

Macroscopic Congestion Intensity Measurement Model Based on Cumulative Logistic Regression

Lei Yu^{*1}, Menghan Liu², Qinyi Shi³ and Guohua Song⁴

¹Department of Transportation Studies, College of Science and Technology, Texas Southern University, USA; Beijing Jiaotong University, Beijing, China

²Highway Transport Division, Transport Planning and Research Institute, Ministry of Communications, Building No. 2, A6, Shuguang Xili, Chaoyang District, Beijing, 100028, P.R. China

³Graduate Research Assistant, Department of Transportation Studies, Texas Southern University, 3100 Cleburne Avenue, Houston, Texas, 77004, USA

⁴MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, #3 Shangyuancun, Haidian District, Beijing, 100044, P. R. China

Abstract: An efficient and accurate measurement of congestion intensity helps investigate the traffic conditions at different road classes and provides useful information for transportation planning and traffic operation improvement. However, the existing traffic congestion intensity measurement models often suffer from two major problems: one is that there is no generally accepted method that can be used to classify the grades of congestion intensity, and the other is that the variables that have been incorporated into the existing congestion intensity measurement models are often inter-correlated. In order to overcome these two deficiencies, this paper analyzes the ordinal characteristics of congestion intensity, and introduces the cumulative logistic regression into the congestion intensity measurement model.

In the model development process, first it adopts the likelihood ratio test to validate the adaptability of the cumulative logistic regression and Wald test to select the independent variables. Then, it develops the measurement model of congestion intensity by using travel speed as the independent variable. The proposed model shows a determination coefficient (pseudo R^2) higher than 0.77 in the goodness-of-fit test, and a prediction with the accuracy of 73.39% against the field observed data. Therefore, the proposed model can be effectively used to determine the traffic congestion intensity on different road classes.

Keywords: Traffic congestion intensity, cumulative logistic regression, likelihood ratio test, wald test.

INTRODUCTION

There is no single identifiable element that is able to describe all of travelers' concerns about traffic congestion. Pisarski summarized four general components of congestion at the Workshop on National Urban Congestion Monitoring in May 1990 as a way to formulate an overall congestion index, including intensity, extent, duration, and variation [1]. Intensity, extent, and duration form a three dimensional framework to describe the magnitude of congestion, and variation is used to measure the reliability of the traffic network.

Congestion intensity is a measure from an individual traveler's perspective, such as "congested" or "no congestion." It reflects the severity of congestion that affects travel and measures the difference between the desired condition and the conditions being analyzed [1]. In the past years, extensive studies have been conducted in the field of congestion intensity evaluation. As we know, traffic

congestion degrades the levels of mobility by generating increased travel time and delay [2], therefore most of the congestion intensity measurement methods are based on travel time and travel speed. The parameters that have been found fundamentally useful in congestion intensity measurement include travel speed, travel rate, delay rate, delay ratio, and relative delay rate [3][4]. Based on these five congestion-descriptive variables, Vaziri developed Highway Congestion Intensity Index (CII) with fuzzy set theory, which provided an alternative approach to the past normative index development methods and can be used to capture the perception of congestion intensity in future index development [5]. In recognition of the uncertainties associated with the derivation of measures of congestion, Hamad and Kikuchi developed a measure of traffic congestion based on fuzzy-rule inference. The proposed inference process allows the mechanism to combine different measures and also to incorporate the uncertainty in the individual measures so that the composite picture of congestion can be reproduced [6].

However, two critical problems have been identified in the current congestion intensity measurement methods: one is that the determination of the threshold of congestion intensity indicators is usually based on the personal

*Address correspondence to this author at the Department of Transportation Studies, College of Science and Technology, Texas Southern University, 3100 Cleburne Avenue, Houston, Texas 77004, USA; Tel: 713-313-7007; Fax: 713-313-1853; E-mail: YuVLX@TSU.EDU

experience of transportation planners and traffic engineers; there is no generally accepted quantification method that can be used to define levels of congestion. The other is that the variables that have been incorporated in the existing congestion intensity measurement models are often inter-correlated, which somehow compromises the model's accuracy. In order to address the incurring problems, this research proposes to use cumulative logistic regression to analyze the ordinal characteristics of congestion intensity, in which integers from "1" to "5" are used to grade on-road drivers' own perception about the intensity of congestion with "1" indicating "severest congestion" and "5" indicating "no congestion;" and likelihood ratio test (LRT) is applied to validate the adaptability of the cumulative logistic regression. Then Wald test is used to select independent variables. This paper develops a congestion intensity measurement model by using travel speed as the only independent variable.

