A Hybrid Quality Control Strategy for Consumption Data on Petroleum, Oil and Lubricants by Oil Equipment

Bixin Li, Rong Fan*, Ji Tang, Jianghe Qiu, Yinyin Wang

Department of Petroleum Oil Lubricants, Army Logistics University of PLA, Chongqing 401331, China

Corresponding Author Email: frfr1002@126.com

ABSTRACT

Petroleum, oil and lubricants (POL) is critical to the functioning of oil equipment. It is very important to maintain the quality of POL consumption data of oil equipment. High-quality POL consumption data provide reliable datasets for computing the annual POL supply quota of oil equipment and improving the energy efficiency of such equipment. Based on the data analysis platform SPSS, this paper designs a hybrid quality control strategy called SEIB-K for POL consumption data of oil equipment, which couples the expectation-maximization algorithm with multiple imputations (EI), box plot (B), and kappa coefficient (K). The SEIB-K can systematically analyze the quality of actual POL consumption data, identify the missing items, abnormal items and contradictory items. Example analysis shows that our strategy could significantly improve the quality of POL consumption data. The research results provide a technical support for accurately determining the annual POL supply quota of oil equipment.

1. INTRODUCTION

The consumption of petroleum, oil and lubricants (POL) determines the annual POL supply quota for oil equipment. The POL consumption is usually evaluated through statistical modelling based on the big data of the specific type of oil equipment. Apparently, the quality of POL consumption data directly bears on the accuracy of the annual POL supply quota [1-3].

However, the data on POL consumption of oil equipment come from various sources, fall in different classes, and accumulate to an enormous size. What is worse, the data quality is greatly affected by various internal and external factors, ranging from equipment conditions, measurement instruments, to operator skills. In most cases, the POL consumption data are not sufficiently objective, accurate, or complete.

If the raw data are directly used, huge errors will arise in the calculation of POL supply quota of oil equipment [4, 5], posing a severe threat to equipment operation. To solve the problem, the quality of POL consumption data must be improved, providing a scientific, accurate and applicable basis for the computation of POL supply quota for oil equipment [6, 7].

Despite the abundant research into data quality control, there is little report on the quality of POL consumption data of oil equipment. To make up for the gap, this paper puts forward a quality control strategy called SEIB-K for POL consumption data of oil equipment. The SEIB-K supplements the missing items by the expectation-maximization algorithm with multiple imputations (EI), and relies on the SPSS platform to control the relevant data on POL consumption of oil equipment. The research results promote the energy efficiency of oil equipment.

2. QUALITY CONTROL OF POL CONSUMPTION DATA

To accurately calculate the POL supply quota for oil equipment, the raw data on POL consumption by oil equipment were collected by POL Service Information System and POL Management Information System. The collected dataset covers many attributes of oil equipment, including serial number, model, rated power, weight, working hour, main tank capacity, auxiliary tank capacity, fuel density, annual fuel consumption, to name but a few [8, 9].

As mentioned before, the numerous data in the original dataset belong to various sources, and different classes, and suffer from the interference of various factors. It is inevitable for the original dataset to contain missing items, duplicate items, abnormal items, and contradictory items. Obviously, it is impossible to directly derive POL supply quota from the original dataset. To solve the problem, the four kinds of items were treated in the following manner.

(1) Missing items: the POL consumptions of some oil equipment are not reported. These missing items were supplemented by the EI algorithm.

(2) Duplicate items: the duplicate items stem from lack of objectivity in data collection. In many cases, the POL consumptions of equipment in the same category are assumed to be the same, which clearly goes against the reality. The resulting duplicate items were detected on SPSS platform.

(3) Anormal items: under special conditions, the POL consumptions of oil equipment are inaccurate. The abnormal items were processed based on probability distribution and box plot.

(4) Contradictory items: some POL consumption data are illogical [10, 11]. The contradiction was solved in the light of the kappa coefficient.

