



## Estimation of Tehran Daily Water Demand Using Time Series Analysis

M. Tabesh<sup>1</sup>, M. Dini<sup>2</sup>  
A.J. Khoshkholgh<sup>3</sup> and B. Zahraie<sup>4</sup>

### Abstract

In recent years many researches have been performed on the subject of optimal usage of water resources. One of the issues in this regard is preparing short term water demand forecasting in conjunction with the optimum water demand management. Considering the effects of climatologic conditions on the short term water demand and according to the similarity of consumption trends in the consecutive days, two conventional and advanced models have been developed in this paper using time series method. These models have been used to predict the short term water demand in Tehran, Iran. In the conventional model the time series of Tehran daily water consumption is divided into different components of trend, seasonal variations, and random variations which are obtained by regression method. In the advanced model the consumption pattern in the past is recognized using combined methods of Auto Regressive (AR) and Seasonal Moving Average. Assuming that this pattern will continue in the future, it is then applied to predict daily water demand. In the conventional models it is assumed that different components of time series are independent and dividable. Therefore, its components are determined by different methods such as regression, moving average, etc. In the advanced models all the components are supposed to be correlated to each other and are therefore analyzed together. Comparing the results with the real data showed the ability and accuracy of the both time series models to predict the Tehran daily water demand. The advanced models produced better results than the conventional methods.

**Keywords:** Short term water demand forecasting, Time series, Average temperature, Number of sunny hours, Daily water consumption.

## برآورد مصرف روزانه آب تهران با استفاده از سری‌های زمانی

مسعود تابش<sup>۱</sup>، مهدی دینی<sup>۲</sup>  
علی جعفر خوش‌خلق<sup>۳</sup> و بنفشه زهرائی<sup>۴</sup>

### چکیده

در سال‌های اخیر بررسی‌های گسترده‌ای بر روی مساله استفاده بهینه از منابع آب انجام شده است. یکی از موضوعات مورد بررسی در این زمینه تهیه مدل پیش‌بینی مصرف کوتاه مدت آب جهت مدیریت مناسب تقاضا می‌باشد. در این مقاله با توجه به تأثیرپذیری مصرف کوتاه مدت آب از عوامل آب و هوایی و همچنین وجود الگوی مصرف مشابه در روزهای متوالی، دو مدل به روش سری‌های زمانی جهت پیش‌بینی کوتاه مدت مصرف آب شهری تهران تهیه شده است. در مدل اول با تجزیه سری زمانی مصرف روزانه آب به مولفه‌های روند، تغییرات فصلی، تغییرات تصادفی و تعیین هر یک از آنها با استفاده از روش رگرسیون، پیش‌بینی انجام می‌شود. در مدل دوم با استفاده از مدل‌های ترکیبی اتورگرسیو و میانگین متحرک فصلی، الگوی مصرف در گذشته شناسایی شده و با فرض اینکه در آینده نیز این الگو ادامه داشته باشد، از آن برای پیش‌بینی مصرف آب استفاده می‌شود. در نهایت نتایج دو مدل تهیه شده با نتایج واقعی مقایسه شده است که توانایی این مدل‌ها را برای پیش‌بینی مصرف کوتاه مدت آب شهری تهران نشان می‌دهد. از بین این دو، مدل پیشرفته دارای خطای کمتری برای داده‌های شهر تهران می‌باشد.

**کلمات کلیدی:** برآورد تقاضای کوتاه مدت، سری زمانی، درجه حرارت متوسط، تعداد ساعات آفتابی، مصرف روزانه آب

تاریخ دریافت مقاله: ۸ آبان ۱۳۸۵

تاریخ پذیرش مقاله: ۲۳ مرداد ۱۳۸۷

1- Associate Professor, Center of Excellence for Engineering and Management of Infrastructures, School of Civil Engineering, University of Tehran, P.O.Box 11155-4563, Tehran, Iran, Email: [mtabesh@ut.ac.ir](mailto:mtabesh@ut.ac.ir)

2- Graduated Student, School of Civil Engineering, University of Tehran, Email: [mm\\_dini@yahoo.com](mailto:mm_dini@yahoo.com)

3- Graduated Student, School of Civil Engineering, University of Tehran

4- Associate Professor, School of Civil Engineering School, University of Tehran, Iran, Email: [bzahraie@ut.ac.ir](mailto:bzahraie@ut.ac.ir)

۱- دانشیار و عضو قطب علمی مهندسی و مدیریت زیرساختها، دانشکده مهندسی عمران، پردیس دانشکده‌های فنی، دانشگاه تهران

۲- فارغ التحصیل کارشناسی ارشد عمران آب، پردیس دانشکده‌های فنی، دانشگاه تهران

۴- دانشیار دانشکده مهندسی عمران، پردیس دانشکده‌های فنی، دانشگاه تهران

## ۱- مقدمه

به طور کلی پیش‌بینی مصرف آب در سه دسته اصلی صورت می‌گیرد. پیش‌بینی دراز مدت معمولاً بازه‌ای از یک تا چند سال را در برمی‌گیرد. پیش‌بینی میان‌مدت، مربوط به نوسانات مصرف آب در طول سال می‌باشد. پیش‌بینی کوتاه مدت معمولاً بازه زمانی بین یک ساعت تا چند روز را پوشش می‌دهد. در مدیریت آب شهری دانستن تقاضای کوتاه مدت بسیار مهم و حیاتی است و در بهره‌برداری بهینه از شبکه توزیع و تاسیسات آب نقش به‌سزایی دارد.

تحقیقات انجام گرفته در زمینه پیش‌بینی مصرف آب از سال‌های دور مورد توجه بوده است ولی روش‌های پیش‌بینی و انگیزه آن در طول این دوران متفاوت می‌باشد. اکثر تحقیقات و مقالاتی که به بحث در مورد پیش‌بینی مصرف آب شهری پرداخته‌اند، مصرف ماهانه یا سالانه و به عبارتی مصرف میان‌مدت و بلندمدت را در نظر گرفته‌اند و تنها تعداد اندکی به بررسی مصرف کوتاه مدت پرداخته‌اند. اولین تحقیقات در زمینه پیش‌بینی مصرف آب با انگیزه تعیین حداکثر تقاضای روزانه، جهت انجام اقدامات لازم برای برآورده نمودن این تقاضا، صورت گرفته است. (Graser (1958) با مطالعه مقادیر آب مصرفی در دالاس تگزاس، دریافت که حداکثر مصرف روزانه آب تابع تعداد روزهای قبلی با حداکثر درجه حرارت بیش از ۳۸ درجه سانتیگراد و تعداد هفته‌های قبلی با بارش ۲/۵۲ سانتی‌متر می‌باشد. (Wong (1972, Young (1973) و Willsie and Pratt (1974) سری زمانی مصرف سالانه آب شهری را به متغیرهای جمعیت، درآمد خانوار، قیمت آب، بارش، درجه حرارت هوا و تبخیر وابسته کردند و توابعی را جهت تخمین مصرف سالیانه آب ارائه دادند.