An accurate measurement of congestion intensity helps investigate the traffic conditions at different road classes and road segments. The results can not only be used in other aspects of congestion quantification, but also provide useful information for transportation planning and traffic operation improvement. The rest of this paper will discuss the development of this model in the sequence of data source and preparation, model variables, algorithm design, algorithm implementation, model validation and application, and conclusions.

MATERIALS AND METHODOLOGY

Data Source and Preparation

As mentioned above, congestion intensity is a measure from an individual traveler's perception, therefore the source of data utilized in this research are collected by travelers driving on the real-world traffic network. The variance between the operating conditions on different road classes often results in people's different sense of congestion, so the testing route is designed to cover different road classes, including expressways, major arterials, and minor arterials (including collectors).

In this research, the data collection was conducted between 6:00 to 19:30 in Beijing, China. Fig. (1) shows the testing routes in GIS map. The testing vehicle was driven on the pre-designed route repeatedly. During the experiment, on-board Global Positioning System (GPS) device is used to collect vehicle's second-by-second speed data. The research inspectors sitting in the vehicle manually record their

position and the number of roadway entrances/exits (signalized intersections), and grade the congestion intensity levels on different road classes during different time periods based on their own perception using integers from "1" to "5", in which 1= severest congestion, 2= heavy congestion, 3= moderate congestion, 4= slight congestion, and 5= no congestion. Drivers were solicited from different people groups.

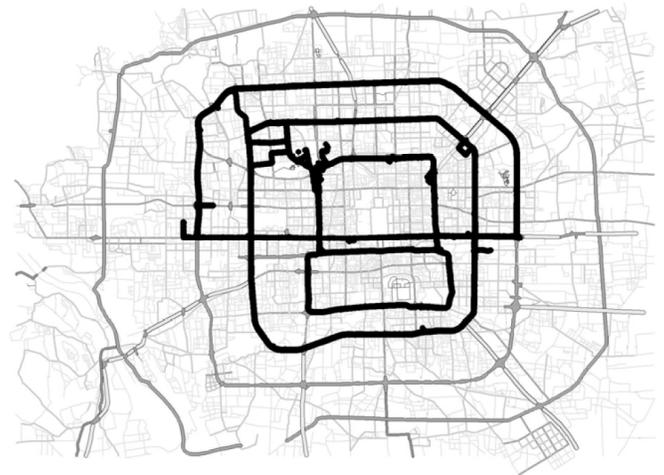


Fig. (1). Testing route for manual determination of congestion intensity.

More than 4,000 data records about congestion intensity are collected during the two-day's on-road test. In order to have an effective evaluation of the model's accuracy, the collected data are divided into two parts: one is used for model development, and the other is used for model validation. The modeling data sample size for the expressways, major arterials, and minor arterials and collectors are 1575, 390, and 305 respectively, which compose a sufficient sample size for conducting cumulative logistic regression analysis. Table 1 shows the modeling data sample size for the congestion intensity measurement.

Model Variables

In this research, individual traveler's personal perception on congestion intensity is considered as the dependent variable, which equals the integer from "1" to "5" with each of the number representing one level of congestion intensity.

After a quality control and mathematical analysis of the collected GPS data, 5 independent variables are determined to reflect real-world traffic operating conditions: (1) travel speed (km hr⁻¹): which refers to the average speed within 1 minute; (2)

Table 1. Modeling Data Sample Size for Congestion Intensity Measurement

Road Class	Sample Size	Severe Congestion	Heavy Congestion	Moderate Congestion	Slight Congestion	No Congestion	Total
Express-Ways	The Number of Samples	377	195	199	501	266	1575
	Proportion	24.51%	12.68%	12.94%	32.57%	17.30%	100%
Major Arterials	The Number of Samples	91	55	91	84	48	390
	Proportion	24.7%	14.9%	24.7%	22.8%	13.0%	100%
Minor Arterials & Collectors	The Number of Samples	68	75	48	61	53	305
	Proportion	22.30%	24.59%	15.74%	20.00%	17.38%	100%

delay ratio: which refers to the ratio of the difference between the travel rate per minute and the travel rate in the real-world traffic flow to the travel rate within that minute; (3) stop time (s km⁻¹): which refers to the ratio of the number of seconds with 0 speed within 1 minute to the travel distance within that time period; (4) the number of stops (the number of times km⁻¹): which refers to the ratio of the number of times the speed changing from greater than 0 to 0 within 1 minute to the travel distance within that time period; and (5) the density of the expressways' entrances/exits or the major and minor arterials' signalized intersections (number km⁻¹), which refers to the ratio of the number of entrances/exits or signals within 1 minute to the travel distance within that time period. During the model development process, the relationship between these 5 independent variables and congestion intensity are evaluated by Wald test, and only the variables that show significant contribution to the modeling results will be selected.

