



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Avanzi, B;Lavender, M;Taylor, G;Wong, B

Title:

Detection and treatment of outliers for multivariate robust loss reserving

Date:

2024-03-24

Citation:

Avanzi, B., Lavender, M., Taylor, G. & Wong, B. (2024). Detection and treatment of outliers for multivariate robust loss reserving. *Annals of Actuarial Science*, 18 (1), pp.102-125.  
<https://doi.org/10.1017/S1748499523000155>.


Persistent Link:

<https://hdl.handle.net/11343/344351>

License:

[cc-by](#)

# Detection and treatment of outliers for multivariate robust loss reserving

Benjamin Avanzi<sup>1</sup> , Mark Lavender<sup>2</sup>, Greg Taylor<sup>2</sup> and Bernard Wong<sup>2</sup>

<sup>1</sup>Department of Economics, Centre for Actuarial Studies, University of Melbourne, Melbourne, VIC, Australia; and <sup>2</sup>School of Risk and Actuarial Studies, UNSW Australia Business School, UNSW Sydney, Sydney, NSW, Australia

**Corresponding author:** Benjamin Avanzi; Email: [b.avanzi@unimelb.edu.au](mailto:b.avanzi@unimelb.edu.au)

(Received 22 December 2022; revised 15 June 2023; accepted 30 June 2023; first published online 24 August 2023)

## Abstract

Traditional techniques for calculating outstanding claim liabilities such as the chain-ladder are notoriously at risk of being distorted by outliers in past claims data. Unfortunately, the literature in robust methods of reserving is scant, with notable exceptions such as Verdonck & Debruyne (2011, *Insurance: Mathematics and Economics*, 48, 85–98) and Verdonck & Van Wouwe (2011, *Insurance: Mathematics and Economics*, 49, 188–193). In this paper, we put forward two alternative robust bivariate chain-ladder techniques to extend the approach of Verdonck & Van Wouwe (2011, *Insurance: Mathematics and Economics*, 49, 188–193). The first technique is based on Adjusted Outlyingness (Hubert & Van der Veen, 2008, *Journal of Chemometrics*, 22, 235–246) and explicitly incorporates skewness into the analysis while providing a unique measure of outlyingness for each observation. The second technique is based on *bagdistance* (Hubert *et al.*, 2016, *Statistics: Methodology*, 1–23) which is derived from the bagplot; however, it is able to provide a unique measure of outlyingness and a means to adjust outlying observations based on this measure.

Furthermore, we extend our robust bivariate chain-ladder approach to an N-dimensional framework. The implementation of the methods, especially beyond bivariate, is not trivial. This is illustrated on a trivariate data set from Australian general insurers and results under the different outlier detection and treatment mechanisms are compared.

**Keywords:** Robust loss reserving; bagplot; adjusted outlyingness; bagdistance; multivariate

**2020 Mathematics Subject Classification:** Primary: 91G70; 62P05

## 1. Introduction

### 1.1 Motivation

At any moment, the number, timing, and severity of future claims payments for a general insurer is shrouded in uncertainty. Reserves are set up to ensure, with a given degree of confidence, that necessary claims payments are met as they arise. The reserving problem is one of solvency; however, it is also one of capital efficiency; if an insurer holds reserves over and above what is needed, they forfeit the opportunity to utilize this capital elsewhere, and insurance becomes more expensive than necessary. It is critical that the results from models and techniques applied to the loss-reserving problem are as accurate as possible when tasked to a range of different data sets.

Some of these data sets may include abnormal observations; outliers; or deviations from model assumptions. Full inclusion of these data points in an analysis may prove detrimental to the accuracy of reserve estimates and the resulting inference. This is an issue that needs to be addressed if these models and techniques are going to reflect reality and be used to inform decisions.

*Robustness* refers to the ability of a model or estimation procedure to not be overtly influenced by outliers in the data set under investigation and/or deviations from the underlying assumptions of the model. Understanding how outliers may impact the results of a model and having statistically sound procedures to detect and treat abnormal observations will improve the robustness of reserving techniques and ultimately lead to more informed and reliable decisions.

While some authors have explored the issue of robustness in reserving, the body of literature in this area is relatively scant. Of particular importance for this paper is the robust GLM chain-ladder of Verdonck & Debruyne (2011) and the robust bivariate chain-ladder of Verdonck & Van Wouwe (2011); however, there has been notable work that moves away from the chain-ladder technique. For robust non-chain-ladder methods, please also see for example Chan & Choy (2003), Chan *et al.* (2008), Pitselis *et al.* (2015).

Verdonck *et al.* (2009) provide a two-stage deterministic robust chain-ladder technique which fundamentally relies on the analysis of residuals given after fitting an over-dispersed Poisson (ODP) GLM to the cumulative and then incremental claims data as described in England & Verrall (1999). A boxplot is employed on the Pearson residuals after fitting the ODP GLM at each stage to detect outlying observations. Under their robust technique, development factors are calculated as medians which are known to be much more robust than means. Unfortunately, in its original formulation, results were poor as reserves were still being heavily influenced by outliers.

The approach of Verdonck *et al.* (2009) was refined by Verdonck & Debruyne (2011). In particular, it was found that the standard threshold value of 1.345 to identify outliers was often too low for triangular loss data. To combat this, an additional stage was added to the methodology whereby the threshold point was taken to be the 75%-quantile of the residuals after an initial fit using the threshold value of 1.345 which led to better results in terms of robustness. Verdonck & Debruyne (2011) also showed that the influence function for reserves with respect to incremental claims is unbounded when assuming a Poisson GLM specification. This provides a formal basis for the non-robustness of the chain-ladder technique and related techniques. A comprehensive study of impact functions of central estimates, mean squared errors, and quantiles of reserves can be found in Avanzi *et al.* (2023).

### 1.2 Statement of contributions

The refined robust GLM chain-ladder technique is an integral component of the robust bivariate chain-ladder as developed by Verdonck & Van Wouwe (2011). In particular, the residuals given after fitting the robust GLM chain-ladder are used to generate a bagplot (Rousseeuw *et al.*, 1999) which may be considered as a bivariate boxplot. The second approach employed to detect and treat outliers is the minimum covariance determinant (MCD) (Rousseeuw, 1984) Mahalanobis distance, also applied to the residuals.

Each of these two techniques has shortcomings that this paper aims to address. In particular, they fail to take skewness of the data sufficiently into account, which may lead to the misclassification of data points between regular and outliers.

We introduce two alternative methodologies that offer a consistent and structured approach to the detection, measurement and treatment of outlying observations with statistical backing. Our methods are based on Adjusted Outlyingness (AO) (Hubert & Van der Veeken, 2008) and *bagdistance* (Hubert *et al.*, 2016). AO provides a unique measure of outlyingness for each observation and explicitly incorporates a robust measure of skewness into the detection process. *Bagdistance* is derived from the bagplot and provides a measure of outlyingness for each observation. Through calculation of the *bagdistance*, a greater variety of alternative treatments for outliers becomes available than when simply using the bagplot. These methodologies are applied and compared on real data. Improvements over the techniques introduced in Verdonck & Van Wouwe (2011) are shown to be material and significant.

Additionally, we extend the methodologies beyond the bivariate case. Multivariate reserving techniques primarily benefit from the consideration of the dependence structure between loss triangles. Such consideration should provide a more accurate estimation of reserves; however, the impact of outliers can still be significant in this setting. Furthermore, information about other lines of business may inform the identification and adjustment of outliers. Hence, appropriate detection techniques and adjustment procedures help ensure that the benefits of multivariate reserving analysis are not lost due to a lack of robustness. This is also a focus of this paper, which develops a multivariate robust reserving technique. We extend the bivariate chain-ladder to an N-dimensional framework and illustrate the implementation of all four outlier detection and treatment techniques on a trivariate example.

Implementation of the methodology is not trivial, and R codes are freely available on GitHub for any interested user; see section 5.

## 2. Outlier Detection Techniques

In section 2.1, we start by reviewing some basic concepts related to robust statistics, which are required to introduce four detection techniques in section 2.2.

### 2.1 Preliminary: general robustness concepts

#### 2.1.1 Influence functions and breakdown points

When considering robust methodologies, two common concepts are influence functions and breakdown points, which we define here for later reference:

**Influence functions** consider an estimation technique and calculate the potential ‘influence’ a single data point could have on that estimator. For robustness, it is desirable to have a bounded influence function such that no single data point can have an unlimited impact on estimators. An example of an estimator with an unbounded influence function is the mean, whereas the median has a bounded influence function.

**Breakdown points** describe the proportion of the data set required to be contaminated (or outlying, however this may be defined) before an estimator provides inaccurate results. For example, when calculating the mean only one data point has to be contaminated to render the technique invalid, however, when calculating the median, 50% of data points need to be contaminated to invalidate the technique.

#### 2.1.2 Halfspace depth and Tukey median

Both the bagplot (Rousseeuw *et al.*, 1999) and the bagdistance (*bd*) (Hubert *et al.*, 2016) are based on the concept of halfspace depth (Tukey, 1975).

The **halfspace depth** (Tukey, 1975) of a point is defined as the minimum number of points (from the data sample) in a closed halfplane through the point of interest. We refer to Fig. 1 which illustrates the concept of halfspace depth in the bivariate case (note that this is scalable to higher dimensions). In Fig. 1, to calculate the halfspace depth of the point marked with a green asterisk, we would consider numerous lines through this point (such as the lines denoted as  $L_1$  and  $L_2$  in the Figure) and look for the minimum number of points on either side of each of these lines. Romanazzi (2001) showed that halfspace depth has a *bounded influence function*. This means that the impact an outlier may have on the halfspace depth of a given observation is limited and highlights the robustness of this statistic.

The **depth median (Tukey median)  $T^*$**  is the point with the greatest halfspace depth. It represents a central point of the data and is marked by the red asterisks in Fig. 2. If the point is not unique, the center of gravity of the deepest region is used.

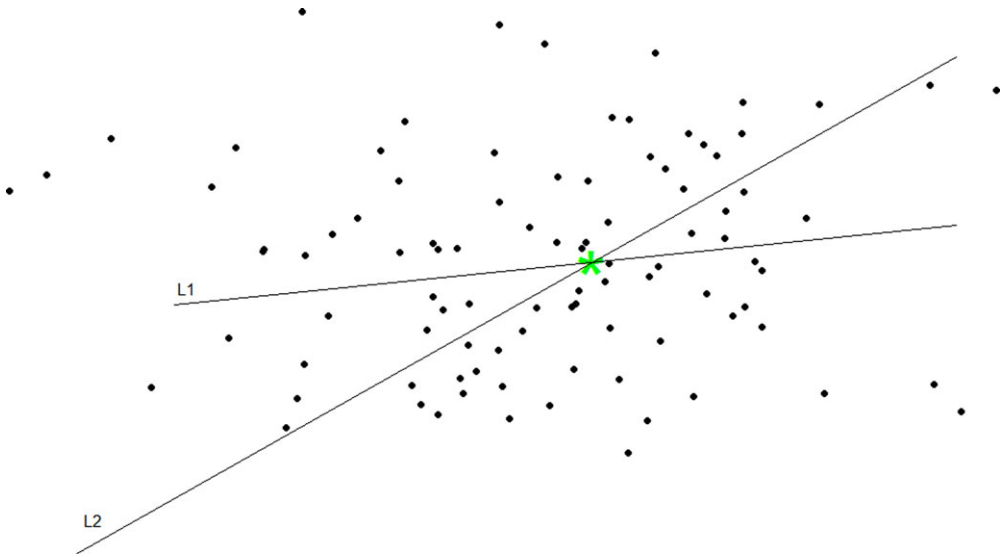


Figure 1. Halfspace depth illustration in 2 dimensions.