(Maidment and Parzen (1984 سری زمانی مصرف ماهیانه (میان‌مدت) آب در تگزاس را توسط چهار معادله مدل کردند که هرکدام از این معادلات، مؤلفه‌هایی از سری زمانی اصلی را تخمین می‌زد. در این تحقیق آنها ابتدا مقادیر سری زمانی را به دو مؤلفه قطعی و تصادفی تقسیم نمودند. مؤلفه‌های قطعی خود شامل مؤلفه‌های روند و فصلی می‌شد که مؤلفه روند را به وسیله رگرسیون بین میانگین مصرف آب سالیانه و جمعیت شهر مدل کردند و مؤلفه فصلی را با استفاده از سری‌های فوریه تخمین زدند. آنها برای مدل کردن مؤلفه تصادفی از دو معادله استفاده نمودند. یکی از این معادلات، همبستگی مصرف آب با مقادیر مصرف شده در روزهای قبل و معادله دیگر همبستگی مصرف آب با متغیرهای آب و هوایی نظیر حداکثر درجه حرارت ماهانه، تبخیر و بارش را بررسی می‌کرد.

(Quevedo et al. (1986 مدلی را جهت پیش‌بینی کوتاه‌مدت و بلندمدت مصرف آب برای شهر بارسلونا تهیه کردند. آنها با استفاده از آنالیز سری‌های زمانی مربوط به مصرف آب در شهر بارسلونا، مدل‌های ماهیانه و روزانه را جهت پیش‌بینی ارائه کردند. در مدل روزانه آنها سه مؤلفه قطعی، تصادفی و مؤلفه مربوط به آنالیز تداخلی دیده می‌شد، ولی در مدل ماهیانه تنها مؤلفه قطعی و تصادفی وجود داشت. مؤلفه‌های قطعی به وسیله توابع مثلثاتی و مؤلفه‌های تصادفی توسط مدل‌های ARIMA غیر فصلی تعیین شدند. در مؤلفه مربوط به آنالیز تداخلی که به وسیله متغیرهای مجازی بیان می‌شد اثرات روزهای هفته و روزهای تعطیل بر مصرف آب بررسی شد.

(Chun Chen (1986 با استفاده از آنالیز سری‌های زمانی، مصرف روزانه و ساعتی آب را توسط یک مدل ساده اتورگرسیو پیش‌بینی نمود. (Zhou et al. (2000 مصرف روزانه آب (کوتاه‌مدت) شهر ملبورن را به صورت مجموعه‌ای از معادلات که شامل چهار مؤلفه روند، فصلی، همبستگی آب و هوایی و خود همبستگی می‌شد، مدل‌سازی کردند. آنها مصرف کل را به دو بخش مصرف اصلی و مصرف فصلی تقسیم نمودند. مصرف اصلی در مطالعه آنها برابر با کمترین مصرف ماهیانه در هر سال تعریف شد و آن را به وسیله یک مدل چندجمله‌ای درجه سه از زمان بیان کردند. مصرف فصلی به صورت تفاضل مصرف روزانه از مصرف اصلی معرفی گردید و بوسیله روش رگرسیون با عوامل آب و هوایی و رگرسیون با مقادیر مصرف آب در گذشته مدل شد. در مؤلفه آب و هوایی مدل آنها متغیرهای ماکزیمم درجه حرارت، بارش، تبخیر دیده می‌شود.

در رابطه با پیش‌بینی مصرف با استفاده از سیستم‌های خبره، تابش (۱۳۸۵) و تابش و همکاران (۱۳۸۶) شبکه‌های عصبی را برای پیش‌بینی کوتاه مدت تقاضای آب شهری تهران به کار بردند که در آنها از پارامترهای هواشناسی و مصرف آب شهری به عنوان ورودی‌های مدل استفاده شده است. کریمی (۱۳۸۰) سیستم فازی را برای پیش‌بینی کوتاه مدت مصرف آب تهران بکار برده که در این تحقیق نیز از داده‌های هواشناسی و داده‌های مصرف آب روزانه برای مدل‌سازی استفاده شده است. همچنین تابش و دینی (۱۳۸۵) روش‌های فازی و نروفازی را برای پیش‌بینی کوتاه مدت مصرف آب بکار بردند که در این مورد نیز داده‌های مصرف و داده‌های هواشناسی برای مدل‌سازی جهت پیش‌بینی مصرف روزانه مورد استفاده قرار گرفته است.

در این مقاله بعد از جمع‌آوری داده‌های هواشناسی و داده‌های مصرف با استفاده از روش‌های سنتی و پیشرفته سری‌های زمانی مدل‌هایی

برای پیش‌بینی مصرف آب شهر تهران ساخته شده و با یکدیگر مقایسه شده‌اند. همچنین برای نشان دادن کارایی مدل ساخته شده بهترین مدل‌های ساخته شده به روش سری‌های زمانی با مدل‌های عصبی و نروفازی مقایسه شده است. از مزیت‌های این مقاله می‌توان به بکارگیری روش‌های سنتی و پیشرفته پیش‌بینی مصرف آب در کنار هم اشاره نمود که در اکثر مقالات اشاره شده در بالا تنها از یکی از روش‌ها استفاده شده است.