After the above quality control process, the preprocessed dataset was adopted to calculate the POL supply quota of oil equipment. The entire workflow from the input of raw data to the output of POL supply quota is explained in Figure 1 below.
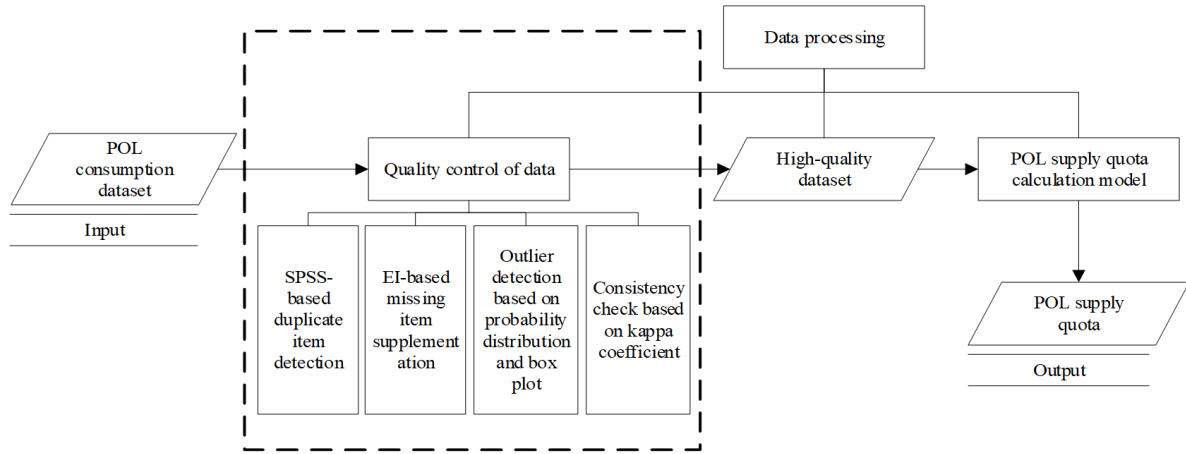

Figure 1. Workflow of data quality control and calculation of POL supply quota

3. SEIB-K STRATEGY

3.1 EI-based supplementation of missing items [12, 13]

Proposed by Dempster et al., expectation maximization (EM) algorithm estimates the maximum likelihood through iteration [1]. Each iteration consists of an expectation (E) step and a maximum likelihood estimation (M) step. Here, the EM algorithm is coupled with multiple imputations (MI) algorithm to supplement the missing items in the original dataset of POL consumption data [2].

3.1.1 The EM algorithm

Let $P(\theta|Y)$ be the posteriori distribution of parameter θ, $P(\theta|Y,Z)$ be the added posteriori distribution of parameter θ, $P(\theta|Z,Y)$ be the conditional distribution of data Z, and $\theta^{(t)}$ be the θ value estimated in the t-th iteration. Then, the E and M steps of the $t+1$-th iteration can be expressed as:

(1) E step

The expected value in the $t+1$-th iteration can be calculated as:

$$Q(\theta \mid \theta^{(t)}, Y) = E(\log P(\theta \mid Y, Z) \mid \theta^{(t)}, Z)$$
$$= \int \log P(\theta \mid Y, Z) \cdot P(Z \mid \theta^{(t)}, Y) d_z$$

(2) M step

The $\theta^{(t+1)}$ obtained in the E step can be validated by:

$$Q(\theta^{(t+1)} \mid \theta^{(t)}, Y) = \max Q(\theta \mid \theta^{(t)}, Y)$$

E and M steps are repeatedly executed until $|Q(\theta^{(t+1)} \mid \theta^{(t)}, Y) - \max Q(\theta \mid \theta^{(t)}, Y)| \leq \varepsilon$, where ε is a sufficiently small positive integer.

3.1.2 The MI algorithm [14, 15]

Inspired by Bayes estimation, the MI algorithm assumes that the missing items are random and the result of observed value. The MI algorithm is generally implemented in three steps:

Step 1. Multiple possible imputation values are generated for each missing item, resulting in several complete datasets.

Step 2. Each dataset is subjected to statistical analysis.

Step 3. The dataset with the best analysis result is selected for supplementing the missing items.

3.1.3 The EI algorithm

The EI algorithm is the combination of the EM algorithm and the MI algorithm. Let Y be an $n \times m$ matrix with m variables, Y_{obs} be the observed part of the matrix, and Y_{mis} be the missing part of the matrix. Then, the EI algorithm can supplement the missing data through the following procedure:

Step 1. The vector θ^* is randomly extracted from the target parameter vector θ.

Step 2. The EM algorithm is called to obtain $Y^*_{i(mis)}$ in conditional distribution $P(Y^*_{i(mis)} \mid Y^*_{i(obs)}, \theta^*)$.