Algorithm Design

Cumulative logistic regression is applied in developing the congestion intensity measurement model. Since the real-world congestion intensity is divided into 5 categories ($j=1, 2, \dots, 5$), when $y=1, y=2, \dots, y=5$, the Logistic model contains 4 logit functions as expressed by Equation (1):

$$\begin{aligned} \ln\left(\frac{p_1}{1-p_1}\right) &= \beta_{01} - \sum_{k=1}^5 \beta_k x_k, \\ \ln\left(\frac{p_1+p_2}{1-p_1-p_2}\right) &= \beta_{02} - \sum_{k=1}^5 \beta_k x_k, \\ \ln\left(\frac{p_1+p_2+p_3}{1-p_1-p_2-p_3}\right) &= \beta_{03} - \sum_{k=1}^5 \beta_k x_k, \\ \ln\left(\frac{p_1+p_2+p_3+p_4}{p_5}\right) &= \beta_{04} - \sum_{k=1}^5 \beta_k x_k. \end{aligned} \tag{1}$$

where

p_1, p_2, p_3, p_4, p_5 = The probability for severe congestion, heavy congestion, moderate congestion, slight congestion, and no congestion, p_5 is used as the benchmark, $p_1+p_2+p_3+p_4+p_5=1$;

x_1 = Travel speed;

x_2 = Delay ratio;

x_3 = The number of stops within a certain time period;

x_4 = Stop time within a certain time period;

x_5 = The density of entrances/exits or signalized intersections;

β_{0k} = Regression intercept; and

β_k = Regression coefficient

The solution of the proposed congestion intensity measurement model can be determined by the method of maximum likelihood. However, in order to use this method, the model has to be further investigated from two aspects [7]: (1) evaluate the model's proportional characteristics, which uses statistical methods to analyze the hypothesis of the proportional characteristics of the cumulative logistic regression model for the purpose of validating the model's applicability; and (2) evaluate the significance of the

independent variables to the modeling results, which examines the relationship between the 5 pre-defined independent variables and congestion intensity and determines the variables that will be included in the model development. The following section presents the results of the evaluation work.

RESULTS

Algorithm Evaluation Results

Evaluation of the Model's Proportional Characteristics

G-statistics (Likelihood Ratio Test) is used to evaluate the model's proportional characteristics, as expressed by Equation (2), which reflects the errors occurred after all the independent variables are incorporated into the model and indicates the part of the change that cannot be explained by the independent variables [7].

$$G = -2(\ln L_p - \ln L_k) \tag{2}$$

where

L_p = Log likelihood of Model p ,

L_k = Log likelihood of Model k , and

Model p and model k refer to the models included in Equation (1) in turn.

The likelihood ratio test (LRT) computes x^2 to determine if the model meets the proportional requirement for conducting a cumulative regression analysis. If x^2 shows a low level of significance, that is to say, if the calculated p-value is less than 0.05, then cumulative logistic model can be used in this analysis.

The five independent variables incorporated in this model are all continuous variables, which usually cause little data loss; therefore the original data are input into the model directly. Table 2 shows the evaluation results of the proposed model's proportional characteristics for different road classes. It indicates that for all the three examined road classes, when the degree of freedom is 12, all the resulted p-values are less than 0.05. Therefore the proposed cumulative logistic regression model is suitable to conduct the congestion intensity measurement.

Table 2. Likelihood Ratio Test on the Model with Five Independent Variables

Road Class	χ^2	Degree of Freedom	P-Value
Expressways	1188.892	12	0.000
Major Arterials	92.936	12	0.000
Minor Arterials and Collectors	83.867	12	0.000

Evaluation of the Significance of the Pre-defined Independent Variables

Wald test is used to determine if an independent variable should be included in a Logistic regression model [7, 8]. Under the Wald test, the bigger the Wald value is, or the smaller the p-value is, the more important the independent

variable means to the regression model; otherwise, the independent variable should be eliminated. The statistics of the Wald test is often denoted by u , which can be calculated using Equation (3), in which β_i stands for the sample value of the independent variable x_i 's regression coefficient, based on the design of the model algorithm, there are 5 independent variables, therefore i refers to the integers from 1 to 5; S_{β_i} stands for the standard deviation of x_i 's regression coefficient.

$$u = \frac{\beta_i}{S_{\beta_i}} \tag{3}$$

Table 3 shows the results of the Wald test on the 5 independent variables when the degree of freedom is set at a value of 1, in which each independent variable's Wald value and its corresponding p-value are listed. Based on the results, the evaluation on the contribution of each independent variable to each road class's congestion intensity can be conducted, from which we can determine the variables that should be included in the model development.