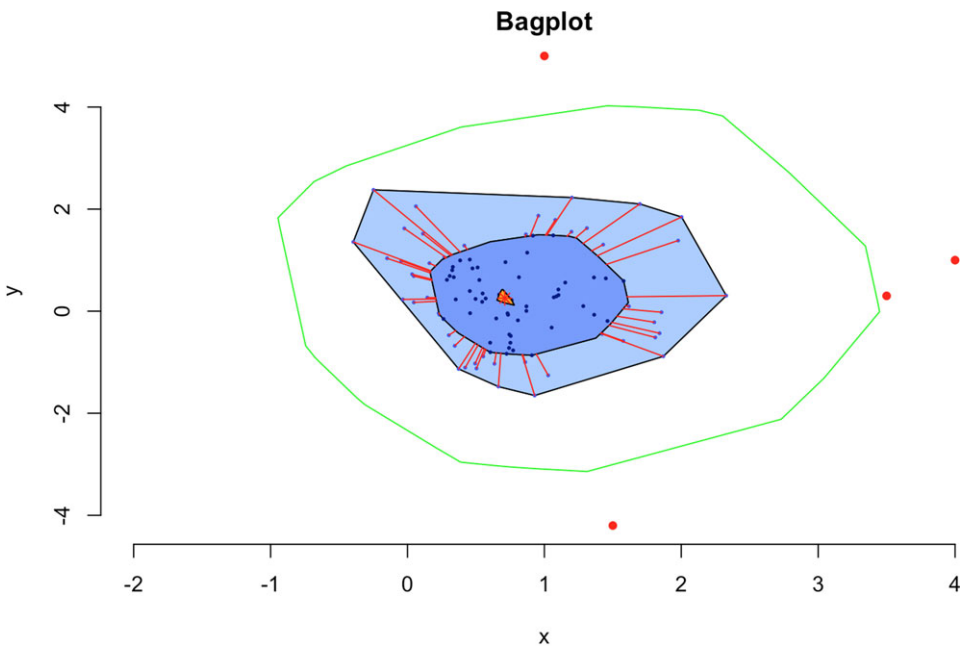


Figure 2. Bagplot with fence drawn.

2.1.3 MCD estimation procedure

A popular technique to estimate location and dispersion parameters in a robust fashion is the MCD (Rousseeuw, 1984) procedure, whereby one finds the

$$\left\lfloor \frac{n+p+1}{2} \right\rfloor \leq h \leq n \tag{1}$$

observations whose classical estimator of the covariance matrix has the smallest determinant where  $n$  represents the number of observations and  $p$  represents the dimension of the data. The variable  $h$  represents the number of observations that are ultimately used, and the procedure puts conditions around the minimum number of observations to avoid selection of a small set that does not truly represent the data. The location vector is then the arithmetic mean of these points, and the scale matrix is taken as a multiple of this covariance matrix. Heuristically, this procedure identifies a ‘core’ set of data points which are mostly related to one another.

### 2.1.4 Swamping and masking

Inappropriate outlier detection techniques may lead to swamping or masking. **Swamping** occurs when one classifies too many observations as outliers, that is, the impact outliers have on the technique being applied results in regular observations incorrectly appearing to be outliers. **Masking**, on the other hand, refers to the failure to classify outliers as such, that is, outliers have impacted the technique being applied in a way that ‘masks’ them from being detected as outliers.

## 2.2 Outlier detection techniques

A robust bivariate chain-ladder technique hinges on the detection and adjustment of bivariate outliers. Verdonck & Van Wouwe (2011) put forward two techniques to be used for this task: the MCD (Rousseeuw, 1984) **Mahalanobis Distance**, and the **bagplot** (Rousseeuw *et al.*, 1999). In this section, we outline these approaches and two different techniques; the **bagdistance** and **AO** – that can be applied to the multivariate reserving problem and which address some of the shortcomings of the former two.

### 2.2.1 MCD mahalanobis distance

A standard method used to detect outliers in multivariate analysis is to calculate some measure of distance of each data point from the center of the data. A popular measure is the **Mahalanobis Distance (MD)**, given by,

$$MD(\mathbf{x}_i) = \sqrt{(\mathbf{x}_i - \hat{\boldsymbol{\mu}})' \hat{\boldsymbol{\Sigma}}^{(-1)} (\mathbf{x}_i - \hat{\boldsymbol{\mu}})}, \quad (2)$$

where  $\hat{\boldsymbol{\mu}}$  represents the sample location vector and  $\hat{\boldsymbol{\Sigma}}$  represents the sample scale matrix. If these are not estimated in a robust manner, then outliers may fail to be detected due to masking and swamping effects. Essentially this means that outlying observations will influence the estimate of the central point and dispersion of the data leading to outliers themselves having small MD (masking) and/or non-outlying observations having high MD (swamping). To combat this, the MCD estimates of location and dispersion are used, and outliers are then flagged as observations that have a MD greater than a certain threshold. Note that  $MD^2 \sim \chi_p^2$  if the underlying data is normal.

### 2.2.2 Bagplot

We refer to Fig. 2 to outline the components of a **bagplot**.

The **bag (B)** is shown in Fig. 2 by the darker inner area surrounding  $\mathbf{T}^*$  and is constructed as follows. First, denote  $D_k$  as the region of all data points that have halfspace depth greater than  $k$  and  $\#D_k$  as the number of data points contained in this region. The bag is given by the linear interpolation with respect to  $\mathbf{T}^*$  of the two regions that satisfy  $\#D_k \leq \lfloor \frac{n}{2} \rfloor < \#D_{k-1}$ . Here, linear interpolation refers to the method of constructing new data points that are in the range of both of  $\#D_k$  and  $\#D_{k-1}$ .

The **fence** is then given by multiplying the bag by some factor relative to  $\mathbf{T}^*$ . Typically this factor is three and Rousseeuw *et al.* (1999) point out that this value was chosen based on simulations. However, in their recent R package, *mrfDepth*, Segaert, Hubert, Rousseeuw, Raymaekers, and Vakili (2017) have instead opted to use the square root of the 99th percentile of a chi-squared distribution where the degrees of freedom is equal to the number of dimensions being considered. This new fence factor was chosen after derivation of the distribution of the *bagdistance* (Hubert *et al.*, 2015); see also the next section. In the two-dimensional case, it is very close to three. We will use this newly developed fence factor throughout the paper.

When displaying a bagplot, the fence is generally not drawn (Rousseeuw *et al.*, 1999). Fig. 2, however, shows the same bagplot with the fence drawn in green, and we can see that the four outlying points marked in red are outside this fence. Note that the choice of this fence factor will directly influence both the number of outliers detected and how they are adjusted. Data points outside the fence are considered outliers and are adjusted to facilitate the application of loss-reserving techniques. The adjustment may be done in a purely graphical manner, that is, bringing outliers back to the fence or loop or a weighting function based on *bagdistance* may be employed.

The perimeter of the lighter blue area is known as the **loop** and is given by the convex hull of all non-outlying points. This is analogous to the whiskers in a univariate boxplot (Tukey, 1977).

### 2.2.3 Bagdistance

Now we present the **bagdistance** (*bd*) (Hubert *et al.*, 2016). This statistic provides a measure of outlyingness for each observation and hence does not rely solely on graphical representations. It can also handle skewness often found in loss data. However, similarly to the bagplot, the *bd* utilizes the bag to capture the shape of the data and subsequently detect outliers. As the bag is formulated based on circa 50% of the data points, there is potential that it does not fully encapsulate the skewness in the set. When this is the case, the ensuing outlier detection results may be flawed. To calculate the *bd*, firstly define  $\mathbf{c}_x$  as the intersections of the boundary of the bag ( $B$ ) and the ray from the Tukey median  $\mathbf{T}^*$  through the points  $\mathbf{x}$  (red lines in Fig. 2). The *bd* is defined as follows

$$bd(\mathbf{x}; P_n) = \begin{cases} 0 & \text{if } \mathbf{x} = \mathbf{T}^*; \\ \frac{\|\mathbf{x} - \mathbf{T}^*\|}{\|\mathbf{c}_x - \mathbf{T}^*\|} & \text{elsewhere,} \end{cases} \quad (3)$$

where  $P_n$  represents the distribution of the data set and  $\|\cdot\|$  is the Euclidean norm such that  $\|\mathbf{x}\| = \sqrt{x_1^2 + \dots + x_n^2}$ . The denominator scales the distance of the data point ( $\mathbf{x}$ ) to  $\mathbf{T}^*$ , relative to the dispersion of the bag (in that direction of the projected ray). A cutoff point is then set, such that data points with a *bd* beyond this threshold are considered outliers and are adjusted back to an appropriate point on the ray emanating from  $\mathbf{T}^*$  passing through  $\mathbf{x}$ . Note that if the cutoff value is chosen to be the same as the fence factor under the bagplot technique, then the same outliers will be detected.

### 2.2.4 Adjusted Outlyingness

**AO** (Hubert & Van der Veen, 2008) is based on an adjustment of the Stahel-Donoho estimate of outlyingness (Donoho, 1982) to explicitly incorporate skewness. Furthermore, it is based on robust estimates of location, scale, and skewness such that it achieves a theoretical breakdown point of 25%. Further highlighting its robustness, the Adjusted Outlyingness (AO) technique has a bounded influence function (Hubert & Van der Veen, 2008). The technique is applied by firstly considering a  $p$ -dimensional sample  $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$  where  $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$  and  $\mathbf{a} \in \mathbb{R}^p$ . The measure of AO for  $\mathbf{x}_i$  is given by

$$AO_i = \sup_{\mathbf{a} \in \mathbb{R}^p} AO^{(1)}(\mathbf{a}'\mathbf{x}_i, \mathbf{X}_n\mathbf{a}), \quad (4)$$

where

$$AO^{(1)}(x_i, X_n) = \begin{cases} \frac{x_i - \text{med}(X_n)}{w_2 - \text{med}(X_n)}, & \text{if } x_i > \text{med}(X_n), \\ \frac{\text{med}(X_n) - x_i}{\text{med}(X_n) - w_1}, & \text{if } x_i < \text{med}(X_n), \end{cases} \quad (5)$$

where  $w_1$  and  $w_2$  are the lower and upper whiskers of the skew-adjusted boxplot (Hubert & Vandervieren, 2008) applied to the data set  $X_n$  and  $\text{med}(X_n)$  denotes the median of that data set. The skew-adjusted boxplot for all the AO values is then constructed and those that are beyond the cutoff value

$$\text{cut-off} = Q_3 + 1.5e^{3MC}IQR \quad (6)$$

are declared as outliers, where  $Q_3$  represents the third quartile of the data, MC represents the medcouple (a robust measure of skewness; see Brys *et al.*, 2004), and where IQR is the interquartile range (i.e.,  $Q_3 - Q_1$ , where  $Q_1$  represents the first quartile of the data). We are only concerned with the upper cutoff value as we are performing the skew-adjusted boxplot technique on the measures of outlyingness, and hence, small results in this context are not of interest. An alternative AO cutoff value of  $\sqrt{\chi_{[99,p]}^2} \cdot \text{median}(AO)$  is given in *mrfDepth* (Segaert *et al.*, 2017), where  $p$  represents the dimension of the data,  $\chi_{[99,p]}^2$  is the 99th percentile of the chi-squared distribution with  $p$  degrees of freedom and AO represents the set of all AO values for the data set. This cutoff value aligns closely with the fence factor used for the bagplot approach. In this paper, we consider both cutoff values.