## ۲- روش تحقیق

سری زمانی مجموعه‌ای از مشاهدات است که برحسب زمان و یا هر کمیت دیگر، مرتب شده و معمولاً آن را به صورت مجموعه  $\{z_{t_1}, z_{t_2}, \dots, z_{t_n}\}$  نشان می‌دهند. سری‌های زمانی را معمولاً به صورت گسسته یا پیوسته بررسی می‌کنند. اگر مشاهدات به طور پیوسته برحسب زمان در نظر گرفته شوند، سری زمانی حاصل را پیوسته می‌نامند. اگر مشاهدات به صورت منظم در فاصله‌های مساوی ثبت شود یک سری گسسته به دست می‌آید معمولاً برای تحلیل یک سری زمانی فرض می‌شود که تغییرات موجود در الگوی سری زمانی نتیجه چهار مؤلفه اصلی مؤلفه روند<sup>۱</sup>، مؤلفه فصلی<sup>۲</sup>، مؤلفه دوره‌ای<sup>۳</sup> و مؤلفه تصادفی<sup>۴</sup> می‌باشد.

مؤلفه روند عبارت از تغییرات دراز مدت در میانگین سری زمانی است. به عبارت دیگر سیر طبیعی سری زمانی در دراز مدت را روند می‌گویند و با  $T_t$  نشان می‌دهند. یک سری زمانی می‌تواند دارای روند مثبت (صعودی)، روند منفی (نزولی) و یا اساساً بدون روند باشد. مؤلفه فصلی با  $S_t$  نمایش داده می‌شود و مربوط به عواملی است که به طریقی منظم و متناوب روی یک دوره کمتر از یک سال عمل می‌کنند. مؤلفه دوره‌ای حرکت نوسانی در یک سری زمانی با دوره نوسان بیشتر از یک سال را شامل می‌شود که با  $C_t$  نشان داده می‌شود. مؤلفه تصادفی تغییرات نامنظم یا تصادفی سری زمانی می‌باشد که با  $E_t$  نشان داده می‌شود. این تغییرات کاملاً تصادفی بوده و نتیجه عوامل غیرقابل پیش‌بینی هستند که به طریقی نامنظم عمل می‌کند.

با توجه به افزایش جمعیت و کمبود منابع آب موجود در شهر تهران و لزوم بهره‌گیری بهینه از سیستم با شرایط فوق در این مقاله دو روش سنتی و پیشرفته برای تحلیل سری‌های زمانی بکار رفته و نتایج آنها با یکدیگر مقایسه گردیده است. در روش سنتی فرض بر این است که مؤلفه‌های مختلف یک سری زمانی از یکدیگر مستقل و قابل تفکیک می‌باشند. لذا با استفاده از روش‌های مختلف همچون

رگرسیون، میانگین متحرک و ... مؤلفه‌های مختلف را تعیین می‌کنند. در روش‌های پیشرفته فرض بر این است که مؤلفه‌ها با یکدیگر همبستگی دارند، لذا همه مؤلفه‌ها با یکدیگر مورد بررسی و تحلیل قرار می‌گیرند.

## ۲-۱- روش‌های سنتی

اگر در سری زمانی مؤلفه‌ها مستقل فرض شوند سری زمانی  $Z_t$  را به صورت زیر می‌توان نوشت.

$$Z_t = T_t + C_t + S_t + E_t \quad (۱)$$

برای تعیین هر یک از این مؤلفه‌ها می‌توان از روش رگرسیون استفاده نمود. شکل کلی یک مدل رگرسیونی خطی چندگانه به صورت زیر می‌باشد:

$$y_j = B_0 + B_1x_{1j} + B_2x_{2j} + \dots + B_kx_{kj} + e_j \quad (۲)$$

که در آن  $y_j$ : مقدار متغیر وابسته که برای عنصر  $j$  اندازه‌گیری شده است،  $x_j$ : ورودی‌های مدل به عنوان متغیر مستقل،  $B_j$ : ضرایب رگرسیون مربوط به متغیر مستقل  $x_j$  و  $e_j$ : متغیری تصادفی با میانگین صفر و واریانس  $\sigma^2$  می‌باشند. برای تعیین معادله باید ضرایب رگرسیون از داده‌های گذشته برآورد گردند.

## ۲-۲- روش‌های آماری پیشرفته

در این روش‌ها تمامی مؤلفه‌های گفته شده در مدلسازی باهم در نظر گرفته می‌شوند و یک مدل کلی برای تمامی آنها بدست می‌آید. این روشها که توسط دو محقق به نام باکس و جنکینس<sup>۵</sup> توسعه یافته‌اند، دارای کارایی قابل توجهی در مدل‌سازی و پیش‌بینی سری‌های زمانی می‌باشند (باکس و جنکینس، ۱۳۷۱). در این مدل‌ها ایستایی<sup>۶</sup> مدل نقش اساسی در مدل‌سازی دارد که به صورت رفتار متعادل سری‌های زمانی پیرامون یک سطح میانگین تعریف می‌شود. آن دسته از سری‌ها که دارای روند و یا تغییرات فصلی می‌باشند و در حقیقت در حول سطح ثابتی تغییر نمی‌کنند، سری‌های نالیستا نام دارند. اینگونه سری‌ها باید توسط تبدیلات مناسب به مدل‌های ایستا تبدیل شوند تا فرآیند مدل‌سازی را بتوان بر روی آنها انجام داد.

## ۲-۲-۱- مدل‌های خودهمبسته (AR)<sup>۷</sup>

در این مدل‌ها مقدار فعلی فرآیند به صورت یک مجموع خطی متناهی از مقادیر قبلی فرآیند و یک ضریب  $a_t$  بیان می‌شود. اگر مقادیر یک فرآیند در زمان‌های با فاصله‌های مساوی  $I, t, t-2, \dots, t-1, z_{t-2}, z_{t-1}, \dots$  نشان داده شوند و

فرآیند خودهمبسته (AR) از مرتبه  $p$  با معادله زیر نمایش داده می‌شود.

$$\tilde{z}_t = \varphi_1 \tilde{z}_{t-1} + \varphi_2 \tilde{z}_{t-2} + \dots + \varphi_p \tilde{z}_{t-p} + a_t \quad (3)$$

که در آن مقدار خود همبستگی جزئی با تأخیر  $k$  می‌باشد.  $a_t$  یک ضریب است که نمونه‌ای تصادفی از یک توزیع ثابت است، که معمولاً توزیع نرمال با میانگین صفر و واریانس  $\sigma_a^2$  فرض می‌شود.

### ۲-۲-۲ مدل‌های میانگین متحرک (MA)

در این مدل  $z_t$  به صورت یک معادله خطی به یک تعداد متناهی  $q$  از  $a_t$  های پیشین وابسته می‌شود. معادله زیر رابطه ریاضی یک فرآیند میانگین متحرک از مرتبه  $q$  را نشان می‌دهد.

