

پیش‌بینی تعداد مسافران قطار سبک شهری مشهد به روش سری‌های زمانی فصلی

نوید کلانتری^۱، رامین خاورزاده^۲، امیر قیاسی^۳

۱- دکترای برنامه ریزی حمل‌ونقل، گروه پژوهشی حمل‌ونقل شاران

۲- دانشجوی دکترای آمار کاربردی، گروه پژوهشی حمل‌ونقل شاران

۳- کارشناس ارشد برنامه ریزی حمل‌ونقل، گروه پژوهشی حمل‌ونقل شاران

چکیده

امروزه تمرکز بیشتر در برنامه ریزی حمل و نقل بر روی برنامه زیری های پویا و در نتیجه آنها مدلسازی های کوتاه مدت و پویای ترافیکی است. مدل سازی و پیش‌بینی تعداد مسافران قطار شهری یکی از کارهای اساسی برای رسیدن به مدیریت بهینه و پویا در قطار شهری می‌باشد. یکی از راه‌های پیش‌بینی تعداد مسافران قطار شهری استفاده از تکنیک‌های سری‌های زمانی است. هدف از این پژوهش بررسی کارایی تکنیک سری‌های زمانی در پیش‌بینی تعداد مسافران می‌باشد. به این منظور از داده‌های تعداد مسافران قطار شهری، شهر مشهد از بدو تاسیس تا ابتدای سال ۱۳۹۲ به صورت روزانه استفاده شد. روند تعداد مسافران در طول این زمان برآورد شد و همچنین اثر روزهای هفته و روزهای خاص پرسفر و کم سفر و همچنین روزهای تعطیل محاسبه شد و در نهایت با استفاده از مدل فصلی ساریما مدلی برای برآورد تعداد مسافران ساخته شده است. در نهایت این مدل با پیش‌بینی تعداد مسافران به روش شبکه عصبی، مورد مقایسه قرار گرفته است.

کلید واژه: سری‌های زمانی، مدل‌های پیش‌بینی، همبستگی جزئی، سری‌های زمانی فصلی، شبکه عصبی.

۱ گروه پژوهشی حمل‌ونقل شاران، ایمیل: kalantari@iust.ac.ir

۲ گروه پژوهشی حمل‌ونقل شاران، ایمیل: r.khavarzade@modares.ac.ir

۳ گروه پژوهشی حمل‌ونقل شاران، ایمیل: amir_ghiasi@alum.sharif.edu



۱- مقدمه

در دهه‌های گذشته با پیشرفت علوم ریاضی و کامپیوتری، برنامه‌ریزان حمل‌ونقل سعی در پیش‌بینی پدیده‌های مربوط به مهندسی حمل‌ونقل و ترافیک داشته‌اند. از جمله این موارد، مدل‌های بیشمار پیش‌بینی حجم شبکه‌های حمل‌ونقل است که از روش‌های مختلف بهره‌جسته و هر یک از آن‌ها دارای مزایا و معایبی هستند. یکی از روش‌های آماری که به بررسی ارتباط زمانی پدیده‌های مختلف می‌پردازد، سری‌های زمانی است. مدل‌هایی که با استفاده از سری‌های زمانی به پیش‌بینی وقایع مختلف می‌پردازند، در واقع از اطلاعات وقایع مشابه پیشین به منظور پیش‌بینی وقایع آینده استفاده می‌کنند. گادری^۱ در سال ۱۹۷۵ با استفاده از روش‌های سری زمانی و داده‌های ماهانه در بازه سال-های ۱۹۵۶ تا ۱۹۷۱، تقاضای سیستم حمل‌ونقل همگانی شهر مونترال کانادا را پیش‌بینی می‌کند. متغیرهای در نظر گرفته شده در این مطالعه متغیرهای وابسته به مدهای حمل‌ونقل (همگانی و شخصی) و ویژگی‌های اقتصادی-اجتماعی ساکنان این شهر بوده است [۱]. در مطالعه‌ای که در سال ۱۹۸۸ توسط کایت و همکاران^۲ ارائه شد، پارامترهای تاثیرگذار در تعداد مسافران سیستم حمل‌ونقل همگانی شهر پرتلند آمریکا در بازه سال‌های ۱۹۷۱ تا ۱۹۸۲ بررسی شد. از جمله این پارامترها، متغیرهای مربوط به سطح سرویس و هزینه‌های سفر بوده است [۲]. در سال ۲۰۱۱، چیانگ و همکاران^۳ از روش‌های سری زمانی به منظور پیش‌بینی و بررسی عوامل تاثیرگذار در سیستم حمل‌و-نقل همگانی شهر تالسای^۴ آمریکا استفاده کرده‌اند. محققان مذکور با شناسایی عوامل تاثیرگذار و آزمایش آن‌ها با مدل‌های رگرسیونی، شبکه‌های عصبی و آریمای^۵ به این نتیجه رسیدند که استفاده ترکیبی از مدل‌های متنوع، دقت بالاتری نسبت به استفاده از هر یک از مدل‌ها به صورت مجزا، به ارمغان می‌آورد [۳].

دیویس و نیهان^۶ روش‌های سری زمانی را با یک روش رگرسیون بدون پارامتر مقایسه کرده‌اند. نتیجه مقایسه این محققان، برتری نسبی روش رگرسیون اشاره شده نسبت به روش‌های سری زمانی بوده است [۴]. در این راستا، تلاش‌های مشابه دیگری به منظور مقایسه روش‌های سری زمانی با روش‌های شبکه عصبی انجام شده است [۵، ۶]. یکی از معروف‌ترین روش‌های سری زمانی، آریمایست. حامد و

¹ Gaudry

² Kyte et.al

³ Chiang et.al

⁴ Tulsa

⁵ Autoregressive Integrated Moving Average (ARIMA)

⁶ Davis and Nihan



الماسعید^۱ از یک مدل ARIMA(0,1,1) برای پیش‌بینی یک دقیقه آینده جریان ترافیک استفاده کردند [۷]. همچنین، ویلیامز و هوئل^۲ از مدل آریمای بدون روند هفتگی برای پیش‌بینی جریان ترافیک بهره بردند [۸].

