مجله دانش آب و خاک/ جلد ۱۹ شماره ۱/ سال ۱۳۸۸

پیش بینی جریان سالانه رودخانه با استفاده از مدل خودهمبسته تجمعی میانگین متحرک و رگرسیون فازی

لاله پرویز[1*]، مجید خلقی[2] و احمد فاخری فرد[3]

تاریخ پذیرش: 87/6/10

1- دانشجوی دکتری، مهندسی منابع آب، دانشگاه تهران
2- گروه مهندسی آب، دانشگاه تهران
3- گروه مهندسی آب، دانشگاه تبریز
*مسئول مکاتبه E-mail:lparviz@ut.ac.ir

چکیده

رشد روزافزون جمعیت و محدودیت منابع آب سطحی در کشور، لزوم پیش‌بینی دقیق‌تر مقدار آورد رودخانه را به دلیل اهمیت در برنامه‌ریزی و مدیریت منابع آب از جمله بهره‌برداری از مخازن و طراحی سازه‌های کنترل سیلاب با استفاده از ابزارها و روش‌های نوین مدلسازی می‌طلبد. در این راستا، مدل‌های سری زمانی از دیرباز مورد توجه هیدرولوژیست‌ها بوده‌اند. هدف این تحقیق، ارزیابی کارآیی دو رهیافت کلی مدل سری زمانی و رگرسیون فازی در پیش‌بینی جریان سالانه رودخانه می‌باشد. در مدل خودهمبسته تجمعی میانگین متحرک از رهیافت سری زمانی، کارآیی روش‌های درست‌نمایی شرطی و درستنمایی غیر شرطی در تخمین پارامترهای مدل مورد بررسی قرار گرفت. در مدل رگرسیون فازی، به منظور در نظرگرفتن عدم قطعیت حاکم بر سیستم‌های طبیعی، از تابع عضویت مثلثی متقارن و نامتقارن استفاده شد. به منظور مقایسه کارآیی دو مدل مذکور در پیش‌بینی جریان سالانه، آمار آبدهی برخی از ایستگاه‌های حوضه آبریز دریاچه ارومیه بکار گرفته شد. نتایج نشان دادند که در بین روش‌های تخمین پارامترها، روش درستنمایی غیرشرطی به عنوان روش کارآمد در تخمین پارامترهای مدل *ARIMA* می‌باشد. مقایسه جریان‌های سالانه پیش‌بینی شده توسط مدل‌های *ARIMA* و رگرسیون فازی براساس معیارهایی مانند *RMSE*، دلالت بر عملکرد بهتر رهیافت رگرسیون فازی نسبت به مدل سری زمانی داشت. عملکرد بهتر تابع عضویت مثلثی متقارن نسبت به نوع نامتقارن آن از حیث درنظر گرفتن عدم قطعیت حاکم بر مسئله مدلسازی از دیگر نتایج این تحقیق می‌باشد.

واژه های کلیدی: تخمین پارامتر، حوضه آبریز دریاچه ارومیه، رگرسیون فازی، ماکزیمم درست‌نمایی، مدل*ARIMA*

Forecasting Annual Streamflow Using Autoregressive Integrated Moving Average Model and Fuzzy Regression

L Parviz[1*], M Kholghi[2] and A Fakherifard[3]

[1]Ph.D Student, Water Resources Engineering, University of Tehran
[2]Department of Water Engineering University of Tehran
[3]Department of Water Engineering, University of Tabriz
*Corresponding author: E-mail:lparviz@ut.ac.ir

Abstract

Steadly increase in the world's population and the lack of sufficient water resources, call for accurate forecasting of streamflow using efficient methods in planning and management of water the available resources such as reservoirs. For this aim, time series models have been long considered by the hydrologists. The goal of this research was to investigate stochastic model (ARIMA) and fuzzy regression performance for the annual streamflow forecasting. The parameter estimation methods of ARIMA model are conditional and unconditional likelihood. Fuzzy regression has been used symmetric and non-symmetric triangular membership for regarding the uncertainties of real systems. For the comparison of ARIMA and fuzzy regression performance, streamflow data of some tributaries of Urmia lake basin were employed. Results indicated that unconditional likelihood is the best method as compared to the other parameter estimation methods. Comparison between the forecasted and observed streamflow series using the two models revealed that fuzzy regression has the best fit to the observed streamflow data. The performance of symmetric triangular membership was better than non- symmetric.

