Migration Letters

Volume: 21, No: S5 (2024), pp. 274-286

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

Improving The Control of Inventory Management Systems Using Robust Estimators in the Presence of Outliers

Hussien shakir Razzaq[1] and Nerda Zura Zaibidi[2]

Abstract

Inventory management systems aim to control inventory levels in the best way while reducing costs to a minimum. However, inventory management faces many challenges that lead to the deficit or increase in storage costs, and among these challenges is the presence of outliers in demand data, therefore this paper seeks to building a new model that uses robust estimators (Median and Median Absolute Deviation) instead of classical estimators that are highly sensitive to outliers, and examining this model on real data for the demand for raw materials used in the cement industry in one of Iraq's factories. The proposed model was able to reduce the deficit ratio by 96% and contributed to reducing the total cost of storage by 75%, in addition to that, the proposed model contributed to reducing the need for safety stock by 30%.

Keywords*: inventory management, outliers, robust estimators, boxplots, reorder point, safety stock.*

1. Introduction

Inventory management is a series of processes that follow a system of control and issuance of orders to ensure the continued availability of quantities upon demand and at the same time reduce unnecessary investments that may be caused by excess inventory (Paramasivan & Subramanian, 2020)

Inventory management is one of the most important aspects of management in institutions that are directly related to real-life situations, just as hardly any organization is free from the problems of controlling inventory, especially in manufacturing and commercial operations, so that spending on inventory management has become at the top of operational expenses (Shekarian et al., 2016)[1]

Inventory control and ensuring control of the flow of materials in stores are among the major problems facing the organizational management and finding optimal solutions to this problem leads to the smooth operation of production and marketing activities because it forms a link between them and ensures the provision of resources and materials entering production when needed in a timely manner.

The optimal management of the inventory avoids the organization pumping and investing large sums of money in the inventory and reduces the possibility of a deficit in the demand for

[1]Department of Decision Science, School of Quantitative Sciences, Universiti Utara Malaysia, Sintok, Kedah, Malaysia.
[2]Department of Decision Science, School of Quantitative Sciences, Universiti Utara Malaysia, Sintok, Kedah, Malaysia.

materials, thus avoiding the loss of continuity of production and consumer confidence in light of the intense competition that the market is currently witnessing.

Since inventory management relies on previous data in the analysis and future decisions regarding the quantity of demand and the size of the safety stock and determining the point of re-order, so the inventory management faces a great challenge in reaching the ideal decisions that ensure reducing inventory costs to a minimum.

But at the same time, the data that is adopted in making decisions related to the control of inventory is affected by a number of factors and circumstances, such as severe fluctuations in the supply and demand market or the spread of epidemics, as happened in the demand data during the spread of the Corona epidemic, moreover, human errors in data entry and Its classification is one of the reasons for the emergence of data that is inconsistent with the other data set.

May one of the most important risks facing data analysis and decision-making is the presence of outliers within the data, and this may cause the decision to deviate from optimization and be negatively affected by the extreme data, but sometimes the outliers may pull decisions to the wrong direction, which causes unexpected losses.

The real data usually contains observations that are distinct from the rest of the data, such as being large or relatively small and move away from the data center, and the presence of these outliers leads to a negative impact on the data analysis (Seo, 2006).

Outliers usually arise due to malfunctions in the measuring devices, human errors in recording, or they may be the result of wrong assumptions established by the researcher (Fatima & Kurmi, 2018). The problem is that the presence of some outliers may distort the group's results and negatively influence the analysis (Cousineau, 2010)

In order to obtain the proper analysis of the data and then take the optimal decision in the management of the inventory, the effect of outliers must be eliminated or reduced to a minimum.

Determining the order quantity, the optimum reorder point, and the safety stock quantity must be preceded by relying on real data that is free from errors or negative effects of extreme data. Therefore, this article focuses mainly on reducing the impact of extreme values if they are present within the demand data in order to obtain decisions optimum inventory management.

2. Literature review

P. O.Agada and Ogwuche (2017) and Nemtajela and Mbohwa (2017) pointed out that most inventory management systems in real life have stochastic demands, and that the probabilistic EOQ model basically seeks to answer two questions: the quantity to be requested and when to order, in other words determining the level of inventory at which the new order is issued.

according to Agarwal (2015), the EOQ model has been used for many years in various industrial sectors for ease of use and help in understanding the behavior of the inventory system because it contributes to companies and retailers in calculating the quantities that reduce the total cost of inventory. So it can be said that EOQ is one of the most used models in inventory management (Personal et al., 2018).

