3rd Conference of Transportation Research Group of India (3rd CTRG)

Studying the effectiveness of using fuel sales as a proxy for estimating seasonal factors of traffic

Pranamesh Chakraborty[a,*], Partha Chakroborty[b]

[a]Former Graduate Student, Department of Civil Engineering, Indian Institute of Technology Kanpur, Kanpur - 208016, India, pranameshbesu@gmail.com
[b]Professor, Department of Civil Engineering, Indian Institute of Technology Kanpur, Kanpur - 208016, India,partha@iitk.ac.in
* Corresponding Author

Abstract. This study is aimed at evaluating the effectiveness of using fuel sales as a surrogate for measuring seasonal variations in traffic flow. Since there are few permanent traffic counters (PTCs) in India and no literature available on seasonal factors, often the practice is to directly use seasonal variations in fuel sales obtained from sales data of nearby pumping stations as a surrogate for measuring seasonal variations in traffic flow. Even though use of such surrogate measures seem logical in the absence of data, one must analyse how good or valid such a practice is. The objective here is to see whether fuel sales data can be used to predict seasonal factors of traffic with reasonable accuracy.

Statistical analysis is done on a large data set containing data on both traffic variations and fuel sales. The following conclusions can be drawn from the analysis: (i) in the absence of traffic data, fuel sales data, if used appropriately can act as a surrogate for seasonal factors of traffic and (ii) seasonal variations in fuel sales directly using as seasonal factors of traffic often result in large errors; however the errors can be reduced significantly if these seasonal variations are used in statistical relations to obtain the seasonal factors. One such relation is also developed here.

Keywords: Fuel Sales; Annual Average Daily Traffic; Seasonal Factors

Introduction

Permanent traffic counters (PTCs) collect traffic volume data throughout the year. However, it is not economically feasible to install PTC in each and every road section. So, the general practice is to install PTCs in selected sites and conduct short period traffic counts (SPTCs) on remaining segments for 7 days, 3 days, etc. The SPTC from a given site is then adjusted using seasonal factors obtained from PTC data to predict the annual average daily traffic (*AADT*) at that site.

Since there are few PTCs in India and no literature available on seasonal factors, often the practice is to directly use seasonal variations in fuel sales obtained from sales data of nearby pumping stations as a surrogate for measuring seasonal variations in traffic flow. Even though use of such surrogate measure seems like a logical decision in the absence of data, one must analyse how good or valid such a practice is. This study is aimed at evaluating the effectiveness of using fuel sales as a surrogate.

So, an attempt has been made to predict seasonal factors of traffic from seasonal factors of fuel sales in petrol pumps through statistical relations. Finally, comparison has been done between seasonal factors estimated from fuel sales data with the actual seasonal factors obtained from the traffic count data.

Literature Review

Seasonal factors are primarily estimated based on traffic-count based methods [1, 2, 3, 4]. Fuel sales or other surrogate measures are also used for estimation of traffic demand related parameters (e.g. vehicle miles travelled (*VMT*), *AADT*, *etc.*)

Xia et al. [5] used roadway characteristics (e.g. number of lanes, functional classification of roads, *etc.*) to directly estimate *AADT* for certain roads in urbanized area.

Fuel sales data has also been widely used for estimation of *VMT*. *VMT* estimation based on fuel sales was proposed by Erlbaum [6]. He uses information on retail fuel sales (gasoline and diesel) from the study area and estimates of fleet fuel efficiency. Given the difficulty in estimating fleet fuel efficiency, Kumapley and Fricker [7] suggest that this method be used only as a preliminary estimate for *VMT*.

The review of literature indicate that fuel sales have not studied as a surrogate for estimation of seasonal factors of traffic. On the other hand, discussions with practitioners in India indicate that often fuel sales are used to get an idea of the traffic demand on different roads. This study therefore tries to understand the efficacy of using fuel sales as a surrogate for seasonal factor estimation. Note that, seasonal factors are used to ultimately determine *AADT* from SPTC data.