در این مطالعه با استفاده از یکی از مدل‌های سری زمانی، مدلی ارائه می‌گردد که سعی در پیش‌بینی تعداد مسافران قطار شهری مشهد دارد. داده‌های استفاده شده در این مقاله مربوط به گزارش داده‌های تعداد مسافران قطار شهری شهر مشهد است که به صورت روزانه از بدو تاسیس تا ابتدای سال ۱۳۹۲ گزارش شده است. در انتها، مدل دیگری با استفاده از روش شبکه عصبی ارائه شده که پیش‌بینی حاصل از آن با پیش‌بینی صورت گرفته توسط روش سری زمانی مذکور مقایسه و نتایج آن بیان شده است.

۲- سری‌های زمانی

سری زمانی مجموعه مشاهداتی است که بر حسب زمان یا مرتبه تکرار، مرتب شده‌اند. تحلیل سری‌های زمانی به‌طور نظری و عملی از سال ۱۹۷۰ برای پیش‌بینی و کنترل، به سرعت توسعه پیدا کرده است. این نوع از تحلیل معمولاً در مورد داده‌هایی به کار برده می‌شود که مستقل نبوده و به‌طور متوالی به هم وابسته‌اند. در واقع، همین وابستگی بین مشاهدات متوالی است که مورد توجه قرار گرفته و سبب کاربرد تحلیل سری‌های زمانی در پیش‌بینی آن‌ها می‌شود. توابع سری‌های زمانی برای یک مجموعه از داده که به‌طور متوالی به دست آمده‌اند به صورت زیر نمایش داده می‌شود:

$$y_t = \alpha_0 + \alpha_1 T^1 + \alpha_2 T^2 + \dots + \alpha_n T^n \quad (1)$$

که در آن متغیر وابسته y_t مشاهدات حاصل از سری‌های زمانی، α ها ضرایب توابع خودهمبستگی و T متغیر مستقل زمانی می‌باشد.

از طرف دیگر، خانواده‌ی اصلی الگوهای سری زمانی یعنی الگوی آریما که با مرتبه‌های p (اتورگرسیون)، d (درجه تمایز) و q (ترتیب میانگین‌های متحرک) تعیین می‌شود، در الگوسازی از وضعیت‌های دنیای واقعی از اهمیت زیادی برخوردار است. در بسیاری از زمینه‌هایی که سری زمانی مورد استفاده قرار می‌گیرد، گرایش‌های دوره‌ای معمول است. در چنین مواردی جهت تحلیل داده‌ها از مدل‌های فصلی ضربی $ARIMA(p,d,q)(P,D,Q)$ استفاده می‌شود که در آن p,q,d مرتبه‌های غیر فصلی و P,Q,D مرتبه‌های فصلی است. در تحلیل سری‌های زمانی، شناسایی الگو، یعنی تعیین

¹ Hamed and Al-Masaeid

² Williams and Hoel



مقادیر مرتبه‌های فصلی و غیر فصلی، بر اساس رفتار نمودارهای ACF^1 و $PACF^2$ صورت می‌گیرد [۹]- [۱۲].

با توجه به نمودارهای ACF و $PACF$ مقادیر p و q تعیین می‌شود و پس از آن می‌توان الگوی فصلی را مشخص نمود. باقیمانده‌های مدل و داده‌های واقعی باید این خصوصیات را دارا باشند که میانگین آن‌ها صفر باشد، واریانس ثابت باشد، برای همه تاخیرها همبستگی وجود نداشته باشد و توزیعشان نرمال باشد. لذا، مدل باید آنقدر تغییر کند تا به مدل مطلوب برسیم [۱۴]. در صورتی که الگوی فصلی بیز موجود باشد، مدل بصورت $SARIMA(p,d,q)(P,D,Q)T$ در می‌آید که در آن T معرف الگوی فصلی است. برای ساختن مدل‌های Box-Jenkins مراحل باید طی شود که شامل: شناسایی، تخمین، چک کردن و در نهایت پیش‌بینی است. این مراحل آنقدر تکرار می‌شود تا نهایتاً بتوان به بهترین مدل رسید [۱۰]. مدل‌های آریمای علاوه بر پیش‌بینی، دامنه‌ی اطمینان هم برای پیش‌بینی ارائه می‌دهند [۱۴].

برای تعیین مناسب بودن^۳ مدل سری زمانی از معیار AIC^4 استفاده می‌شود. ضریب AIC یک معیار مستقل و عینی برای انتخاب مناسب‌ترین مدل است. در این روش، مناسب بودن مدل بر اساس مجموع مربعات باقیمانده صورت می‌گیرد و با توجه به تعداد پارامترهای مستقل کاهش می‌یابد [۱۳]. اگرچه با افزایش تعداد پارامترهای مستقل، مدل انعطاف‌پذیرتری حاصل می‌شود، اما با توجه به آنکه انتخاب مدل ساده‌تر، همواره منطقی‌تر است، این موضوع عموماً مورد قبول است که بهترین مدل مدلی است که علاوه بر ارائه پیش‌بینی مناسب، از تعداد کمتری پارامتر استفاده کرده باشد. یکی از راه‌های رسیدن به چنین مدلی استفاده از معیار AIC است [۱۵]. در این مطالعه، برای رسیدن به مدل نهایی از مقادیر ACF و $PACF$ در سطح معنی داری ۰,۰۵ مدل‌های مناسب اولیه شناسایی شده و پس از ساختن مدل، با ضریب AIC بهترین مدل انتخاب شد.

