

Estimating Bus Travel Time Using Survival Models

Amir Reza Mamdoohi¹

| armamdoohi@modares.ac.ir

Amin Delfan Azari²

| a.azari@imps.ac.ir

Mehrdad Alomoradi³

| m.alimoradi@imps.ac.ir

Abstract The prevailing model in the studies that estimate bus travel time is the linear regression which assumes the limit of the normal distribution for all observations. Besides, survival models can calculate that the probability of an event can change over time. Thus, examining event probabilities that change over time is ideal for risky basic models such as survival ones. Although these kinds of models are used less in the research of bus travel time, in this study Accelerated Failure Time (AFT) survival models and linear regression models are compared in the form of two modeling approaches, link-based, and section-based. As for modeling the Automated Vehicle Location (AVL), data of 32 buses in line number 313 in Tehran (from Sepah Sq. to Enqelab Sq.) is used, including the information for one week for May, August, and November 2015. According to the results, the accuracy of survival models is better than the linear regression model in both modeling approaches. Furthermore, the performance of the linear regression model is unfavorable for both observations of short (less than 100 seconds) and long (more than 900 seconds) travel time. In addition, the particular lane that has been built in the opposite direction in this route reduces the bus travel time by an average of about 15.7 percent.

Keywords: Bus Travel Time, Link, Segment, Survival Models, Accelerated Failure Time.

JEL Classification: R4, R41, R42.

1. Associate Professor, Faculty of Environment and Civil Engineering, Tarbiat Modares University, Tehran, Iran (Corresponding Author).

2. M.A. Student in Economics, Institute for Management and Planning Studies (IMPS), Tehran, Iran.

3. Lecturer, Institute for Management and Planning Studies, Tehran, Iran.

برآورد زمان سفر اتوبوس شهری با کاربری مدل‌های بقا

armamdoohi@modares.ac.ir |

امیررضا ممدوحی

دانشیار، دانشکده مهندسی عمران و محیط زیست، دانشگاه
تربیت مدرس (نویسنده مسئول).

a.azari@imps.ac.ir |

امین دلفان آذری

دانشجوی کارشناسی ارشد اقتصاد، موسسه عالی آموزش و
پژوهش مدیریت و برنامه ریزی.

m.alimoradi@imps.ac.ir |

مهرداد علیمردادی

مربی، موسسه عالی آموزش و پژوهش مدیریت و برنامه ریزی.

مقاله پژوهشی

پذیرش: ۱۳۹۹/۰۴/۱۷

دریافت: ۱۳۹۹/۰۲/۰۶

چکیده: مدل غالب در پژوهش‌های پیشین برای برآورد زمان سفر اتوبوس، رگرسیون خطی است که فرض محدودکننده توزیع نرمال را برای همه مشاهده‌ها دارد. از طرفی، مدل‌های بقا امکان محاسبه احتمال پیشامد رویدادهایی را دارند که می‌توانند در گذر زمان تغییر کنند. بنابراین، بررسی احتمال‌های رخدادی که با گذشت زمان تغییر می‌کنند، برای مدل‌های خطرپایه مانند بقا ایده‌آل هستند. در حالی که از این مدل‌ها در حوزه زمان سفر اتوبوس کمتر استفاده می‌شود. در این پژوهش، مدل‌های بقا زمان شکست شتابیده (AFT)، و رگرسیون خطی در قالب دو رویکرد کمان پایه و بخش پایه مقایسه می‌شوند. برای مدلسازی، داده‌های موقعیت خودکار خودروها (AVL)، مربوط به ۳۲ اتوبوس فعال در خط ۳۱۳ شهر تهران (میدان سپاه تا میدان انقلاب)، شامل اطلاعات یک هفته از هر یک از ماه‌های اردیبهشت، مرداد، و آبان در سال ۱۳۹۴ بکار می‌رود. بر اساس نتایج، دقت مدل‌های بقا در هر دو رویکرد مدلسازی بهتر از مدل رگرسیون خطی است. همچنین، برای مشاهده‌های زمان سفر کوتاه (کم‌تر از ۱۰۰ ثانیه)، و بلند (بیش از ۹۰۰ ثانیه)، عملکرد مدل رگرسیون خطی بسیار نامطلوب است. علاوه بر این‌ها، خط ویژه خلاف جهت که در این مسیر وجود دارد، به‌طور میانگین حدود ۱۵/۷ درصد از زمان سفر اتوبوس را می‌کاهد.

کلیدواژه‌ها: زمان سفر اتوبوس، کمان، بخش، مدل بقا، زمان شکست شتابیده.
طبقه‌بندی JEL: R42, R41, R4.

مقدمه

بهبود عملکرد حمل‌ونقل همگانی یکی از راه‌های افزایش تعداد کاربران این سامانه‌هاست. به‌طور معمول، کاربران حمل‌ونقل همگانی قابلیت اطمینان (پیش‌بینی زمان سفر و وجود جای کافی برای سوار شدن)، رامعیار کلیدی برای سنجش کیفیت و عنصر اصلی برای بکارگیری آن می‌دانند (Brownstone & Small, 2005). برای جبران بی‌ثباتی ذاتی زمان سفر اتوبوس‌ها در یک مسیر راهکارهای گوناگونی ارائه می‌شود: حرکت اتوبوس‌ها با سرعت کم‌تر تا اتوبوس تاخیردار به فاصله زمان‌بندی برسد؛ گذر اتوبوس از برخی ایستگاه‌ها بدون سوار کردن مسافر برای رسیدن به زمان‌بندی؛ یا محدود کردن تعداد مسافرانی که در هر ایستگاه می‌توانند سوار شوند. در مقابل، به‌جای تلاش برای جبران مستقیم بی‌ثباتی ذاتی، شرکت‌ها می‌کوشند با ارائه اطلاعات با کیفیت بالا (مانند موقعیت کنونی اتوبوس‌ها یا پیش‌بینی زمان رسیدن آن به ایستگاه)، عملکرد سامانه خود را بهبود دهند. این کار باعث افزایش قابلیت اطمینان درک‌شده^۱ از سوی مسافران می‌شود و به آن‌ها اجازه می‌دهد تصمیم‌های بهتری برای سفر خود بگیرند (Gayah *et al.*, 2016). منابع اطلاعات واقعی شامل سامانه‌های موقعیت‌یاب جهانی (GPS)^۲، موقعیت خودکار خودروها (AVL)^۳، شمارشگر خودکار مسافران (APC)^۴، و گردآوری خودکار کرایه (AFCs)^۵ هستند. برخی اطلاعات مانند تعداد اشغال صندلی‌ها نیز با بکارگیری سامانه‌های موبایل، مستقیم از مسافران به‌دست می‌آیند.

با استفاده از نتایج برآورد زمان سفر اتوبوس، گردانندگان سامانه‌های اتوبوسرانی می‌توانند زمان‌بندی حمل‌ونقل همگانی را برای همخوانی با زمان واقعی طی مسیر اتوبوس اصلاح کنند و زمان انتظار مسافران را بکاهند. همچنین، برآورد زمان سفر اتوبوس می‌تواند در ارزیابی سایر سامانه‌های حمل‌ونقل از قبیل سامانه‌های اولویت‌دهی به حمل‌ونقل همگانی (PTPS)^۶، و سامانه کنترل چراغ در تقاطع‌های چراغ‌دار بکار رود (Kieu *et al.*, 2012). بیش‌تر مدل‌های پیش‌بینی زمان سفر اتوبوس نیز به‌طور مستقیم یا به صورت ترکیبی با مشاهده‌های بهنگام، از اطلاعات مدل‌های برآورد برای ارائه پیش‌بینی‌های خود بهره می‌گیرند. در واقع برآورد، پایه پیش‌بینی برای آینده است. بیش‌تر پژوهش‌های انجام‌شده برای برآورد زمان سفر اتوبوس، مدل‌های رگرسیون را برای مدلسازی بکار

1. Perceived Reliability
2. Global Positioning System
3. Automated Vehicle Location
4. Automated Passenger Counter
5. Automated Fare Collection systems
6. Public Transport Priority Systems

می‌برند (Yu et al., 2017). فرض اصلی و محدودکننده این مدل‌ها آن است که توزیع جزء خطای مشاهده‌ها، نرمال در نظر گرفته می‌شود. در واقع، در مدل‌های رگرسیون الگوی تغییرهای یکسانی (بر اساس توزیع نرمال)، برای توضیح رفتار داده‌های مشاهده‌شده در نظر گرفته می‌شود. در این پژوهش، مدل‌های بقا^۱ بکار می‌رود. این مدل‌ها توزیع‌های گوناگون شامل نمایی، وایبل، نرمال، لجستیک، و گامای تعمیم‌یافته را برای جزء خطا در نظر می‌گیرند. با توجه به توانایی هر مدل در توضیح مشاهده‌ها، مدل برتر برای مدلسازی انتخاب و توزیع جزء خطای آن برای داده‌ها پذیرفته می‌شود. بنابراین، اهداف اصلی این پژوهش را می‌توان این‌گونه نام برد: توسعه مدل تخمین زمان سفر اتوبوس با بکارگیری مدل‌های بقا؛ ارزیابی دقت مدل‌های بقا در تخمین زمان سفر اتوبوس؛ تحلیلی بر توزیع داده‌های مشاهده‌شده زمان سفر اتوبوس؛ و تحلیلی بر سهم عوامل اصلی در مدت زمان سفر اتوبوس. پس از مقدمه، در بخش دوم، مبانی نظری پژوهش تبیین می‌شود و پژوهش‌های انجام‌شده در حوزه مدلسازی زمان سفر اتوبوس مورد بازبینی قرار می‌گیرند. در بخش سوم، روش انجام پژوهش بیان می‌شود. در بخش چهارم، مطالعه موردی و یافته‌های مدل‌ها بررسی می‌شوند. در بخش پنجم نیز نتیجه‌گیری و پیشنهادهای اجرایی و پژوهشی مطرح می‌شود.