As to the expressways, Table 3 shows that the p-value for travel speed and delay ratio are both less than 0.05, which means that these two independent variables have significant contribution to the modeling results; however the p-value for stop time, the number of stops, and the density of entrances/exits or signals are all greater than 0.05, which means that these three independent variables do not have significant importance to the modeling results and should be excluded from the model. As to the major arterials, minor arterials and collectors, the p-value for delay ratio, stop time, the number of stops, and the density of entrances/exits or signals are all greater than 0.05, therefore they are not statistically important to the modeling results, which leads to the conclusion that in this analysis only travel speed should be included in the congestion intensity measurement model.

Based on the above illustration, travel speed and delay ratio are recommended to be included in the congestion intensity measurement for expressways, and travel speed becomes the only variable that should be included in the congestion intensity measurement for major arterials, minor arterials and collectors. In this research, we further investigate the necessity of including the two variables in the congestion intensity measurement for expressways. Input travel speed and delay ratio into the proposed regression

model for a second time to check the model's proportional characteristics, the LRT shows χ^2 equals 1048.6 at the 6 degree of freedom, and the resulted p-values are both less than 0.05, therefore, the logistic regression model is valid to be used in this analysis. Under Wald test at 95% confidence level, the Wald values for travel speed and delay ratio are 178.517 and 21.174 respectively, both of which are greater than 3.841 (The 0.05 level of significance for the Wald test is 3.841). The p-values for both of these two variables are 0.00, which indicates that travel speed and delay ratio both have important influence on the modeling results. However, as we can see, the Wald value of delay ratio (21.174) is far smaller than that of travel speed (178.517), which demonstrates that delay ratio has far less contribution to the expressway congestion intensity measurement than travel speed does. In addition, the standard deviation of the regression coefficient for delay ratio is 18.768, which is much bigger than that of travel speed and will result in a broader confidence interval under 95% confidence level. Due to the delay ratio's comparatively smaller contribution to the modeling results and bigger standard deviation of its regression coefficient, this variable is ultimately eliminated from the variables that are to be included in the proposed logistic regression model for expressways.

In summary, travel speed is determined to be the independent variable that will be incorporated into the proposed cumulative logistic regression based congestion intensity measurement model for all road classes. Table 4 shows the LRT results on the model based on travel speed.

The above table illustrates that at the degree of freedom 3, χ^2 equals 1478.806 under LRT for expressways, and the p-value is less than 0.05. Therefore, the travel-speed based model meets the proportional requirement of cumulative logistic regression model and can be used safely in this analysis. Moreover, comparing with the χ^2 value under LRT in the travel-speed and delay-ratio based model, which equals 1048.6 at degree of freedom 6, the χ^2 generated in this analysis has a bigger value, which implies that the travel speed based model has a better goodness-of-fit. The above results also show that the travel speed based cumulative logistic regression model is valid to conduct congestion intensity measurement for major arterials, minor arterials, and collectors as well.

Table 3. Wald Test on the Five Independent Variables

Road Class	Statistics	Travel Speed	Delay Ratio	Stop Time	The Number of Stops	The Density of Entrances/ Exits or Signals
Expressways	Wald	182.45	16.56	0.17	1.32	0.97
	p -Value	0.00	0.00	0.68	0.25	0.35
Major Arterials	Wald	130.40	0.12	0.04	0.03	0.06
	p -Value	0.00	0.73	0.85	0.86	0.84
Minor Arterials & Collectors	Wald	120.12	0.10	0.03	0.02	0.04
	p -Value	0.00	0.53	0.55	0.63	0.64

Table 4. Likelihood Ratio Test on the Model Based on Travel Speed

Road Class	χ^2	Degree of Freedom	P-Value
Expressways	1478.806	3	0.000
Major Arterials	18.057	3	0.000
Minor Arterials and Collectors	23.847	3	0.000

Algorithm Implementation Results

The method of maximum likelihood is utilized to calculate the results of the travel-speed based cumulative logistic regression model for traffic congestion intensity on different road classes. Equations (4), (5), and (6) show the calculation algorithm, in which speed refers to the average travel speed within 1 minute with the unit of km/hr, p_1 , p_2 , p_3 , p_4 , and p_5 stand for severe congestion, heavy congestion, moderate congestion, slight congestion, and no congestion respectively.