Although the previous studies did not address the impact of outliers in inventory management, in general, they opened the door widely to ways to detect and treat outliers, in addition to using multiple methods to reduce the impact of outliers. Also, many studies have dealt with the negative impact of outliers on the classical statistical parameters used mainly in inventory management models.

Some justifications in the branches of applied mathematics assumed that a simple error in the mathematical model leads to a slight defect in the final results. However, this assumption is wrong when dealing with outliers because sometimes a few deviations may lead to a large defect in the results due to the extreme sensitivity of some statistical parameters (Huber & Wiley, 1981).

Bollen (1987) stated that the potential effect of outliers in the field of statistics is well known and that the existence of these values is one of the factors that cause improper solutions. Nevertheless, researchers neglect this factor when discussing improper solutions.

According to Tukey (1960), outliers should be given special attention because the presence of a single outlier can lead to significant negative effects on the classical estimates of location and scale.

Since the outliers are located three times the standard deviation plus the mean from the data center, these values may have a strong influence on parameter estimation and may cause harmful effects on statistical analyzes, weakening the strength of the tests, and contribute to increasing the variance of error (Osborne et al., 2004; Seo, 2006).

Pankratz, (2000) and Kaiser, R., and Maravall Herrero (1999) stated that outliers may cause defects in the construction of models and biases in statistical parameters, as well as may cause errors in forecasts.

In general, the detection of outliers has been widely applied in many areas of research such as statistics, data mining and sensor networks, and has recently gained a lot of interest in industrial and financial applications due to its importance in detecting values that indicate the presence of a specific defect, machine malfunctions, or fraudulent transactions. (C. T. Lu et al., 2003; Lazarevic & Kumar, 2005; Gupta et al., 2014; Garces & Sbarbaro, 2011; Benjelloun et al., 2019; Pokrajac et al., 2007; Zhao et al., 2014).

Although the detection of outliers has become an essential enabling technology for a wide range of scientific applications such as industry, business, security and engineering, there is a very large difference in the algorithms and techniques that are applied to detect outliers and are highly dependent on the nature and characteristics of the data being handled (Zhang, Q., Segall, R. S., & Cao, 2010; J. Zhang, 2013) .

Since this research mainly focuses on finding improvement in inventory management decisions, any method for detecting outliers will suffice, and from here, the box plot method will be used, which is one of the most popular ways to detect outliers.

According to Y. Zhang et al. (1996) and Lavrac et al. (2000), Box plots are most suitable for exploring both symmetric and skewed quantitative data. Torgo and Ribeiro (2003) stated that Box plots are visualization tools that are often used to identify outliers. Also, Fatima and Kurmi (2018) mentioned that this method identifies outliers in a more appropriate way.

On the other hand, the classical statistics found in inventory models are very sensitive to outliers, so suggesting more robust alternatives would be more appropriate in resisting the influence of outliers and may lead to better decisions.

It is no secret that the median is one of the measures of central tendency, which is more resistant to outliers than the arithmetic mean, and it can be a substitute for it in expressing the mean of the data.

Although many robust estimators of location exist, the sample median is still the most widely known, and it has a breakdown point 50% (Rouss & Christophe, 1993). And the most popular robust estimator of location is the sample median (Picek, 2012). Cousineau, (2011) stated that a more robust estimate of the population central tendency is the median. A few outliers will have a limited impact on this statistic.

That is, the median will be resistant to the presence of outliers even if they constitute 50% of the data. It is also characterized by its ease of calculation and does not require a lot of complications that other statistics need when calculating, Therefore, it may be an ideal substitute for the arithmetic mean in inventory management models.

Hampel, (1974) suggested the Median Absolute Deviation as robust estimate of the spread. However, it differs from the standard deviation in that it has a breakdown point of approximately 50% (Olewuezi, 2011).