مبانی نظری پژوهش

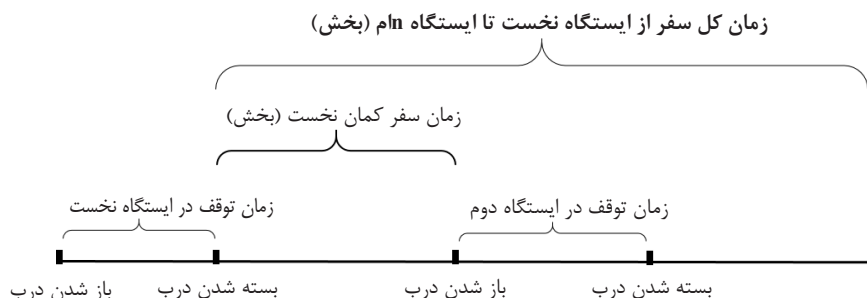
زمان سفر، بر پایه تعریفی که در کتابچه راهنمای ظرفیت حمل‌ونقل همگانی و کیفیت خدمات^۲ اشاره می‌شود، کل زمانی است که از یک مبدا به یک مقصد صرف می‌شود. در این پژوهش، بازه زمانی خروج اتوبوس از ایستگاه مبدا و ورود آن به ایستگاه مقصد، زمان سفر اتوبوس را بین آن دو ایستگاه نشان می‌دهد. ورود و خروج اتوبوس بر پایه باز و بسته شدن در اتوبوس در هر ایستگاه مشخص می‌شود. فاصله بین دو ایستگاه متوالی، کمان^۳ نامیده می‌شود. بخش^۴ نیز عبارت است از فاصله بین دو ایستگاه که ایستگاه مقصد می‌تواند متوالی ایستگاه مبدأ نیز نباشد. کمان در واقع، بخشی است که شامل تنها دو ایستگاه مبدا و مقصد می‌شود (شکل ۱).

1. Survival Models

2. Transit Capacity and Quality of Service Manual: TRB 2003

3. Link

4. Segment/ Section



شکل ۱: تعریف عملیاتی زمان سفر اتوبوس در این پژوهش

برآورد زمان سفر بر اساس مشاهده‌های زمان‌های ورود و خروج اتوبوس در زمان گذشته انجام می‌شود که در آن‌ها شرایط ترافیکی مشخص هستند. در حالی که پیش‌بینی عبارت از حدس زمان سفر اتوبوس در شرایط نامشخص ترافیک است. یادآور می‌شود که در ادبیات، واژه‌های برآورد و پیش‌بینی به جای هم بکار می‌روند و یک تعریف یکسان جهانی برای آن‌ها وجود ندارد (Rahman *et al.*, 2018).

در ادبیات، برآورد زمان سفر اتوبوس هم بر پایه اطلاعات خود اتوبوس و هم در ارتباط با شیوه‌های حمل‌ونقل شخصی انجام می‌شود. این دو رویکرد به عنوان دو طبقه‌بندی برای برآورد زمان سفر اتوبوس مطرح می‌شوند. در حالت نخست، بیش‌تر پژوهش‌های پیشین از تحلیل رگرسیون، روش آماری یا روش شبیه‌سازی برای برآورد زمان سفر اتوبوس بهره می‌برند (Levinson, 1983; McKnight *et al.*, 1997). اطلاعات اتوبوسی که در این پژوهش‌ها بکار می‌روند، عواملی هستند که بر زمان سفر اتوبوس اثر می‌گذارند. برای نمونه، تعداد مسافرانی که سوار یا پیاده می‌شوند، میانگین زمان سفر، زمان توقف در ایستگاه، و سایر عوامل اثرگذار. در حالت دوم، بیش‌تر پژوهش‌ها به تفاوت‌های بین سفرهای اتوبوس و سواری می‌پردازند. آن‌ها می‌گویند که زمان توقف و زمان کاهش سرعت/شتاب را از زمان سفر اتوبوس جدا کنند تا نسبت بین زمان سفر واقعی اتوبوس و زمان سفر سواری را فرمول‌بندی کنند. برای این کار نیز روش رگرسیون خطی به دلیل سادگی و توانایی برای یافتن ارتباط بین آن دو برگزیده می‌شود (Kieu *et al.*, 2012).

نخستین بار آلفا و همکاران^۱ (۱۹۸۸)، یک مدل رگرسیون خطی برای نواحی حومه شهر کانادا ارائه می‌کنند. کیتی و دیلی^۲ (۲۰۰۳)، با بکارگیری داده‌های موقعیت خودکار خودرو (AVL)، توصیفی کلی

1. Alfa *et al.*

2. Cathey & Dailey

برای برآورد زمان ورود/خروج خودروهای حمل‌ونقل همگانی به ایستگاه‌ها ارائه می‌کنند. شلابی و فرهان^۱ (۲۰۰۴)، با افزودن داده‌های APC به داده‌های AVL پویایی برآوردها را می‌افزایند. این پژوهش، زمان توقف اتوبوس در ایستگاه را جداگانه به عنوان یک متغیر در نظر می‌گیرد. آن‌ها همچنین با پیش‌بینی تعداد مسافران شبانه‌روزی ایستگاه، زمان ثابتی برای سوار شدن هر یک از مسافران در نظر می‌گیرند. پاتنایک و همکاران^۲ (۲۰۰۴)، دسته‌ای از مدل‌های رگرسیون چندمتغیره را توسعه می‌دهند. در مدل آن‌ها از میان تعداد زیادی از عوامل، متغیرهای مسافت مسیر، زمان توقف در ایستگاه، تعداد ایستگاه‌ها، و زمان روز به عنوان متغیر مستقل معنادار می‌شوند و متغیری مانند وضعیت آب‌وهوا در مدل معنادار نمی‌شود. چن و همکاران^۳ (۲۰۰۴)، مدلی پویا با بکارگیری ترکیب فیلترینگ کالمن^۴ و شبکه‌های عصبی مصنوعی ارائه می‌کنند. یافته‌های پژوهش بیانگر عملکرد بهتر مدل ترکیبی از مدل ANN تنها بود. جونگ و ریل^۵ (۲۰۰۵)، با بکارگیری داده‌های AVL، مدل پیش‌بینی زمان سفر اتوبوس را برای کاربردهای بهنگام توسعه می‌دهند. آن‌ها سه مدل میانگین داده‌های تاریخی، رگرسیون، و شبکه‌های عصبی مصنوعی را مقایسه می‌کنند و نتیجه می‌گیرند که مدل شبکه‌های عصبی عملکرد بهتری از دو مدل دیگر دارد. چین و همکاران^۶ (۲۰۰۲)، با بکارگیری داده‌های APC و AVL، یک مدل شبکه‌های عصبی مصنوعی را توسعه می‌دهند. در زمینه اطلاعات آب‌وهوا نیز از میان سه شاخص باد، بارش، و میزان دید، تنها میزان بارش رابطه معناداری با میزان دیرکرد اتوبوس نشان می‌دهد. محدوده کار واناچاکشی و همکاران^۷ (۲۰۰۹)، بکارگیری داده GPS برای برآورد زمان سفر اتوبوس در شرایط ترافیک ناهمگون (کامیون، اتوبوس، ون، خودرو، سه‌چرخ، موتور، و دوچرخه) در هند است و از الگوریتم فیلترینگ کالمن بهره می‌برند.