For expressways:

$$\begin{aligned} \ln\left(\frac{p_1}{1-p_1}\right) &= 6.803 - 0.348 \times Speed, \\ \ln\left(\frac{p_1 + p_2}{1-p_1-p_2}\right) &= 10.234 - 0.348 \times Speed, \\ \ln\left(\frac{p_1 + p_2 + p_3}{1-p_1-p_2-p_3}\right) &= 13.552 - 0.348 \times Speed, \\ \ln\left(\frac{p_1 + p_2 + p_3 + p_4}{p_5}\right) &= 18.858 - 0.348 \times Speed. \end{aligned} \tag{4}$$

For major Arterials:

$$\begin{aligned} \ln\left(\frac{p_1}{1-p_1}\right) &= 2.340 - 0.235 \times Speed, \\ \ln\left(\frac{p_1 + p_2}{1-p_1-p_2}\right) &= 4.263 - 0.235 \times Speed, \\ \ln\left(\frac{p_1 + p_2 + p_3}{1-p_1-p_2-p_3}\right) &= 7.381 - 0.235 \times Speed, \\ \ln\left(\frac{p_1 + p_2 + p_3 + p_4}{p_5}\right) &= 10.664 - 0.235 \times Speed. \end{aligned} \tag{5}$$

For Minor arterials and collectors:

$$\begin{aligned} \ln\left(\frac{p_1}{1-p_1}\right) &= 2.354 - 0.308 \times Speed, \\ \ln\left(\frac{p_1 + p_2}{1-p_1-p_2}\right) &= 5.225 - 0.308 \times Speed, \\ \ln\left(\frac{p_1 + p_2 + p_3}{1-p_1-p_2-p_3}\right) &= 7.408 - 0.308 \times Speed, \\ \ln\left(\frac{p_1 + p_2 + p_3 + p_4}{p_5}\right) &= 10.715 - 0.308 \times Speed. \end{aligned} \tag{6}$$

Fig. (2) demonstrates the estimated values generated by the travel-speed based cumulative logistic regression model on different road classes. It shows the probability of different levels of congestion that may occur under different travel speeds.

Model Validation Results

In order to evaluate the accuracy and reliability of the developed model, two methods are used in model validation: (1) evaluate the model’s goodness-of-fit by Cox and Snell’s pseudo R^2 and Nagelkerke’s pseudo R^2 ; and (2) evaluate the model’s accuracy by comparing the modeling results with the real-world observation results.

Validate the Model by Using Statistical Tests

Pearson’s chi-square test cannot be applied when the dependent variables are categorical variables, therefore in this analysis, Cox and Snell’s pseudo R^2 and Nagelkerke’s pseudo R^2 are used to evaluate the model’s goodness-of-fit [7]. When the independent variable and the dependent variable only have a weak correlation, the pseudo R^2 value tends to be close to 0; if the pseudo R^2 value is close to 1, it means that the measurement model has a high level of goodness-of-fit.