Chung et al., (2008) noted that the standard deviation is often inflated due to the presence of outliers, in contrast, MAD is more robust to outliers. MAD is a robust measure of scale that is simple to implement and easy to interpret (Arachchige & Prendergast, 2019), and it is a robust estimator of dispersion that is not influenced by outliers (Voloh et al., 2018),

3. The Developed Model of Inventory Management

The traditional inventory management model is based on several concepts, the most important of which is the reorder point, the safety stock, as well as the order quantity. Each of these concepts has a special mathematical equation to calculate it, as follows:

order quantity $Q = \sqrt{\frac{2KB}{h}}$
Safety stock $SS = Z \cdot \sigma_{DLT}$
reorder point $ROP = \mu + Z \cdot \sigma_{DLT}$
where, μ = Demand rate per period,
σ_{DLT} = Standard deviation of demand during lead time,
Z = Normal table value for the given service level.
B = Annual Demand.
K = Fixed Cost Per Quantity.
h = Annual Holding Cost.

However, these three elements may be affected or collapsed when there are outliers in the data, assuming that there are outliers in the data in special cases, it is necessary to develop special models for inventory management in such cases.

Therefore, the desired goal of optimal inventory management may be achieved by replacing classical statistics with robust statistics to reduce or eliminate the impact of outliers on decision-making.

Hence, this paper proposes to make changes in the equations for calculating the re-order point, the safety stock and the economic order quantity by replacing the classical arithmetic mean and

standard deviation with the Median and Median Absolute Deviation, which are considered solid statistics that are not affected by the presence of outliers, and then reformulating the mathematical equations for the elements of inventory management and tested to show the effectiveness of the new model in the face of the impact of outliers.
Thus, the mathematical equations for the proposed model will be as follows:

$$Q_{\text{robust}} = \sqrt{\frac{2K(\text{med} * 365)}{h}}$$

$$\text{Safety stock}_{\text{robust}} = Z * \text{MAD}$$

$$\text{ROP}_{\text{robust}} = \text{med} + Z * \text{MAD}$$

Where :
Z = Normal table value for the given service level.
Med is the median of demand data
MAD is Median Absolute Deviation

4. Numerical Experiments

In order to prove the extent of the impact of the outliers in the demand data on the decisions taken regarding inventory management, as well as to prove the effectiveness of the proposed model to reduce the impact of outliers and compare it with the original model, the two models were applied to the demand data for the raw materials involved in cement production in one of Iraq's factories.

The demand data mentioned for 730 working days was taken directly from the demand data for materials (oil, limestone, dust and iron dust) to test the effectiveness of the two models on this data and compare the total cost of both models to indicate the preference between them.

Demand data tested for the purpose of detecting outliers by the box-plot method, the data test results were as follows:

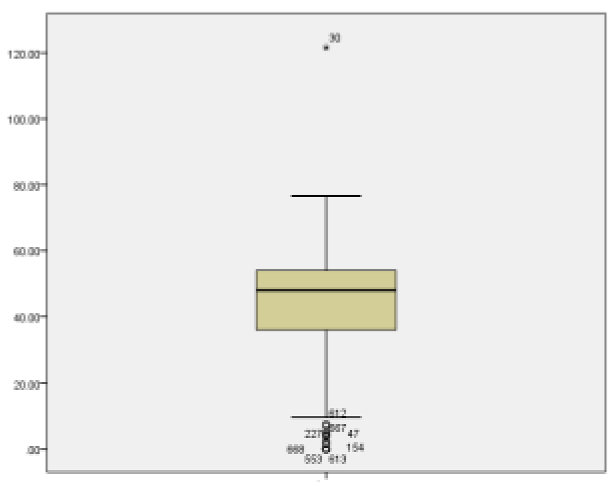

Figure 1: Detection of outliers of iron data

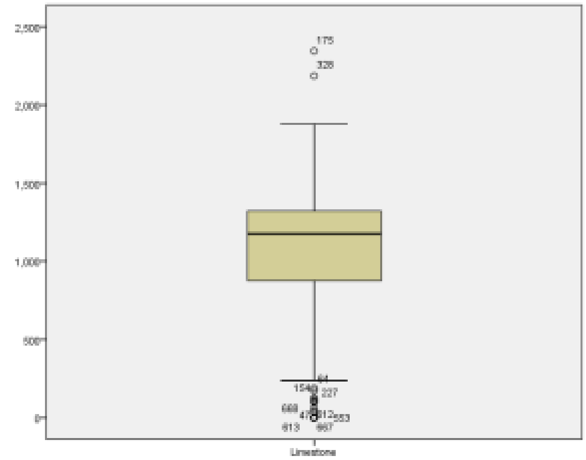

Figure 2: Detection of outliers of limestone data

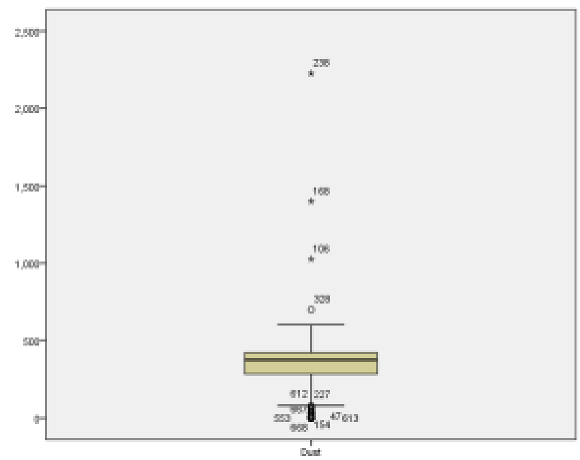

Figure 3: Detection of outliers of dust data

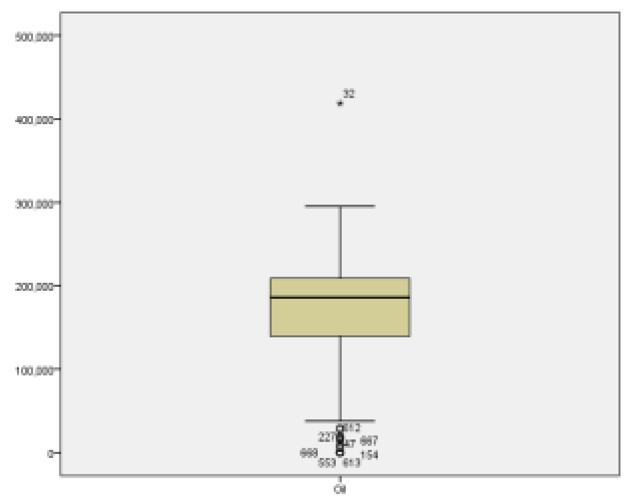

Figure 4: Detection of outliers of iron data

Through the demand data and costs related to storage, the order quantity, re-order point and safety stock were calculated, with the calculation of the total cost of the stock without calculating the deficit, for the original model and the proposed model, as in the table below:

Table 1: original and new model

		Q	Average	SS	ROP	Annual demand	Ordering cost	Holding cost	TC. for 730 days without shortage
Oil	original model	9009858.9	166803.2	181080	5185176	60,883,168	2000000	3	54059153.38
	proposed model	9509734.4	185825.44	127740	5702503	67,826,286	2000000	3	57058406.12
Limestone	original model	87763.12	1055.1	1149.23	64455.23	385,112	4000000	400	70209879.65
	proposed model	92658.14	1176.1	807.27	71373.27	429,277	4000000	400	74126508.08
Iron	original model	5948.75	43.09	47.05	5217.85	15,728	4500000	4000	47590024.16
	proposed model	6277.23	47.98	33.16	5790.76	17,513	4500000	4000	50217813.57
Dust	original model	44518.11	339.36	388.97	20750.57	123,866	4000000	500	44518113.17
	proposed model	46882.22	376.36	260.32	22841.92	137,371	4000000	500	46882218.38

And the following charts showed the comparison between the behavior of the original model and the proposed model during a period of 730 days:

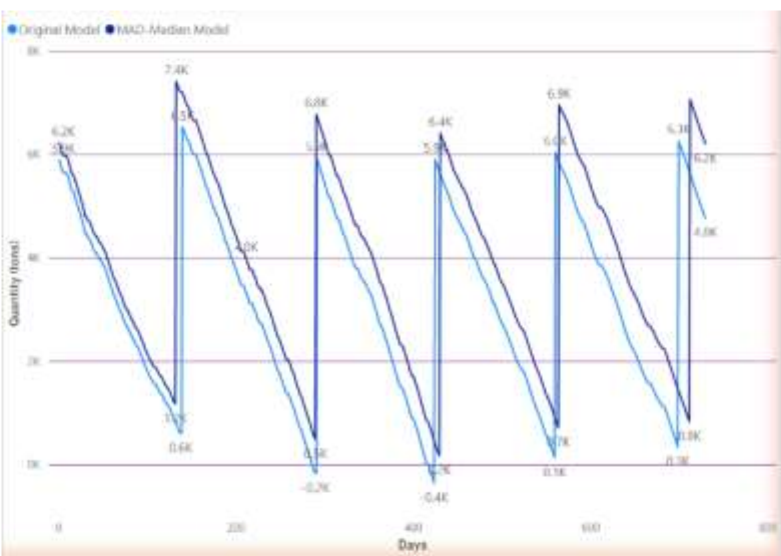

Figure 5:Comparing the proposed model with the original model within 730 days for iron