Step 3. The imputation parameter α can be computed based on the complete dataset (Y_{obs}, Y^*_{mis}):

$$\hat{\alpha} = \hat{\alpha}(Y) = \hat{\alpha}(Y_{obs}, Y^*_{mis})$$

The variance between the imputations is $\hat{U} = \text{var}(\hat{\alpha})$.

Step 4. The above steps are repeated M times to obtain $\hat{\alpha}_k$. In $\hat{U}_k$, $k = 1, 2, \cdots, M$.

Step 5. The imputation parameter α for the EI algorithm can be finalized as:

$$\bar{\alpha} = P(\hat{\alpha}) \cdot \hat{\alpha}$$

3.2 Outlier detection based on probability distribution and box plot

Outlier detection aims to look for items that significantly deviate from the other useful items in the massive POL consumption data, which are noisy, stochastic, and incomplete [16]. The procedure of outlier detection is presented in Figure 2 below.

The POL consumption data of oil equipment are a discrete random dataset. Let $x_1, x_2, \cdots, x_n$ be the values of variable X, and $p_1, p_2, \cdots, p_n$ be the probabilities corresponding to the X values. Then, the probability distribution of the discrete random variable X can be expressed as:

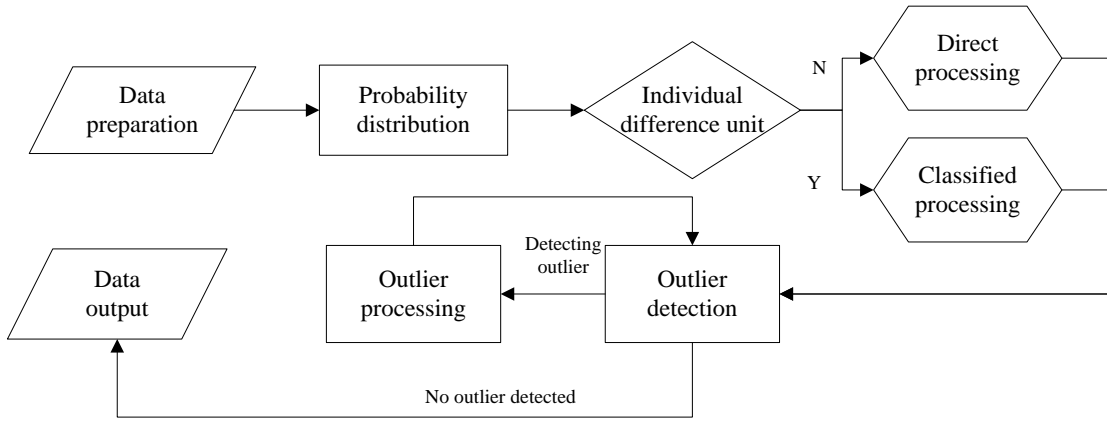$$P(X = x_i) = p_i, i = 1, 2, \cdots, n$$

Figure 2. The workflow of outlier detection

The probability p_i satisfies the condition $\sum_{i=1}^{n} p_i = 1$. Hence, the probability distribution function of discrete random variable X can be established as:

$$F(x) = \sum_{x_i < n} p_i$$

The outliers in POL consumption data were located through box plot analysis, a popular outlier detection method. The structure of the box plot is presented in Figure 3 below.
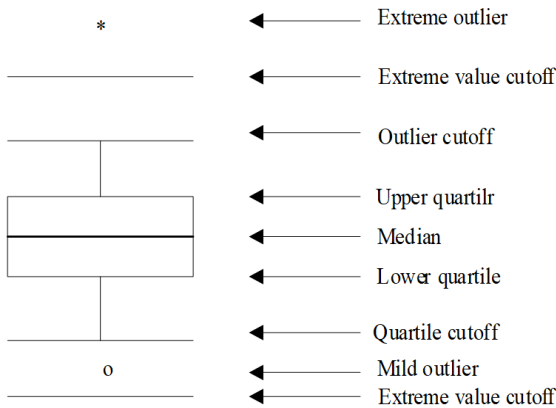

Figure 3. Box plot structure

For simplicity, the upper quartile, lower quartile, and the difference between them are denoted as U, L and IQR (IQR=U-L), respectively. Then, the upper outlier cutoff, lower outlier cutoff, upper bound, and lower bound can be determined by U_1=U+1.5IQR, L_1=L-1.5IQR, U_2=U+3IQR, and L_2=L-3IQR, respectively. The outliers between upper cutoff and upper bound are defined as mild outliers; those beyond the upper bound are defined as extreme outliers.