۳- برآورد تعداد مسافران قطار شهری

از آنجایی که یکی از اهداف اصلی در تحلیل داده‌های مسافران قطار شهری مشهد، سنجش روند و عوامل تاثیرگذار بر تعداد مسافران می‌باشد، در گام اول با بررسی روند داده‌ها و عوامل مختلف تاثیرگذار، اثر هر کدام از سری داده‌ها کم شده است. با توجه به شکل تعداد مسافران روزانه قطار شهری،

¹ Autocorrelation Function

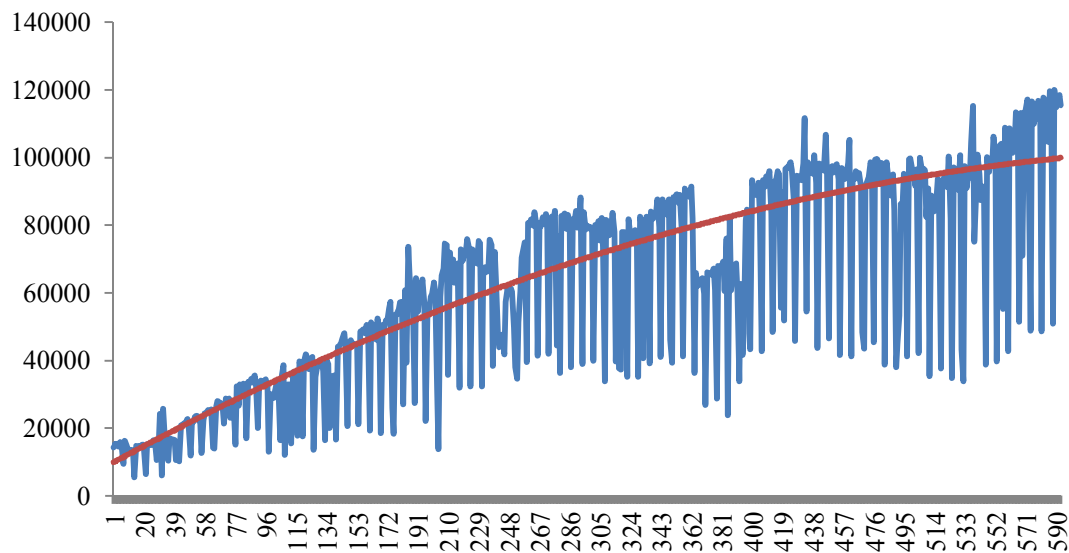
² Partial Autocorrelation Function

³ Goodness Of Fit

⁴ Akaike Information Criterion



می توان روندی درجه دوم برای روند صعودی تعداد مسافران در نظر گرفت. مدل برآورد شده در زیر آمده است:



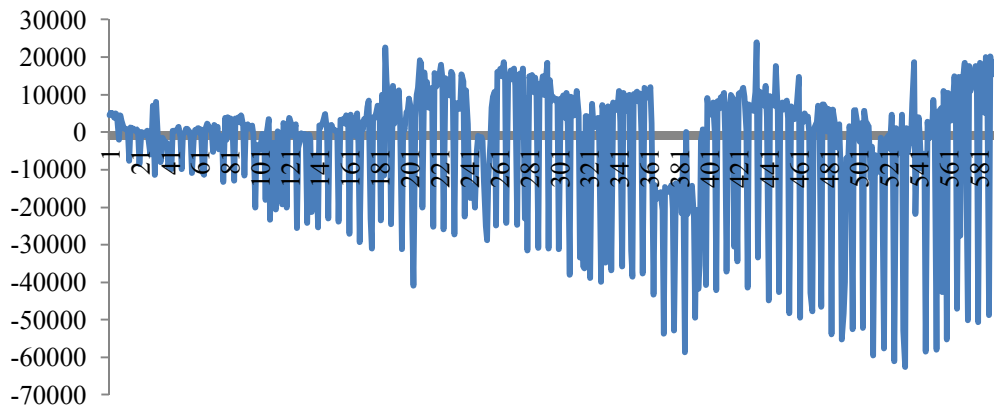
شکل ۱: برآورد روند تعداد مسافران قطار شهری

$$C = 9642 + 257.1 \times T - 0.177 \times (T)^2 \quad (2)$$

که در آن، C تعداد مسافران قطار شهری در هر روز و T روزهای سپری شده از بدو تاسیس قطار شهری است. در این مرحله تعداد مسافران برآورد شده بر اساس مدل روند برای هر یک از روزها محاسبه شده و در هر روز تفاضل مقدار واقعی مسافران با تعداد برآوردی بر اساس مدل روند، به عنوان داده‌های بدون روند (روندزوده^۱) در نظر گرفته می‌شود.

¹ No trend



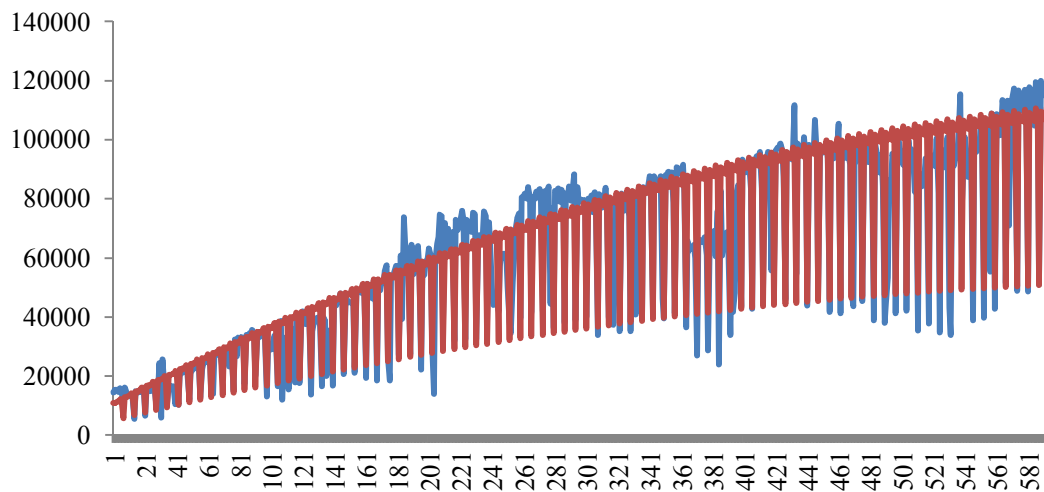


شکل ۲: داده‌های بدون روند تعداد مسافران قطار شهری

همانگونه که در شکل ۲ مشخص است، تعداد مسافران در یک دوره زمانی ۷ روزه از الگوی خاصی پیروی می‌کند. در مرحله سوم با برآورد اثر هر یک از روزهای هفته الگوی هفتگی تعداد مسافران را برآورد و سپس این الگوی هفتگی از داده‌ها حذف می‌شود.