Under the AO methodology, not all univariate vectors  $\mathbf{a}$  can be considered; however, Hubert & Van der Veen (2008) note that taking  $m = 250p$  directions provides a good balance between ‘efficiency’ and computation time. From here, if  $p = 2$ , to visualize the bivariate data a version of the bagplot based on the AO values (rather than halfspace depth) may be constructed (Hubert & Van der Veen, 2008).

When constructing an AO-based bagplot, the bag is given by the convex hull of the 50% of data points with the smallest AO (note this is different from Fig. 2, which was based on halfspace depth). There are three possible approaches, which differ in the mechanism used to detect and hence treat outliers:

1. A fence is drawn by multiplying the AO-based bag by 3 (or some other factor) relative to the point with the lowest AO. Outliers are then flagged as those observations outside the fence, and the loop is the convex hull of all points within the fence.  
Since the fence is generated from the bag which only considers 50% of the data points, it may fail to fully capture the shape of the data and in particular the skewness in the set.
2. Utilize the alternative cutoff value given in *mrfDepth* where no fence is drawn.  
Here, the cutoff does not incorporate a robust measure of skewness and instead relies on the median value of the AOs and a quantile of the chi-squared distribution.
3. Outliers are flagged using the traditional cutoff value given by equation (6). In this case, the loop will be generated by the convex hull of all points with AO less than this value and no fence will be generated.

The traditional AO cutoff value incorporates a robust measure of skewness known as the medcouple which considers the whole data set. Hence, it is more equipped to capture the total skewness, and the loop that is generated by this approach more fully captures the skewness in the data in comparison to the fence methodology.

For the reasons explained above and unless stated otherwise, we will use the third approach.

### 3. An Extension of the Robust Bivariate Chain-ladder

In this section, we implement the AO and *bagdistance* (see section 2) outlier detection and treatment techniques in a bivariate reserving setting. We show that the four techniques often lead to different results (although the bagplot and *bagdistance* will be the same if you use the fence factor as the cutoff distance). The differences are not only in the number of outliers flagged but also in which observations are detected and hence treated. In each case, the different techniques should be implemented and results compared.

This section is motivated by some of the shortcomings of the current robust bivariate chain-ladder methodology. Firstly, the MCD Mahalanobis distance approach assumes elliptical symmetry of the multivariate data. If this assumption is not met, we may fail to detect outlying observations as well as falsely declare regular observations as outliers due to masking and swamping effects. The bagplot is better able to effectively visualize bivariate data, highlighting any correlation, skewness, and tail behavior.

We now present comparisons between these four outlier detection techniques when applied to real non-life insurance data.

#### 3.1 The robust bivariate chain-ladder

The robust bivariate chain-ladder technique was put forward by Verdonck & Van Wouwe (2011), and the general steps involved are as follows:

1. Apply the robust Poisson GLM chain-ladder technique (Verdonck & Debruyne, 2011) to each triangle separately. Obtain residuals from each triangle given by

$$r_{ij} = \frac{X_{ij} - \hat{\mu}_{ij}}{V^{\frac{1}{2}}(\mu_{i,j})}, \quad (7)$$

where  $X_{ij}$  represents the incremental claim for accident year  $i$  and development year  $j$ .

2. Store residuals from each triangle as bivariate observations in a  $n \times 2$  matrix  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)'$  where  $\mathbf{x}_i = (r_{kj1}, r_{kj2})$ . We have that  $n = I(I + 1)/2$ .
3. Apply either the bagplot or MCD Mahalanobis Distance (see section 2) to these bivariate residuals.
4. Adjust outliers. For the bagplot, outlying observations are brought back to the fence or loop. For the MCD Mahalanobis distance technique, observations are brought back to the tolerance ellipse representing the 95% quantile of the  $\chi^2_2$  distribution.
5. Backtransform adjusted residuals to obtain robust incremental claims  $X_{i,j}^{Rob}$ .
6. Apply the multivariate time series chain-ladder (Merz & Wüthrich, 2008) to the robust (adjusted) observations.

We now provide an example of the above framework on real bivariate data. We however include two alternative outlier detection and treatment techniques: the AO and *bagdistance*.

#### 3.2 Bivariate example using belgian data

The data for this example are from Belgian non-life insurers and are given in Appendix A.1 (from Shi *et al.*, 2012).

##### 3.2.1 Bagplot

We begin by illustrating the major shortcoming of the bagplot approach in that it does not provide a measure of outlyingness that is unique for each observation. This is because the bagplot approach is based on halfspace depth whereby multiple different observations can have the same

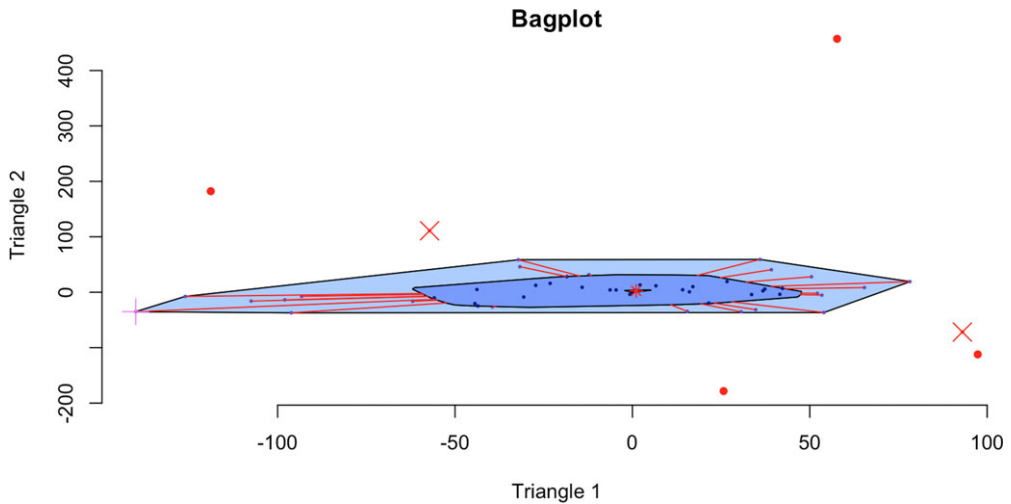


Figure 3. Bagplot.

halfspace depth even if they are at different levels of outlyingness. Fig. 3 shows the bagplot for this data, where 6 outliers have been detected. Four of these observations have the lowest halfspace depth of 1 which may be expected for outliers. However, two of these outliers,  $X_{6,5}$  and  $X_{7,3}$ , have a halfspace depth of 2 and 3, respectively. These points are marked with red crosses. Additionally, one non-outlying observation,  $X_{9,2}$ , has a halfspace depth of 1 and is marked with a purple ‘+.’ Hence without the corresponding bagplot, little inference can be made about how outlying the outliers are and importantly about whether an observation is outlying or not. This will become an even bigger concern when extending this technique to higher dimensions where graphical representations become less available. Indeed, the bagplot is purely based on a form of ranking (halfspace depth) (see section 2.2.2). As a result, without having the graph available (such as for higher dimensions as considered in the next section) it is difficult to communicate how outlying an observation is.

### 3.2.2 MCD mahalanobis distance

When employing the MCD Mahalanobis Distance technique, a graphical representation is also available in the bivariate case. In particular, we can plot each data point as well as tolerance ellipses of which have a squared distance to the central estimate of the data equal to a quantile of the  $\chi^2_2$  distribution. The classical tolerance ellipse is constructed when the Mahalanobis Distance is calculated using the classical estimators of the location vector and scale matrix rather than their robust counterparts from the MCD procedure outlined in section 2.1.3. Outliers are adjusted by a technique known as **bivariate Winsorization** such that an outlying observation  $\mathbf{x}$  is adjusted according to

$$\min \left( \sqrt{\frac{c}{MD^2(\mathbf{x})}}, 1 \right) \cdot \mathbf{x}, \tag{8}$$

where  $c$  is equal to the 95% quantile of a  $\chi^2_2$  distribution ( $\chi^2_{0.95,2}$ ). Under this methodology, 7 observations are detected as outliers. These are the same observations as was flagged under the bagplot approach as well as  $X_{9,2}$ , the only other observation with a halfspace depth of 1. Fig. 4 gives these plots before and after outliers have been adjusted.

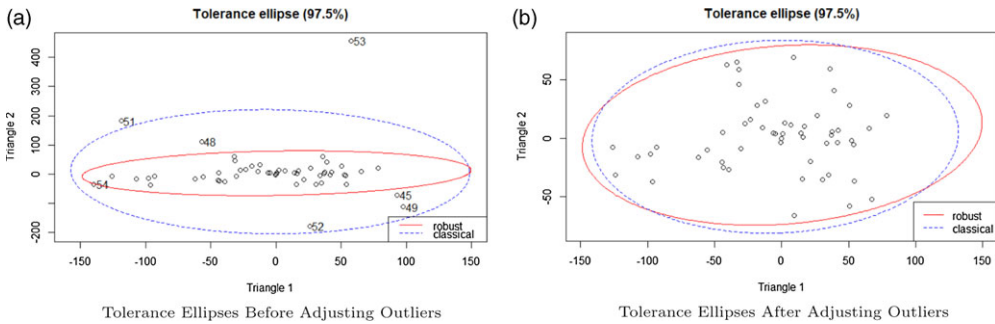


Figure 4. Tolerance ellipses.