چن و همکاران (۲۰۱۲)، با بکارگیری رگرسیون زمان سفر اتوبوس، توانایی مدل‌های کمان‌پایه و بخش‌پایه را مقایسه می‌کنند. بر اساس یافته‌ها، عملکرد مدل‌های بخش‌پایه از مدل کمان‌پایه بهتر است. هدف پژوهش اسلام و همکاران^۸ (۲۰۱۵)، مطالعه عوامل اثرگذار بر قابلیت اطمینان مجموعه سامانه حمل‌ونقل همگانی است. بر پایه نظریه زنجیره مارکوف و با بکارگیری داده‌های عددی دو مسیر در سنگاپور و استکهلم، یافته‌های شبیه‌سازی آن‌ها نشان می‌دهد که واریانس انحراف از زمان‌بندی تعیین‌شده برای اتوبوس‌ها در طول یک

1. Shalaby & Farhan
2. Patnaik *et al.*
3. Chen *et al.*
4. Kalman Filter
5. Jeong & Rilett
6. Chien *et al.*
7. Vanajakshi *et al.*
8. Islam *et al.*

مسیر مشخص افزایشی است و این افزایش ارتباطی با اندازه اتوبوس ندارد. گنجی و همکاران (۱۳۹۵)، در پی شناسایی متغیرهای اثرگذار بر زمان سفر اتوبوس و سپس ارائه یک مدل برآورد معتبر هستند. آن‌ها از مدل جنگل تصادفی در پژوهش خود بهره می‌برند. برای سادگی و همچنین به دلیل در دسترس نبودن اطلاعات، داده‌های بلندمدت مانند آب‌وهوا و فصل، ثابت دیده می‌شوند. بر اساس تحلیل حساسیت انجام شده، شیب مسیر و روز هفته به عنوان اثرگذارترین متغیرها شناخته می‌شوند. هاشمی و البدوی (۱۳۹۸)، سعی در پایه‌گذاری نگاه علمی به زمان سفر اتوبوس دارند و به دنبال کاربرد سامانه‌های هوشمند حمل‌ونقل در این کار هستند. به این منظور، با کاربرد داده‌های AVL شهر قزوین، سه مدل شبکه‌های عصبی مصنوعی، سری‌های زمانی، و ماشین بردار پشتیبان (SVM)^۱، را برای قیاس پیش‌بینی زمان ورود اتوبوس به ایستگاه بکار می‌برند. بر اساس یافته‌ها، عملکرد مدل شبکه‌های عصبی مصنوعی بهتر از دو مدل دیگر است. یو و همکاران (۲۰۱۷)، برای پیش‌بینی زمان سفر اتوبوس از مدل جنگل تصادفی بر پایه نزدیک‌ترین همسایه (RFNN)^۲ بهره می‌برند. آن‌ها دقت مدل خود را با چهار مدل دیگر شامل رگرسیون خطی، K-امین همسایگان نزدیک (KNN)^۳، ماشین بردار پشتیبان، و جنگل تصادفی کلاسیک (RF) قیاس می‌کنند که دقت مدل ارائه شده آن‌ها از سایر مدل‌ها بیش‌تر است. یو و همکاران (۲۰۱۸)، مدل‌های برآورد زمان سفر اتوبوس را توسعه می‌دهند و برای نخستین بار از مدل‌های بقا در این کار بهره می‌برند. اما تنها یک زوج ایستگاه (یک بخش) را بررسی می‌کنند که در بر دارنده هفت ایستگاه در یک مسیر است و متغیرهای محدودی را در نظر می‌گیرند.

در پژوهش حاضر همه زوج ایستگاه‌های یک مسیر (کمان‌ها و بخش‌ها) دیده می‌شوند. همچنین، به جز نرخ رسیدن مسافران، سایر متغیرهایی که در پژوهش‌های پیشین بکار می‌روند (جدول ۱)، در این پژوهش نیز بررسی می‌شوند.

روش‌شناسی پژوهش

مدل بقا برای مطالعه داده‌های مدت‌زمانی بکار می‌رود. این مدل که بر پایه احتمال خطر ساخته می‌شود، احتمال شرطی پیشامد یک رویداد را در زمان t بررسی می‌کند، به شرط آن‌که پیش از آن رخ نداده باشد. مدل بقا امکان محاسبه احتمال پیشامد رویدادهایی را دارد که می‌توانند در گذر زمان تغییر کنند. بنابراین، بررسی احتمال‌های رخ داده که با گذشت زمان تغییر می‌کنند، برای مدل‌های خطرپایه ایده‌آل هستند (Washington *et al.*, 2020). تابع توزیع تجمعی مدل‌های

1. Support Vector Machines
2. Random Forests Based on the Nearest Neighbor
3. K-Nearest Neighbors

مدت زمانی خطرپایه عبارت است از $F(t) = P(T < t)$ که در آن P نشانگر احتمال پیشامد رویداد، T متغیر تصادفی زمان، و t یک زمان مشخص است. اگر t متغیر تصادفی پیوسته باشد، F نیز پیوسته و افزایشی است. در نتیجه، تابع چگالی مدل‌های خطرپایه به عنوان مشتق یکم F نسبت به زمان برابر است با $f(t) = \frac{dF(t)}{dt}$ ، و تابع خطر برای وقوع رویداد به شکل $h(t) = \frac{f(t)}{1-F(t)}$ خواهد بود. $h(t)$ احتمال شرطی پیشامد رویداد بین زمان t و $t+dt$ است، به شرطی که آن رویداد تا زمان t رخ نداده باشد.

جدول ۱: متغیرهای بکاررفته در ادبیات برآورد و پیش‌بینی زمان سفر اتوبوس

پژوهش	مسافت	زمان توقف	تعداد ایستگاه	موقعیت لحظه‌ای	ویژگی خاص مسیر	تعداد تقاطع	تعداد چراغ	نرخ رسیدن مسافر	سرفاصله زمانی	آبوهوا	روز هفته	زمان روز	فصل سال
آلفا و همکاران (۱۹۸۸)	✓		✓		×		✓					✓	
کیتی و دیلی (۲۰۰۳)				✓									
شلابی و فرهان (۲۰۰۴)		✓	✓	✓			✓						
پاتنایک و همکاران (۲۰۰۴)	✓	✓	✓				✓			×		✓	
چن و همکاران (۲۰۰۴)	✓									✓	✓	✓	
جونگ و ریل (۲۰۰۵)	✓	✓											
چین و همکاران (۲۰۰۲)		✓					✓			✓	✓	✓	
واناجاکشی و همکاران (۲۰۰۹)				✓									
چن و همکاران (۲۰۱۲)	✓									✓	✓		
اسلام و همکاران (۲۰۱۵)	✓	✓	✓				✓						
گنجی و همکاران (۱۳۹۵)	✓				✓		✓			✓	✓		
هاشمی و البدوی (۱۳۹۸)									✓	✓	✓		
یو و همکاران (۲۰۱۷)									✓	×	✓	✓	✓
یو و همکاران (۲۰۱۸)	✓												✓
دسترسی به متغیر در این پژوهش	*	*	*	*	*	*	*	*	-	*	*	*	*
توضیح متغیرها:	✓			معناداری متغیر			×		بی‌معنایی متغیر				
دسترسی به متغیر در این پژوهش:	*			متغیر وجود دارد			-		متغیر وجود ندارد				

در واقع، $h(t)$ نرخ پیشامد رویداد را در لحظه t نشان می‌دهد (Allison, 2014)، یعنی:

$$\lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (1)$$

در رویکرد زمان شکست شتابیده (AFT)^۱، برای مدل‌های بقا فرض می‌شود که در یک تابع بقای پایه، متغیرهای کمکی به صورت مستقیم بر زمان بقای رویداد اثرگذارند (شتاب‌دهنده یا کندکننده). تابع بقای پایه حالتی است که ضرایب همه متغیرهای کمکی صفر هستند. برای سادگی فرض می‌شود که متغیرهای کمکی با تابع $EXP(BX)$ که در آن β بردار پارامترهای برآوردپذیر است، بر تابع بقای پایه اثر می‌گذارند. در مدل AFT فرض می‌شود که لگاریتم زمان بقا (T)، به صورت خطی با متغیرهای کمکی در ارتباط است و می‌توان آن را به شکل رابطه (۲) نوشت:

$$\ln(T_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + u_i \quad (2)$$

T_i : زمان باقیمانده تا پیشامد رویداد در مشاهده نام، x_i : متغیر نام، u_i : جزء خطای مشاهده نام، و β_i : ضریب متغیر نام هستند. جزء خطا می‌تواند توزیع‌های گوناگونی داشته باشد. توزیع‌های پارامتری معمول در ادبیات ارزش حدی^۲، ارزش حدی تعمیم‌یافته^۳، نرمال، و لجستیک هستند که به ترتیب توزیع‌های وایبل، گامای تعمیم‌یافته، لاگ‌نرمال، و لاگ‌لجستیک^۴ را برای متغیر وابسته در نظر می‌گیرد. تابع بقا برای هر یک از این توزیع‌ها به صورت تابع (۳) است (Allison, 2014):

$$\begin{aligned} \text{wei bull:} \quad & S(w) = \exp\left(-\exp\left(\frac{w-\mu}{\sigma}\right)\right) \\ \text{Generalized Gamma:} \quad & S(w) = \begin{cases} 1 - \frac{r(\delta^{-2}\delta^{-2}\exp(\delta\frac{w-\mu}{\sigma}))}{r(\delta^{-2})} & \text{if } \delta > 0 \\ \frac{r(\delta^{-2}\delta^{-2}\exp(\delta\frac{w-\mu}{\sigma}))}{r(\delta^{-2})} & \text{if } \delta < 0 \end{cases} \\ \text{log - normal:} \quad & S(w) = 1 - \Phi\left(\frac{w-\mu}{\sigma}\right) \\ \text{log - logistic:} \quad & S(w) = (1 + \exp(\frac{w-\mu}{\sigma}))^{-1} \end{aligned} \quad (3)$$