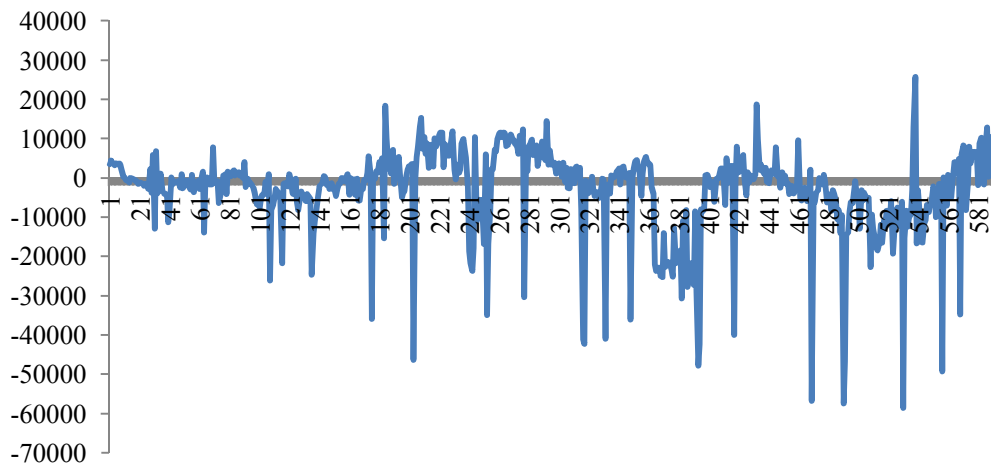
جدول ۱: ضریب برآورد شده برای هر یک از روزهای هفته

ضریب برآورد شده	روزهای هفته
۱/۰۹۹	شنبه
۱/۰۷۲۸	یکشنبه
۱/۸۱۳۹	دوشنبه
۱/۰۶۹۲۱	سه شنبه
۱/۱۱۰۷۱	چهارشنبه
۱/۰۵۷۷۳	پنج شنبه
۱/۵۰۹۱۶	جمعه و روزهای تعطیل



شکل ۳: برآورد هم زمان اثرات روزهای هفته و ایام تعطیل و روند بر تعداد مسافران قطار شهری

در شکل ۴ داده های روند زدوده که میزان اثر هر یک از روزها از آن حذف شده، نمایش داده شده است.



شکل ۴: داده های روند زدوده

با توجه به نمودار داده های بدون روند و بدون اثر هفتگی که در شکل ۴ نمایش داده شده است، می توان این گونه برداشت کرد که در روزهای خاصی از سال تعداد مسافران بیشتر و یا کمتر از حد انتظار است. بر این اساس، دو دسته ی روزهای پر سفر و روزهای کم سفر در نظر گرفته شده است، که در جدول ۲ و ۳ نشان داده شده اند.

جدول ۲: روزهای پر سفر

مناسبت	تاریخ هجری	تاریخ قمری
وفات حضرت معصومه	۳ اسفند ۱۳۹۱ و ۱۳ اسفند ۱۳۹۰	۱۰ ربیع الثانی
پیروزی انقلاب اسلامی	۲۲ بهمن	
ولادت حضرت رسول اکرم	۵ بهمن ۱۳۹۱ و ۱۶ بهمن ۱۳۹۰	۱۲ ربیع الاول
رحلت حضرت رسول اکرم	۲۲ دی ۱۳۹۱ و ۲ بهمن ۱۳۹۰	۲۸ صفر
شهادت امام رضا	۲۳ دی ۱۳۹۱ و ۴ بهمن ۱۳۹۰	۲۹ صفر و یا ۳۰ صفر
تاسوعا و عاشورای حسینی	۴-۵ آذر ۱۳۹۱ و ۱۴-۱۵ آذر ۱۳۹۰	۹-۱۰ محرم
روز عرفه	۴ آبان ۱۳۹۱ و ۱۵ آبان ۱۳۹۰	۹ ذی الحجه
ولادت امام رضا وروز قبلش	۶-۷ مهر ۱۳۹۱ و ۱۶-۱۷ مهر ۱۳۹۰	۱۰-۱۱ ذی القعدة
۱۰ روز اول ماه مهر	۵-۱ مهر	
ولادت حضرت معصومه	۲۸ شهریور ۱۳۹۱ و ۷ مهر ۱۳۹۰	۱ ذی القعدة
عید فطر	۲۹ مرداد ۱۳۹۱ و ۹ شهریور ۱۳۹۰	۱ شوال
ولادت حضرت قائم	۱۵ تیر ۱۳۹۱ و ۲۶ تیر ۱۳۹۰	۱۵ شعبان
مبعث رسول اکرم	۲۹ خرداد ۱۳۹۱ و ۹ تیر ۱۳۹۰	۲۷ رجب
ولادت حضرت فاطمه زهرا	۲۳ اردیبهشت ۱۳۹۱ و ۳ خرداد ۱۳۹۰	۲۰ جمادی الثانی
روز طبیعت	۱۳ فروردین	
عید نوروز	۴-۱ فروردین	
اواخر اسفند ماه	۲۹-۲۰ اسفند	

جدول ۳: روزهای کم سفر

مناسبت	تاریخ هجری	تاریخ قمری
ماه مبارک رمضان	۲۸-۱ مرداد و ۳۱ تیر ۱۳۹۱	۲۹-۱ رمضان
روزهای غیر تعطیل نوروز	۵-۱۲ فروردین	
شب یلدا	۳۰ آذر	

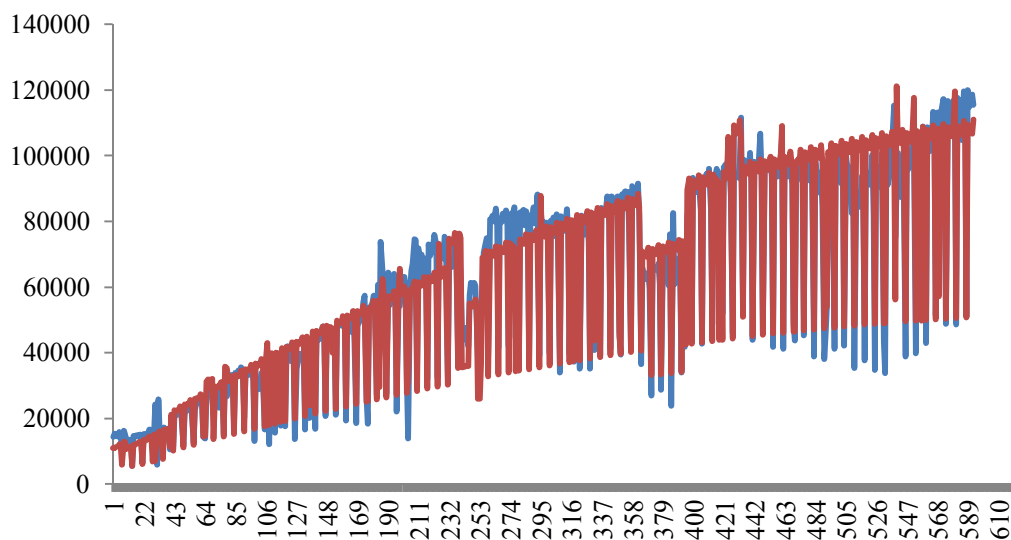
بعد از مشخص نمودن روزهای پرسفر و کم سفر، ضرایب برآورد شده برای این دو دسته در جدول ۴ نشان داده شده است:



جدول ۴: برآورد پارامترهای مدل

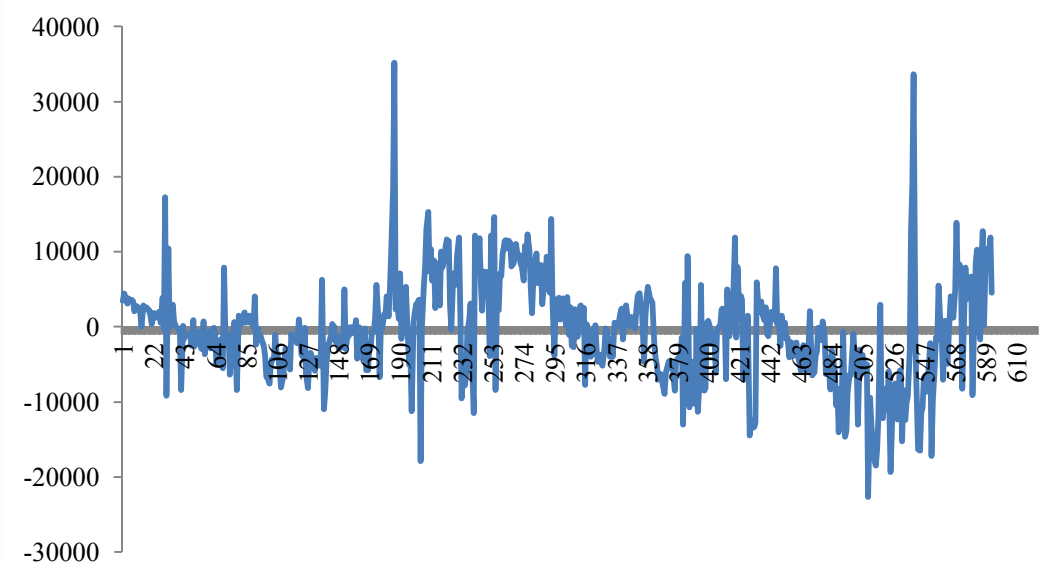
پارامترها	برآورد	انحراف معیار برآورد
پر سفر	۱/۱۳۸۸۲۸	۰/۰۱۶
کم سفر	۰/۸۰۶۵۴۱	۰/۰۱۹

با کمک این ضرایب و همچنین مدل روند و ضرایب روزهای هفته و روزهای تعطیل، برآورد مدل کلی (نمودار قرمز رنگ) در مقایسه با مقادیر واقعی (نمودار آبی رنگ) در شکل ۵ نمایش داده شده است.



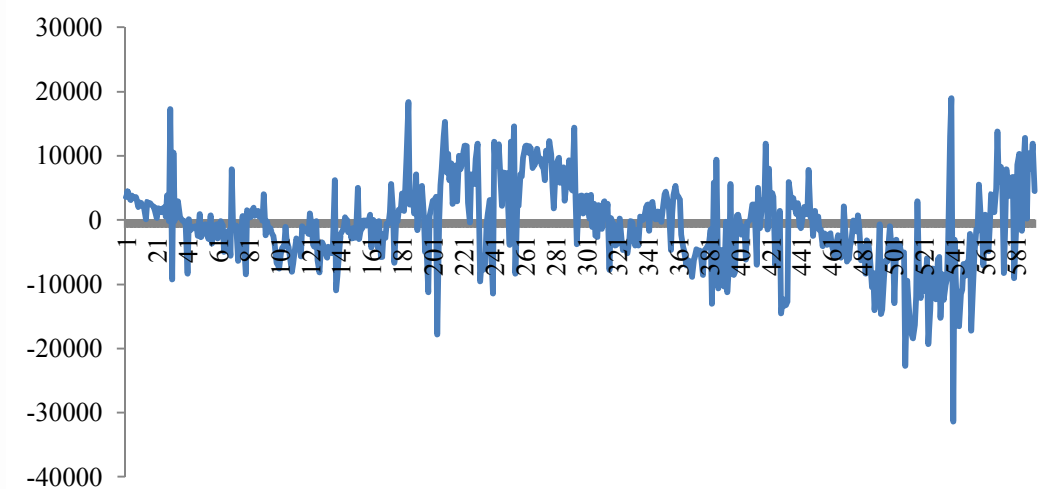
شکل ۵: برآورد هم زمان روند و اثر روزهای هفته و روزهای پرسفر و کم سفر

در شکل ۶ داده‌های روند زدوده که میزان اثر هر یک از روزها و همچنین اثر روزهای کم سفر و پر سفر از آن حذف شده نمایش داده شده است.