By the standard of the box plot, the items below L-1.5IQR, or above U+1.5IQR are outliers. Albeit a bit of arbitrariness, the outlier detection is empirical and suitable for processing the data needing special attention. Previous studies have shown the superiority of box plot in outlier detection [17, 18].

3.3 Consistency check based on kappa coefficient [19-21]

Kappa coefficient is calculated based on crosstab to measure data consistency:

$$K = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}$$

where, r is the number of rows (columns) in the crosstab; x_{ii} is the number of type combinations along the diagonal; x_{i+} is the total number of observations in row i; x_{+i} is the total number of observations of column i; N is the number of medium observations.

The calculated value of kappa coefficient fell between -1 and 1, but the normal value range is [0,1] [22, 23]. The consistency between two items is divided into the following levels based on the kappa coefficient [6]:

If kappa < 0.2, the logical consistency between the two items is strongly weak;

If $0.2 \leq$ kappa < 0.4, the logical consistency between the two items is slightly weak;

If $0.4 \leq$ kappa < 0.6, the logical consistency between the two items is neutral;

If $0.6 \leq$ kappa < 0.8, the logical consistency between the two items is slightly high;

If $0.8 \leq$ kappa < 1, the logical consistency between the two items is strongly high [24, 25].

4. EXAMPLE ANALYSIS

To verify its effectiveness, the proposed hybrid strategy S-EIBK was applied to process the example dataset randomly selected from POL consumption data of oil equipment. Based on the SPSS data analysis platform, the data quality was controlled using EI algorithm, box plot analysis and kappa testing.

4.1 EI-based supplementation of missing items

Based on the SPSS platform, the missing items in the example dataset were separately analyzed and calculated by mean estimation, regression prediction, the EM algorithm, and the EI algorithm. The maximum number of iterations of the EM algorithm was set to 25, and that of the MI part of the EI algorithm was set to 20.

The original dataset contains 100 items, which were randomly selected from the POL consumption data of oil equipment. The error of each algorithm was calculated by

$$\Delta = \frac{|original - estimate\ value|}{original\ value} \times 100\%.$$

Table 1 shows some of the missing items estimated by the contrastive algorithms. The errors of the four algorithms are compared in the scatter plot of Figure 4.

Table 1. Estimates of missing items

Algorithm	$X_{(4,7)}$		$X_{(4,10)}$		$X_{(10,8)}$		$X_{(15,10)}$		$X_{(19,8)}$	
	Data	Error	Data	Error	Data	Error	Data	Error	Data	Error
Original dataset	5,135		49.5		75.3		36.9		49.4	
Mean estimation	5,204.7	1.4%	50.2	2.2%	75.5	0.3%	50.2	36.0%	75.5	52.8%
Regression prediction	5,352	4.2%	50.9	2.8%	59.8	20.6%	45.3	22.8%	58.3	18.0%
EM algorithm	5,195	1.2%	50.1	1.2%	62.3	17.3%	36.9	0.0%	56.6	14.6%
EI algorithm	4,956	3.5%	47.8	3.4%	68.2	9.4%	36.9	0.0%	54.8	11.0%

Algorithm	$X_{(21,7)}$		$X_{(30,10)}$		$X_{(32,7)}$		$X_{(54,8)}$		$X_{(62,10)}$	
	Data	Error	Data	Error	Data	Error	Data	Error	Data	Error
Original dataset	6,603		52		4052		86.7		42.5	
Mean estimation	5,204.7	21.2%	50.2	3.5%	5204.7	28.4%	75.5	12.9%	50.2	18.1%
Regression prediction	4,829	26.9%	48.5	6.7%	4861	20.0%	81.9	5.5%	39.2	7.8%
EM algorithm	6,604	0.0%	53.5	2.9%	4056	0.1%	64.9	25.1%	42.5	0.0%
EI algorithm	6,605	0.0%	52	0.0%	4054	0.0%	78.5	9.5%	42.5	0.0%