که در آن‌ها: $W = \log(T)$ ، $\mu = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$ ، σ = واریانس T ، δ = پارامتر

1. Accelerated Failure Time
2. Extreme Value
3. Generalized Extreme Value
4. Wei bull, Generalized Gamma, Log-Normal, Log-Logistic

شکل در توزیع گامای تعمیم‌یافته، $\Gamma(a)$ = تابع گامای کامل، $\Gamma(a, z)$ = تابع گامای ناکامل، و ϕ = تابع توزیع تجمعی توزیع نرمال هستند. از رابطه اصلی مدل به دست می‌آید که $T = e^{\beta X} e^u$. اگر مقدار متغیر X_k به اندازه δ تغییر یابد، آن گاه نسبت زمان‌های بقا برابر است با:

$$\frac{T(X_k + \delta)}{T(X_k)} = e^{(X_k + \delta - X_k)\beta_k} = e^{\delta\beta_k} \quad (4)$$

اگر مقدار δ برابر ۱ واحد باشد، آن گاه: $\frac{T(X_k + 1)}{T(X_k)} = e^{\beta_k}$ است. بنابراین، اگر $\beta_k > 1$ باشد، یعنی متغیر X_k شتاب پیشامد رویداد را کند و زمان وقوع آن را طولانی‌تر می‌کند. اگر $\beta_k = 1$ باشد، یعنی متغیر X_k اثری روی شتاب پیشامد رویداد ندارد. اگر $\beta_k < 1$ باشد، یعنی متغیر X_k شتاب پیشامد رویداد را افزایش می‌دهد و زمان وقوع آن را کوتاه‌تر می‌کند. رویدادی که در این پژوهش مدت زمان آن در نظر گرفته می‌شود، «رسیدن اتوبوس به ایستگاه مقصد» است.

در ادبیات برای ارزیابی مدل‌های پارامتری، معیارهای اطلاعات آکائیکه (AIC)^۱، و بیزین (BIC)^۲ بکار می‌رود (Jenkins, 2005). در این پژوهش، دو معیار دیگر یعنی مجذور میانگین مربع‌های خطا (RMSE)^۳، و میانگین خطای مطلق (MAE)^۴، برای قیاس بین مدل‌ها بکار می‌رود. هر دو معیار RMSE و MAE بر پایه اندازه‌گیری تفاوت مقدار پیش‌بینی‌شده و مقدار واقعی عمل می‌کنند. با این تفاوت که RMSE نسبت به داده‌هایی با اختلاف زیاد بین پیش‌بینی و مقدار واقعی حساس است، ولی MAE این گونه نیست (Kennedy, 2003). پس از قیاس مدل‌های گوناگون بقا با کاربرد چهار شاخص بالا، بهترین مدل از دید قابلیت برآورد به عنوان مدل برتر برگزیده می‌شود. با توجه به ضرایب متغیرهای معنادار در مدل، روشن می‌شود که هر متغیر چه اثری بر زمان سفر اتوبوس دارد. سپس این مدل با مدل رگرسیون خطی قیاس می‌شود که مدل غالب در برآورد زمان سفر اتوبوس است تا دقت آن‌ها ارزیابی شود.

مطالعه موردی

شبکه اتوبوسرانی تهران در سال ۱۳۹۴ دارای ۴۶۳ خط رفت و برگشتی و ۴۷۵۰ ایستگاه بود و ۶۵۰۰ دستگاه اتوبوس معمولی و تندرو در این شبکه فعال بودند. بر اساس آمار شرکت مطالعات جامع

1. Akaike's Information Criterion
2. Bayesian Information Criterion
3. Root Mean Square Errors
4. Mean Absolute Error

حمل‌ونقل و ترافیک تهران^۱ در سال ۱۳۹۴ حدود ۴/۵ میلیون نفر در روز در شهر تهران با ناوگان اتوبوسرانی و مینی‌بوسرانی جابه‌جا شده‌اند و سهم این بخش از جابه‌جایی روزانه شهر تهران حدود ۲۰ درصد بوده است. مطابق شکل (۲)، مسیری که در این پژوهش به عنوان نمونه موردی بررسی می‌شود، خط اتوبوسرانی ۳۱۳ شهر تهران با درازای حدوداً ۵/۵ کیلومتر بین میدان سپاه تا میدان انقلاب است. از بین خطوط فعال شبکه اتوبوسرانی تهران، در این پژوهش خطی مد نظر است که نخست در مرکز شهر و محدوده طرح ترافیک واقع شده باشد. از طرفی، بر اساس گنجی و همکاران (۱۳۹۵)، متغیر شیب مسیر یکی از سه متغیر اصلی اثرگذار بر زمان سفر اتوبوس است. بنابراین، خطوطی مورد توجه قرار می‌گیرند که مسیر حرکت اتوبوس‌ها در آن با توجه به جغرافیای شهری تهران، در محور شرقی-غربی (و نه شمالی-جنوبی) باشد. از دیگر شاخص‌های انتخاب خط اتوبوسرانی در این پژوهش، وجود خط ویژه اتوبوس و تقاطع‌های چراغ‌دار و غیرچراغ‌دار کافی در طول مسیر است. در نهایت، از بین دو خط اتوبوسرانی ۲۹۷ (میدان خراسان-میدان انقلاب)، و ۳۱۳ (میدان سپاه-میدان انقلاب)، خط ۳۱۳ انتخاب می‌شود. با این انتخاب، داده‌های خط ۲۹۷ در مسیر مشترک با خط ۳۱۳ نیز در مدل‌ها بکار می‌روند تا از مشاهدات بیش‌تری برای مدلسازی استفاده شود. خط ۳۱۳ در محدوده مرکزی تهران واقع است و دربردارنده ۱۲ ایستگاه، ۸ تقاطع، ۸ چراغ راهنمایی، خط ویژه اتوبوس، و یک مسیر بدون شیب است. در این پژوهش، اطلاعات AVL مربوط به یک هفته از هر یک از ماه‌های اردیبهشت، مرداد، و آبان سال ۱۳۹۴ بکار می‌رود. این اطلاعات دربرگیرنده کد شناسایی اتوبوس، تاریخ و زمان دقیق سفر اتوبوس، موقعیت جغرافیایی اتوبوس (هر دو دقیقه یک‌بار)، و موقعیت‌های دقیق جغرافیایی باز و بسته شدن در اتوبوس (فقط درب عقب) هستند.



شکل ۲: نقشه نمونه موردی، مسیر خط اتوبوسرانی میدان سپاه تا میدان انقلاب تهران

1. <http://trafficorg.tehran.ir/Default.aspx?tabid=229>

علاوه بر این، اطلاعات هواشناسی شامل دمای هوا و میزان بارندگی نیز با بکارگیری داده‌های ایستگاه هواشناسی مهرآباد تهران (نزدیک‌ترین ایستگاه هواشناسی به خط اتوبوسرانی) در مدل‌ها بررسی می‌شود. به دلیل کم بودن مشاهده‌های برخی کمان‌ها، تنها فاصله بین ایستگاه‌های شماره ۳ تا ۱۰ (جدول ۲)، در مدلسازی بکار می‌رود تا یکنواختی مشاهده‌ها حفظ شود.

جدول ۲: مشخصات ترافیکی کمان‌های خط اتوبوسرانی میدان سپاه تا میدان انقلاب تهران (شماره ۳۱۳)

شماره کمان	ایستگاه مبدا	ایستگاه مقصد	فاصله (متر)	میانگین زمان سفر (ثانیه)	انحراف معیار زمان سفر (ثانیه)	تعداد سواره‌ای	تعداد چهارراه چراغ‌دار	سهم خط ویژه* (درصد)
۳	سواره‌تالقانی	بهار	۴۲۶	۸۵/۶	۳۱	۱	۱	۱۰۰
۴	بهار	مفتح	۳۶۶	۸۸/۷	۳۳	۱	۱	۱۰۰
۵	مفتح	ایران‌شهر	۴۲۹	۹۴	۲۳	۰	۱	۱۰۰
۶	ایران‌شهر	ویلا	۵۲۰	۱۱۵/۷	۳۴	۰	۳	۱۰۰
۷	ویلا	شرکت نفت	۶۰۱	۶۴/۷	۱۷	۰	۱	۱۰۰
۸	شرکت نفت	میدان فلسطین	۳۴۷	۱۵۱/۲	۳۱	۱	۱	۵۰
۹	میدان فلسطین	دبستان فردوسی	۴۱۶	۶۶/۳	۱۲	۰	۰	۰

* سهمی از آن کمان است که به صورت خط ویژه ساخته می‌شود. این نسبت، بازهٔ صفر (یعنی تداخل با ترافیک) تا صد (کل کمان به شکل خط ویژه برای اتوبوس) را در برمی‌گیرد.