شکل ۶: داده‌های روند زدوده

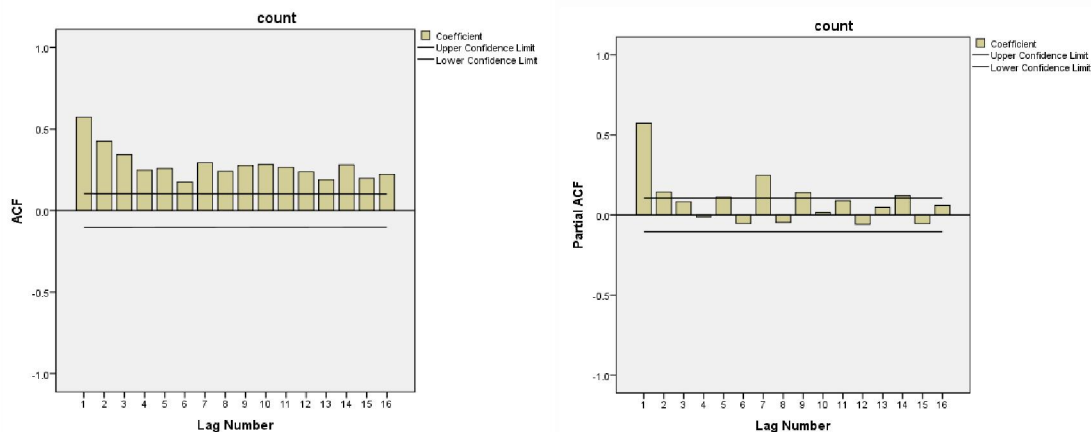
با توجه به شکل ۶، در روز شهادت امام رضا (ع) در هر دو سال، تعداد مسافران همچنان بیشتر از حد انتظار است که حتی با وجود لحاظ نمودن این روز در دسته‌ی روزهای پرسفر، همچنان مدل تعداد سفر کنندگان را به درستی برآورد نمی‌کند. یک دلیل دیگر برای این موضوع، تعطیل بودن این روز است که در مدل تعداد مسافران در این روز به علت تعطیلی کم برآورد می‌شود. برای حل این مشکل می‌توان این روز را با در نظر گرفتن به عنوان یک روز خاص، تعطیل محسوب نکرد که در این حالت اثر تعطیلی در مدل در این روز خاص تاثیر گذار نخواهد بود. در شکل ۷ مقادیر باقیمانده مشاهده می‌شود.



شکل ۷: باقیمانده‌های داده‌های روند زدوده

۴- استفاده از مدل ساریما^۱ در برآورد تعداد مسافران قطار شهری

بعد از این که روند تعداد مسافران قطار شهری و اثر هر یک از روزهای هفته و تعطیلی و روزهای خاص (پرسفر و کم سفر) از داده ها کم شد، می توان با استفاده از مدل های ساریما پیش بینی قابل قبولی از برآورد تعداد مسافران قطار شهری در آینده داشت. از این رو با محاسبه ی میزان همبستگی تعداد مسافران قطار شهری در هر روز با روزهای قبلی نمودار همبستگی ACF و همبستگی جزئی PACF رسم شده است که در شکل ۸ نشان داده شده است.



شکل ۸: همبستگی و همبستگی جزئی تعداد مسافران قطار شهری

همانطور که در شکل ۸ مشاهده می شود، تعداد باقیمانده ها در هر روز با تعداد مسافران در روز قبل از آن و دو روز قبل از آن و با ۷ روز قبل از آن همبستگی معنی داری دارد. پس بر این اساس می توان با محاسبه ی همبستگی ها و محاسبه ی مدل ساریما پیش بینی دقیقی از تعداد مسافران داشت. در جدول ۵ مقادیر برآورد شده، انحراف معیار، مقادیر، آماره ی آزمون t-student و سطح معنی داری هر کدام از پارامترهای مدل نشان داده شده است.

جدول ۵: برآورد پارامترهای مدل

پارامترهای مدل ساریما						
سطح معنی داری	آماره t	انحراف معیار	برآورد پارمتر			
۰/۰۰۰	۷/۲۰۲	۰/۰۴۱	۰/۳۹۷	Lag 1	AR	مدل
۰/۰۰۰	۴/۶۸۰	۰/۰۴۳	۰/۲۷۶	Lag 2		
۰/۰۱۰	۲/۵۹۵	۰/۰۴۱	۰/۱۷۵	Lag 7		

¹ SARIMA



برای ارزیابی نتایج، از دو معیار میانگین مربعات خطا و خطای مطلق استفاده شده است، که آماره $RMSE^1$ یا مقدار جذر میانگین مربعات خطا، بیانگر میزان خطای برآورد متغیرها نسبت به مقدار مشاهداتی است که براساس ریشه‌ی مربعات خطا مورد محاسبه قرار می‌گیرد و رابطه آن به شرح زیر است:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n [T_{es} - T_{ob}]^2} \quad (3)$$

که در آن T_{es} مقدار محاسبه شده به وسیله مدل، T_{ob} مقدار مشاهده شده و همچنین N تعداد کل مشاهدات می‌باشد. هر چه مقدار این آماره کوچک‌تر باشد به معنای این است که مدل پیش بینی مناسبی برای مقدار مشاهدات برآورد کرده است. همانگونه که در جدول ۶ نشان داده شده است مقدار $RMSE$ برابر ۱۹,۲۲۵ محاسبه شده است که با توجه به مدل‌های دیگری که برای داده‌ها ساخته شده است، بهترین مدل $SARMA(2,0,0)(7,0,0)$ خواهد بود. یکی دیگر از معیارهای که در مدل سازی به آن توجه خواصی می‌شود مقدار R^2 مدل است که نماد دیگری از دقت مدل برازش شده بر مقادیر مشاهده شده است که برای مدل ساخته شده مقدار R^2 برابر ۰,۹۰۲ حاصل شده که با توجه به اینکه این معیار، عددی بین ۰ و ۱ بوده و به ۱ نزدیکتر است، مدل ساخته شده از کیفیت مناسبی برخوردار است.