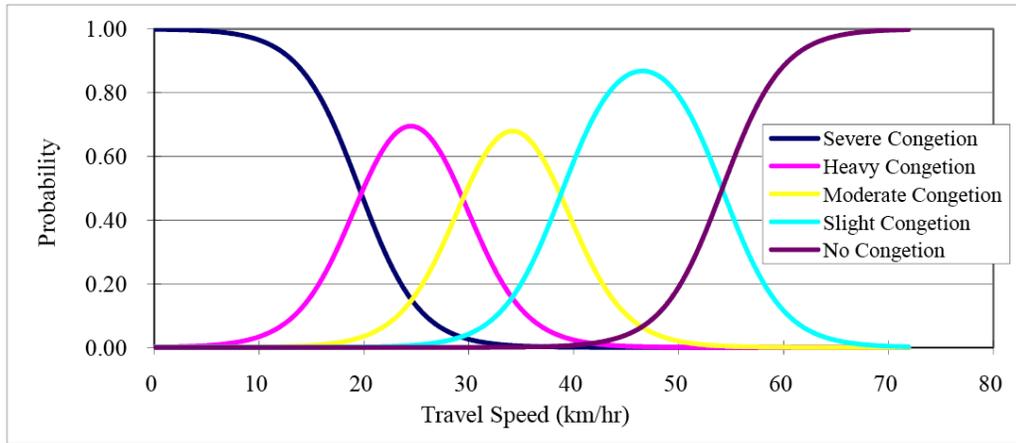
Table 5 shows the results of the goodness-of-fit test on the proposed congestion intensity measurement model for different road classes. It demonstrates that Cox and Snell’s pseudo R^2 and Nagelkerke’s pseudo R^2 for expressways are 0.875 and 0.898 respectively, and the corresponding pseudo R^2 for major arterials, and minor arterials and collectors are all greater than 0.77. The resulted pseudo R^2 values indicate that the proposed congestion intensity measurement model has a high level of goodness-of-fit. The expressways usually have a more stable traffic flow, therefore the congestion intensity measurement results for this road class tend to be more consistent, which results in a higher pseudo R^2 value.

Table 5. Goodness-of-Fit Test on the Congestion Intensity Measurement Model

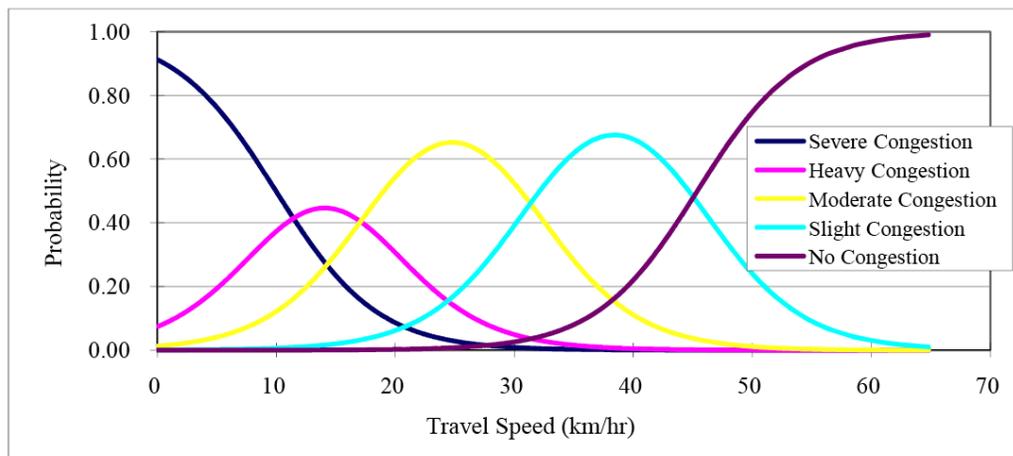
Road Class	Cox and Snell’s Pseudo R^2	Nagelkerke’s Pseudo R^2
Expressways	0.875	0.898
Major Arterials	0.770	0.805
Minor Arterials and Collectors	0.780	0.813

Validate the Model by Comparing the Model Estimation with Field Observed Data

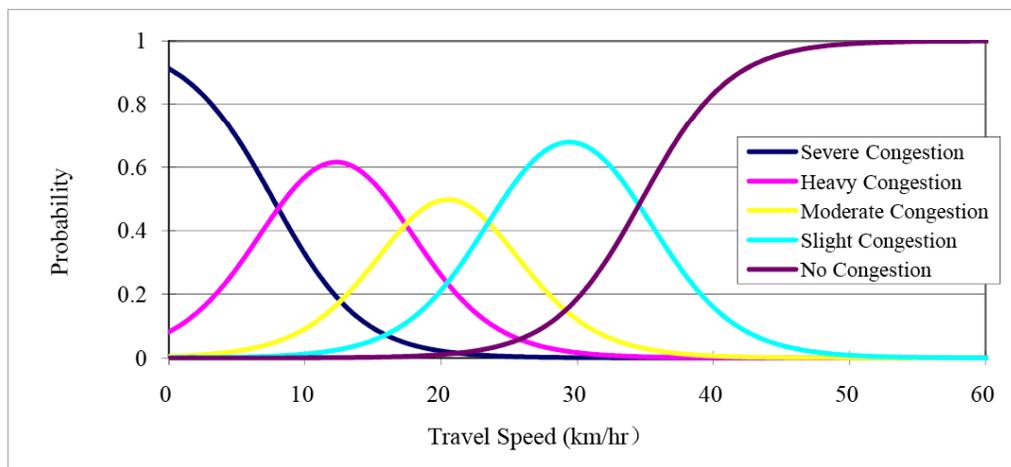
As has been explained in the section of data source and preparation, the data collected from the real-world traffic network are divided into two parts, one for model development and the other for model validation. In this step, the part of the field observed data that are used for model development are input into the congestion intensity measurement model. Table 6 shows the comparison between the modeling results and the rest of the field observed data. It demonstrates that the average accuracy rate for expressways, major arterials, and minor arterials and collectors are 78.41%, 64.36%, and 59.02% respectively. In general, the