با مطالعه ادبیات، متغیرهای زیر برای این پژوهش انتخاب می‌شوند که به دو دسته متغیرهای پیوسته و مجازی تقسیم می‌شوند (جدول ۳).

جدول ۳: متغیرهای مدل‌سازی زمان سفر اتوبوس در خط میدان سپاه - میدان انقلاب برای سه ماه گوناگون از سال ۱۳۹۴

دسته	نماد	عنوان	توضیحات	واحد
پیوسته	Trip	زمان سفر اتوبوس	فاصله زمانی بسته شدن درب اتوبوس در ایستگاه مبدا تا باز شدن درب آن در ایستگاه مقصد، برای هر بخش	
	Dwell	زمان توقف در ایستگاه‌های بین راه	مجموع فاصله زمانی بین باز و بسته شدن درب اتوبوس در ایستگاه‌های بین راه	ثانیه
	Headway	سرفاصله زمانی	فاصله زمانی بین هر دو اتوبوس متوالی در طول خط	
	Distance	مسافت	فاصله بین ایستگاه مبدا و مقصد هر بخش	کیلومتر
	Stops	نسبت تعداد ایستگاه‌های بین راه	تعداد ایستگاه‌های بین راه به ازای هر کیلومتر فاصله	
	Sign	نسبت تقاطع چراغ‌دار در مسافت	تعداد تقاطع چراغ‌دار در هر کیلومتر از مسافت بخش	بر کیلومتر
	Way	نسبت تعداد سهراهی در مسافت	تعداد سهراهی در هر کیلومتر از مسافت بخش	
	Temp	دمای هوا	میانگین دمای هوا به صورت روزانه	سانتی‌گراد
	Rain	بارش باران	میزان بارش باران به صورت روزانه	میلی‌متر
	Dedicated	نسبت خط ویژه	عددی بین صفر تا یک برای نسبتی از مسافت که به صورت خط ویژه است (صفر به معنای هیچ و یک به معنای همه مسافت است)	-

ادامه جدول ۳: متغیرهای مدلسازی زمان سفر اتوبوس در خط میدان سپاه - میدان انقلاب برای سه ماه گوناگون از سال ۱۳۹۴

دسته	نماد	عنوان	توضیحات	واحد
بازه زمانی	Ordi	اردیبهشت	سفرهای اتوبوس در ماه اردیبهشت	-
	Mord	مرداد	سفرهای اتوبوس در ماه مرداد	-
	Aban	آبان	سفرهای اتوبوس در ماه آبان	-
	T1	بازه زمانی اول صبح	از ۵:۳۰ تا ۶:۳۰ روزهای شنبه تا چهارشنبه	-
	T2	بازه زمانی اوج صبح	از ۶:۳۰ تا ۹ صبح روزهای شنبه تا چهارشنبه	-
	T3	بازه زمانی میانه روز	از ۹ تا ۱۶:۳۰ روزهای شنبه تا چهارشنبه	-
	T4	بازه زمانی اوج عصر	از ۱۶:۳۰ تا ۱۹:۳۰ روزهای شنبه تا چهارشنبه	-
بازه زمانی آخر عصر	T5	بازه زمانی آخر عصر	از ۱۹:۳۰ تا ۲۰:۳۰ روزهای شنبه تا چهارشنبه	-
	T6	بازه زمانی پنج‌شنبه	از ۵:۳۰ تا ۲۰:۳۰ روز پنج‌شنبه	-
	T7	بازه زمانی جمعه	از ۶ تا ۲۰ روز جمعه	-

یافته‌های مدلسازی

مدل‌های رگرسیون خطی و بقا برای هر دو رویکرد کمان‌پایه و بخش‌پایه در نرم‌افزار STATA.14 ساخته می‌شوند. علامت متغیرها برای هر دو مدل و در هر دو رویکرد مطابق چشم‌داشت هستند. برای مدل‌های بقا، با توجه به شاخص‌های تعریف‌شده RMSE، BIC، AIC، و MAE، در رویکرد کمان‌پایه مدل لاگ‌نرمال، و در رویکرد بخش‌پایه مدل لاگ‌جستیک به عنوان مدل‌های برتر برگزیده می‌شوند. سپس این مدل‌ها با مدل رگرسیون خطی قیاس می‌شوند (جدول ۴). در رویکرد کمان‌پایه دقت هر دو مدل مطلوب نیست (ضرایب برازش ۰/۴۵ و ۰/۵۶)، اما دقت مدل بقا (لاگ‌نرمال) بیش‌تر است. در مدلسازی بخش‌پایه دقت هر دو گونه مدل‌های رگرسیون خطی و بقا مطلوب است (ضرایب برازش هر دو ۰/۹۳)، اما با توجه به قیاس برآورد‌مشاهده، باز هم دقت مدل بقا (لاگ‌جستیک) بیش‌تر است. در این رویکرد، اگرچه شاخص‌های MAE و RMSE رگرسیون خطی بهتر از بقا هستند، اما باید توجه داشت که اصولاً تابع هدف مدل‌های رگرسیون، کمینه‌سازی RMSE و MAE است. اما تابع هدف مدل‌های بقا (و همه مدل‌هایی که با روش MLE^۱ پارامترها را برآورد می‌کنند)، بیشینه‌سازی تابع درست‌نمایی (Likelihood) است. بنابراین، برای نتیجه‌گیری بهتر، شاخص برآورد‌مشاهده دو مدل

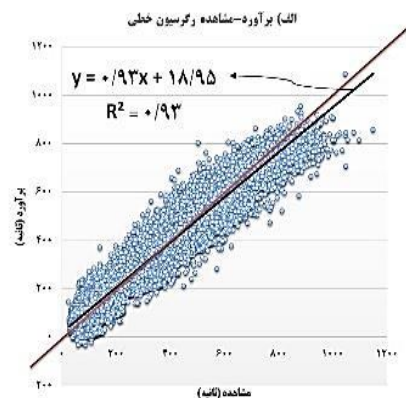
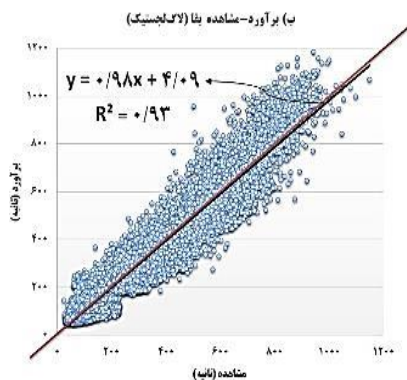
1. Maximum Likelihood Estimation

نیز مقایسه می‌شوند (شکل‌های ۳ و ۴). با توجه به شکل (۴)، و ضرایب خطوط برازش‌شده، مشخص است که مدل لاگ‌لجستیک بقا عملکرد بهتری دارد. همچنین، برای مشاهده‌های بیش‌تر از ۹۰۰ ثانیه، مدل رگرسیون خطی عملکرد نامناسبی دارد. برای مقادیر کوچک، مشاهده‌های زمان سفر (کم‌تر از ۱۰۰ ثانیه) نیز، مدل رگرسیون خطی مقادیر منفی برای زمان سفر اتوبوس برآورد می‌کند که همگی جزو نقاط ضعف مدل‌های رگرسیون خطی برداشت می‌شوند.

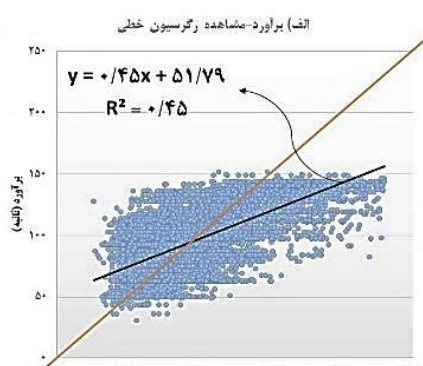
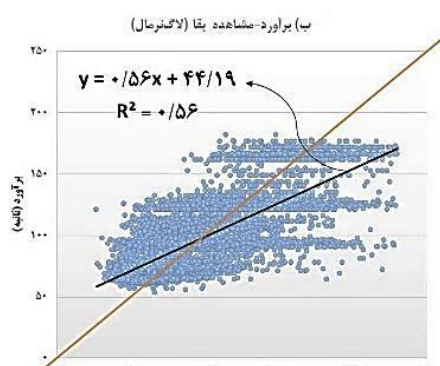
جدول ۴: قیاس دقت مدل‌سازی کمان‌پایه و بخش‌پایه برای مدل‌های رگرسیون خطی و بقا برآورد زمان سفر اتوبوس برای سه ماه گوناگون از سال ۱۳۹۴ در خط میدان سپاهمیدان انقلاب