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (4)$$

که در آن  $\theta(B)$  عملگر میانگین متحرک یا ضرایب میانگین متحرک می‌باشد.

### ۳-۲-۲ مدل‌های مخلوط خودهمبسته- میانگین

#### متحرک (ARMA)

به منظور نیل به انعطاف‌پذیری بیشتری در برازش مدل به سری‌های زمانی واقعی گاهی مدل خودهمبسته و مدل میانگین متحرک، هر دو در یک مدل گنجانده می‌شوند. این امر منجر به مدل مخلوط خودهمبسته- میانگین متحرک یا ARMA می‌شود که به صورت معادله زیر نوشته می‌شود.

(5)

$$\tilde{z}_t = \varphi_1 \tilde{z}_{t-1} + \dots + \varphi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

که در آن  $\varphi(B)$  عملگر خودهمبسته و  $\theta(B)$  عملگر میانگین متحرک می‌باشد.

### ۴-۲-۲ مدل‌های خودهمبسته- میانگین متحرک

#### تفاضلی (ARIMA)

بسیاری از سری‌هایی که در عمل وجود دارند، رفتار نایستایی را به نمایش می‌گذارند و پیرامون یک میانگین ثابت تغییر نمی‌کنند. با این وجود آنها از این لحاظ همگنی نشان می‌دهند که صرفنظر از مؤلفه روند، یک قسمت از سری تا حد زیادی مشابه قسمتی دیگر رفتار

می‌کند. مدل‌هایی که این چنین رفتار همگن نایستایی را توصیف می‌کنند، با فرض اینکه تفاضل مناسبی از فرآیند ایستاست، محاسبه می‌شوند. این مدل‌ها را فرآیندهای خودهمبسته- میانگین متحرک تفاضلی (ARIMA) می‌نامند و به طور کلی این مدل‌ها به صورت  $ARIMA(p,d,q) \times (P,D,Q)^s$  بوده و با معادله زیر بیان می‌شوند.

$$(1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps}) (1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) (1 - B^s)^D (1 - B)^d z_t = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_q B^q) a_t \quad (6)$$

که در آن  $p$  مرتبه خودهمبستگی غیر فصلی  $P$  مرتبه خودهمبستگی فصلی،  $q$  مرتبه میانگین متحرک غیر فصلی،  $Q$  مرتبه میانگین متحرک فصلی،  $d$  مرتبه تفاضل گیری غیر فصلی،  $D$  مرتبه تفاضل گیری فصلی و  $S$  طول فصل مدل سازی می‌باشد. همچنین  $\Phi(B^s)$  و  $\Theta(B^s)$  چندجمله‌ای‌هایی برحسب  $B^s$  می‌باشند که به ترتیب از درجه  $P$   $Q$  می‌باشند.  $\Phi_p$   $\Theta_p$  ثابت‌های چندجمله‌ای،  $\varphi_p$  ضریب خودهمبسته و  $\theta_p$  ضریب میانگین متحرک می‌باشد. در ضمن  $B_j$  ضرایب رگرسیون مربوط به متغیر مستقل ( ورودی‌های مدل) می‌باشد. برای ارزیابی مدل‌ها از شاخص‌های آماری استفاده شده است که در ادامه به آنها اشاره می‌شود.

**ضریب تشخیص ( $R^2$ ):** ضریب  $R^2$  همان توان دوم ضریب همبستگی می‌باشد. برای حالتی که مقادیر داده‌های واقعی و داده‌های پیش‌بینی شده کاملاً به همدیگر انطباق داشته باشند، مقدار ضریب بسته به روند صعودی و نزولی داده‌ها ۱ یا -۱ خواهد بود. در صورتی که هیچ انطباقی بین داده‌های واقعی و پیش‌بینی وجود نداشته باشد مقدار ضریب  $R^2$  صفر خواهد بود.

**میانگین خطای مطلق (MAPE):** خطای مطلق به صورت تفاضل داده‌های واقعی از داده‌های پیش‌بینی، تقسیم بر داده‌های واقعی تعریف می‌شود. میانگین خطای مطلق یا خطای میانگین مطلق نسبی از میانگین گیری تمامی خطای‌های مطلق بدست می‌آید و به صورت زیر تعریف می‌شود:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_{actual\ i} - y_{forecast\ i}|}{y_{actual\ i}} \quad (7)$$

که در آن  $y_{actual}$  و  $y_{forecast}$  به ترتیب برابر مقادیر واقعی و مشاهداتی می‌باشند.

### ۳- مراحل ساخت مدل

در این قسمت نحوه انتخاب پارامترهای مؤثر و ساختار مدل‌های سنتی و پیشرفته تهیه شده معرفی می‌شوند.

#### ۳-۱- انتخاب پارامترها

داده‌های مورد استفاده در مدل شامل اطلاعات اندازه گیری (ثابت) شده پارامترهای هواشناسی و اطلاعات مربوط به مصرف روزانه آب در شهر تهران می‌باشد. داده‌های هواشناسی که از طریق سازمان هواشناسی کل کشور تهیه شده است شامل درجه حرارت خنک، درجه حرارت تر، درجه حرارت مینیمم، درجه حرارت ماکزیمم، درجه حرارت میانگین روزانه، مقدار بارش روزانه، تعداد ساعات آفتابی در روز، سرعت باد، رطوبت هوا و فشار هوا می‌باشد. این اطلاعات مربوط به ده سال میلادی (۱۹۹۹-۱۹۹۰) بوده است. داده‌های مربوط به مصرف روزانه آب شهر تهران نیز از طریق شرکت مهندسی آب و فاضلاب کشور تهیه گردیده است. این داده‌ها مربوط به مصرف روزانه آب از سال ۱۳۷۰ تا ۱۳۷۹ می‌باشد. براساس اطلاعات موجود، داده‌های هواشناسی مصرف آب از ابتدای سال ۱۳۷۰ تا دهم دی ماه ۱۳۷۸ یکدیگر را پوشش می‌دهند. بنابراین عملیات مربوط به تجزیه و تحلیل و مراحل مدل‌سازی بر روی داده‌های سال‌های ۱۳۷۰ تا ۱۳۷۷ انجام می‌گردد و داده‌های سال ۱۳۷۸ برای تست مدل در نظر گرفته می‌شود.