Description of data used

This study requires year long traffic volume data and fuel sales data. Traffic volume data from toll plazas located in different parts of India have been used in this analysis. Since these toll plazas collect the data throughout the year, they are considered as PTCs. Fuel sales data are obtained from petrol pumps located along the project road (road on which toll plaza is present). The petrol pumps are selected such that there are no access/egress points to/from the project road in between the toll plaza and the selected petrol pumps. This is done to ensure that same traffic volume passes through each of the petrol pumps and the toll plaza.

Traffic and fuel sales data of seven years from four sites are used in this study. Since twelve seasonal factors can be extracted from one year data of each site, so in total, $12 \times 7 = 84$ data points are used for this analysis. Fuel sales data are obtained from nine petrol pump for these four sites. The details of the data are given in Table 1. The first column gives the location of the site *i.e.* the state in which the toll plaza is located. The second column gives the duration (in years) for which traffic volume data and fuel sales data are available. The third column gives number of petrol pumps from which fuel sales data are obtained for the respective site. Each data set has been individually studied to

Table 1: Details of data used in this study.

Location (State)	Duration (Years)	Number of petrol pumps
Maharashtra	3	5
Madhya Pradesh	2	1
Gujarat	1	2
Tamil Nadu	1	1

identify and remove possible outliers or abnormalities either due to data recording errors or due to occurrence of out-of-the ordinary events.

Methodology and Results

This section describes the procedure for estimation of seasonal factors of traffic from seasonal factors of fuel sales. Before presenting the analysis, some preliminary discussion on various terms used here are presented in this section.

Let, the monthly traffic volume at the *j*-th PTC for month *m* of year *k* be denoted by $MTV_j^{m,k}$. Then monthly average daily traffic ($MADT_j^{m,k}$) and annual average daily traffic ($AADT_j^k$) can be defined using Equations 1 and 2, respectively.

$$MADT_j^{m,k} = \frac{MTV_j^{m,k}}{N^{m,k}} \quad (1)$$

$$AADT_j^k = \frac{1}{12}\sum_{m=1}^{12} MADT_j^{m,k} \quad (2)$$

where, $N^{m,k}$ is the number of days in month *m* of year *k*.

Seasonal factor of traffic ($SFT_j^{m,k}$), as shown in Equation 3, is defined as the ratio of $MADT_j^{m,k}$ to $AADT_j^k$. It depicts the seasonal variation of traffic over the year.

$$SFT_j^{m,k} = \frac{MADT_j^{m,k}}{AADT_j^k} \quad (3)$$

Similarly, let the monthly fuel sales volume for month *m*, year *k* at petrol pump *i* of site *j* be denoted as $MFS_{i,j}^{m,k}$. Then, total monthly fuel sales $TMFS_j^{m,k}$, monthly average daily fuel sales $MADFS_j^{m,k}$, annual average daily fuel sales $AADFS_j^k$ and seasonal factor of fuel sales $SFF_j^{m,k}$ are calculated as shown in Equations 4, 5, 6 and 7, respectively.

$$TMFS_j^{m,k} = \sum_{i=1}^{P_j} MFS_{i,j}^{m,k} \quad (4)$$

$$MADFS_j^{m,k} = \frac{TMFS_j^{m,k}}{N^{m,k}} \quad (5)$$

$$AADFS_j^k = \frac{1}{12}\sum_{m=1}^{12} MADFS_j^{m,k} \quad (6)$$

$$SFF_j^{m,k} = \frac{MADFS_j^{m,k}}{AADFS_j^k} \quad (7)$$

where, P_j denotes number of petrol pumps used for site *j* (given in Table 1) and $N^{m,k}$ is the number of days in month *m* of year *k*.

In this context, it may be pointed out that Hallenbeck and Kim [8] had noted that, seasonal factors for truck traffic and automobile traffic are not necessarily the same. Hence, it makes sense to calculate their seasonal factors separately. In this study, however, instead of considering truck traffic and car traffic separately, seasonal factors for truck traffic and total traffic are obtained separately. The reason for doing this is that predictions on truck traffic data are often required for pavement design and maintenance studies, while information of total traffic is required for traffic engineering and planning purpose.