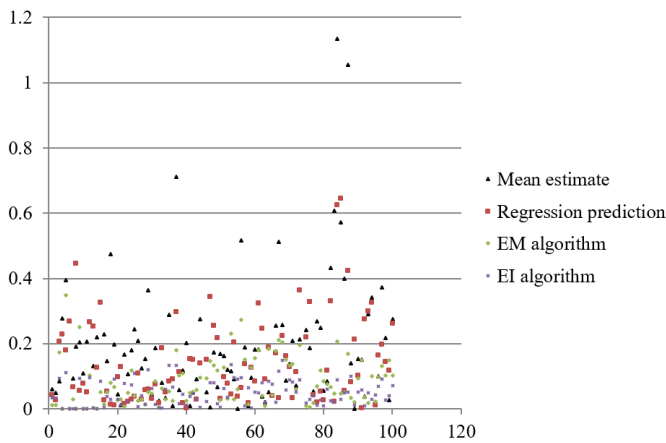| Algorithm	$X_{(71,7)}$		$X_{(77,8)}$		$X_{(86,10)}$		$X_{(88,7)}$		$X_{(96,8)}$	
	Data	Error	Data	Error	Data	Error	Data	Error	Data	Error
Original dataset	6,083		94.2		41.6		6,192		31.6	
Mean estimation	5,024.7	14.4%	75.5	19.9%	47.1	13.2%	5,204.7	15.9%	75.5	138.9%
Regression prediction	5,772	5.1%	69.2	26.5%	52.1	25.2%	6,972	12.6%	21.3	32.6%
EM algorithm	6,087	0.1%	84.6	10.2%	41.5	0.2%	6,192	0.0%	33.2	5.1%
EI algorithm	6,092	0.1%	88.9	5.6%	41.5	0.2%	6,190	0.0%	34.2	8.2%

Figure 4. Comparison of errors of the four algorithms

As shown in Figure 4, the mean estimation and regression prediction had much larger errors than EM algorithm and EI algorithm. The EI algorithm achieved a smaller error than EM algorithm. It can be concluded that EI algorithm is the most effective method in supplementing the missing items in the example dataset.

The excellence of the EI algorithm, which combines the merits of EM algorithm and MI algorithm, is attributable to the following factors: First, multiple imputations are performed to increase the accuracy of the supplementary items; Second, the uncertainty of the missing items are fully considered in the algorithm, making the prediction closer to the reality; Third, the algorithm can simulate the distribution of the missing items, without changing the relationship between variables.

4.2 Outlier detection based on probability distribution and box plot

The annual fuel consumptions were selected from the POL consumption data of oil equipment as an example. Based on the SPSS platform, the probability distribution and scatter plot of the annual fuel consumptions were generated as Figures 5 and 6, respectively.

As shown in Figures 5 and 6, the annual fuel consumptions in the example dataset are normally distributed around 5,000. There were no individual difference unit, but some obvious outliers. Based on the SPSS, the outlier detection parameters were configured as follows: U=5,699; L=4,300; IQR=1,399; U_1=7,797.5; L_1=2,201.5; U_2=9,896; L_2=103.

Then, a box plot of the annual fuel consumptions in the example dataset was prepared on the SPSS platform (Figure 7). On this basis, the locations and sizes of outliers were obtained as in Table 2.

As shown in Figure 7 and Table 2, there were many outliers in the annual fuel consumptions. These outliers were checked carefully, and then supplemented by the EI algorithm as missing items. The supplemented dataset of annual fuel consumptions was checked again for outliers, and supplemented again if there were still outliers. This process was repeated until no outlier is left.
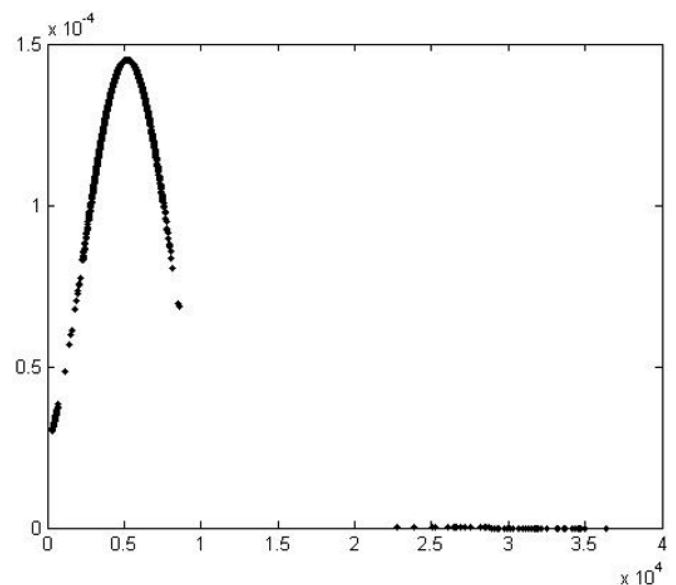