(a) Expressways



(b) Major Arterials



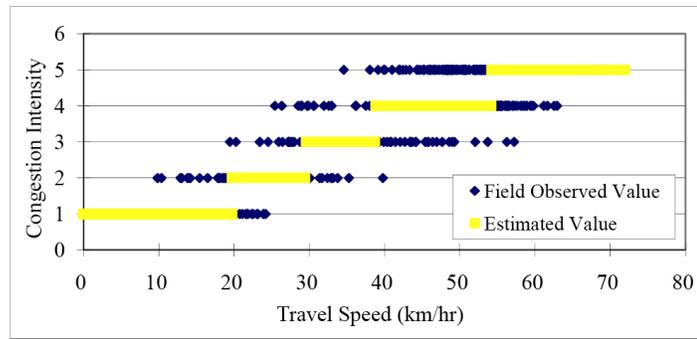
(c) Minor Arterials and Collectors

Fig. (2). Estimated values generated by travel-speed based cumulative logistic regression model on different road classes.

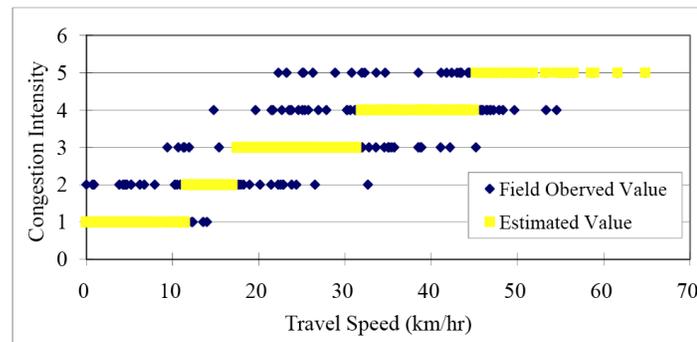
accuracy rate of the developed congestion intensity measurement model reaches 73.39%.

Fig. (3) shows the comparison of the estimated traffic congestion intensity versus the observed values on different road classes. As shown in Fig. (3a), the modeling results on expressways display a high level of accuracy, in which the

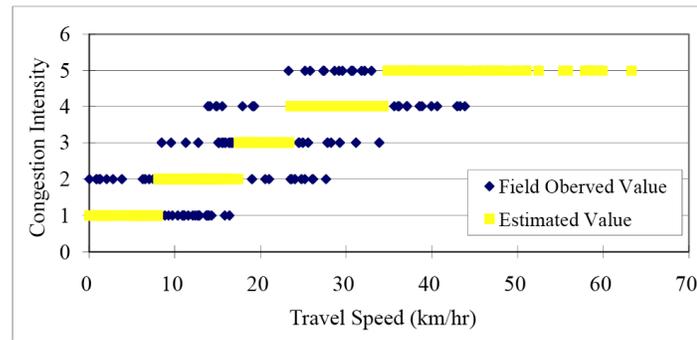
estimation for “severe congestion” reaches the highest accuracy rate, and the estimation for “no congestion” has a comparatively lower accuracy rate. This can be explained by the fact that usually vehicles have to drive constantly at a very low speed under severe congestion, and therefore the drivers tend to have similar perception on the congestion intensity. Fig. (3b, c) show the comparison results between



(a) Expressways



(b) Major Arterials



(c) Minor Arterials and Collectors

Fig. (3). Comparison of the estimated traffic congestion intensity versus the field observed values on different road classes.