در این مقاله برای شناسایی عوامل مؤثر بر مصرف آب شهر تهران تحلیل همبستگی بر روی داده‌های موجود انجام گرفته و سپس با توجه به مقدار ضریب همبستگی عواملی که دارای بیشترین مقدار همبستگی با مصرف آب می‌باشند جهت مدل‌سازی انتخاب گردیده‌اند. در این مقاله از فرمول پیرسن<sup>۱۳</sup> (رابطه ۱) برای محاسبه مقدار همبستگی مثبت و منفی بین پارامترها استفاده شده است.

$$R = \frac{\sum (x - \bar{x})(y - \bar{y})}{\left[ \sum (x - \bar{x})^2 \cdot \sum (y - \bar{y})^2 \right]^{1/2}} \quad (8)$$

که در آن  $x$  و  $y$  متغیرهای مستقل (درجه حرارت) و وابسته (مصرف روزانه آب) و  $\bar{x}$  و  $\bar{y}$  مقادیر میانگین حسابی آنها می‌باشد.

در نهایت پارامترهای زیر بعنوان عوامل مؤثر بر مصرف آب تهران در نظر گرفته شده‌اند:

درجه حرارت میانگین: به دلیل داشتن ضریب همبستگی بالا با مصرف آب و امکان پیش‌بینی آن توسط سازمان هواشناسی با دقت بالا برای چند روز آینده

تعداد ساعات آفتابی: به دلیل وجود همبستگی نسبی مثبت (حدود ۰/۶)

درصد رطوبت: به دلیل وجود همبستگی نسبی منفی (حدود ۰/۶-)

فشار هوا: به دلیل وجود همبستگی نسبی منفی (حدود ۰/۵-)

اثر ماه: به دلیل تکرار الگوی مصرف ماهانه در سال‌های مورد بررسی  
اثر روزهای هفته: به دلیل تکرار الگوی مصرف هفتگی در سال‌های مختلف

#### ۳-۲- مدل سنتی

در روش سنتی مؤلفه‌های مختلف سری زمانی یعنی مؤلفه روند، مؤلفه تغییرات فصلی، مؤلفه دوره‌ای و مؤلفه تصادفی هر یک به صورت جداگانه تعیین می‌گردد و در نهایت مدل به صورت مجموعه‌ای از این مؤلفه‌ها مشخص می‌شود. با توجه به طول دوره آماری مورد بررسی (۸ سال) عملاً مؤلفه دوره‌ای قابل تشخیص نخواهد بود. بنابراین، این مطالعه به تعیین مؤلفه روند و مؤلفه تغییرات فصلی و تغییرات تصادفی محدود می‌شود که این مؤلفه‌ها با استفاده از برقراری رگرسیون بین ورودی‌ها با خروجی‌ها تعیین می‌شوند.

در مجموع ۶ مدل سنتی ساخته شد که نتایج آنها در جدول ۱ بطور خلاصه آورده شده است. همانطوری که از جدول مشخص است مدل اول دارای یک ورودی شامل درجه حرارت میانگین روزانه می‌باشد و در مدل‌های بعدی ورودی‌های دیگر نیز وارد شده است. نتایج مدل‌ها با استفاده از مقایسه ضریب تشخیص ( $R^2$ ) با هم مقایسه گردیده است. از بین این مدل‌ها، مدل ۶ دارای بهترین ضریب تشخیص می‌باشد که ساختار آن در جدول ۲ نشان داده شده است. همانطور که مشاهده می‌شود مؤلفه روند به صورت تابعی از زمان می‌باشد. مؤلفه فصلی تابعی از پارامترهای هواشناسی، نوع فصل و روزهای هفته است. از بین پارامترهای هواشناسی، درجه حرارت میانگین روزانه و تعداد ساعات آفتابی در این مدل وارد شده است. سایر پارامترها به دلیل وابستگی به پارامترهای فوق‌تاثیر چندانی در بهبود نتایج ایجاد نمی‌کردند. البته استفاده از آنها می‌توانست به دلیل خطای پیش‌بینی پارامترهای هواشناسی برای روزهای بعد، خطای مدل را نیز افزایش دهد. در این بین مشاهده شد که وارد کردن پارامترهای هواشناسی، مانند درصد رطوبت نسبی روزانه و فشار هوا به دلیل وابستگی بین پارامترهای هواشناسی، در مدل‌سازی تاثیر مثبتی ایجاد نمی‌کنند. همچنین برای تک تک ماه‌ها مدل ۵ ساخته شد که در نهایت در مدل ۶ برای متغیر نوع ماه، تیر- مرداد- شهریور و آبان- آذر- دی- بهمن به دلیل رفتار یکسان مدل در تک تک ماه‌ها به صورت یک پارامتر در نظر گرفته شده‌اند. برای روزهای هفته نیز

### ۳-۳- مدل پیشرفته

در ساخت مدل‌های پیشرفته باید نرمال بودن داده‌ها تست می‌شد که با محاسبه ضریب چولگی این کار انجام شده است. ضریب چولگی داده‌های مربوط به مصرف روزانه آب شهر تهران برابر  $\hat{\gamma} = -0.105$  محاسبه شد که با قبول ریسک برابر با  $\alpha = 0.01$  فاصله اطمینان برای پذیرش فرض صفر بودن ضریب چولگی برابر  $[-0.114, 0.114]$  می‌شود. با توجه به اینکه ضریب چولگی محاسبه شده درون این فاصله قرار می‌گیرد. می‌توان با احتمال ۹۹ درصد فرض نرمال بودن داده‌ها را پذیرفت. بنابراین احتیاجی به انجام تبدیل برای نرمال کردن داده‌ها نمی‌باشد. مدل‌های پیشرفته ساخته شده شامل مدل‌های خود همبسته (AR)، میانگین متحرک (MA)، خود همبسته میانگین متحرک (ARMA) و خود همبسته میانگین متحرک تفاضلی (ARIMA) می‌باشند. در مجموع با استفاده از نرم‌افزار Statistica ده‌ها مدل با استفاده از این روش‌ها ساخته شده

روزهای شنبه تا چهارشنبه و روزهای پنج شنبه و جمعه مشابه هم عمل می‌کردند که بصورت یک پارامتر در نظر گرفته شده‌اند. تاثیر نوع ماه و روزهای هفته در مدل‌سازی به صورت صفر و یک می‌باشد بدین ترتیب که مدل ۵ برای پارامترهای درجه حرارت و ساعات آفتابی در اسفند ماه ساخته شده است و برای ماه‌های دیگر میزان تصحیح با علامت مثبت یا منفی وارد می‌شود. مدل ۶ نیز برای پارامترهای درجه حرارت و ساعات آفتابی در ماه اسفند و روزهای پنج شنبه و جمعه ساخته شده است و برای ماه‌های دیگر یا روزهای دیگر میزان تصحیح با علامت مثبت یا منفی وارد می‌شود.