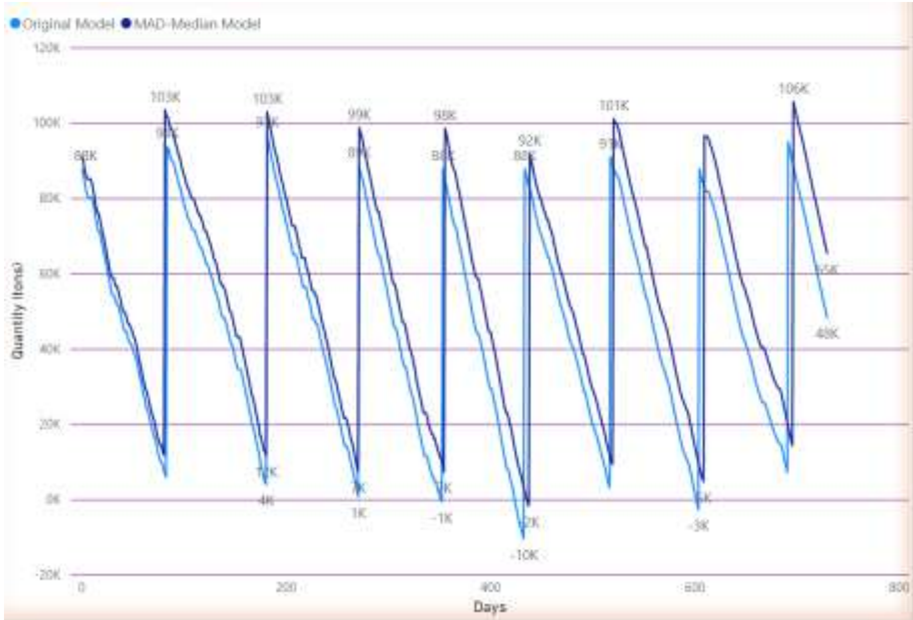

Figure 6: Comparing the proposed model with the original model within 730 days for limestone

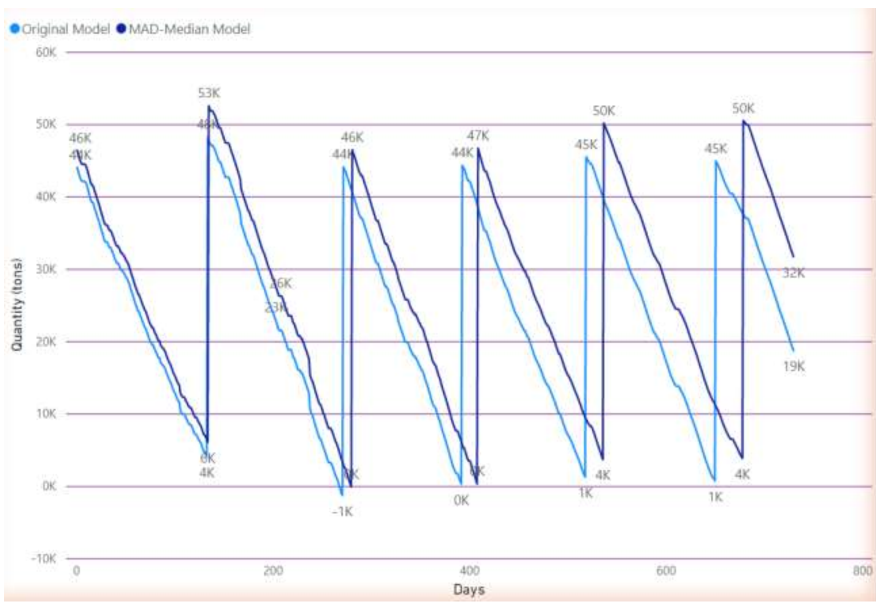

Figure 7: Comparing the proposed model with the original model within 730 days for dust