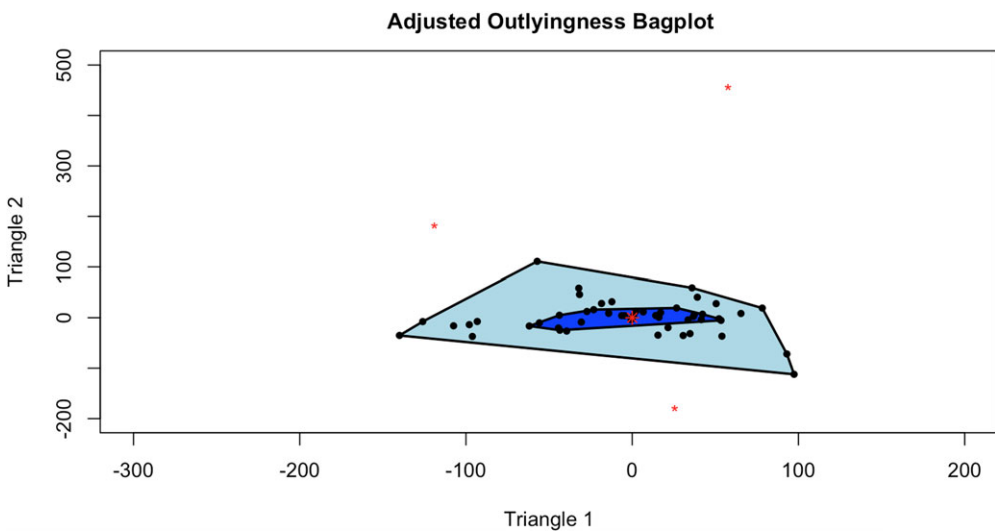


Figure 5. Adjusted outlyingness bagplot using traditional cutoff value (Approach 3).

### 3.2.3 AO

We now explore AO as an outlier detection and adjustment technique. AO explicitly incorporates a robust measure of skewness and provides a measure of outlyingness for each multivariate observation, alleviating the aforementioned issue with the bagplot based on halfspace depth. Additionally, it may be better equipped to handle more extreme levels of skewness. A bagplot based on AO is also available in the two-dimensional case as shown in Figs. 5 and 6, which illustrate all three methods outlined in section 2.2.4.

For these plots, the red asterisk is the point with the lowest AO and represents a central point of the data analogous to the Tukey median, the dark blue area represents the bag which contains the 50% of points with the lowest AO, and the light blue area represents the loop whose perimeter is constructed by the convex hull of all points not declared as outliers when using the traditional cutoff method (Hubert & Van der Veen, 2008).

In Fig. 5, we have used approach 3 with the traditional cutoff value given in Hubert & Van der Veen (2008) ( $\text{cut-off} = Q_3 + 1.5e^{3MC}IQR$ ) which incorporates a robust measure of skewness calculated from the whole data set and hence more fully considers the shape of the data to declare outliers, which are shown in red. This cutoff value was calculated to be 5.5813. The light blue area

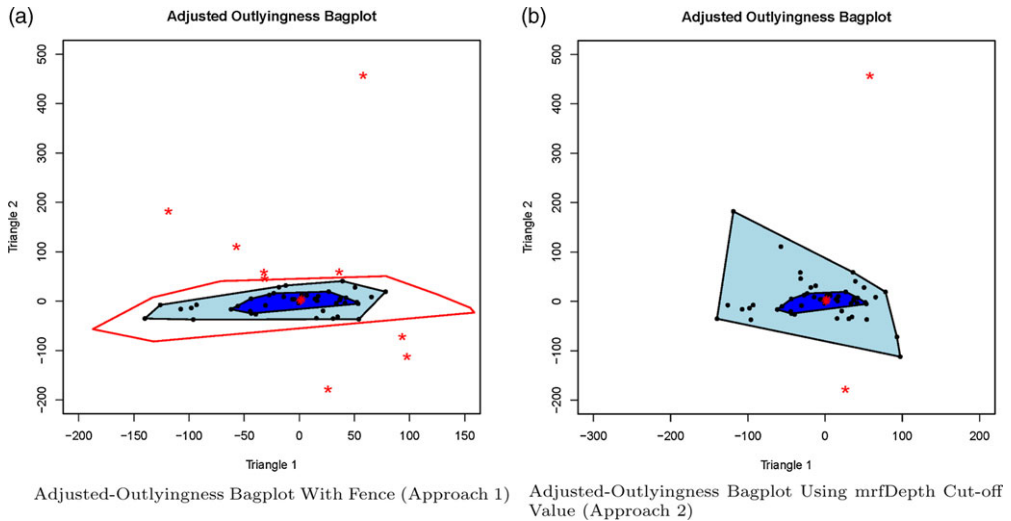


Figure 6. Adjusted outlyingness bagplots (Approaches 1–2).

represents the convex hull of all points not declared as outliers under this methodology and may be considered as an AO-based loop.

The fence approach (approach 1) only captures the shape of the data from the 50% of observations determined to be least outlying. Fig. 6(a) shows the AO bagplot with the fence drawn, and we see that under this approach an additional 6 observations are detected as outliers. A further issue that presents itself in this situation is the adjustment to the fence or the loop.

Finally, we may use the alternative cutoff value as given in the *mrfDepth* package (Segaert *et al.*, 2017) (Approach 2). The cutoff value under this approach is calculated to be 7.6227 compared with 5.5813 under the traditional calculation. An AO bagplot for this approach is shown in Fig. 6(b) where no fence is drawn. In this case, we detect 1 less outlier in comparison to the traditional case. Based on our previous arguments regarding the potential lack of total consideration of skewness under the two alternative approaches, we recommend use of the traditional cutoff value (Approach 3).

### 3.2.4 Bagdistance

The second outlier detection approach we propose is the *bagdistance* (*bd*) (see section 2.2.3) which utilizes the bag to capture the shape of the data and provides a distance measure for each observation to represent its outlyingness. A cutoff point is set such that observations with a *bd* beyond this threshold are classified as outliers and are adjusted back to an appropriate point on the ray emanating from  $T^*$  passing through  $x$ . An illustration of the *bd* is given in Fig. 7 where the *bd* for the outliers is given by taking the ratio of the orange line to the white line (noting that the orange line continues through to  $T^*$ ). If we choose the *bd* cutoff distance to be uniformly the same as the fence factor for a corresponding bagplot, we will detect the same outliers under the two approaches.

### 3.2.5 Summary of detection results

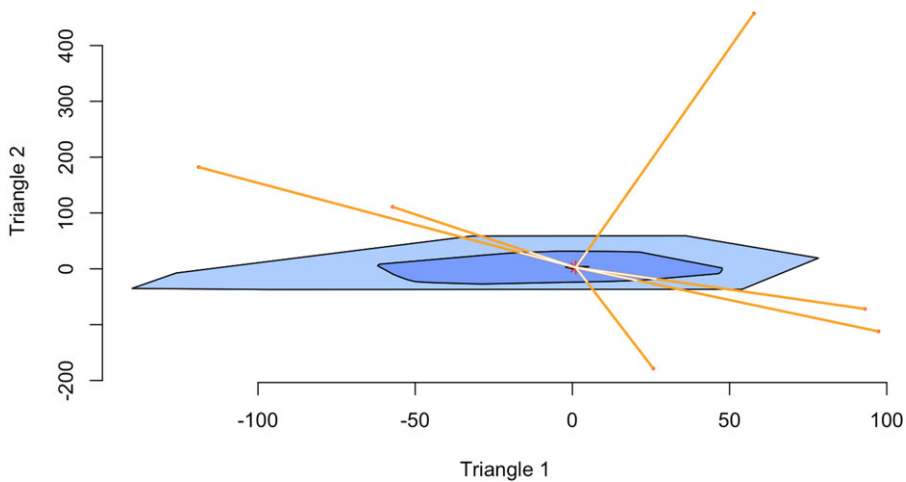
Table 1 summarizes the observations that were detected as outliers by the four techniques discussed here where a tick indicates that the observation was flagged under the relevant technique and a cross indicates that it was not. All observations not listed here were not flagged by any of the techniques. Further, the results in brackets correspond to the halfspace depth, MCD Mahalanobis distance, AO, and *bd* values for each observation, respectively.

**Table 1.** Outlier detection results.

Outliers	Bagplot	MCD	AO*	bagdistance**
$X_{6,5}$	✓ (2)	✓ (11.4205)	X (3.8184)	✓ (3.5744)
$X_{7,3}$	✓ (3)	✓ (17.3909)	X (3.7012)	✓ (3.9989)
$X_{7,4}$	✓ (1)	✓ (22.3643)	X (5.5598)	✓ (4.9284)
$X_{8,2}$	✓ (1)	✓ (50.4314)	✓ (6.2001)	✓ (6.8217)
$X_{8,3}$	✓ (1)	✓ (43.4263)	✓ (7.9189)	✓ (7.0153)
$X_{9,1}$	✓ (1)	✓ (262.9189)	✓ (15.7057)	✓ (15.7549)
$X_{9,2}$	X (1)	✓ (7.6540)	X (1.6197)	X (4.4161)

\*As detected by Approach 3 (traditional AO cutoff value).

\*\*Using cutoff distance of  $\sqrt{\chi_{0.99,2}^2}$ .



**Figure 7.** Bagdistance illustration.

### 3.3 Adjustment of outliers

In the previous section, we outlined how outliers may be detected. Of course, an equally crucial task is to determine how to treat those outliers. In this section, we discuss different methods for such treatment.

In some robust analyses, if a data point is extremely outlying, it is simply removed. Unfortunately, this option is not available in most triangular reserving techniques (which require a fully populated triangle), and hence, we must formulate and choose appropriate adjustment mechanisms.

#### 3.3.1 Bagplot based adjustments

Under the bagplot approach, Verdonck & Van Wouwe (2011) suggest that outlying observations should be brought back to the fence however upon inspecting their plots they appear to have brought observations back to the loop.

Fig. 8(a) and (b) show the bagplot with the fence drawn in green and the bagplot after adjusting residuals back to the fence, respectively. In Fig. 8(b), all of the adjusted outliers are within the loop; however, this is not always the case as after adjustment of the outlying observations, the bagplot methodology is performed again and with some data sets, some of the observations will

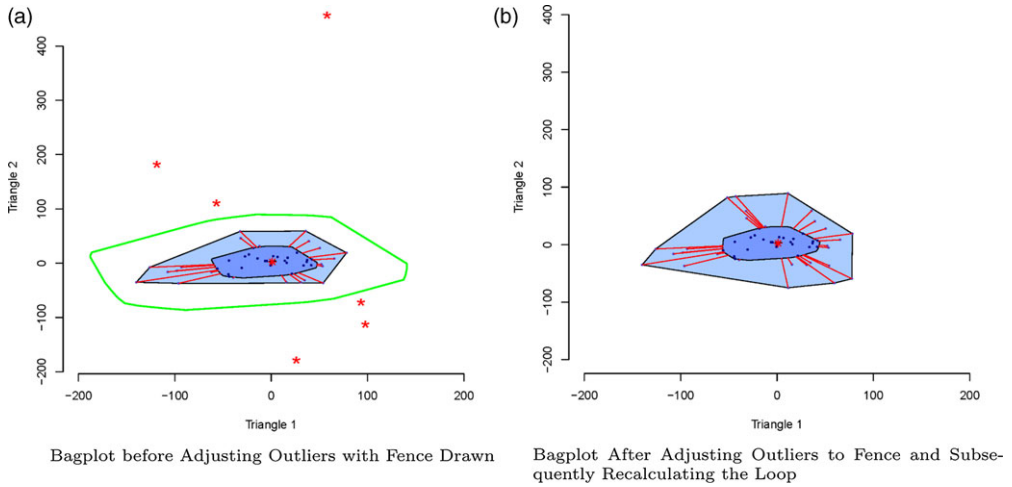


Figure 8. Bagplot before and after adjusting outliers to fence.