Table 6. Comparison Between the Congestion Intensity Measurement Modeling Results and the Field Observed Data

Road Class	Sample Size (the Number of Times)	Severe Congestion	Heavy Congestion	Moderate Congestion	Slight Congestion	No Congestion	Total
Expressways	Field Observed Data	413	210	179	500	273	1575
	Modeling Results	387	162	115	407	164	1235
	Accuracy Rate	93.70%	77.14%	64.25%	81.40%	60.07%	78.41%
Major Arterials	Field Observed Data	98	65	85	88	54	390
	Modeling Results	87	28	59	50	27	251
	Accuracy Rate	88.78%	43.08%	69.41%	56.82%	50.00%	64.36%
Minor Arterials & Collectors	Field Observed Data	74	66	55	56	54	305
	Modeling Results	47	38	26	33	36	180
	Accuracy Rate	63.51%	57.58%	47.27%	58.93%	66.67%	59.02%

the estimated congestion intensity and the field observed data on major arterials, minor arterials and collectors, from which a decreased level of accuracy for the modeling results is observed. This is mainly because on arterials and collectors, the traffic conditions are more complex and the traffic flow is easier to be interfered by various traffic management measures, which to some extent complicates the traveler's determination of the level of congestion.

DISCUSSION AND CONCLUSION

An efficient and accurate measurement of congestion intensity helps investigate the traffic conditions at different road classes and provides useful information for transportation planning and traffic operation improvement. This paper developed a macroscopic traffic congestion intensity measurement model based on cumulative logistic regression model. In order to improve the model's reliability and accuracy, congestion intensity measurement data collected in the real-world traffic network in Beijing, China were used in the model development and validation, and specific algorithms were designed for estimating the congestion intensity on different road classes.

In the model development process, LRT is used to evaluate the proposed model's proportional characteristics, which ensured a valid usage of cumulative logistic regression in this analysis. Wald test is applied to evaluate the significance of the 5 predefined independent variables (including travel speed, delay ratio, stop time, the number of stops, and the density of entrances/exits or signals) to congestion intensity measurement, the results indicate that travel speed has far more contribution to the modeling results, therefore it is determined as the independent variable included in the developed model.

In the model validation process, the statistical analysis shows an approximate 80% goodness-of-fit. By comparing the modeling results with the field observed congestion intensity data, the accuracy rate of the developed congestion intensity measurement model reaches 73.39%. Therefore we can see that the model developed in this research is accurate and useful for real-world congestion intensity measurement.

This research also discovered that the developed model has a better performance when it is used to estimate the congestion intensity on expressways, especially in terms of determining severe congestions; while more variance occurs when the model is used to measure the congestion intensity on arterials and collectors for a lower level of congestion intensity determination. Therefore it is recommended to further improve the congestion intensity measurement model by incorporating the parameters such as the factors that may influence traveler's perception, roadway conditions, transportation facilities, and traffic management strategies into the development of the model.

ACKNOWLEDGEMENTS

The authors acknowledge the supports of this paper by Major Project of Beijing Municipal Science and Technology Commission # D07050600440704, Major Project of Beijing Transportation Research Center # CZ200704, and National Basic Research Program of China # 2006CB705500.

REFERENCES

- [1] FHWA. *Summary of the Recommendations of the Workshop on National Urban Congestion Monitoring*. FHWA-PL-90-029. Federal Highway Administration, U.S. Department of Transportation, Washington, D.C., 1990.
- [2] FTA. *Joint FHWA/FTA Regulations, Management and Monitoring Systems, 49 C.F.R. Part 614*. Federal Transit Administration (FTA). Washington, D.C., 1993.
- [3] T. Lomax, S. Turner, G. Shunk, *et al.* "Quantifying congestion. Volume 1: Final Report". NCHRP Report, Transportation Research Board, Texas Transportation Institute, 1997.
- [4] S. Turner, "Examination of indicators of congestion level", *Transportation Research Record 1360*, Transportation Research Board, Washington, D.C., 1992, pp. 150-157.
- [5] M. Vaziri, "Development of highway congestion index with fuzzy set models", *Transportation Research Record 1802*, Transportation Research Board, Washington, D.C., 2002, pp. 16-22.
- [6] K. Hamad, and S. Kikuchi, "A measure of traffic congestion: A Fuzzy Inference Approach", In 81st Transportation Research Board Annual Meeting, Washington, D.C., 2002.
- [7] J. Wang, and Z. Guo, *Logistic Regression Models-Methods and Applications*, Beijing: Higher Education Press, 2001.
- [8] X. Chen, X. Dong, N. Wang, and J. Fang, "Logistic regression based on principal component analysis in resolving the co-linearity", *Chinese Journal of Health Statistics*, vol. 20, no. 4, pp. 21-24, 2003.

Received: October 17, 2009

Revised: November 11, 2009

Accepted: November 11, 2009

© Yu *et al.*; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.