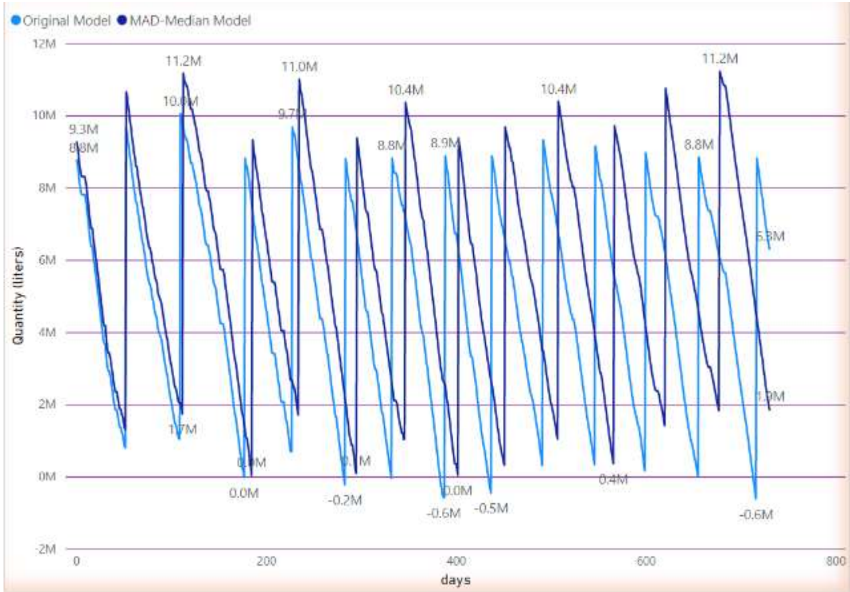

Figure 8:Comparing the proposed model with the original model within 730 days for oil

By testing each model, the following table showed the quantities related to the inability to meet the demand for raw materials and the lost quantities of cement production due to this deficit, as follows:

Table 2: shortage in original and developed models

Materials with the tow models		shortage first 365 day	shortage in cement production	second 365 days	shortage in cement production	total shortage	total of shortage in cement production
Oil	original model	296553.41	1602.99	1642810.2	8880.05	1939363.6	10483.05
	proposed model	0	0	0	0	0	0
Limestone	original model	582.29	529.36	13214.73	12013.39	13797.02	12542.74
	proposed model	0	0	1723.32	1566.66	1723.32	1566.66
Iron	original model	166.83	5772.32	362.43	12540.17	529.26	18312.49
	proposed model	0	0	0	0	0	0
Dust	original model	1318.75	3878.67	0	0	1318.75	3878.67
	proposed model	120.83	355.38	0	0	120.83	355.38

By comparing the results obtained, it can be noted that the deficit caused by the lack of oil supply in the original model, which amounted to approximately 1603 tons of cement production , decreased to zero for the first year when applying the proposed model, and from 8880 tons to zero also in the second year.

As for the limestone processing deficit caused by the original model, it caused a decrease in cement production of 529 tons of cement in the first year and 12013 tons in the second year, while the proposed model reduced the deficit by 100% for the first year and 87% in the second year.

While the iron deficiency in the original model caused a deficit of 18,312 tons of cement within two years, and the proposed model was able to reduce the percentage of deficit to zero during the same period.

Also, the lack of dust in the original model caused a deficit in cement production of 3878 tons in two years. As for the proposed model, the deficit was reduced to 355 tons during the same period.

Referring to table 1, it can be seen that Q, ROP and SS have differed in their values in the original model and the proposed model for all the four raw materials (oil, limestone, dust and iron dust), since the method of calculating Q value that is calculated has been changed, it was based on the annual demand, while the proposed model relied on the median value multiplied by 365 days to extract the value of Q because the annual demand was definitely affected by the presence of outliers.

Also, the median was re-used in the proposed model to extract ROP to get rid of outliers, while the original model was based on the arithmetic median, which is highly affected by any outliers, so the reorder points turned from (5185176, 64455.23, 5217.85, 20750.57)
to (5702503, 71373.27, 5790.76, 22841.92) for raw materials (oil, limestone, iron, and dust), respectively, and this change in re-order points has given the decision maker an additional amount in stock to avoid shortages and get rid of the influence of outliers.