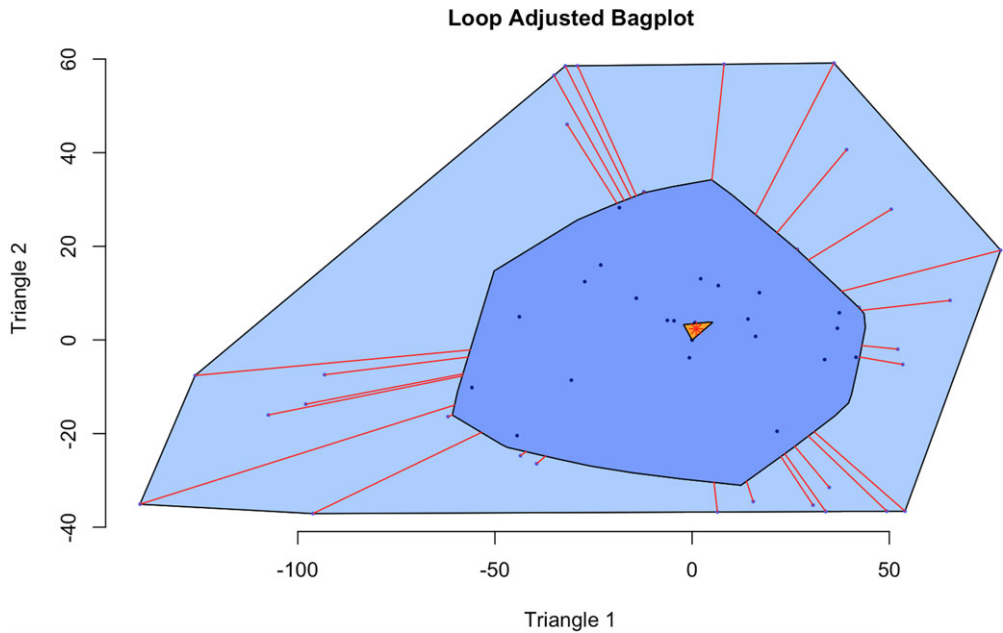


Figure 9. Bagplot after adjusting outliers to loop.

be outlying because the overall shape of the data has changed which will lead to a different bag and hence fence.

Fig. 9 shows the bagplot after adjusting outliers to the loop. Note that the adjustment in this case is more drastic in comparison to the fence adjustment.

### 3.3.2 MCD mahalanobis distance based adjustments

The adjustment methodology under the MCD Mahalanobis Distance approach is outlined in section 3.2.2.

### 3.3.3 Bagdistance based adjustments

Through calculation of the  $bd$  for each observation, a similar adjustment technique as was used in the MCD Mahalanobis example can be employed in that we adjust an outlying observation  $\mathbf{x}$  according to

$$\min\left(\frac{f}{bd}, 1\right)(\mathbf{x} - \mathbf{T}^*) + \mathbf{T}^*, \quad (9)$$

where  $f$  represents the factor of the bag that we wish to adjust outliers back to. For adjustment back to the bagplot-fence, we choose  $f = \sqrt{\chi_{0.99,2}^2}$ . Note that other adjustment functions may be employed.

For illustration purposes, we have considered two other adjustment functions such that, dependant on the level of outlyingness as measured by  $bd$ , observations are adjusted to differing degrees.

In the first case, all observations ( $\mathbf{x}$ ) are adjusted according to

$$\mathbf{x}^{rob} = \begin{cases} \mathbf{x}, & \text{if } bd_{\mathbf{x}} \leq f; \\ \min\left(\frac{f}{bd_{\mathbf{x}}}, 1\right)(\mathbf{x} - \mathbf{T}^*) + \mathbf{T}^*, & \text{if } f < bd_{\mathbf{x}} \leq \frac{1}{2}(2f + \sqrt{4f + 1} + 1); \\ \frac{f + \sqrt{bd_{\mathbf{x}}}}{bd_{\mathbf{x}}}(\mathbf{x} - \mathbf{T}^*) + \mathbf{T}^*, & \text{if } bd_{\mathbf{x}} > \frac{1}{2}(2f + \sqrt{4f + 1} + 1). \end{cases} \quad (10)$$

This means that moderate outliers will be brought back to the fence, whereas more extreme outliers will be adjusted to points beyond the fence according to their levels of outlyingness. Specifically, the  $\frac{1}{2}(2f + \sqrt{4f + 1} + 1)$  constraint is chosen as it is the value of  $bd_{\mathbf{x}}$  where the adjustment function in case 3 of equations 10 and 11 (i.e.,  $\frac{f + \sqrt{bd_{\mathbf{x}}}}{bd_{\mathbf{x}}}$ ) is equal to 1. For  $bd_{\mathbf{x}} > \frac{1}{2}(2f + \sqrt{4f + 1} + 1)$ ,  $\frac{f}{bd_{\mathbf{x}}} < \frac{f + \sqrt{bd_{\mathbf{x}}}}{bd_{\mathbf{x}}} < 1$  which means that the points that fall into case 3 will be adjusted back towards the center of the data, however, to a point beyond the fence. For completeness,  $\frac{f + \sqrt{bd_{\mathbf{x}}}}{bd_{\mathbf{x}}} > 1$  for  $bd_{\mathbf{x}} < \frac{1}{2}(2f + \sqrt{4f + 1} + 1)$ , which means that if the constraint was set at a lower value, then case 3 could lead to outliers being adjusted further away from the central point of the data rather than closer to it. This takes one view point on outliers whereby if an observation is very far outlying then it should still represent moderately rare events in the data. Another approach would be to consider outliers beyond a certain limit as misleading or errors and adjust them further towards the center of the data, limiting their influence. Further investigation will be required based on the data set being analyzed, and the adjustment functions can then be designed depending on the conclusions drawn.

The second approach we consider is given by

$$\mathbf{x}^{rob} = \begin{cases} \mathbf{x}, & \text{if } bd_{\mathbf{x}} \leq f; \\ \min\left(\frac{f}{bd_{\mathbf{x}}}, 1\right)(\mathbf{x} - \mathbf{T}^*) + \mathbf{T}^*, & \text{if } f < bd_{\mathbf{x}} \leq \frac{1}{2}(2f + \sqrt{4f + 1} + 1) \text{ OR } bd_{\mathbf{x}} > u > f; \\ \frac{f + \sqrt{bd_{\mathbf{x}}}}{bd_{\mathbf{x}}}(\mathbf{x} - \mathbf{T}^*) + \mathbf{T}^*, & \text{if } \frac{1}{2}(2f + \sqrt{4f + 1} + 1) < bd_{\mathbf{x}} \leq u. \end{cases} \quad (11)$$

Under this adjustment function, we set an additional limit  $u$  such that if outliers are beyond this point they are also brought back to the fence. The rationale for this approach is that if an observation is this far outlying, it should be given full adjustment back to the fence as it either represents a data entry error or the information contained in that observation is too irrelevant for the current task to include more fully. Note that  $u$  is an arbitrary selection that will be dependent

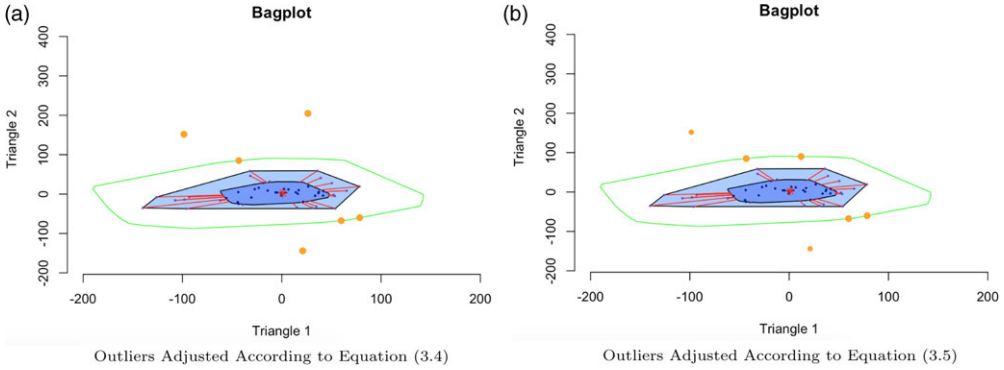


Figure 10. Bagdistance adjusted bagplots.

on the data set under review. The key takeaway is that with unique measures of outlyingness for each observation, there is significantly greater flexibility in how observations can be adjusted.

Fig. 10(a) shows the outliers after adjustment according to equation (10), and Fig. 10(b) shows the outliers after adjustment according to equation (11) where we have set  $f = \sqrt{\chi_{0.99,2}^2}$  and  $u = 15$ . Under the latter approach with  $u = 15$ , observation  $X_{9,1}$  which has bagdistance of 15.8 as shown in Table 1 is brought back to the fence whereas it remains beyond the fence in Fig. 10(a).

**Remark 3.1.** Additionally, we may use the bd values as a residual term when fitting a model with a robust loss function (e.g., robust M-estimation). For example, we could utilize the bd in Huber’s loss function such that we have

$$L_c^{bd}(\mathbf{x}) = \begin{cases} \frac{1}{2}bd^2, & \text{if } bd \leq c; \\ c(|bd| - \frac{1}{2}c) & \text{otherwise.} \end{cases} \tag{12}$$

3.3.4 AO-based adjustments

The AO approach provides a measure of outlyingness for each observation. This measure is based on numerous one-dimensional projections of the individual multivariate observations and the total multivariate sample. For each of these projections, the univariate measure of AO is calculated. The maximum of these univariate AO values after all the projections is taken to be the AO measure in the multivariate case. Without knowing which direction has led to the final AO, we are seemingly unable to backtransform the AO measure to lead to adjusted residuals and ultimately adjusted claim amounts. A solution is to use the AO-based bagplot such that we adjust outliers back to the fence or loop. Consideration must be given to which cutoff value or whether a fence has been drawn to detect outliers under the AO approach. In each case, the loop may be different. We recommend using the traditional cutoff value (Approach 3) as it explicitly incorporates skewness.

3.3.5 Summary of adjustment mechanisms

Tables 2 and 3 summarize the claim values for the outliers detected and their adjusted values for Triangle 1 and 2, respectively. Note that in the case of the AO-Fence and AO-Loop adjustments, we are only considering those observations that were flagged as outliers based on the traditional cutoff.