نحوه عملکرد مدل برای یک روز پنجشنبه از مهرماه سال ۱۳۷۸ که درجه حرارت میانگین آن برابر ۱۹ درجه سانتیگراد و تعداد ساعات آفتابی آن روز برابر ۴ و مقدار اختلاف مصرف واقعی و مصرف پیش‌بینی شده هفت روز قبل آن برابر ۵۰۰۰۰ مترمکعب است، در جدول ۲ آورده شده است.

جدول ۱- مقایسه نتایج مدل‌های سنتی ساخته شده

مدل	M1	M2	M3	M4	M5	M6
ورودی‌ها	درجه حرارت	درجه حرارت	درجه حرارت	درجه حرارت	درجه حرارت	درجه حرارت
	ساعات آفتابی	ساعات آفتابی	ساعات آفتابی	ساعات آفتابی	ساعات آفتابی	ساعات آفتابی
		درجه حرارت	درصد رطوبت	فشار هوا	ماه‌های سال	ماه‌های سال
$R^2$	0.66	0.68	0.69	0.69	0.83	0.84

جدول ۲- روش استفاده از مدل تهیه شده به روش سنتی برای پیش‌بینی مصرف آب.

نوع مؤلفه (۱)	متغیر مستقل (۲)	ضریب متغیر مستقل (۳)	مقدار متغیر مستقل (۴)	(۳) * (۴)
روند	مقدار ثابت	-9.33E+07	-----	-9.33E+07
	سال	69471	1378	95.73E+06
تغییرات فصلی	مقدار ثابت	-189714	-----	-189714
	درجه حرارت میانگین	13639	19	259141
	تعداد ساعات آفتابی	8968	4	35872
	فروردین	-303487	0	0
	اردیبهشت	-207025	0	0
	خرداد	-65830.5	0	0
	تیر- مرداد - شهریور	69341.57	0	0
	مهر	-82483.7	1	-82483.7
	آبان - آذر - دی - بهمن	-115654	0	0
	شنبه تا چهارشنبه	-54214.3	0	0
تغییرات تصادفی	خطای هفت روز قبل	0.56	50000	28000
	مقدار مصرف آب (متر مکعب)			$\sum = 2480815$

جدول ۳ - نتایج حاصل از برآورد پارامترهای مدل  
ARIMA(0,1,2)(3,1,1)<sup>7</sup>

مقدار پارامتر	نوع الگو	علامت پارامتر
0.536	میانگین متحرک غیر فصلی از مرتبه اول	$\theta_1$
0.104	میانگین متحرک غیر فصلی از مرتبه دوم	$\theta_2$
0.104	خودهمبسته فصلی از مرتبه اول	$\Phi_1$
0.129	خودهمبسته فصلی از مرتبه دوم	$\Phi_2$
0.088	خودهمبسته فصلی از مرتبه سوم	$\Phi_3$
0.967	میانگین متحرک فصلی از مرتبه اول	$\Theta_1$

جدول ۴ - مقایسه نتایج حاصل از شبیه سازی مصرف به دو

روش سنتی و پیشرفته

روش تحلیل پیشرفته	روش تحلیل سنتی	معیار ارزیابی
۲/۲۸٪	۳/۳۳٪	میانگین خطای مطلق (MAPE)
۰/۹۴۵	۰/۹۳	ضریب تشخیص (R <sup>2</sup> )

با مشاهده شکل‌های ۱ و ۲ و جدول ۵ مشخص می‌شود که روش سنتی در ماه‌های گرم سال پیش‌بینی‌های دقیق‌تری را نسبت به روش پیشرفته ارائه می‌دهد و برعکس در ماه‌های سرد، روش پیشرفته پیش‌بینی‌های دقیق‌تری را ارائه می‌دهد.

برای مقایسه توانمندی‌های انواع روش‌های مدل‌سازی جهت مصرف آب، نتایج بدست آمده برای پیش‌بینی مصرف کوتاه مدت آب شهر تهران بوسیله سری‌های زمانی با نتایج مدل‌های نروفازی و شبکه عصبی ساخته شده برای پیش‌بینی مصرف آب شهری تهران که توسط تابش و دینی (۱۳۸۵) و تابش (۱۳۸۵) بدست آمده‌اند مقایسه می‌شوند. البته لازم به تذکر است که تعداد داده در این مدل‌ها با مدل سری زمانی متفاوت می‌باشد. بطور مثال داده‌های تست برای دو مدل نروفازی و شبکه عصبی در حدود سه سال می‌باشد در هر صورت مقایسه نتایج در ارزیابی مدل‌های سری زمانی موثر است.

داده‌های ورودی مدل نروفازی شامل درجه حرارت متوسط روزانه، درصد رطوبت نسبی، مصرف یک روز قبل، یک هفته قبل و یک سال قبل و خروجی مدل نیز مصرف یک روز بعد می‌باشد. همچنین ورودی‌های مدل شبکه عصبی عبارتند از: درجه حرارت متوسط روزانه، درصد رطوبت نسبی و مصرف یک روز قبل تا هفت روز قبل. جدول ۶ نتایج این مقایسه را نشان می‌دهد. همانطور که مشاهده می‌شود

و نتایج آنها با استفاده از نمودار خودهمبستگی سری‌ها با یکدیگر مقایسه شده است. در نهایت از میان کلیه مدل‌های پیشرفته، مدل ARIMA(0,1,2)(3,1,1)<sup>7</sup> به عنوان مدل بهینه انتخاب شده است. مدل فوق یک مدل خودهمبسته میانگین متحرک تفاضلی با مرتبه تفاضل‌گیری فصلی و غیرفصلی یک است که مرتبه خودهمبستگی فصلی و غیرفصلی آن به ترتیب سه و صفر می‌باشد. مرتبه میانگین متحرک فصلی و غیرفصلی مدل نیز به ترتیب یک و دو می‌باشد. عدد هفت نماینده طول فصل مدل‌سازی است که برابر هفت روز می‌باشد. مقادیر پارامترهای مدل بصورت جدول ۳ می‌باشد.