So, separate analysis has been done to estimate seasonal factors of total traffic and truck traffic. It is assumed that variation in traffic flow in a given area will be reflected in the total fuel sales at all the pumping stations in that area. Since all trucks in India use diesel as fuel, seasonal factors for truck traffic are estimated from diesel sales data, while total fuel sales (sum of petrol and diesel sales) are used for estimation of seasonal factors for total traffic.

Cluster analysis and multiple regression analysis are used to classify the sites into different groups. However, for the current data set, it is observed that all the four sites, used in this study, can be broadly classified as a single group, which can be termed as commuter routes. Thus, data from all the four sites are combined together for further analysis.

Linear regression analysis is used to develop a relation between seasonal factors obtained from traffic count ($SFT_j^{m,k}$) and from fuel sales ($SFF_j^{m,k}$) data. The regression model proposed here is:

$$SFT_j^{m,k} = a_0 + a_1 \times SFF_j^{m,k} + \epsilon \quad (8)$$

where, a_0, a_1 are constants and ε is a normally distributed, homoscedastic error term. Alternatively, this model can be written as:

$$\overline{SFT}_j^{m,k} = a_0 + a_1 SFF_j^{m,k} \quad (9)$$

Note that, $\overline{SFT}_j^{m,k}$ is the estimated value of $SFT_j^{m,k}$.

Figures 1 and 2 show the relationship between $SFT_j^{m,k}$ and $SFF_j^{m,k}$ for total and truck traffic, respectively. The best fit straight lines and their equations are also shown in the figures along with the *t*-values (in parenthesis) and R^2 value. As indicated by the *t*-values, all the coefficients are statistically significant at 95% confidence level. The plot of data points in the figures indicate that higher $SFF_j^{m,k}$ implies, in general, a higher $SFT_j^{m,k}$. This further implies that seasonal factors in total fuel sales (diesel fuel sales) can be used, in the absence of traffic data, to get an idea about the seasonal factors in total traffic (truck traffic).

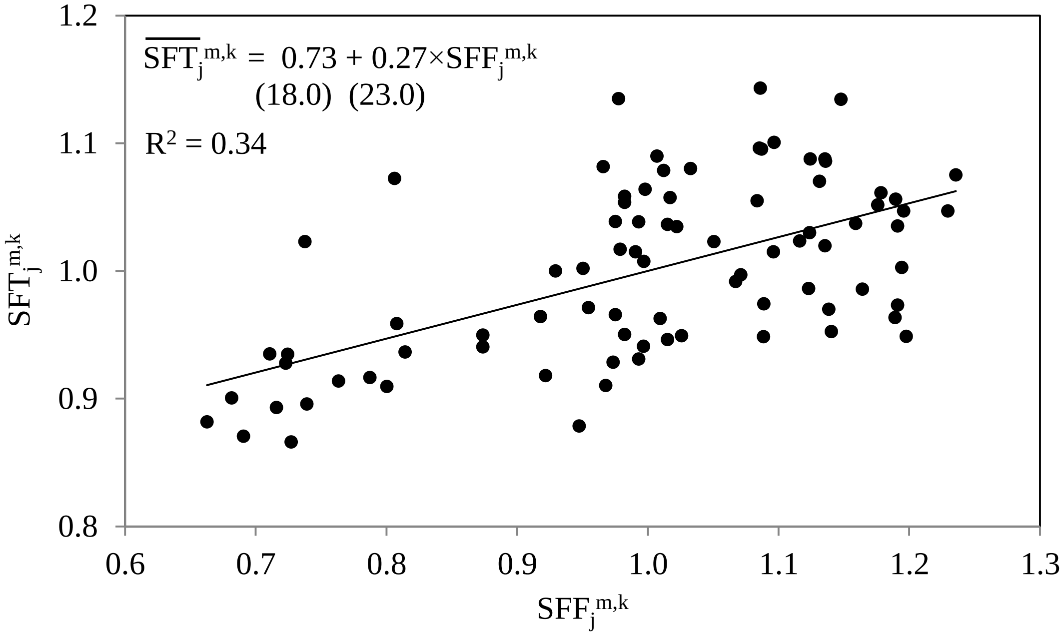

Figure 1: Estimation of seasonal factors of total traffic from that of total fuel sales.