**Table 2.** Triangle 1 outlier adjustment results.

Outliers	Initial	MCD	BP-Fence	BP-Loop	AO-Fence	AO-Loop	AO-mrfDepth	<i>bd</i> -(10) (no limit)	<i>bd</i> -(11) (limit)
$X_{6,5}$	238,375	226,990	231,580	219,269	-	-	-	232,213	232,213
$X_{7,3}$	633,500	652,977	645,490	656,826	-	-	-	645,081	645,081
$X_{7,4}$	432,257	403,580	409,118	393,805	-	-	-	409,668	409,668
$X_{8,2}$	1,458,541	1,557,335	1,543,798	1,565,203	1,802,824	1,588,895	-	1,484,742	1,484,742
$X_{8,3}$	727,098	713,520	715,180	710,780	702,174	700,640	700,177	723,168	723,168
$X_{9,1}$	2,210,754	2,139,306	2,143,332	2,137,794	2,119,751	2,133,590	2,133,589	2,164,880	2,144,073
$X_{9,2}$	1,517,501	1,538,538	-	-	-	-	-	-	-

**Table 3.** Triangle 2 outlier adjustment results

Outliers	Initial	MCD	BP-Fence	BP-Loop	AO-Fence	AO-Loop	AO-mrfDepth	<i>bd</i> -(10) (no limit)	<i>bd</i> -(11) (limit)
$X_{6,5}$	6,650	9,124	8,215	11,050	-	-	-	8,049	8,049
$X_{7,3}$	86,734	75,546	80,162	73,947	-	-	-	80,344	80,344
$X_{7,4}$	18,109	29,028	27,279	33,348	-	-	-	26,990	26,990
$X_{8,2}$	132,208	98,329	103,742	96,595	17,257	88,685	-	123,419	123,419
$X_{8,3}$	20,923	49,957	47,709	57,597	76,939	80,387	81,427	29,809	29,809
$X_{9,1}$	112,103	43,470	47,024	41,679	24,263	37,621	37,620	67,438	47,179
$X_{9,2}$	33,250	34,062	-	-	-	-	-	-	-

For the AO-mrfDepth approach, we are considering outliers flagged under this methodology and outliers are adjusted back to the loop in this case. As we adjust residuals back in the direction of the central estimate of the data, this may lead to an upward or downward adjustment for each residual and similar adjustments will be seen for the corresponding claim observations. Note that ‘-’ values are provided where outliers were not detected under that methodology.

### 3.4 Final reserves and discussion

Once the outliers have been appropriately identified and treated, the data can be used to compute reserves estimates. Here, we used the multivariate time series chain-ladder technique as described in Merz & Wüthrich (2008).

Table 4 summarizes the final reserve estimates and their associated RMSE for each individual triangle and total reserves under each outlier detection and adjustment technique as well as when we simply apply the multivariate chain-ladder without adjusting any observations. However, we consider the last three development periods as separate univariate triangles. This is because there are few data points for these development periods and applying the multivariate chain-ladder to such periods often leads to highly volatile results or potentially failure in that elements of the estimated correlation matrices may have absolute values greater than one. This in turn may lead to a lack of convergence. Note that some authors (including Merz & Wüthrich, 2008) suggest extrapolation of these correlation variables from the previous periods however as our main focus is on outlier detection and adjustment we have not pursued this option.

The robust reserves in this example are always modestly less than the original reserves. This phenomenon is well known in the robust literature and is obviously related to all adjustments being *reductions*; losses are thus decreased and so are the associated calculated reserves. The greatest adjustment is for the *bagdistance* technique where a limit is set for the adjustment function.

In each case, we saw a much greater reduction in the RMSE than reserves. For example, the bagplot-loop methodology leads to a 0.91% reduction in reserves and a 12.17% reduction in

**Table 4.** Bivariate example reserves.

	No. of outliers detected	Triangle 1		Triangle 2		Total		Difference (%)	
		Reserve	RMSE	Reserve	RMSE	Reserve	RMSE	Reserve	RMSE
Original	–	6,435,951	322,573	489,028	90,542	<b>6,924,978</b>	<b>337,001</b>	–	–
MCD	7	6,438,541	283,054	438,497	45,552	<b>6,877,037</b>	<b>293,870</b>	–0.69	–12.80
Bagplot-Fence	6	6,427,705	291,416	441,754	50,822	<b>6,869,460</b>	<b>301,923</b>	–0.80	–10.41
Bagplot-Loop	6	6,416,549	283,191	445,097	41,669	<b>6,861,647</b>	<b>295,999</b>	–0.91	–12.17
AO-Fence	3	6,509,807	294,572	389,602	62,491	<b>6,899,409</b>	<b>288,668</b>	–0.37	–14.34
AO-Loop	3	6,456,446	292,021	444,764	47,660	<b>6,901,210</b>	<b>297,914</b>	–0.34	–11.60
AO-mrfDepth	2	6,456,381	291,999	445,496	47,722	<b>6,901,876</b>	<b>297,901</b>	–0.33	–11.60
<i>bd</i> -(10) (no limit)	6	6,417,482	304,232	456,439	68,329	<b>6,873,921</b>	<b>315,131</b>	–0.74	–6.49
<i>bd</i> -(11) (limit)	6	6,412,956	301,127	438,432	61,425	<b>6,851,388</b>	<b>306,010</b>	–1.07	–9.20

RMSE. This suggests that even if the change in reserve estimates is only minor, the accuracy of such estimates may be enhanced considerably as a result of these robust techniques.

## 4. Detection and Treatment of Outliers in N-Dimensional Reserving

### 4.1 An N-dimensional framework

We now put forward a framework for N-dimensional robust chain-ladder reserving. The methodology is as follows. Firstly, perform the robust Poisson GLM chain-ladder technique on each triangle separately to obtain residuals. These residuals are then stored in an N-dimensional matrix. It is these N-dimensional residuals that outlier detection and adjustment techniques are applied to. The techniques are based on AO, halfspace depth, MCD Mahalanobis distance, and *bagdistance*.

Under the AO approach outliers are detected using the standard cutoff value (see section 2.2.4). In the three-dimensional case, a graphical representation is also constructed. Firstly, declare the point with the lowest AO as the AO-median. Next, form the convex polyhedron that contains the 50% of points with the lowest AO which can be considered as a N-dimensional AO-based bag. We then construct an N-dimensional AO-based loop by forming the convex polyhedron that contains all non-outlying points.

Outliers are adjusted by bringing them back to the intersection point of the ray from the AO-median to the outlying point and the AO-based loop. This intersection is found using the parametric line clipping algorithm (Cyrus & Beck, 1978; Liang & Barsky, 1984). Adjusted residuals are then backtransformed to give robust incremental claims. The multivariate time series chain-ladder (Merz & Wüthrich, 2008) is applied on these robust claims.

For the halfspace depth approach, we calculate residuals in the same fashion; however, the outlier detection and treatment methodology is altered. The halfspace depth methodology outlined in section 2.1.2 can be applied in N dimensions with Fig. 1 providing an illustration in 2 dimensions. Firstly, calculate the halfspace depth of each observation. The Tukey Median is then the observation with the greatest halfspace depth or the center of gravity of the deepest region if this is not a unique point. Next, formulate the N-dimensional bag. Firstly, let  $D_s$  represent the region that contains all points with halfspace depth greater than or equal to  $s$ . Let  $\#D_s$  be the number of data points in  $D_s$ . The bag is constructed by linearly interpolating the two convex regions that satisfy  $\#D_s \leq \lfloor \frac{n}{2} \rfloor < \#D_{s-1}$  with respect to the Tukey Median. This interpolation is done in the same way as for the bivariate bagplot as described in Miller *et al.* (2003) however now applicable for any polytype rather than only two-dimensional polygons. Firstly, calculate

$$\lambda = \frac{\frac{n}{2} - \#D_s}{\#D_{s-1} - \#D_s}. \tag{13}$$

The bag is given by  $\lambda \cdot$  outer polytype  $+ (1 - \lambda) \cdot$  inner polytype. Next, the N-dimensional fence is constructed by multiplying the N-dimensional bag by a fence factor with respect to the Tukey Median. Outliers are declared as points beyond the fence. The fence factor we use is  $\sqrt{\chi_N^2}$ . For all  $N > 1$ , this is greater than the previously used arbitrary value of 3 in the two-dimensional case, and hence, less outliers will be detected and adjusted. Then, construct the N-dimensional loop by forming the convex hull of all non-outlying points. Outliers are treated by bringing them back to either the fence or loop. In three dimensions, the bag, loop, and fence are convex polyhedrons and graphical representation is available. Adjusted residuals are backtransformed and the multivariate time series chain-ladder applied.

The MCD Mahalanobis distance and *bagdistance* approaches are implemented in the same fashion as they were for the bivariate case. For the trivariate case, a graphical representation of the tolerance ellipsoid for the MCD Mahalanobis technique may be displayed. We now show the implementation of these four techniques on real data and compare the results.

## 4.2 Illustration with the AUSI dataset

For the purposes of illustration, we use data from three different lines of business of two Australian insurers. A short discussion of this data is provided in Appendix A.2.

### 4.2.1 Data adjustments

The fitting of the robust Poisson GLM stage of the methodology is only possible if there are no negative incremental claims in each triangle. However, of the 820 observations for each line, the CTP line from insurer 1, the CTP line from insurer 2, and the Home line from insurer 1 have 1, 21, and 57 negative claims, respectively. We have set all these claims to zero. To ensure that this did not significantly impact results, we calculated development factors under the classical chain-ladder technique before and after the adjustment. The largest change in a development factor was 0.231%, and hence, the effect of the adjustment is small and is deemed acceptable.

### 4.2.2 Detection of outliers

Firstly, we perform the AO technique on these residuals. 34 outliers were flagged under this approach. We then construct a 3D version of the AO-based bagplot such that the 50% of data points with the lowest AO are contained in a convex polyhedron analogous to the bag, and an additional convex polyhedron is formed for all non-outlying observations analogous to the loop. Fig. 11(a) shows the residuals with the 3D AO bag drawn in yellow, and Fig. 11(b) shows the residuals with the 3D AO-loop in orange and the corresponding bag contained inside. Outliers are shown in red in each case.

We note that this loop extends in the various directions of the data's dispersion and seemingly captures its general shape. For completeness, we have also investigated the use of the cutoff value for AO of  $\sqrt{\chi_{(99,N)}^2} \cdot \text{median}(\text{AO})$  as suggested in Segaeert *et al.* (2017). Under this approach, 48 outliers are detected.

We now illustrate the halfspace depth-based approach. Fig. 12(a) shows the residuals with the bag drawn in yellow, and Fig. 12(b) shows the residuals with the bag loop and fence all drawn. The loop is again in orange and is within the light blue fence. Outliers are declared as the points outside the fence, and under this methodology, we detect 48 outliers. Interestingly, 29 of the 34 observations flagged as outliers under the traditional AO methodology are also detected under the halfspace depth approach. This highlights some agreement between the two methodologies. Note that the bags under both the AO and halfspace depth approaches are relatively similar; however, the fence which is used to declare outliers under the halfspace depth methodology appears unable

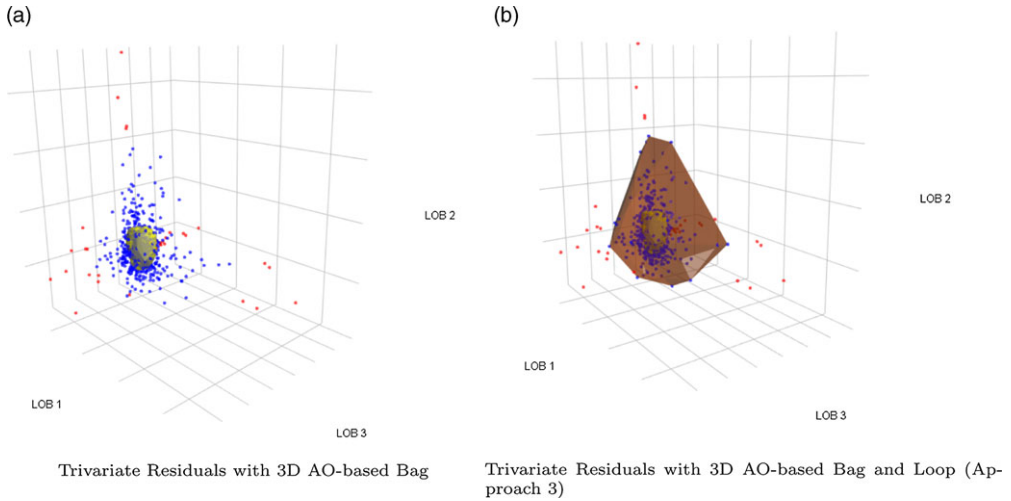


Figure 11. AO based detection.

