Empirical Study on CNG Refilling Behaviors of Different Timescale Based on Boxplot Method

Yang Li¹,²,*, Yulian Zhao¹, Jian Shao²
¹College of Engineering and Technology, Tianjin Agricultural University, Tianjin 300384, China
²Tianjin Huabei Gas & Heating Engineering Design Co. LTD,Tianjin 300384,China
Email address: liyang005@163.com(Y. Li)

Abstract. The number of CNG secondary filling stations is limited in cities because of their security requirement. However, the number of CNG vehicles is increasing. It is important to improve the efficiency of CNG secondary filling stations, as critical infrastructures in CNG industry, to meet the increasing demand. As a first step to find possible improvement methods, the periodical CNG refueling behaviors are empirically studied in this article. By boxplot method, the intraday, day-of-the-week and monthly CNG refueling behaviors are studied. Based on analysis of more 250 thousand records of a CNG secondary filling station in Tianjin, the hourly, daily, and monthly volume, initial pressure, final pressure and the time spent in each refueling behavior are presented. From the central tendency and outliers analysis, it is found that one possible way to improve efficiency is to consider factors in intraday and monthly timescale. Another way to improve efficiency is to reduce outliers, which are independent of time.

1. Introduction
As a transportation fuel, Composed natural gas (CNG) is stored in high-pressure (20MPa in China) tanks on the vehicle. During the refilling process, CNG vehicles (CNGV) connect to the high pressure reservoir tanks at refueling stations through compressed natural gas dispenser for vehicle. Compared to traditional transportation fuel like diesel and gasoline, CNG, which mainly consisted of methane, used as an alternative to traditional transportation fuel has increased in the world in recent years for its environmental and economic merits. But for security limits, the CNG filling station, as a critical infrastructure in CNG industry, is not enough to meet the increasing number of CNG vehicles. Therefore, the CNGV drivers usually need to wait for hours[1-3] at the CNG secondary filling stations. Thus it is important to improve the efficiency of the CNG secondary filling stations. However, the first step for this improvement is to find the inefficient behaviors of these stations. In this paper, the refueling behaviors in different timescales, especially the intraday, intra-week, and monthly, are studied.

Given the fixed number of CNG dispensers, the way to reduce the time waiting for refilling is to improve the efficiency of each dispenser. Therefore, the periodic behaviors of a CNG secondary filling station in Tianjin are studied based on boxplot method through analyzing more than 250 thousand refueling records from the station. As reference [4] has pointed that the study on efficiency of CNG filling station are mostly based upon mathematical modeling[5] or simulation[6,7]. In this study, the
The empirical method of boxplot is used to mine the periodic behaviors in different scales, such as hour, day and month.

The article is arranged as follows: In Sec. 2, the data and the method of boxplot used in this study are described. Sec. 3 includes the empirical analysis results by figures through visualization. And the Sec. 4 concludes the study.

2. Dataset and method

2.1. Dataset

There are 254,841 records of a year refilling log of a CNG filling station in Tianjin in this study. Each log has 24 attributes. But only 5 attributes are used in this analysis. The 5 variables used in this study are occurring time(t in this paper), volume(v), initial pressure(pi), final pressure(pf) and the time total used for each refilling action(dt). The dataset was collected from January to December in 2014. Table 1 gives a description of the 5 attributes of this dataset.

| Index | Time         | Volume | Initial Pressure | Final Pressure | Time used for refilling |
|-------|--------------|--------|------------------|----------------|-------------------------|
| 1     | 1-1-2014 0:00 | 17.53  | 4.2              | 17.7           | 123                     |
| 2     | 1-1-2014 0:01 | 13.53  | 7.2              | 18.6           | 88                      |
| 3     | 1-1-2014 0:03 | 8.58   | 10.2             | 18.4           | 66                      |
| 4     | 1-1-2014 0:05 | 14.47  | 6.4              | 18.3           | 96                      |
| 5     | 1-1-2014 0:08 | 14.23  | 6.4              | 18.1           | 116                     |

The “Index” describes the order of the refilling behavior in this dataset. The attribute “Time” indicates when the refilling behavior happened. “Volume” is the amount of CNG filled into the on-board storage in the refilling behavior measured by mass flow meter of CNG dispenser under standard state, the “Initial Pressure” and “Final Pressure” provide the value of pressure of the on-board storage before and after the refilling behavior. And “Time used for refilling” describes the time used by each refilling action, which is calculated by the time interval between the time beginning refilling behaviors of each dispenser gun and the time finishing the refilling behavior, which is described as dt in this study.

2.2. Method

Since the number of stations is not easily increased in a certain time such as one year, one available way to improve refilling efficiency is to study the periodic behaviors of different time scales to optimize the equipment of each CNG filling station. As a common sense, the refilling behaviors in different hour in a day is different. But there is no empirical results. And the refilling behaviors in each day of a week, or each month of a year may be not the same yet. And there is short of related quantified results.

Therefore, for each time scale, boxplot method is used in this article for attributes of volume, initial and final pressure, and time used for each refilling action. For each attribute, the boxplot method gives its five statistics[8,9], such as the median, or the second quartile(Q2),the first quartile(Q1),the third quartile(Q3),the upper fence and the lower fence. The upper fence is defined as

\[ Q2 + 1.5 \times (Q3 - Q1) \]

And the lower fence is defined as

\[ Q2 - 1.5 \times (Q3 - Q1) \]
The calculation of these statistics is provided by python[10]. The analysis results of each attribute are also visualized by python language. Data with the value of attribute outside its upper and lower fences are considered as abnormal behaviors.