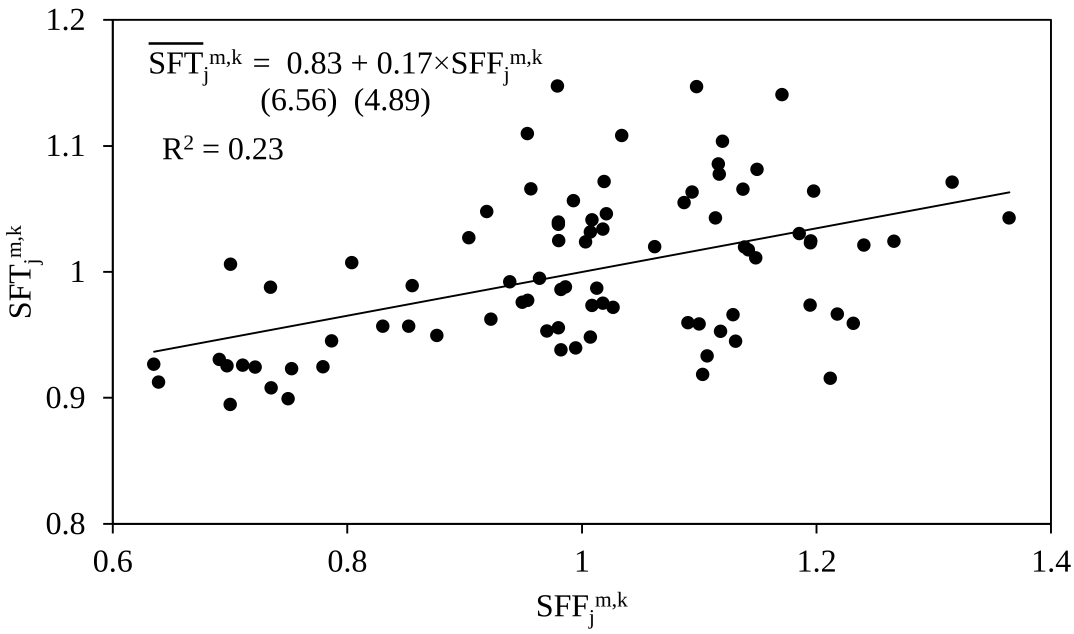

Figure 2: Estimation of seasonal factors of truck traffic from that of diesel sales.

Errors in estimation of seasonal factors from fuel sales

In this section, the errors in estimation of seasonal factors of traffic from fuel sales is evaluated in terms of root mean square error for site *j* for year *k*, given by Equation 10.

$$RMSE_1^{j,k} = \sqrt{\frac{1}{12}\sum_{m=1}^{12}(\overline{SFT}_j^{m,k} - SFT_j^{m,k})^2} \tag{10}$$

In order to see whether using Equation 9 provides significant benefits in terms of predicting seasonal factors of traffic, the errors when raw fuel sales are directly used as estimates of $SFT_j^{m,k}$ are also calculated as $RMSE_2^{j,k}$.

$$RMSE_2^{j,k} = \sqrt{\frac{1}{12}\sum_{m=1}^{12}(SFF_j^{m,k} - SFT_j^{m,k})^2} \tag{11}$$

The reason for calculating $RMSE_2^{j,k}$ is that the current practice is to use $SFF_j^{m,k}$ as the estimates of $SFT_j^{m,k}$. The $RMSE_1^{j,k}$ is compared with $RMSE_1^{j,k}$ to determine the extent of benefit, if any, that is obtained by using Equation 9.