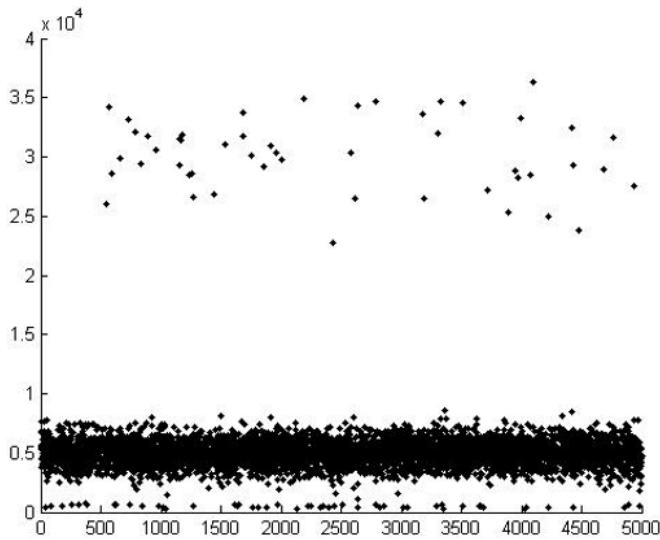**Figure 5.** The probability distribution of annual fuel consumptions
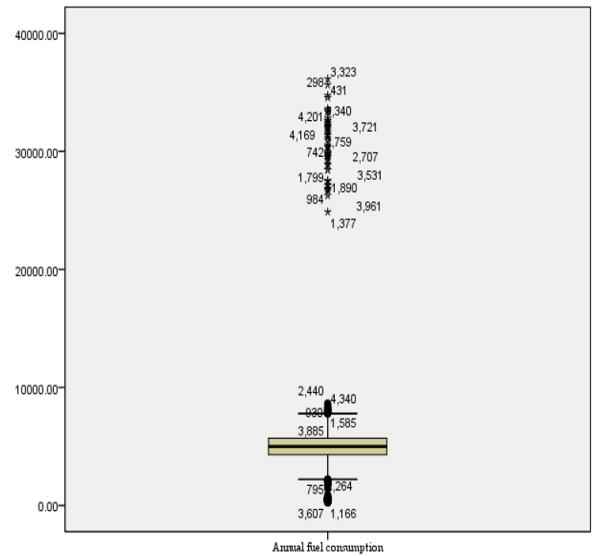
Figure 6. Distribution of annual fuel consumptions

Figure 7. The box plot of annual fuel consumptions

Table 2. The outliers of annual fuel consumptions

Upper outliers				Lower outliers			
Location	Size	Location	Size	Location	Size	Location	Size
132	8249.2	2209	32849.8	60	393.2	2264	343
146	30004	2440	8615.4	369	1472	2268	587.5
213	27465.6	2707	29280.5	388	495.6	2395	545.5
227	29621.4	2759	29485.2	399	396.7	2517	398.6
...	...	...	...	...	...	...	...
1959	8226.7	4392	8079.3	2050	524	4580	493.3
2086	31162.7	4684	28448.3	2105	2190.2	4662	404.8
2135	8043	4708	27544.5	2222	516.9	4856	2062.8
2147	33464	4949	30336.1	2257	572.6		

4.3 Consistency check based on kappa coefficient

Through a survey on 100 units, indices A1 and A2 were selected to check the consistency between the annual fuel consumptions with the actual data. The results of the crosstab-based consistency check are recorded in Table 3.

Then, kappa test was carried out using the functional logic consistency of analysis-description statistics-crosstab in the SPSS. The test results are displayed in Table 4.

As shown in Table 4, kappa = 0.514 and P < 0.00 1 manifest the presence of logical consistency between indices A1 and A2.

Since the kappa coefficient fell between 0.4 and 0.6, the logical consistency between the two indices is neutral. This means the POL consumption dataset, after being processed by the proposed hybrid strategy, has high logical consistency, without any contradictory items. In other words, the processed dataset is highly effective, requiring no removal of contradictory items.