As for the safety stock, it has been reduced by approximately 30% in the proposed model for the four raw materials, and this in itself will contribute to reducing inventory costs and a significant saving in the amounts invested in the capital in the safety stock, as MAD was used instead of the standard deviation to avoid the impact of outliers in calculating safety stock quantities.

It can also be noted that the increase in the quantity of Q has led to a slight increase in the total cost of inventory (without calculating the cost of the deficit), this is very natural, because an increase in the element of safety necessarily leads to an increase in the quantity of inventory and thus an increase in the cost. However, this increase is a very slight increase if compared to the cost of the deficit caused by the lack of supply of raw materials on demand and the halt in production that leads to the loss of a percentage of the profits, this is in isolation regarding the moral losses due to the suspension of cement production, the low level of customer satisfaction and the increase in production costs, and just calculating the size of the financial profits lost as a result of the production deficit gives an indication of the need to use exceptional models for inventory management when there are exceptional data for demand or the presence of extreme values within the data.

In summary, the table below shows the actual value of the lost profits due to the deficit in cement production, which was caused by the lack of raw materials on demand for the purpose of production, noting that the cement production plant under study records a profit of (20000) ID for each ton sold of cement.

Table 3: total profit lost, inventory total cost with and without shortage

	original model				proposed model			
Materials	total of shortage in cement production	lost profits	TC. of inventory without shortage	TC. of inventory with shortage	total of shortage in cement production	lost profits	TC. of inventory without shortage	TC. of inventory with shortage
oil	10483	209661000	54059153	263720153	0	0	57058406	57058406
limestone	12543	250854800	70209880	321064680	1567	31333200	74126508	105459708
iron	18312	366249800	47590024	413839824	0	0	50217814	50217814
dust	3879	77573400	44518113	122091513	355	7107600	46882218	53989818
total	45217	904339000	216377170	1120716170	1922	38440800	228284946	266725746

As the model succeeded in reducing the losses resulting from the deficit and production stoppage from 45,217 tons of cement to 1922 tons and reduce the loss of profits amounting to 904339000 dinars to 38,440,800 dinars, which means that it has contributed to reducing losses by 96%, also the proposed model has reduced the total cost of inventory by 75% as a result of reducing the impact of the presence of negative outliers and the use of robust statistics instead of classical statistics.

5. Conclusion

In this paper, the proposal of a new model for inventory management was discussed in the event that the data contained outliers and recalculated the amount of Q, ROP and SS after replacing the classical statistics with robust statistics with a breakdown point of up to 50%, which is the median and MAD, and then the model was examined the proposed model is based on real demand data for the raw materials used in the cement industry in one of the cement factories in Iraq. The proposed model succeeded in reducing the impact of outliers, improving the decisions taken to manage inventory in the factory, and reducing costs, in addition to reducing the safety stock.

References

Agarwal, S. (2015). ECONOMIC ORDER QUANTITY MODEL : A REVIEW REVIEW PAPER ECONOMIC ORDER QUANTITY MODEL : A REVIEW. December 2014.

Arachchige, C. N. P. G., & Prendergast, L. A. (2019). Confidence intervals for median absolute deviations. 2, 1–13.

Benjelloun, F. Z., Oussous, A., Bennani, A., Belfkih, S., & Ait Lahcen, A. (2019). Improving outliers detection in data streams using LiCS and voting. Journal of King Saud University - Computer and Information Sciences, xxxx. https://doi.org/10.1016/j.jksuci.2019.08.003

Bollen, K. A. (1987). Outliers and improper solutions: A confirmatory factor analysis example.

Chung, N., Zhang, X. D., Kreamer, A., Locco, L., Kuan, P., Bartz, S., Linsley, P. S., Ferrer, M., & Strulovici, B. (2008). Genome-Scale RNAi Screens. 149–158. https://doi.org/10.1177/1087057107312035

Cousineau, D. (2010). Outliers detection and treatment : A review Outliers detection and treatment : a review . March 2015.

Cousineau, D. (2011). Outliers detection and treatment : a review . 3(1), 58–67.

Fatima, M., & Kurmi, J. (2018). Comparative Analysis of Outlier Detection Methods. 5(4), 1–5.