Table 2 shows the root mean square error for every site in different years when $\overline{SFT}_j^{m,k}$ is obtained using Equation 9 ($RMSE_1^{j,k}$) and when the $SFF_j^{m,k}$ are directly used as seasonal factors of traffic ($RMSE_2^{j,k}$). The results show that,

1. $RMSE_1^{j,k}$ errors are of the order of 3% to 7% of $SFT_j^{m,k}$.
2. $RMSE_2^{j,k}$ errors are of the order of 6% to 20% of $SFT_j^{m,k}$.

This indicates that (i) using $\overline{SFT}_j^{m,k}$ instead of $SFF_j^{m,k}$ directly provides substantial gains in terms of accuracy and (ii) $SFT_j^{m,k}$ is a reasonably good estimator of the seasonal factors for both total and truck traffic.

Table 2: $RMSE_1^{j,k}$ and $RMSE_2^{j,k}$ values for total and truck traffic.

Site in	Year	Total Traffic		Truck Traffic	
		$RMSE_1^{j,k}$	$RMSE_2^{j,k}$	$RMSE_1^{j,k}$	$RMSE_2^{j,k}$
Tamil Nadu	2011-12	0.05	0.05	0.04	0.03
Madhya Pradesh	2010-11	0.07	0.13	0.07	0.15
	2011-12	0.06	0.14	0.08	0.17
Gujarat	2012	0.07	0.14	0.04	0.15
Maharashtra	2010-11	0.05	0.11	0.05	0.15
	2011-12	0.04	0.15	0.04	0.19
	2012-13	0.04	0.12	0.03	0.14

Conclusions

This study is aimed at evaluating the effectiveness of using seasonal variations in fuel sales as a surrogate for measuring seasonal variations in traffic flow. Though in India, variation in fuel sales is widely and directly used as a surrogate measure to estimate seasonal factors of traffic, this study is believed to be the first to find out its potency. The results show that, seasonal factors of traffic for the project road cannot be taken to be equal to the seasonal factors obtained from nearby fuel sales data. Rather, it is better to predict seasonal factors of traffic from that of the fuel sales, using regression constants determined in this study. The results show that, on an average, errors from such prediction are in the range of 3% to 7% and are quite acceptable. However, this study is based on data from only four sites. Although the sites are geographically distributed, more sites are required to make more robust estimates of the regression parameters.

Acknowledgement

The data used in this study was provided gratis by V. R. Techniche Consultants Pvt. Ltd. The authors gratefully acknowledge this support.

References

1. Ritchie, S. G. (1986). A statistical approach to statewide traffic counting. Transportation Research Record, Vol. 1090, 1986, pp. 1421.
2. Faghri, A. and Chakroborty, P. Development and evaluation of a statistically reliable traffic counting program. Transportation Planning and Technology, Vol. 18 (3), 1994, pp. 223237.
3. Lingras, P. and Adamo, M. Average and peak traffic volumes: neural nets, regression, factor approaches. Journal of Computing in Civil Engineering, Vol. 10(4), 1996, pp. 300306.

4. Yang, S., Lu, C., Zhao, F., Reel, R., and OHara, J. D. Estimation for seasonal factors of similarity-based traffic for urban roads in Florida. Transportation Research Record: Journal of the Transportation Research Board, Vol. 2121(1), 2009, pp. 7480.
5. Xia, Q., Zhao, F., Chen, Z., Shen, L. D., and Ospina, D. Estimation of annual average daily traffic for nonstate roads in a Florida county. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1660(1), 1999, pp. 3240.
6. Erlbaum, N. Estimated county level vehicle miles of travel. Report No. 603690, Planning Division, New York State Department of Transportation, Albany, 1989.
7. Kumapley, R. K. and Fricker, J. D. Review of methods for estimating vehicle miles traveled. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1551(1), 1996, pp. 5966.
8. Hallenbeck, M. E. and Kim, S. G. Final technical report for task A : Truck loads and flows. Report No. 9233, Washington State Department of Transportation, 1993.