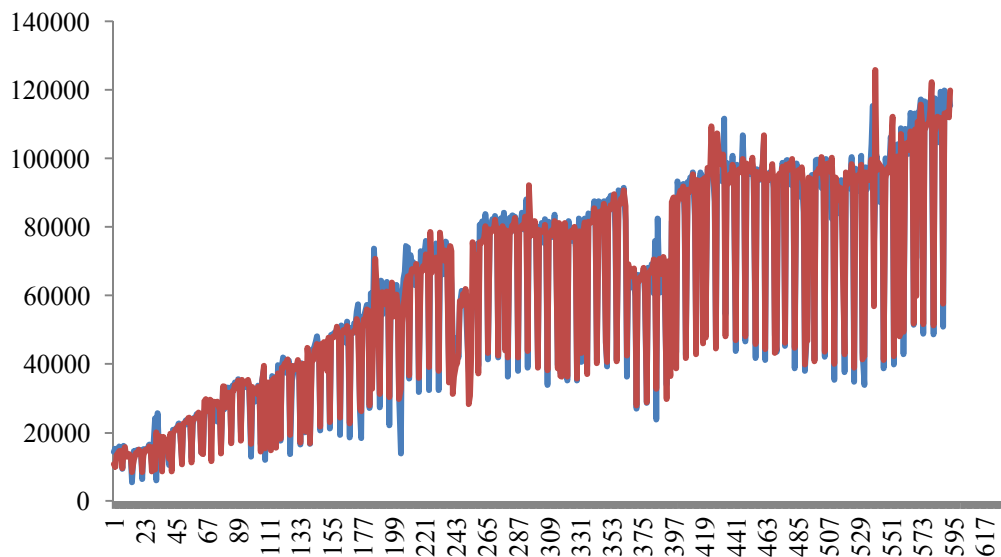
جدول ۶: برآورد پارامترهای مدل

Ljung-Box			نیکویی برازش مدل			مدل
سطح معنی داری	درجه آزادی	آماره	AIC	RMSE		
۰/۱۲۹	۱۶	۱۳/۸۲۱	۱۸/۴۰۷	۱۹/۲۲۵	۰/۹۰۲	مقادیر

در شکل ۹ تعداد مسافران واقعی قطار شهری (نمودار آبی رنگ) و تعداد مسافران حاصل از پیش‌بینی مدل (نمودار قرمز رنگ) مشاهده می‌شود. همانگونه که از نمودار مشخص است مدل از برازندگی بسیار خوبی برخوردار است.

¹ Root Mean Square Error





شکل ۹: تعداد مسافران واقعی و تعداد پیش‌بینی حاصل از مدل

در نهایت مدل کلی برای برآورد تعداد مسافران قطار شهری در رابطه ۳ نشان داده شده است.

(۴)

$$c = a * b * (9642 + 257.1 * t - 0.177 * t^2) + (0.397 * e_1 + 0.276 * e_2 + 0.175 * e_7)$$

که در آن، a ضریب روزهای هفته (جدول ۱)، b ضریب روزهای پرسفر یا کم سفر (جدول ۴)، t تعداد روزهای تاسیس قطار شهری، e_1 باقیمانده‌ی برآورد یک روز قبل، e_2 باقی‌مانده‌ی برآورد دو روز قبل، e_7 باقیمانده‌ی برآورد هفت روز قبل است.

با توجه به اینکه مدل روند به صورت مدل درجه دوم برآورد شده است، زمانی که ظرفیت برای مسافران کامل شود، مدل درجه دوم به نقطه‌ی بیشینه‌ی خود خواهد رسید. می‌توان گفت برآورد مدل برای این نقطه ۱۰۳۰۰۰ نفر است، که با توجه به این رقم با رسیدن برآورد روند مقدار روند بی‌اثر می‌شود.

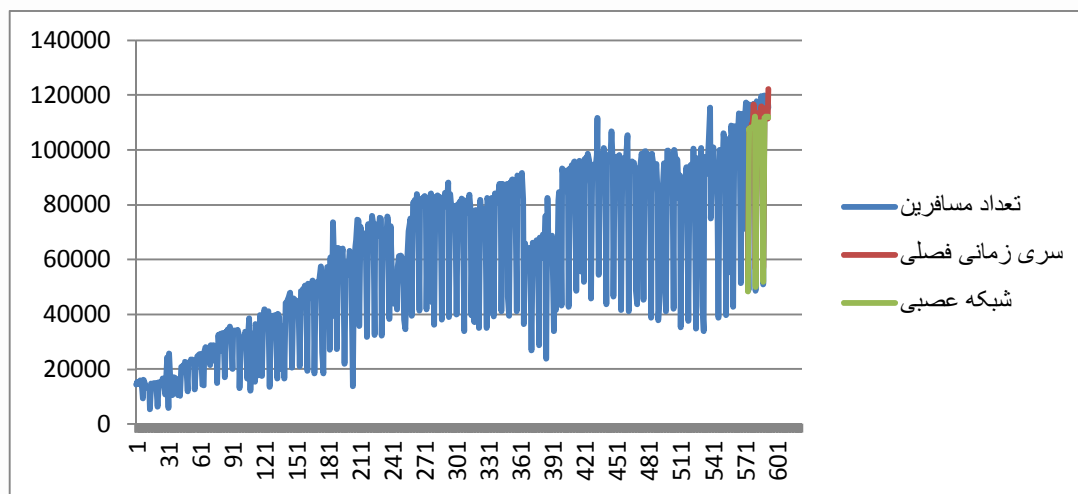
۵- ارزیابی مدل سری زمانی فصلی در مقایسه با شبکه عصبی

در این مطالعه برای پیش‌بینی تعداد مسافران قطار شهری از شبکه‌های عصبی نیز استفاده شده، تا بتوان این روش را با روش اصلی یعنی روش سری زمانی فصلی مورد مقایسه قرار داد. به این منظور از روش پس‌انتشار خطا که یک روش سیستماتیک برای آموزش شبکه‌های چندلایه است، استفاده شده است. منظور از آموزش یک شبکه انتخابی بر مبنای اطلاعات موجود، تنظیم مقادیر وزن‌ها و آریبی یا



مقادیر ثابت اولیه است. این الگوریتم مبتنی بر قانون یادگیری اصلاح و خطا است. از قانون یادگیری پس انتشار برای آموزش شبکه‌های عصبی مصنوعی چندلایه پیش‌خور که عموماً شبکه‌های چندلایه پرسپترون هم نامیده می‌شوند، استفاده می‌شود. پارامترهای شبکه طوری تنظیم می‌شوند که پاسخ واقعی شبکه هرچه بیش‌تر به سمت پاسخ مطلوب نزدیک‌تر شود. مدل شبکه عصبی مصنوعی داده‌ها به سه قسمت تست، آزمون و اعتبارسنجی تقسیم شده‌اند. ۷۰ درصد داده‌ها برای آموزش (مهم‌ترین بخش مدل)، ۱۵ درصد داده‌ها برای تست و ۱۵ درصد داده‌ها برای اعتبارسنجی بکار برده می‌شود که به این منظور از پیش فرض خود مدل استفاده گردیده است.