Garces, H., & Sbarbaro, D. (2011). Outliers detection in environmental monitoring databases. Engineering Applications of Artificial Intelligence, 24(2), 341–349. https://doi.org/10.1016/j.engappai.2010.10.018

Gupta, M., Gao, J., Aggarwal, C. C., & Han, J. (2014). Outlier Detection for Temporal Data: A Survey. IEEE Transactions on Knowledge and Data Engineering, 26(9), 2250–2267. https://doi.org/10.1109/TKDE.2013.184

Hampel, F. R. (1974). The influence curve and its role in robust estimation. Journal of the American Statistical Association, 69(346), 383–393.

Huber, P. J., & Wiley, J. (1981). Robust Statistics.

Kaiser, R., & Maravall Herrero, A. (1999). SEASONAL OUTLIERS IN TIME SERIES SERIES.

Lavrac, N., Miksch, S., & Kavsek, B. (2000). 5 th International Workshop on Intelligent Data Analysis in Medicine and Pharmacology A workshop at.

Lazarevic, A., & Kumar, V. (2005). Feature bagging for outlier detection. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 157–166. https://doi.org/10.1145/1081870.1081891

Lu, C. T., Chen, D., & Kou, Y. (2003). Algorithms for spatial outlier detection. Proceedings - IEEE International Conference on Data Mining, ICDM, 597–600. https://doi.org/10.1109/icdm.2003.1250986

Nemtajela, N., & Mbohwa, C. (2017). Relationship between inventory management and uncertain demand for fast moving consumer goods organisations. Procedia Manufacturing, 8(October 2016), 699–706. https://doi.org/10.1016/j.promfg.2017.02.090

Olewuezi, N. P. (2011). Note on the Comparison of Some Outlier Labeling Techniques. Journal of Mathematics and Statistics 7, 7(4), 353–355.

Osborne, J. W., Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should ALWAYS check for them). 9.

P. O.Agada, & Ogwuche, E. H. (2017). A probabilistic economic order quantity (eoq) model for inventory management of drugs and hospital consumables. 2(2), 737–742.

Pankratz, A. E. (2000). Outliers in multivariate time series. 789–804.

Paramasivan, C., & Subramanian, T. (2008). financial management.

Personal, M., Archive, R., Pan, J., Shachat, J., & Wei, S. (2018). Cognitive stress and learning Economic Order Quantity (EOQ) inventory management: An experimental investigation. 93214.

Picek, J. (2012). Robust Estimation of Location and Regression. Methodology in Robust and Nonparametric Statistics, 69–116. https://doi.org/10.1201/b12681-4

Pokrajac, D., Lazarevic, A., & Latecki, L. J. (2007). Incremental local outlier detection for data streams. Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2007, April, 504–515. https://doi.org/10.1109/CIDM.2007.368917

Rouss, P. J., & Christophe, C. (1993). Alternatives to the Median Absolute Deviation. 88(424).

Seo, S. (2006). A review and comparison of methods for detecting outliersin univariate data sets. In Department of Biostatistics, Graduate School of Public Health. http://d-scholarship.pitt.edu/7948/

Shekarian, E., Olugu, E. U., Abdul-rashid, S. H., & Kazemi, N. (2016). Analyzing optimization techniques in inventory models : the case of fuzzy economic order quantity problems. 1229–1240.

Torgo, L., & Ribeiro, R. (2003). Predicting Outliers. 447–458.

Tukey, J. W. (1960). A survey of sampling from contaminated distributions. Contributions to Probability and Statistics 448-485., 1960.

Voloh, B., Watson, M. R., & König, S. (2018). MAD saccade : statistically robust saccade threshold estimation via the median absolute deviation. 12(8), 1–12.

Zhang, Q., Segall, R. S., & Cao, M. (2010). Visual Analytics and Interactive Technologies : Data , Text and Web Mining Applications.

Zhang, J. (2013). Advancements of Outlier Detection : A Survey. 13(01), 1–26. https://doi.org/10.4108/trans.sis.2013.01-03.e2

Zhang, Y., Luo, A., & Zhao, Y. (1996). Outlier detection in astronomical data. 1980.

Zhao, X., Liang, J., & Cao, F. (2014). A simple and effective outlier detection algorithm for categorical data. 469–477. https://doi.org/10.1007/s13042-013-0202-4