Table 3. The results of the crosstab-based consistency check

		A1			
		Satisfy	Normal	Dissatisfy	Total
	Satisfy	21	7	2	30
A2	**Normal**	4	48	3	55
	Dissatisfy	3	8	4	15
	Total	28	63	9	100

Table 4. The results of kappa consistency test

		Value	Asymptotic standard error a	Approximate value Tb	Approximate value Sig.
Consistency measure	**Kappa**	0.514	0.074	6.735	0.000
Invalid case N		100			
a. No null hypothesis;					
b. Null hypothesis with asymptotic standard error.					

5. CONCLUSIONS

The effective supply of the POL to oil equipment directly hinges on the accuracy of the annual POL supply quota, which depends on the quality of the data on POL consumption of oil equipment. If the POL consumption data are of low quality, the annual POL supply quota might deviate from the actual demand, failing to support the normal operation of the oil equipment.

Based on the SPSS platform, this paper designs a hybrid strategy to ensure the quality of the POL consumption data of oil equipment. As its name suggests, the proposed SEIB-K

strategy is a combination of the EI algorithm, the box plot, and the kappa coefficient.

Firstly, the EM and EI algorithms were integrated into the EI algorithm to identify and supplement the missing items in the POL consumption data. With the aid of SPSS platform, the EI algorithm was also adopted to detect and correct the duplicate items, improving the quality of the POL consumption dataset.

Next, the probability distribution and box plot were employed to detect and locate the extreme values in POL consumption dataset. The extreme values were corrected as outliers, eliminating the abnormal data in the dataset of POL consumption.

Finally, the contradictory items in the dataset were identified through kappa consistency test. After the removal of the contradictory items, the POL consumption dataset is suitable for determining accurate annual POL supply quota for oil equipment.

REFERENCES

[1] Hooi Lean, H., Mishra, V., Smyth, R. (2016). Conditional convergence in US disaggregated petroleum consumption at the sector level. Applied Economics, 48(32): 3049-3061. https://doi.org/10.1080/00036846.2015.1133901

[2] Sunnu, A.K., Ayetor, G.K., Gaye, J.M. (2019). Straight vegetable oil fuel performance and exhaust emissions under turbocharged and naturally aspirated conditions. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 1-11. https://doi.org/10.1080/15567036.2019.1677822

[3] Hoesly, R.M., Smith, S.J. (2018). Informing energy consumption uncertainty: an analysis of energy data revisions. Environmental Research Letters, 13(12): 124023. https://doi.org/10.1088/1748-9326/aaebc3

[4] Van Hoang, T.H., Shahzad, S.J.H., Czudaj, R.L., Bhat, J.A. (2019). How do oil shocks impact energy consumption? A disaggregated analysis for the US. The Energy Journal, 40(1): 167-210. https://doi.org/10.5547/01956574.40.SI1.thoa

[5] Brahim, A.O., Abderafi, S. (2017). Pressure effect on the stabilization column in the petroleum refinery. Energy Procedia, 118: 233-237. https://doi.org/10.1016/j.egypro.2017.07.013

[6] Yang, R., Ozer, H., Al-Qadi, I.L. (2016). Regional upstream life-cycle impacts of petroleum products in the United States. Journal of Cleaner Production, 139: 1138-1149. https://doi.org/10.1016/j.jclepro.2016.08.164

[7] Pinos, A.A.R., Badoga, S., Dalai, A.K., Adjaye, J. (2019). Modelling of H2 consumption and process optimization for hydrotreating of light gas oils. The Canadian Journal of Chemical Engineering, 97(6): 1828-1837. https://doi.org/10.1002/cjce.23450

[8] Tolliver, C., Islam, M., Shin, K.J., Managi, S. (2018). The impact of energy security risks on energy consumption. International Journal of Innovation and Sustainable Development, 12(3): 258-270. https://doi.org/10.1504/IJISD.2018.091522

[9] Aravind, M., Nayar, J. (2020). Integration of oil with macroeconomic indicators and policy challenges in regard to Oman. International Journal of Energy Sector Management, 14(1): 172-192.

https://doi.org/10.1108/IJESM-08-2018-0006

[10] Mayer, A., Hazboun, S.O. (2019). Does fracking drive you to drink? Unconventional oil and gas production and alcohol consumption in US counties. The Extractive Industries and Society, 6(3): 823-830. https://doi.org/10.1016/j.exis.2019.04.002