The statistics of different time scale realized by using attribute “Time” to partition each log. For intraday behaviors, if the ith action’s time is represented by t[i], then its intraday action is divided by class attribute of python t[i].hour. Intra-week, and monthly behaviors are detected by class attributes t[i].weekday() and t[i].month respectively.

To compare the difference between median and mean, the mean value of each attribute is computed. And the mean value is presented by dotted line.

3. Empirical analysis results

3.1 Intraday behaviors

Based on attributes of volume, initial pressure of on-board storage, final pressure of on-board storage, and time used for refilling action, the values five statistics of each hour in one day are calculated and visualized by boxplot. Figure 1 presents the results of boxplot analysis. The circle points outside the upper and lower fences are outliers. The dotted curves are the means of each hour in these figures.

Intraday behavior’s mean and median, both being widely used statistic tools to derive central tendency, show in figure 1 that volume, initial pressure and time interval obviously are not constant. The intraday behaviors of final pressure seems steady. By axes zoom effect, figure 2 presents the boxplot results with final pressure value above 16.8Mpa. From figure 2, there still changes in either median or mean.

For intraday behavior’s outliers, except volume, other 3 attributes have outliers in each hour. From pressure, the outliers are refilling actions with higher initial pressure or lower final pressure.
Fig. 1. Boxplot of intraday attributes (color online).

Fig. 2. Boxplot of intraday final pressure with limited values (color online).
Fig. 3. Boxplot of intra-week behaviors (color online).
3.2 Intra-week behaviors
By calculating the five statistics of refilling behavior in Monday to Sunday, the intra-week, or day-of-the-week results are presented in figure 3. The figure shows that all the four attributes nearly keep the same in each day of the week. In other words, there is no changes within a day, and there is no need for increasing CNG refilling stations’ efficiency in this timescale. For outliers, all attributes have abnormal behaviors in each day of the week. And these abnormal behaviors with either higher initial pressure or lower final pressure as the results of intraday behaviors.

3.3 Monthly behaviors
Figure 4 is the boxplot of monthly behaviors of the CNG secondary filling station. For the monthly central tendency, the figure shows that it is obviously inconstant. Four outliers, exception of volume, other attributes show that there are abnormal behaviors in each month of a year. As the same of the previous timescale, outliers have the initial pressure of on-board storage or final pressure of no-board storage with higher or lower value respectively.
4. Discussions and conclusion

4.1 Discussions
Based on boxplot method, the refilling behaviors of three different timescale, i.e. hourly, daily, and monthly, are studied in this article. The standard boxplot method calculates 5 statistics, which can describe the central tendency, dispersion and outliers of the data set. For central tendency, hourly and monthly behaviors are not steady. But daily behaviors within a week seems constant. And the central tendency of volume and initial pressure of the on-board storage changes in the opposite direction. This is understandable that more volume of CNG refilled have to be happen with a lower initial pressure. On the other hand, a smaller volume CNG usually happens with a higher initial pressure. Therefore, the two attributes’ central tendency changes as the way shown in figure 1 and 4. For the outliers, all attributes in the three timescale can detect abnormal behaviors which are presented in small circles from figure 1 to 4. These outliers are reasonable in that a normal refilling action have to happen with larger volume, lower initial pressure, higher final pressure and proper length of refilling time interval. Those actions without one of these values are considered outliers, and are also considered as inefficient refilling behaviors.

4.2 Conclusion
From figure 1 to 4, the boxplot method presents the refilling behaviors under different timescale. Median and mean, the most widely used statistic tools to derive the central tendency, show that the intra-week behaviors seem steady, and the intraday and monthly ones change with time. Therefore one possible way to improve efficiency of the CNG secondary filling station is to consider factors affecting these behaviors in intraday and monthly timescale. The boxplot method provides another possible efficient way which reducing the abnormal behaviors. These outliers happen in any timescale, which means that they are caused by time independent factors.

Acknowledgments
This work was financially supported by Science and Technology Program of Tianjin (14ZXCXGX00-392).

References
[1] http://www.sohu.com/a/211956574_712320
[2] http://news.163.com/14/0724/15/A1U9LMU100014AED.html
[3] http://paper.dzwww.com/dzrb/content/20130911/Articel19004MT.htm
[4] Ríos-Mercado R. Z., Borraz-Sánchez C., Optimization problems in natural gas transportation systems: a state-of-the-art review[J], Applied Energy, 2015, 147:536–555.
[5] Khadem J, Saadat-Targhi M, Farzaneh-Gord M, Mathematical modeling of fast filling process at CNG refueling stations considering connecting pipes[J], Journal of natural gas science and engineering, 2015, 26:176-184.
[6] Khamforoush M, Moosavi R, Hatami T, Compressed natural gas behavior in a natural gas vehicle fuel tank during fast filling process:Mathematical modeling, thermodynamic analysis, and optimization[J], Journal of natural gas science and engineering, 2014, 20: 121-131.
[7] Baratta M, Kheshtinejad H, Laurenzano D, Misul D, Modelling aspects of a CNG injection system to predict its behavior under steady state conditions and throughout driving cycle simulations[J], Journal of natural gas science and engineering, 2015, 24: 52-63.
[8] Tukey TW, Exploratory data analysis [M], Philippines: Addison-Wesley publishing company, 1977. p27-56.
[9] Chen XR, Probability and statistics[M], Beijing: Science press, 2000.
[10] https://matplotlib.org/examples/statistics/boxplot_demo.html