رویکرد مدل‌سازی									
کمان‌پایه					بخش‌پایه				
شاخص‌ها					شاخص‌ها				
خط برآورد - مشاهده					خط برآورد - مشاهده				
مدل	ضریب	عرض از	RMSE	MAE	مدل	ضریب	عرض از	RMSE	MAE
	نکویی	شیب				نکویی	شیب		
	برازش	مبدا				برازش	مبدا		
رگرسیون					رگرسیون				
خطی	۰/۴۵	۰/۴۵	۵۱/۷۹	۳۳/۱۷	خطی	۰/۹۳	۰/۹۳	۱۸/۹۵	۵۲/۸۲
لاگ‌نرمال	۰/۵۶	۰/۵۶	۴۴/۱۹	۳۰/۵۷	لاگ	۰/۹۳	۰/۹۸	۴/۰۹	۵۶/۸۴
بقا					لجستیک‌بقا				

برای بررسی دقیق‌تر اثر هر یک از متغیرها بر زمان سفر اتوبوس، در مدل‌سازی بخش‌ها (دقت مدل کمان‌پایه مناسب نیست)، ضرایب هر یک از متغیرهای معنادار مدل در سه بازه زمانی اوج صبح، میانه روز، و اوج عصر به تفکیک سه ماه اردیبهشت، مرداد، و آبان محاسبه می‌شود و سپس در مقدار میانگین هر یک از متغیرها ضرب می‌شود (متغیر دمای هوا معنادار نشد). این محاسبه از دید مدل‌سازی، اثر هر یک از جمله‌های متغیرهای مستقل را در تشکیل متغیر وابسته (زمان سفر) نشان می‌دهد. در واقع، یافته‌ها نشانگر سهم هر یک از متغیرها در زمان سفر اتوبوس در هر بخش است (جدول ۵).



شکل ۳: قیاس برآورد-مشاهده مدل‌های رگرسیون خطی و بقا (لاگ‌نرمال)، در مدلسازی کمان پایه زمان سفر اتوبوس برای سه ماه گوناگون از سال ۱۳۹۴ در خط میدان سپاه-میدان انقلاب



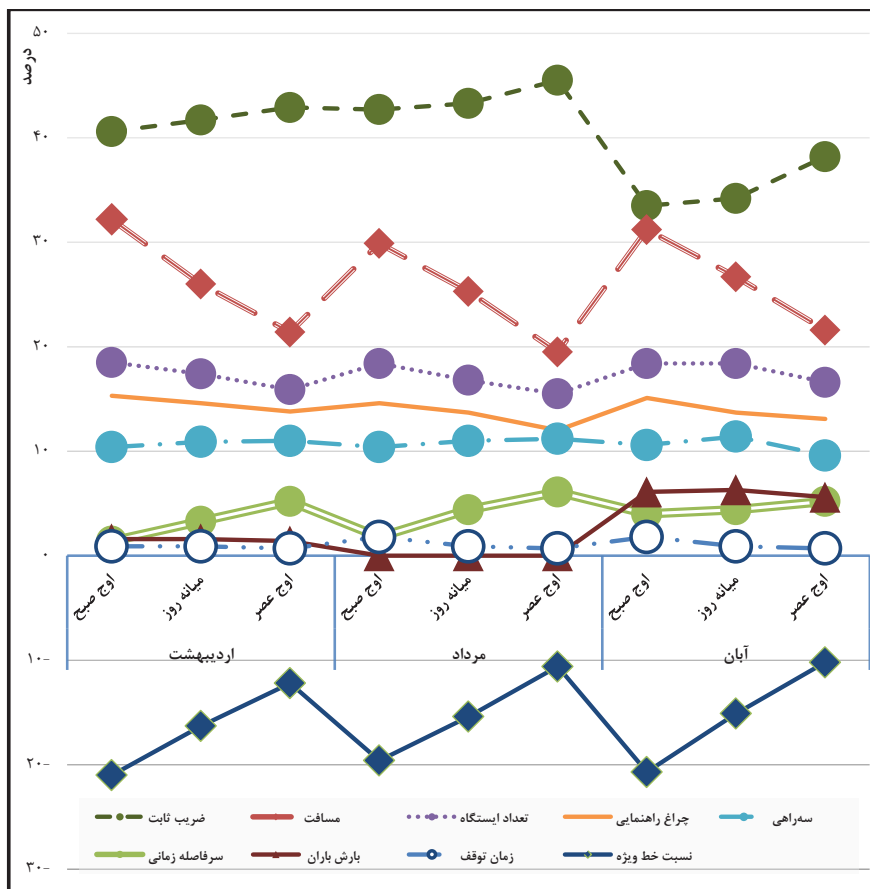
شکل ۴: قیاس برآورد-مشاهده مدل‌های رگرسیون خطی و بقا (لاگ‌لجستیک)، در مدلسازی بخش پایه زمان سفر اتوبوس برای سه ماه گوناگون از سال ۱۳۹۴ در خط میدان سپاه-میدان انقلاب

جدول ۵: سهم متغیرها در توضیح زمان سفر اتوبوس به تفکیک ماه و بازه زمانی روز در مدلسازی بخش پایه برای سه ماه گوناگون از سال ۱۳۹۴ در خط میدان سپاهمیدان انقلاب

بازه روز	زمان سفر (ثانیه)	زمان توقف	مسافت	سرفاصله زمانی	تعداد ایستگاه	سه راهی	چراغ راهنمایی	نسبت خط ویژه	بارش باران	ضریب ثابت	میانگین سهم متغیرها (درصد)	
											۱	۲
اوج صبح	۲۵۸	۰/۹	۳۲/۲	۱/۴	۱۸/۵	۱۰/۴	۱۵/۳	-۲۱/۰	۱/۶	۴۰/۶	۱- پهشت	
میانه روز	۳۱۷	۰/۹	۲۶/۰	۳/۳	۱۷/۴	۱۰/۹	۱۴/۶	-۱۶/۳	۱/۶	۴۱/۷		
اوج عصر	۳۳۰	۰/۷	۲۱/۴	۵/۲	۱۵/۹	۱۱/۰	۱۳/۸	-۱۲/۲	۱/۴	۴۲/۹		
اوج صبح	۲۵۶	۱/۸	۲۹/۹	۱/۸	۱۸/۴	۱۰/۴	۱۴/۶	-۱۹/۶	۰	۴۲/۷	۲- صبح	
میانه روز	۲۹۹	۰/۹	۲۵/۳	۴/۴	۱۶/۸	۱۱/۰	۱۳/۷	-۱۵/۴	۰	۴۳/۳		
اوج عصر	۳۰۸	۰/۷	۱۹/۵	۶/۱	۱۵/۵	۱۱/۲	۱۲/۰	-۱۰/۶	۰	۴۵/۵		
اوج صبح	۲۵۷	۱/۸	۳۱/۲	۴/۰	۱۸/۴	۱۰/۶	۱۵/۱	-۲۰/۷	۶/۱	۳۳/۵	۳- عصر	
میانه روز	۳۰۲	۰/۹	۲۶/۷	۴/۴	۱۸/۴	۱۱/۴	۱۳/۷	-۱۵/۱	۶/۳	۳۴/۲		
اوج عصر	۳۱۴	۰/۷	۲۱/۶	۵/۲	۱۶/۶	۹/۶	۱۳/۱	-۱۰/۲	۵/۶	۳۸/۲		

دیده می شود که به جز میزان بارش باران و ضریب ثابت که در ماه آبان نسبت به دو ماه دیگر تغییر سهم قابل توجهی (به ترتیب حدود ۵ درصد افزایش و ۸ درصد کاهش) دارد، سهم سایر متغیرها در ماه های گوناگون تقریباً میانگین ثابتی دارد و تغییر چندانی ندارد. برای درک بهتر سهم متغیرها و اثر هر یک بر زمان سفر اتوبوس، تغییر سهم ها در سه ماه اردیبهشت، مرداد، و آبان و در بازه های زمانی اوج صبح، میانه روز، و اوج عصر ترسیم می شوند (شکل ۵). بیشترین سهم از زمان سفر اتوبوس در این نمونه، مربوط به ضریب ثابت مدل و حدود ۴۰/۲ درصد است. این ضریب نشانگر میانگین اثر متغیرهایی است که در مدل وارد نمی شوند. با گذر زمان روز از اوج صبح به اوج عصر در هر ماه، سهم ضریب ثابت افزایش می یابد. می توان چنین توجیه کرد که داده های پژوهش مربوط به سال ۱۳۹۴ است و چون حجم ترافیک مسیر در آن زمان در دسترس نیست، در مدل دیده نمی شود. از سویی، خط اتوبوسرانی ۳۱۳ در محدوده طرح ترافیک قرار دارد، و آلوده خودروهایی شخصی از ساعت ۶:۳۰ صبح تا ۱۷ عصر در این منطقه ممنوع است، و پس از آن بر اساس بازدید میدانی، حجم ترافیک به شدت افزایش می یابد. چون سهم ضریب ثابت هم فزاینده است، بنابراین، می توان نتیجه گرفت که سهم عمده ای از این ضریب را حجم ترافیک می سازد. پس از ضریب ثابت، متغیرهای مسافت و تعداد