برای مقایسه و ارزیابی دو مدل پیش‌بینی، ابتدا داده‌های مرتبط به ۲۰ روز آخر داده‌های در اختیار، کنار گذاشته شده و تعداد مسافران قطار شهری در این ۲۰ روز به وسیله دو روش، پیش‌بینی شده است. برای ارزیابی مقادیر پیش‌بینی شده از ملاک RMSE که در بخش ۴ معرفی شد، استفاده شده است. مقدار جذر میانگین مربعات خطا برای دو مدل سری زمانی فصلی و شبکه عصبی برای پیش‌بینی تعداد مسافران قطار شهری در ماه اسفند سال ۱۳۹۱ به ترتیب ۲۸,۳۴۵ و ۳۴,۱۲۴ محاسبه شده است. مقادیر پیش‌بینی شده بوسیله دو مدل، در نمودار (۱-۱۰) نشان داده شده است. از آنجایی که RMSE مربوط به مدل سری زمانی فصلی مقدار کمتری است، می‌توان نتیجه گرفت که در این جا، مدل سری زمانی فصلی از عملکرد بهتری نسبت به مدل شبکه عصبی برخوردار است.



شکل ۱۰: مقادیر پیش‌بینی برای دو روش سری زمانی و شبکه عصبی

۶- نتیجه گیری

همان گونه که پیش تر بیان شد، امروزه رویکرد اصلی مدیریت سیستم های حمل و نقل بسوی مدیریت هوشمند و بر خط است. لذا مدل های بلند مدت امروزه جایگاه خود را به مدل های روندگرا و کوتاه مدت داده اند. این مقاله گامی در این راستا بوده است. بیشتر مقالات قبلی ارائه شده در خصوص مدل های کوتاه مدت بر مبنای مدل های شبکه عصبی بوده اند. در این مقاله از مدل سری زمانی فصلی به منظور مدل سازی تعداد مسافران قطار شهری شهر مشهد استفاده شده است. نتایج پیاده سازی مدل که دارای تطابق بسیار مناسب با داده های برداشت شده بود، با مدل شبکه عصبی مقایسه گردید. بر اساس این مقایسه مشخص گردید که دقت مدل سری زمان بیشتر از مدل شبکه عصبی بوده است. ضمناً باید توجه داشت که مدل های سری زمانی سابقه پیاده سازی چندانی در حمل و نقل و بخصوص مدلسازی آن نداشته و تا آنجا که نویسندگان اطلاع دارند دارای سابقه نیست و می توان آن را به عنوان یکی از اصلی ترین نوآوری های این مطالعات دانست.



۷- مراجع

- 1-M. Gaudry, "An aggregate time-series analysis of urban transit demand: The Montreal case," *Transportation Research*, vol. 9, pp. 249-258, 8// 1975.
- 2-M. Kyte, J. Stoner, and J. Cryer, "A time-series analysis of public transit ridership in Portland, Oregon, 1971–1982," *Transportation Research Part A: General*, vol. 22, pp. 345-359, 9// 1988.
- 3-W.-C. Chiang, R. A. Russell, and T. L. Urban, "Forecasting ridership for a metropolitan transit authority," *Transportation Research Part A: Policy and Practice*, vol. 45, pp. 696-705, 8// 2011.
- 4-G. A. Davis and N. L. Nihan, "Nonparametric regression and short-term freeway traffic forecasting," *Journal of Transportation Engineering*, vol. 117, pp. 178-188, 1991.
- 5-B. L. Smith and M. J. Demetsky, "Short-term traffic flow prediction: neural network approach," *Transportation Research Record*, 1994.
- 6-M. S. Dougherty and M. R. Cobbett, "Short-term inter-urban traffic forecasts using neural networks," *International journal of forecasting*, vol. 13, pp. 21-31, 1997.
- 7-M. M. Hamed, H. R. Al-Masaeid, and Z. M. B. Said, "Short-term prediction of traffic volume in urban arterials," *Journal of Transportation Engineering*, vol. 121, pp. 249-254, 1995.
- 8-B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, pp. 664-672, 2003.
- 9-Box, G. a. (1976). *Time series analysis: forecasting and control*, San Francisco, Holden Day Pulications.
- 10-Duenas, C. F. (2005). *Stochastic model to forecast ground level ozone concentration at urban and rural areas*. . *Chemosphere*, 61(10), pp. 1379-1389.
- 11-Gouvvia, N. a. (2000.). *Time series analysis of air pollution and mortality: effects by causes, age*. *Epidemiology community health*, 54: 750-755.
- 12-Hamilton, J. (1994). *Time series analysis*,. *Time series analysis*,.



- 13-Ingrisch, M. S. (2009). *Model selection in dynamic contrast enhanced MRI: The Akaike Information Criterion*. In Dössel, O. and Schlegel, WC. (Eds.) *IFMBE Proceedings 25/IV*.
- 14-Lumbreras, J. G.-M. (2009). *Computation of uncertainty for atmospheric emission projections from key pollutant sources in Spain*. *Atmospheric Environment*, 43, pp. 1557-1564.
- 15-Wagemakers, E. a. (2004). *AIC model selection using Akaike weights*. *Psychonomic Bulletin and Review*, 11, pp. 192-196.



A Seasonal Time Series Approach to the Prediction of LRT Passengers in Mashhad

Navid Kalantari¹, Ramin Khavarzade², Amir Ghiasi³

1- PhD in Transportation Engineering, Sharan Transportation Research Center (STRC)

2- PhD Student, Department of Statistics, Sharan Transportation Research Center (STRC)

3- MSc in Transportation Engineering, Sharan Transportation Research Center (STRC)

Abstract

Today, transportation researches are more focused towards dynamic and online planning and thus, they are more inclined towards short-term traffic prediction models. Short term modeling of LRT systems is one of the major steps that have to be taken in order to perform dynamic and optimal decisions in this system. Time Series analysis could be used as a tool for short term traffic prediction. This paper aims to investigate the capabilities of time series models for short term passenger modeling. Therefore the passenger count data of Mashhad LRT system has been used. These data ranged from the establishment of the system until 2013. The trend of these data has been analysis and the effect of special event and holidays has been modeled. Finally a SARIMA model has been developed in this paper and was compared with Neural Networks.

Keywords: *Time series modeling, short term prediction models, partial correlation, SARIMA, Neural Networks.*