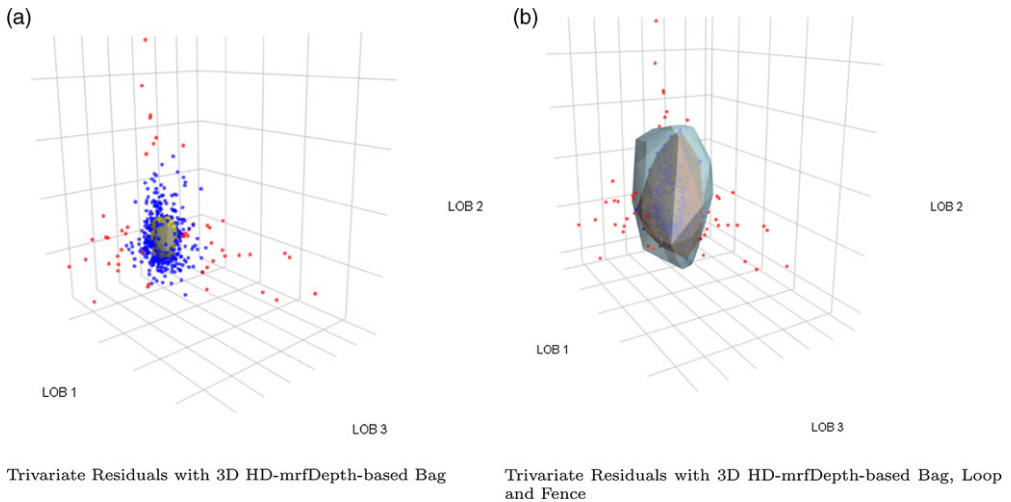


Figure 12. Halfspace depth-based detection.

to capture the three-dimensional skewness in the data. Rather, it remains relatively elliptical in shape whereas the AO-loop stretches in each direction that the data are further reflected in the number of outliers flagged. We conject that the halfspace depth approach may be misclassifying regular observations as outliers in this instance.

To highlight the differences between a fence-based approach and AO approach, we may also draw an AO-based fence by multiplying the AO bag by three and declare as outliers observations outside this fence. Fig. 13(a) shows the residuals with the AO bag drawn in yellow and AO-fence in light blue. In Fig. 13(b), we have drawn the AO bag, loop, and fence. In this case, the convex polyhedrons representing the loop and fence are overlapping, and in particular, the loop seemingly extends further in each direction that the data are dispersed. Notably, the AO-fence appears reasonably elliptical in shape, similar to its halfspace depth counterpart.

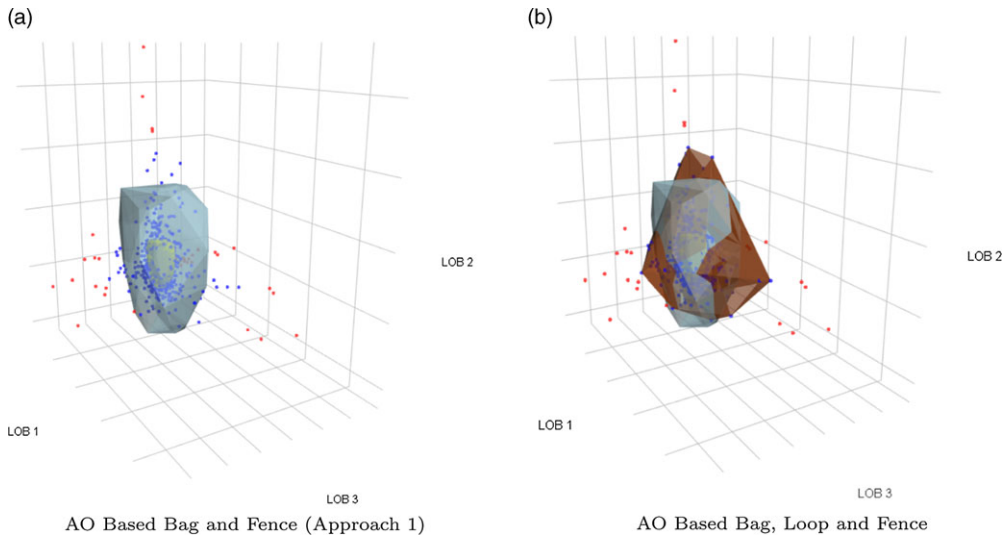


Figure 13. AO based detection (Approach 1 versus Approach 3).

Under this fence-based approach, 66 observations are flagged as outliers. We believe that using an AO-based fence may again fail to fully capture the shape of the data as we suspect has occurred for the halfspace depth approach. In particular, we believe this will be of the greatest concern when there is significant skewness in the data set.

Now, the MCD Mahalanobis Distance technique which does not consider skewness in the data set at all. Analogous to tolerance ellipses as described in section 3.2.2, tolerance ellipsoids have a squared distance to the central estimate of the data equal to a quantile of the  $\chi_3^2$  distribution. This concept is extendable to higher dimensions where the squared distance is equal to a quantile of the  $\chi_p^2$  distribution where  $p$  represents the dimension of the data. Fig. 14(a) shows the residuals with the 97.5% robust tolerance ellipsoid drawn in orange, and Fig. 14(b) shows the non-robust tolerance ellipsoid drawn in green with its robust counterpart encapsulated within it.

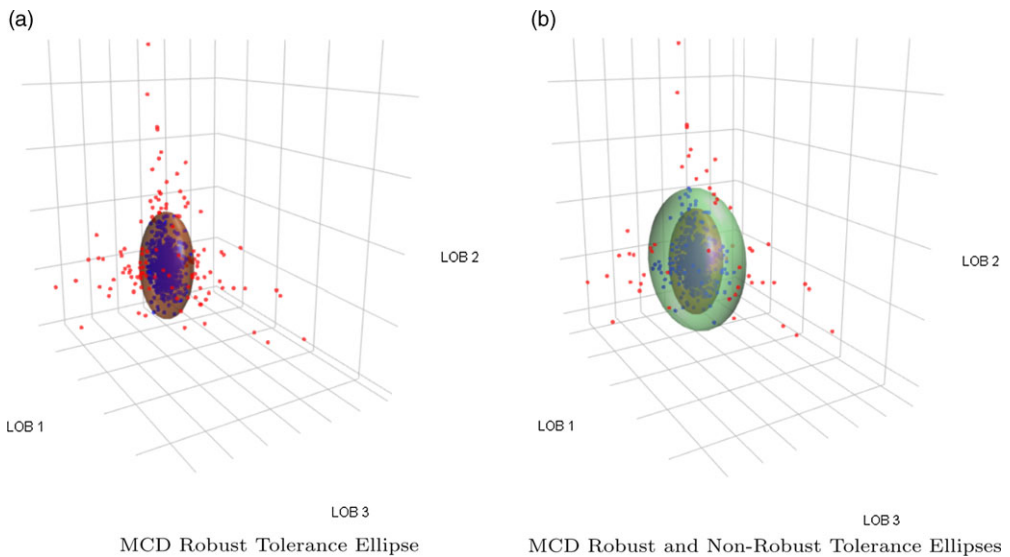
This is analogous to what we saw in the bivariate case (see Fig. 6). Under the MCD Mahalanobis distance approach, 120 observations were flagged as outliers. This approach saw a substantially greater number of observations detected as outliers than the other available techniques as it does not consider skewness, likely leading to misclassifications, for example, swamping as the tolerance ellipsoid does not extend towards the skewed data.

#### 4.2.3 Treatment of outliers

Now we turn to the treatment of outliers. Under the AO, halfspace depth, and MCD Mahalanobis distance approaches, outlying observations are brought back to a relevant convex polyhedron (i.e., loop or fence) or tolerance ellipsoid. Under the *bagdistance* methodology, an alternative approach may be used which weights the degree of adjustment according to the level of outlyingness of each observation. In particular, we have again considered the weighting functions given by equations (10) and (11) and set  $f = \sqrt{\chi_3^2}$  and  $u = 10$ . As noted in section 3.3.3,  $u$  is an arbitrary selection that will be dependent on the data set under review. Adjusted residuals under each methodology are backtransformed to give robust incremental observations.

**Table 5.** Trivariate reserves (Scaled values with different radix).

	No. of outliers detected	LOB 1		LOB 2		LOB 3		Total		Difference (%)	
		Reserve	RMSE	Reserve	RMSE	Reserve	RMSE	Reserve	RMSE	Reserve	RMSE
Original	-	900.00	66.62	1,300.00	112.34	800.00	74.48	<b>3,000.00</b>	<b>176.69</b>	-	-
MCD	120	826.15	42.19	1,241.05	97.84	750.55	39.81	<b>2,802.70</b>	<b>138.34</b>	<b>-6.58</b>	<b>-21.71</b>
HD Fence	48	904.73	56.61	1,134.21	96.41	770.12	47.26	<b>2,843.71</b>	<b>156.06</b>	<b>-5.21</b>	<b>-11.68</b>
HD Loop	48	905.34	56.25	1,131.76	95.82	765.90	46.17	<b>2,841.63</b>	<b>155.07</b>	<b>-5.28</b>	<b>-12.24</b>
AO-Fence	66	903.41	55.39	1,128.94	95.20	759.14	41.51	<b>2,834.21</b>	<b>153.22</b>	<b>-5.53</b>	<b>-13.28</b>
AO-Loop	34	924.92	61.04	1,130.74	96.31	765.70	45.93	<b>2,875.41</b>	<b>162.25</b>	<b>-4.15</b>	<b>-8.17</b>
AO-mrf Depth	48	876.42	57.85	1,256.43	104.58	764.17	42.68	<b>2,909.42</b>	<b>160.23</b>	<b>-3.02</b>	<b>-9.32</b>
<i>bd</i> (no limit)	48	887.55	61.79	1,286.06	108.27	783.01	54.57	<b>2,961.24</b>	<b>167.60</b>	<b>-1.29</b>	<b>-5.15</b>
<i>bd</i> (limit)	48	888.25	61.82	1,286.33	108.29	773.03	54.26	<b>2,960.83</b>	<b>167.75</b>	<b>-1.31</b>	<b>-5.06</b>



**Figure 14.** MCD mahalanobis distance detection.

#### 4.2.4 Final reserves and discussion

The time series multivariate chain-ladder technique (Merz & Wüthrich, 2008) is performed on these robust observations. The results for final reserves and estimated RMSE are given in Table 5. Note that we have scaled the presented results within each LOB to ensure commercial data confidentiality.