ایستگاه‌های بین راه بیش‌ترین سهم را از زمان سفر اتوبوس دارند. اصولاً نقش و اثر این دو متغیر یکسان است و سهم آن‌ها در مجموع حدود $43/3$ درصد است. در هر روز از ماه، با گذشت زمان از اوج صبح تا اوج عصر، اثر این متغیرها کاهش می‌یابد. پس از این متغیرها، تعداد چراغ راهنمایی در مسیر با حدود 14 درصد، بیش‌ترین سهم را داراست. اثر این متغیر نیز با گذر زمان طی روز، کاهشی است. در ادبیات، سهم این متغیر بین 10 تا 20 درصد است. متغیر تعداد سهراهی در مسیر با حدود $10/7$ درصد سهم، در جایگاه بعدی با اثری کم‌تر از چراغ راهنمایی است. چراغ‌های راهنمایی در تقاطع‌های چهارراهی نصب هستند، بنابراین، توزیع سهم بین این دو متغیر منطقی است. ولی برخلاف سه متغیر پیشین (مسافت، تعداد ایستگاه بین راه، و چراغ راهنمایی)، اثر این متغیر طی روز، افزایشی است و از اوج صبح تا اوج عصر (به‌جز در ماه آبان)، حدود یک درصد افزایش می‌یابد. بر اساس بازدید میدانی، افزایش حجم ترافیک در تقاطع‌های سهراهی این خط می‌تواند این مسئله را به‌خوبی توجیه کند. سهم متغیر سرفاصله زمانی هم در روز، افزایشی است. چون در این خط با گذر زمان از اوج صبح تا اوج عصر، متناسب با زمان سفر اتوبوس، سرفاصله زمانی نیز افزایش می‌یابد. برخلاف سایر متغیرها، سهم متغیر زمان توقف در ایستگاه‌های بین راه تقریباً ثابت و اندک است، یعنی تغییرهای زمان توقف اتوبوس‌ها در ایستگاه‌های بین راه اثر چندانی بر زمان سفر کلی اتوبوس در یک بخش ندارد. اصولاً زمان توقف در ایستگاه‌ها تابع نرخ رسیدن مسافران به ایستگاه و نوع خدمت‌رسانی در آن ایستگاه است، و اثر قابل‌توجهی بر زمان سفر اتوبوس ندارد. بر اساس یافته‌های مدل، افزایش میزان بارش باران، زمان سفر اتوبوس را می‌افزاید. افزایش لغزندگی خیابان به سبب بارندگی، افزایش زمان سفر را توجیه می‌کند. هرچه میزان بارش بیش‌تر باشد، اثر آن بر زمان سفر اتوبوس بیش‌تر است. این اثر در ماه آبان (به دلیل بارندگی بیش‌تر از سایر ماه‌ها) کاملاً دیده می‌شود. تنها متغیری که زمان سفر اتوبوس را می‌کاهد، خط ویژه مسیر اتوبوس است. قابلیت این خط برای کاهش زمان سفر اتوبوس در اوج صبح حدود 20 درصد است؛ اما این اثر در روز کاهش می‌یابد و به حدود 11 درصد می‌رسد.



شکل ۵: سهم متغیرها در زمان سفر اتوبوس در سه بازه اوج صبح، میانه روز، و اوج عصر در سه ماه گوناگون از سال ۱۳۹۴ در خط میدان سپاه-میدان انقلاب

بحث و نتیجه گیری

بیشتر پژوهش‌های پیشین در رویکرد تخمین زمان سفر اتوبوس برای شناسایی عوامل اثرگذار بر زمان سفر آن، مدل رگرسیون خطی را در کار خود بکار می‌برند. فرض اصلی این مدل آن است که مشاهدات از توزیع نرمال پیروی می‌کنند. در این پژوهش، از مدل‌های بقا برای تخمین زمان سفر اتوبوس استفاده

می‌شود. این مدل‌ها توزیع‌های جزء خطای متفاوتی را در نظر می‌گیرند که با توجه به توانایی هر نوع از مدل‌های بقا در توضیح مشاهده‌ها، جزء اخلاص آن برای مشاهده‌ها پذیرفته می‌شود، بنابراین، می‌توان توزیع زمان سفر اتوبوس‌ها را نیز بررسی کرد. این هدف در قالب دو نوع مدلسازی کمان‌پایه و بخش‌پایه بررسی می‌شود. کمان عبارت است از فاصله دو ایستگاه متوالی و بخش عبارت است از فاصله بین دو ایستگاه که می‌توانند متوالی نباشند. بر اساس یافته‌ها، در هر دو نوع مدلسازی کمان‌پایه و بخش‌پایه، دقت مدل‌های بقا از مدل رگرسیون خطی بیش‌تر است. در نوع بخش‌پایه، اگرچه اندازه شاخص‌های مجذور میانگین مربع‌های خطا (RMSE)، و میانگین مطلق خطا (MAE)، مدل رگرسیون خطی بهتر از مدل لاگ‌لجستیک بقا است، اما ضرایب خط برازش برآورد-مشاهده مدل بقا بسیار بهتر از مدل رگرسیون خطی است. باید توجه داشت که اصولاً تابع هدف مدل‌های رگرسیون کمینه کردن مقادیر RMSE و MAE است، در حالی که تابع هدف مدل‌های بقا و همه مدل‌هایی که از روش برآورد درست‌نمایی بیشینه (MLE)، برای برآورد پارامترهای خود استفاده می‌کنند، بیشینه کردن تابع درست‌نمایی (Likelihood) است.

همچنین، دقت مدل‌های ساخته‌شده نوع کمان‌پایه مناسب نیست و بسیار کم‌تر از نوع بخش‌پایه است. اصولاً زمان سفر در کمان‌ها دارای نوسان زیادی است و به تقاطع‌های مسیر به‌شدت حساس است، اما با در نظر گرفتن بخش‌ها، می‌توان از این حساسیت کاست. از سوی دیگر، دقت مدل لاگ‌لجستیک بقا (که به عنوان مدل برتر بقا برای توضیح مشاهده‌ها انتخاب می‌شود)، نسبت به مدل رگرسیون خطی بالاتر است. بنابراین، می‌توان نتیجه گرفت که در نظر گرفتن فرض توزیع نرمال برای مشاهده‌های زمان سفر اتوبوس می‌تواند نادرست باشد.

بیش‌ترین سهم هر یک از متغیرها در تشکیل زمان سفر اتوبوس، متعلق به ضریب ثابت مدل است که به‌طور میانگین حدود ۴۰/۲ درصد از زمان سفر اتوبوس را تشکیل می‌دهد. با توجه به لحاظ نشدن متغیر حجم ترافیک در مدل، و نیز روند تغییرهای سهم سایر متغیرها در طول روز می‌توان عنوان کرد که بخش عمده‌ای از این سهم را متغیر حجم ترافیک تشکیل می‌دهد. افزایش حجم ترافیک پس از پایان ساعات طرح ترافیک، بر زمان سفر اتوبوس‌ها به‌شدت اثرگذار است. این افزایش از راه تقاطع‌های مسیر بر حرکت اتوبوس اثر می‌گذارد. با محدود کردن حرکت خودروهای شخصی در تقاطع‌ها می‌توان اثرگذاری این متغیر را کنترل کرد. دو متغیر مسافت و تعداد ایستگاه‌ها در مجموع حدود ۴۳/۳ درصد از زمان سفر اتوبوس را تشکیل می‌دهند. جنس این متغیرها اصولاً یکسان است. متغیرهای چراغ راهنمایی و سهراهی که تقاطع‌های مسیر را شامل می‌شوند، در مجموع حدود ۲۴/۷ درصد از زمان سفر اتوبوس را تشکیل می‌دهند. متغیرهای سرفاصله زمانی، بارش باران، و زمان توقف در ایستگاه‌های بین راه، روی هم‌رفته

حدود ۷/۵ درصد از زمان سفر اتوبوس را تشکیل می‌دهند. سهم عمده بارش باران در ماه آبان مشاهده می‌شود که میزان بارش قابل توجه است. زمان توقف نیز کم‌ترین سهم را از زمان سفر اتوبوس در بین متغیرها دارد. خط ویژه خلاف جهت به‌طور میانگین حدود ۱۵/۷ درصد از زمان سفر اتوبوس را می‌کاهد. سهم این متغیر در اوج صبح حدود ۲۰ درصد است، اما در طول روز کاهش می‌یابد و در اوج عصر به حدود ۱۱ درصد می‌رسد. افزایش حجم ترافیک در اوج عصر، به‌ویژه در تقاطع‌ها و نیز تردد سایر وسایل نقلیه از جمله سواری‌ها و موتورسیکلت‌ها در خط ویژه، از جمله دلایل این کاهش اثر هستند.