با جایگذاری مقادیر پارامترها در رابطه ۶ معادله مدل به صورت زیر بدست می‌آید:

(۹)

$$(1 - 0.104B^7 - 0.129B^{14} - 0.088B^{21})(1 - B^7)(1 - B)z_t = (1 - 0.967B^7)(1 - 0.536B - 0.104B^2)a_t$$

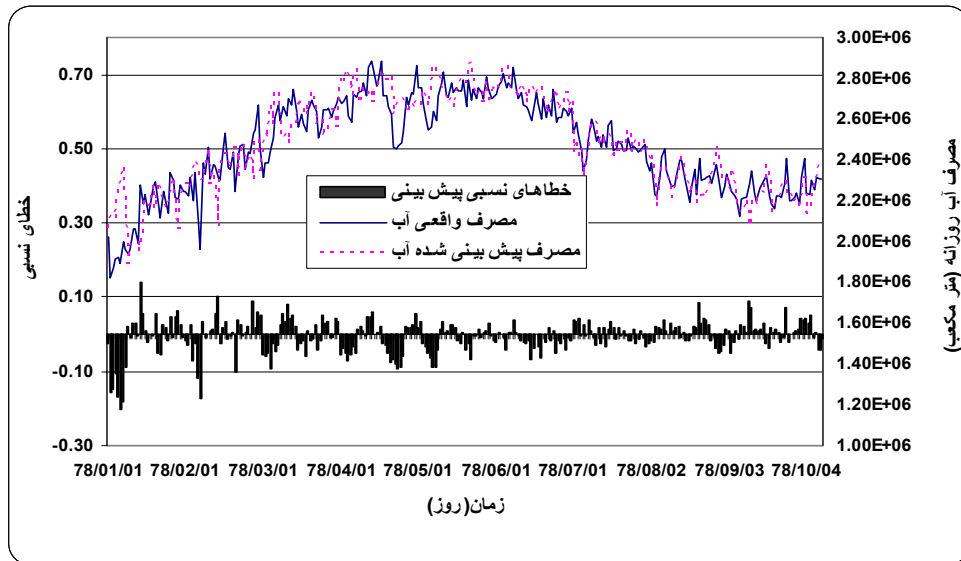
۴-۳ - نتایج مدل‌ها

جدول ۴ مقایسه‌ای از ضرایب تشخیص و درصد خطای میانگین مطلق را برای دو نوع مدل سنتی و پیشرفته نشان می‌دهد. همانطوری که مشاهده می‌شود مدل تهیه شده توسط روش‌های آماری پیشرفته درصد بیشتری از تغییرات موجود در مصرف آب را می‌تواند توضیح دهد. در ضمن در این مدل خطاهای ناشی از پیش‌بینی متغیرهای آب و هوایی که در مدل تهیه شده به روش سنتی وجود دارد، وارد نمی‌شود.

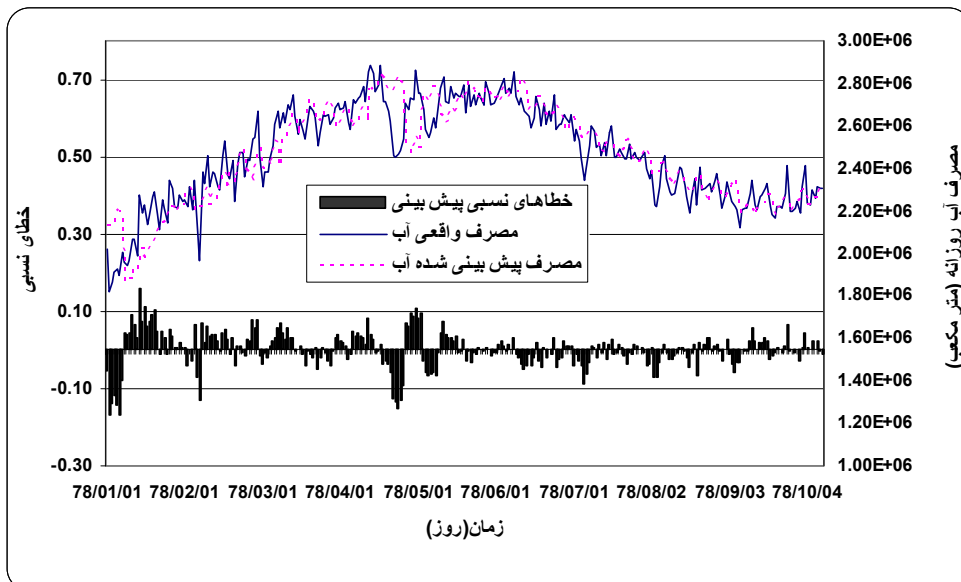
در ادامه، توانایی مدل‌های تهیه شده در پیش‌بینی میزان مصرف آب مورد بررسی قرار می‌گیرد. به دلیل وجود یک دوره فصلی با طول فصل برابر با هفت روز در سری زمانی مصرف آب و امکان پیش‌بینی پارامترهای آب و هوایی مؤثر بر مصرف آب برای هفت روز آینده، گام زمانی پیش‌بینی، برابر با هفت روز در نظر گرفته شده و مصرف برای هفت روز آینده پیش‌بینی می‌شود. پس از هفت روز، مقادیر واقعی مصرف به اطلاعات قبلی اضافه شده و به اصطلاح مدل به هنگام می‌شود و مجدداً برای هفت روز آینده پیش‌بینی صورت می‌گیرد. شکل‌های ۱ و ۲ مقایسه‌ای از مقادیر مصرف واقعی و مصرف پیش‌بینی شده برای سال ۱۳۷۸ را توسط این دو مدل نشان می‌دهد.

در جدول ۵ نیز معیار درصد خطای میانگین مطلق (MAPE) برای مقادیر پیش‌بینی شده توسط هر دو مدل در ماه‌های فروردین تا آذر سال ۱۳۷۸ مقایسه شده‌اند.

