuw PJ, Leroy AM. Robust Regression and Outlier Detection. New York. Wiley. 1987. Available from: https://doi.org/10.1002/0471725382.
16) Lopuhaa HP, Rousseeuw PJ. Breakdown Points of Affine Equivariant Estimators of Multivariate Location and Covariance Matrices. *The Annals of Statistics*. 1991;19(1):229–248. Available from: https://doi.org/10.124/AOS/1176347978.
17) Zuo YJ, Serfling R. General notions of statistical depth function. *The Annals of Statistics*. 2000;28(2):461–482. Available from: https://doi.org/10.124/aos/1016218226.