این پژوهش چند سهم مهم در ادبیات دارد. مدل‌های بقا اصولاً برای مطالعه داده‌های مدت‌زمانی هستند، اما در این پژوهش برای مطالعه تخمین زمان سفر اتوبوس بکار می‌روند. توسعه مدل‌های تخمین زمان سفر اتوبوس با استفاده از مدل‌های بقا، نخستین سهم این پژوهش است. همچنین، در این پژوهش دقت مدل‌های رگرسیون خطی (به عنوان مدل غالب)، و بقا (به عنوان یک مدل جدید)، در حوزه تخمین مشاهده‌های زمان سفر اتوبوس مقایسه می‌شود. سهم دیگر این پژوهش، بررسی توزیع مشاهده‌های زمان سفر اتوبوس‌هاست. با بکارگیری توزیع‌های وایبل، گامای تعمیم‌یافته، لاگ‌نرمال، و لاگ‌لجستیک در کنار استفاده از تعداد ۱۱۱۷۰ مشاهده برای کمان‌ها و ۲۹۲۵۰ مشاهده برای بخش‌ها، قیاس کاملی بین توزیع‌های مختلف زمان سفر اتوبوس‌ها انجام شد.

از جمله پیشنهادهای اجرایی این پژوهش شامل این موارد می‌شود: بکارگیری مدل‌های بقا برای تخمین زمان سفر اتوبوس‌ها در شهر تهران و اصلاح برنامه زمان‌بندی فعلی حرکت آن‌ها، به‌ویژه برای فواصل با زمان سفر کوتاه (کم‌تر از ۱۰۰ ثانیه) و بلند (بیش‌تر از ۹۰۰ ثانیه)؛ توسعه و ایجاد خطوط ویژه اتوبوس در کنار کنترل تقاطع‌های آن با سایر خطوط، به‌ویژه در ساعات اوج ترافیک، برای افزایش قابلیت اطمینان زمان سفر اتوبوس‌ها؛ و کنترل حرکت سایر خودروها و موتورسیکلت‌ها در خطوط ویژه اتوبوس. در پایان چند پیشنهاد پژوهشی برای پر کردن محدودیت‌های پژوهش بیان می‌شود: در نظر گرفتن یک خط اتوبوسرانی که بخشی از آن به شکل خط ویژه (همسو یا خلاف جهت سایر خودروها)، و بخش دیگر به شکل مختلط با ترافیک باشد، تا قیاس بهتری از اثر خطوط ویژه بر حرکت اتوبوس‌ها در یک خط اتوبوسرانی انجام شود (در پژوهش حاضر داده‌های کمی از ایستگاه‌های شامل مسیر مختلط با ترافیک در دسترس بودند که به‌ناچار مشاهده‌های زمان سفر آن ایستگاه‌ها از مدل‌ها حذف شدند)؛ مطالعه زمان سفر اتوبوس در مسیر ویژه (مانند خطوط BRT)، خطوط ویژه (هم‌جهت و خلاف جهت ترافیک)، و خطوط مختلط با ترافیک، سپس قیاس زمان‌ها و سهم متغیرهای اثرگذار در هر یک از این سه مسیر (در پژوهش حاضر به دلیل کمبود زمان و حجم بالای داده‌های هر خط، این کار میسر نشد)؛ در نظر گرفتن حجم

ترافیک مسیر در مدل و محاسبه اثر این متغیر بر زمان سفر اتوبوس (داده‌های پژوهش حاضر مربوط به سال ۱۳۹۴ بودند و حجم ترافیک آن در دسترس نبود)؛ بررسی امکان استفاده از داده‌های سامانه‌های گردآوری خودکار کرایه (AFC)، یا شمارشگر خودکار مسافران (APC)، و موبایل در صورت وجود، برای مدلسازی زمان توقف اتوبوس‌ها در ایستگاه‌ها یا مدلسازی همزمان زمان سفر اتوبوس و زمان توقف آن در ایستگاه‌ها (در پژوهش حاضر این داده‌ها در دسترس نبودند)؛ و استفاده از مدل‌های ترکیبی بقا و سایر مدل‌هایی که برای تخمین و پیش‌بینی زمان سفر اتوبوس بکار می‌روند (مانند فیلترینگ کالمن و شبکه‌های عصبی مصنوعی)، و قیاس دقت آن‌ها با مدل‌های ساده بقا.

منابع

الف) فارسی

- گنجی، محمدرضا؛ ارباب پوربیدگلی، محمد؛ پورابراهیمی، محمدرضا، و گلرو، امیر (۱۳۹۵). مدلسازی زمان سفر برای سامانه اتوبوس تندرو به روش جنگل تصادفی (مطالعه موردی: شهر تهران). شانزدهمین کنفرانس بین‌المللی مهندسی حمل‌ونقل و ترافیک.
- هاشمی، حسین، و البدوی، امیر (۱۳۹۸). پیش‌بینی زمان ورود اتوبوس به ایستگاه با استفاده از داده‌های AVL: مطالعه موردی سیستم اتوبوسرانی شهر قزوین. فصلنامه مهندسی حمل‌ونقل، ۱۰(۳)، ۵۴۲-۵۳۱.

الف) انگلیسی

- Alfa, A. S., Menzies, W. B., Purcha, J., & Mcpherson, R. (1988). A Regression Model for Bus Running Times in Suburban Areas of Winnipeg. *Journal of Advanced Transportation*, 21(3), 227-237.
- Allison, P. D. (2014). *Event History and Survival Analysis: Regression for Longitudinal Event Data* (Vol. 46): SAGE Publications.
- Brownstone, D., & Small, K. A. (2005). Valuing Time and Reliability: Assessing the Evidence from Road Pricing Demonstrations. *Transportation Research Part A: Policy and Practice*, 39(4), 279-293.
- Cathey, F., & Dailey, D. J. (2003). A Prescription for Transit Arrival/Departure Prediction Using Automatic Vehicle Location Data. *Transportation Research Part C: Emerging Technologies*, 11(3-4), 241-264.
- Chen, G., Yang, X., An, J., & Zhang, D. (2012). Bus-Arrival-Time Prediction Models: Link-Based and Section-Based. *Journal of Transportation Engineering*, 138(1), 60-66.
- Chen, M., Liu, X., Xia, J., & Chien, S. I. (2004). A Dynamic Bus-Arrival Time Prediction Model Based on APC Data. *Computer-Aided Civil and Infrastructure Engineering*, 19(5), 364-376.

- Chien, S. I.-J., Ding, Y., & Wei, C. (2002). Dynamic Bus Arrival Time Prediction with Artificial Neural Networks. *Journal of Transportation Engineering*, 128(5), 429-438.
- Gayah, V. V., Yu, Z., Wood, J. S., & Consortium, M. N. T. R. (2016). *Estimating Uncertainty of Bus Arrival Times and Passenger Occupancies*. Mineta Transportation Institute Publications.
- Islam, M. K., Vandebona, U., Dixit, V. V., & Sharma, A. (2015). A Model to Evaluate the Impact of Headway Variation and Vehicle Size on the Reliability of Public Transit. *IEEE Transactions on Intelligent Transportation Systems*, 16(4), 1840-1850.
- Jenkins, S. P. (2005). Survival Analysis. *Unpublished Manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK*, 42(1), 54-56.
- Jeong, R., & Rilett, L. R. (2005). Prediction Model of Bus Arrival Time for Real-Time Applications. *Transportation Research Record*, 1927(1), 195-204.
- Kennedy, P. (2003). *A Guide to Econometrics*: MIT Press.
- Kieu, L. M., Bhaskar, A., & Chung, E. (2012). *Benefits and Issues for Bus Travel Time Estimation and Prediction*. Paper Presented at the Australasian Transport Research Forum 2011 Proceedings.
- Levinson, H. S. (1983). Analyzing Transit Travel Time Performance. *Transportation Research Record*, 915(1), 1-6.
- McKnight, C., Paaswell, R., Ali, M., Kamga, C., & Cruz, W. (1997). Impact of Congestion on New York Bus Service. New York City Transit, New York.
- Patnaik, J., Chien, S., & Bladikas, A. (2004). Estimation of Bus Arrival Times Using APC Data. *Journal of Public Transportation*, 7(1), 1-8.
- Rahman, M. M., Wirasinghe, S., & Kattan, L. (2018). Analysis of Bus Travel Time Distributions for Varying Horizons and Real-Time Applications. *Transportation Research Part C: Emerging Technologies*, 86(1), 453-466.
- Shalaby, A., & Farhan, A. (2004). Prediction Model of Bus Arrival and Departure Times Using AVL and APC Data. *Journal of Public Transportation*, 7(1), 3-10.
- Vanajakshi, L., Subramanian, S. C., & Sivanandan, R. (2009). Travel Time Prediction under Heterogeneous Traffic Conditions Using Global Positioning System Data from Buses. *IET Intelligent Transport Systems*, 3(1), 1-9.
- Washington, S., Karlaftis, M. G., Mannering, F., & Anastasopoulos, P. (2020). *Statistical and Econometric Methods for Transportation Data Analysis*: CRC press.
- Yu, B., Wang, H., Shan, W., & Yao, B. (2018). Prediction of Bus Travel Time Using Random Forests Based on Near Neighbors. *Computer-Aided Civil and Infrastructure Engineering*, 33(4), 333-350.
- Yu, Z., Wood, J. S., & Gayah, V. V. (2017). Using Survival Models to Estimate Bus Travel Times and Associated Uncertainties. *Transportation Research Part C: Emerging Technologies*, 74(1), 366-382.