دقت نتایج نزدیک به یکدیگر بوده و در مورد روش سری زمانی (روش پیشرفته) نتایج اندکی دقیق تر می باشد.



شکل ۱ - مقایسه مصرف واقعی و پیش بینی شده برای سال ۱۳۷۸ در شهر تهران به روش سنتی.



شکل ۲ - مقایسه مصرف واقعی و پیش بینی شده برای سال ۱۳۷۸ در شهر تهران به روش پیشرفته.

0.0235	0.0216	آبان
0.0232	0.0185	آذر

جدول ۶- نتایج مدل نروفازی و شبکه عصبی

MAPE (%)	R <sup>2</sup>	نوع مدل
2.82	0.928	مدل نروفازی
2.94	0.9322	مدل شبکه عصبی
3.33	0.93	سری زمانی (سنتی)
2.28	0.945	سری زمانی (پیشرفته)

جدول ۵ - مقایسه شاخص MAPE دو نوع مدل تهیه شده در پیش بینی مصرف برای ماههای مختلف سال ۱۳۷۸.

ماه	روش مدل های آماری پیشرفته	روش سنتی
فروردین	0.0686	0.0586
اردیبهشت	0.0376	0.0388
خرداد	0.0250	0.0369
تیر	0.0542	0.0411
مرداد	0.0310	0.0273
شهریور	0.0207	0.0203
مهر	0.0210	0.0188



#### ۴- نتیجه‌گیری

پیش‌بینی مطمئن مصرف کوتاه مدت آب، مدیران یک شبکه آبرسانی را قادر می‌سازد که برنامه‌ریزی بهتری جهت مدیریت و بهره‌برداری بهینه از شبکه داشته باشند. در این مقاله با استفاده از روش سری‌های زمانی مدلی جهت پیش‌بینی مصرف کوتاه مدت آب شهر تهران ساخته شده است. پارامترهای ورودی مدل شامل پارامترهای هواشناسی و تاثیر ماه‌های فصل و روزهای هفته می‌باشد. پارامترهای هواشناسی موثر در مدل‌سازی از طریق تحلیل همبستگی بین ورودی‌ها با مصرف انتخاب شده است و در نهایت با وارد کردن آنها در مدل‌سازی تاثیر آنها در مدل نشان داده شده است. در مجموع دو مدل سنتی و پیشرفته با استفاده از سری‌های زمانی ساخته شده است که مقایسه نتایج آنها نشان می‌دهد که مدل‌های پیشرفته در مجموع نتایج بهتری نسبت به روش سنتی بدست می‌دهند. از طرف دیگر بررسی این مدل‌ها در فصلهای سرد و گرم بطور جداگانه نشان می‌دهد که مدل سنتی برای ماههای گرم و مدل پیشرفته برای ماههای سرد مناسب می‌باشد. همچنین مقایسه این مدل‌ها با مدل‌های نروفازی و شبکه عصبی نشان می‌دهد که مدل‌های سری زمانی از قابلیت بالایی در پیش‌بینی مصرف آب برخوردار هستند.

#### پی‌نوشت‌ها

1. Trend Component
2. Seasonal Component
3. Cyclical Component
4. Random Component
5. Box-Jenkins
6. Stationary
7. Autoregressive Model
8. Moving Average Model
9. Autoregressive Moving Average Model
10. Autoregressive Integrated Moving Average Model
11. Mean Absolute Percentage Error
12. Pearson

#### ۵- مراجع

باکس، جی.ای.پی. و جنکینس، جی.ام. (۱۳۷۱). "تحلیل سری‌های زمانی در پیش‌بینی و کنترل"، مشکاتی (مترجم)، چاپ اول، مرکز نشر و انتشارات دانشگاه شهید بهشتی.

تابش، م. و دینی، م. (۱۳۸۵)، "برآورد کوتاه مدت مصرف آب شهری به روش نروفازی"، هفتمین گنگره بین‌المللی مهندسی عمران، دانشگاه تربیت مدرس، اردیبهشت.

تابش، م. (۱۳۸۵)، "گزارش نهایی طرح پژوهشی تخمین تقاضای کوتاه‌مدت آب شهری با استفاده از شبکه‌های عصبی و

سیستم‌های فازی و ترکیبی"، شرکت سهامی مدیریت منابع آب ایران، معاونت پژوهش و مطالعات پایه، وزارت نیرو.

تابش، م.، گوشه، س. و یزدانپناه، م. ج. (۱۳۸۶)، "کاربرد شبکه‌های عصبی مصنوعی در تخمین تقاضای کوتاه مدت آب شهری"، نشریه دانشکده فنی، جلد ۴۱، شماره ۱، صفحه ۱۱ - ۲۴.

کریمی، د. (۱۳۸۰)، "کاربرد منطق فازی در پیش‌بینی کوتاه مدت مصرف آب تهران"، پایان نامه کارشناسی ارشد، گروه مهندسی عمران محیط زیست، دانشگاه تربیت مدرس.

Chun Chen, Y., (1986), "Applications of time series analysis to water demand Prediction", *Proceedings of Water and Data Processing*, Presse Pons et Chausseées, Paris.

Graser, H.J., (1958), "Meter records in system planning", *Journal of American Water Works Association*, Vol. 50, No.11, pp. 1395-1402.

Maidment, D.R., and Parzen, E., (1984), "Cascade model of monthly municipal water use", *Journal of Water Resources Research*, Vol.20, No.1, pp.15-23.

Quevedo, J., and Cembrano, G., (1986), "Water demand forecasting through time series analysis", *Proceedings of Water and Data Processing*, Presse Pons et Chausseées, Paris.

Willsie, R.H., and Pratt, H.L., (1974), "Water use relationships and projection corresponding with regional growth", *Water Resources Bulletin*, Vol. 10, No. 2, pp. 360-371.

Wong, S.T., (1972), "A model on municipal water demand: A case study of northeastern Illinois", *Land Economy*, Vol. 48. No. 1, pp. 34-44.

Young, R.A., (1973), "Price elasticity of demand for municipal water: A case study of Tucson and Arizona", *Journal of Water Resources Research*, Vol. 9, No. 4, pp. 1068-1072.

Zhou, S.L., McMahon, T.A., and Lewis, J., (2000), "Forecasting daily urban water demand: A case study of Melbourne", *Journal of Hydrology*, Vol. 236, No. 2, pp. 153-164.