[11] Jackson, D., Huang, M., Fernando, H., Ansari, G., Howarth, M., Mesaros, C., Penning, T., Elferink, C. (2019). Using precision environmental health principles in risk evaluation and communication of the deepwater horizon oil spill. NEW SOLUTIONS: A Journal of Environmental and Occupational Health Policy, 28(4): 599-616. https://doi.org/10.1177/1048291118815606

[12] Nasyrov, I.A., Ahmadieva, A.I., Fazullin, D.D., Mavrin, G.V., Sokolov, M.P. (2017). Petroleum containing wastewater products purification by carbon-containing wastes pyrolysis products. The Turkish Online Journal of Design, Art and Communication, 7: 1713-28. https://doi.org/10.7456/1070DSE/150

[13] Aldhaidhawi, M., Miron, L., Chiriac, R., Badescu, V. (2018). Autoignition process in compression ignition engine fueled by diesel fuel and biodiesel with 20% rapeseed biofuel in diesel fuel. Journal of Energy Engineering, 144(5): 04018049. https://doi.org/10.1061/(ASCE)EY.1943-7897.0000563

[14] Zhang, C., Fu, J., Pu, Z. (2019). A study of the petroleum trade network of countries along "The Belt and Road Initiative". Journal of Cleaner Production, 222: 593-605. https://doi.org/10.1016/j.jclepro.2019.03.026

[15] Peng, P., Cheng, S., Lu, F. (2020). Characterizing the global liquefied petroleum gas trading community using mass vessel trajectory data. Journal of Cleaner Production, 252: 119883. https://doi.org/10.1016/j.jclepro.2019.119883

[16] Das, R.K., Sharma, S.K. (2017). Fuel characterization and performance parameters analysis of diesel engine using blends of palm biodiesel and tyre pyrolysis oil. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 39(5): 1491-1497. https://doi.org/10.1007/s40430-016-0696-2

[17] Li, S., Wang, Q. (2019). India's dependence on foreign oil will exceed 90% around 2025-The forecasting results based on two hybridized NMGM-ARIMA and NMGM-BP models. Journal of Cleaner Production, 232: 137-153. https://doi.org/10.1016/j.jclepro.2019.05.314

[18] Dolotovskii, I.V., Dolotovskaya, N.V., Larin, E.A. (2018). Software and dataware for energy generation and consumption analysis system of gas processing enterprises. In Journal of Physics: Conference Series 1015(4): 1-5. https://doi.org/10.1088/1742-6596/1015/4/042010

[19] Fushimi, C., Tachibana, C., Tomita, R., Sakurai, M. (2018). Optimization of Energy Consumption and Cost in Solvent Recovery Section for Large-Scale Biofuel Production from Microalgae. Journal of Chemical Engineering of Japan, 51(9): 809-816. https://doi.org/10.1252/jcej.17we402

[20] Wang, M., Guo, Z., Jiao, C., Lu, S., Li, J., Xue, H., Li, J., Li, J., Chen, G. (2019). Exploration progress and geochemical features of lacustrine shale oils in China. Journal of Petroleum Science and Engineering, 178: 975-986. https://doi.org/10.1016/j.petrol.2019.04.029

[21] Liu, G.X., Wu, M., Jia, F.R., Yue, Q., Wang, H.M. (2019). Material flow and spatial data analysis of the

petroleum use to carbon dioxide (CO_2) emissions in Northeast China. Journal of Industrial Ecology, 23(4): 823-837. https://doi.org/10.1111/jiec.12819

[22] Brahim, A.O., Abderafi, S., Bounahmidi, T. (2017). Modeling the stabilization column in the petroleum refinery. Energy Procedia, 139: 61-66. https://doi.org/10.1016/j.egypro.2017.11.173

[23] Güntner, J.H. (2019). How do oil producers respond to giant oil field discoveries? Energy Economics, 80: 59-74.

https://doi.org/10.1016/j.eneco.2018.12.012

[24] Akramov, T.F., Yarkeeva, N.R. (2019). Rational usage of associated petroleum gas. SOCAR Proceedings, (1): 68-77. http://dx.doi.org/10.5510/OGP20190100382

[25] Al-Sahlawi, M., Miah, F., Al-Titi, O. (2018). Gasoline consumption in Saudi Arabia: An analysis of consumer attitudes and perceptions using microsurvey data. OPEC Energy Review, 42(4): 355-386. https://doi.org/10.1111/opec.12135