Under each methodology, reserves have been reduced. The greatest reduction was seen when outliers were detected and adjusted using the MCD Mahalanobis distance approach. This is likely because there was a much greater proportion of observations detected as outliers under this approach. However as this methodology does not consider the skewness in the data set, it may be inappropriate in this situation (likely excessive). Rather, the AO technique based on the traditional cutoff value is likely more equipped to capture the shape of the data in this regard and we believe it to be the most suitable approach for this data set. The smallest reduction in reserves is given for the *bd* approach where there was no upper limit on the weighting function. This is likely because under this approach outliers have been adjusted to a smaller degree than in comparison to the alternative mechanisms. Additionally, we have seen the RMSE of reserves decrease in

each case. The MCD Mahalanobis distance approach saw the greatest reduction in this metric of 21.71%, whereas the traditional AO approach (loop) saw a 8.17% reduction. Notably, the shift in the RMSE for each approach was significantly greater than the corresponding change in reserves. This highlights that even for small adjustments in terms of the point estimate of reserves when using these robust techniques, we may see significant improvements in accuracy.

## 5. R Codes

All relevant R codes used for the bivariate and trivariate illustrations can be found at <https://github.com/agi-lab/reserving-robust>, along with a synthetic dataset (with features inspired by the AUSI data used in section 4) to help replication of results.

## 6. Conclusion

We put forward two alternative robust bivariate chain-ladder techniques. The first technique is based on AO and explicitly incorporates skewness into the analysis while providing a unique measure of outlyingness for each observation. The second technique is based on *bagdistance* which is derived from the bagplot; however, it is able to provide a unique measure of outlyingness and a means to adjust outlying observations based on this measure. Using a real bivariate data set, we illustrated how those techniques compared to existing ones. While the (inevitable) reduction in central estimates is (desirably) modest, the reduction in RMSE is much more significant. This highlights that even for small adjustments in terms of the point estimate of reserves when using these robust techniques, we may see significant improvements in accuracy. We showed that our methodology offers material and significant improvements, compared to the techniques introduced in Verdonck & Van Wouwe (2011).

We then extended our framework to N dimensions and implemented all four outlier detection and treatment techniques on a real trivariate data set. The mathematical generalization of the bivariate methodologies was not trivial. The illustration demonstrated good performance again, and results were qualitatively similar to those discussed in the bivariate example.

Through our expositions and extension to higher dimensions, we have added to the toolbox of techniques available to detect and treat outliers in multivariate reserving. We believe that the new techniques applied in this paper address some of the shortcomings of the previous approaches and should be explored as common practice when implementing robust multivariate reserving techniques in practice.

While we focused on the multivariate time series chain-ladder model of Merz & Wüthrich (2008) for our illustrations, it should be noted that the detection and treatment methodologies developed in this paper can be applied to other reserving techniques, including in most recent machine learning actuarial procedures (see, e.g., Richman, 2018). Notably, our work is concerned with the detection and treatment of outliers rather than finding the best reserving model for the data and we have used the chain-ladder model for illustrative purposes. In the case that one proceeds to model the robustified data, back-testing of results compared to those from unadjusted data may provide useful insights into the effectiveness of the techniques.

**Acknowledgements.** Authors are very grateful for comments from two anonymous referees, which led to significant improvements in the paper. Earlier versions of this paper were presented at the Australian Actuaries Institute *General Insurance Seminar* conference in Melbourne (Australia) as well as at an Australian Actuaries Institute *Insights* session in Sydney (Australia). The authors are grateful for constructive comments received from colleagues who attended this event and from those who read the earlier version of the paper published online to support the presentation. The authors are also grateful to Sharanjit Paddam and Will Turvey for comments on earlier drafts of the paper. Additionally, the authors are grateful for the research assistance of Wilson Cheng, Wanzhang (Simon) Jing, and Yun Wai (William) Ho, who contributed materially to the review of the code related to the examples presented.

This research was supported under Australian Research Council's Linkage (LP130100723, with funding partners Allianz Australia Insurance Ltd, Insurance Australia Group Ltd, and Suncorp Metway Ltd) and Discovery Project funding scheme (DP200101859). Furthermore, Lavender acknowledges support from the UNSW Business School Honours Scholarship. The views expressed herein are those of the authors and are not necessarily those of the supporting organizations.

## References

- Avanzi, B., Lavender, M., Taylor, G.C. & Wong, B. (2023). On the impact of outliers in loss reserving. *European Actuarial Journal* in press.
- Brys, G., Hubert, M. & Struyf, G. (2004). A robust measure of skewness. *Journal of Computational and Graphical Statistics*, **13**(4), 996–1017.
- Chan, J. & Choy, S. (2003). Scale mixtures distributions in insurance applications. *ASTIN Bulletin*, **33**(1), 93–104.
- Chan, J., Choy, S. & Makov, U. (2008). Robust bayesian analysis of loss reserves data using the generalized t-distribution. *ASTIN Bulletin*, **38**(1), 207–230.
- Cyrus, M. & Beck, J. (1978). Generalized two- and three-dimensional clipping. *Computers and Graphics*, **3**(1), 23–28.
- Donoho, D. (1982). Breakdown properties of multivariate location estimators. Boston: Harvard University, Qualifying Paper.
- England, P. & Verrall, R. (1999). Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics*, **25**(3), 281–293.
- Hubert, M., Rousseeuw, P. & Segaert, P. (2015). Multivariate functional outlier detection. *Statistical Methods & Applications*, **24**(2), 177–202.
- Hubert, M., Rousseeuw, P. & Segaert, P. (2016). Multivariate and functional classification using depth and distance, arXiv.org, Statistics: Methodology, 1–23.
- Hubert, M. & Van der Veken, S. (2008). Outlier detection for skewed data. *Journal of Chemometrics*, **22**, 235–246.
- Hubert, M. & Vandervieren, E. (2008). An adjusted boxplot for skewed distributions. *Computational Statistics and Data Analysis*, **52**(12), 5186–5201.
- Liang, Y. & Barsky, B. (1984). A new concept and method for line clipping. *ACM Transaction on Graphics*, **3**(1), 1–22.
- Merz, M. & Wüthrich, M. V. (2008). Prediction error of the multivariate chain ladder reserving method. *North American Actuarial Journal*, **12**(2), 175–197.
- Miller, K., Ramaswami, S., Rousseeuw, P., Antoni Sellarès, J., Souvaine, D., Streinu, I. & Struyf, A. (2003). Efficient computation of location depth contours by methods of computational geometry. *Statistics and Computing*, **13**(2), 153–162.
- Pitselis, G., Grigoriadou, V. & Badounas, I. (2015). Robust loss reserving in a log-linear model. *Insurance: Mathematics and Economics*, **64**, 14–27.
- Richman, R. (2018). AI in actuarial science. SSRN.
- Romanazzi, M. (2001). Influence function of halfspace depth. *Journal of Multivariate Analysis*, **77**(1), 138–161.
- Rousseeuw, P. (1984). Least median of squares regression. *Journal of the American Statistical Association*, **79**(388), 871–880.
- Rousseeuw, P., Ruts, I. & Tukey, J. (1999). The bagplot: a bivariate boxplot. *The American Statistician*, **54**(4), 382–287.
- Segaert, P., Hubert, M., Rousseeuw, P., Raymaekers, J. & Vakili, K. (2017). Package 'mrfdepth'.
- Shi, P., Basu, S. & Meyers, G. G. (2012). A Bayesian log-normal model for multivariate loss reserving. *North American Actuarial Journal*, **16**(1), 29–51.
- Tukey (1977). *Exploratory Data Analysis*. 1st edition. Reading Mass: Addison-Wesley Publishing Co.
- Tukey, J. (1975). Mathematics and the picturing of data\* (pp. 523–531). In Proceedings of the International Congress of Mathematicians.
- Verdonck, T. & Debruyne, M. (2011). The influence of individual claims on the chain-ladder estimates: analysis and diagnostic tool. *Insurance: Mathematics and Economics*, **48**(1), 85–98.
- Verdonck, T. & Van Wouwe, M. (2011). Detection and correction of outliers in the bivariate chain-ladder method. *Insurance: Mathematics and Economics*, **49**(2), 188–193.
- Verdonck, T., Van Wouwe, M. & Dhaene, J. (2009). A robustification of the chain-ladder method. *North American Actuarial Journal*, **13**(2), 280–298.

## Appendix

### A. Data

#### A.1 Bivariate

The data used to illustrate the robust bivariate chain-ladder techniques are provided in this section. It is of incremental claims for Personal (Table A.1) and Commercial (Table A.2) auto insurance from a major U.S. property-casualty insurer from 1988 to 1997. The data are taken from Shi *et al.* (2012).

**Table A.1.** Bivariate data set (a) (Shi *et al.*, 2012).

<i>i/j</i>	1	2	3	4	5	6	7	8	9	10
1	1,376,384	1,211,168	535,883	313,790	168,142	79,972	39,235	15,030	10,865	4,086
2	1,576,278	1,437,150	652,445	342,694	188,799	76,956	35,042	17,089	12,507	
3	1,763,277	1,540,231	678,959	364,199	177,108	78,169	47,391	25,288		
4	1,779,698	1,498,531	661,401	321,434	162,578	84,581	53,449			
5	1,843,224	1,573,604	613,095	299,473	176,842	106,296				
6	1,962,385	1,520,298	581,932	347,434	238,375					
7	2,033,371	1,430,541	633,500	432,257						
8	2,072,061	1,458,541	727,098							
9	2,210,754	1,517,501								
10	2,206,886									

**Table A.2.** Bivariate data set (b) (Shi *et al.*, 2012).

<i>i/j</i>	1	2	3	4	5	6	7	8	9	10
1	33,810	45,318	46,549	35,206	23,360	12,502	6,602	3,373	2,373	778
2	37,663	51,771	40,998	29,496	12,669	11,204	5,785	4,220	1,910	
3	40,630	56,318	56,182	32,473	15,828	8,409	7,120	1,125		
4	40,475	49,697	39,313	24,044	13,156	12,595	2,908			
5	37,127	50,983	34,154	25,455	19,421	5,728				
6	41,125	53,302	40,289	39,912	6,650					
7	57,515	67,881	86,734	18,109						
8	61,553	132,208	20,923							
9	112,103	33,250								
10	37,554									

## A.2 Trivariate

We use data from three different lines of business of two Australian insurers to illustrate our robust multivariate chain-ladder techniques in section 4.2. In particular, we investigate CTP and home insurance lines from one insurer and a CTP line from another. The data are quarterly and run from the beginning of 2004 through 2013.

The complete data set that these three lines of business have been taken from was developed as part of a Linkage Project grant awarded by the Australian Research Council until 2016 for a project titled *Modelling claim dependencies for the general insurance industry with economic capital in view: an innovative approach with stochastic processes*. It is referred to as the AUSI dataset, an acronym of the names of the project partners (Allianz Australia Insurance Ltd, UNSW Australia, Suncorp Metway Ltd, and Insurance Australia Group Ltd).