Keywords: ARIMA model, Fuzzy regression, Maximum likelihood, Urmia lake basin, Parameter estimation

مقدمه

تبیین پدیده‌های هیدرولوژیکی به صورت سری‌های زمانی به علت عدم پیروی این پدیده ها از قوانین مشخص و معین می‌باشد به طوری که هدف تجزیه و تحلیل اطلاعات گذشته سری در جهت تعمیم الگوی برآوردی یا مدل پیشنهادی به آینده می‌باشد. در تمام مراحل مدلسازی هدف ایجاد مدل هایی با فرم ساده، حداقل تعداد پارامترها و رسایی مدل از لحاظ توان نمایش پدیده مورد نظر است. در دستیابی به این هدف، تخمین پارامترهای مدل برای تامین بهترین برازش با داده های واقعی نقش اساسی دارد زیرا که برآورد اشتباه پارامترها منجر به اریبی و پیش‌بینی‌های غیر قابل قبول خواهد شد (مشکانی 1371). در بین مدل‌های معمول سری‌های زمانی هیدرولوژیکی، استفاده از مدل خود همبسته تجمعی میانگین متحرک (*ARIMA*) در پیش‌بینی جریان رودخانه دارای سابقه طولانی می‌باشد.

کارسون و همکاران (1970) برای اولین بار سری‌های سالانه جریان رودخانه را با مدل *ARIMA* مدلسازی کردند که کاربردهای زیادی درهیدرولوژی داشته است. روند پایه‌ای مدل‌های *ARIMA* توسط باکس و جنکینز (1976) و براک ول و دیویس (1987) بیان شده است، سپس تغییرات و تصحیحاتی در مورد این مدل در دهه‌های هفتاد و هشتاد میلادی انجام گرفته است (مونتاناری و همکاران 1997) بطوری که از مدل فصلی خودهمبسته تجمعی میانگین متحرک (*SARIMA*) در مدلسازی سری زمانی جریان رودخانه استفاده شد (دلور و همکاران 1976). تحقیقاتی نیز در زمینه تخمین پارامترهای مدل*ARIMA* انجام گرفت به عنوان نمونه بررسی فرم های مختلف تابع درست‌نمایی مدل *ARMA* با هدف جستجوی نقطه حداکثر با روش های بهینه سازی (اوبداگ و هارنبوچی، 2006) و استفاده از تداوم تخمین گر حداکثر درست‌نمایی (شاین و لی، 2000). مقایسه بین عملکرد دو مدل *ARIMA* و *Thomas Fiering* در پیش بینی جریان رودخانه و کیفیت آب، حاکی از عملکرد بهتر روش *Thomas Fiering* بوده است (کرینس و همکاران 2005). کمالی و همکاران (1385) از مدل *SARIMA* برای پیش‌بینی جریان ماهانه ورودی به مخزن سد شهید عباسپور استفاده کردند. مقایسه نتایج مدل با داده های اندازه گیری شده، حکایت از کارایی قابل قبول مدل دارد (کمالی و همکاران 1385) .

از طرفی دیگر تحلیل رگرسیون یکی از مهم‌ترین ابزار آماری مورد استفاده توسط هیدرولوژیست‌ها می‌باشد. اساس تمام آنها ساختن یک مدل (معادله پیش بینی) بر اساس داده‌های مشاهداتی است. در برخی از موارد به علت پیچیدگی زیاد ساختار سیستم های مورد بررسی و عدم قطعیت متغیرهای مورد استفاده، تئوری مجموعه های فازی در تحلیل رگرسیون ابزار مناسبی جهت مدلسازی با استفاده از توابع عضویت فازی می باشد. روش رگرسیون آماری براساس قوانین احتمالات و رگرسیون فازی هم براساس تئوری احتمالات و هم تئوری مجموعه های فازی است. در تحلیل رگرسیون آماری خطای محاسباتی بین داده های مشاهداتی و مدل رگرسیونی متغیری تصادفی با توزیع نرمال است ولی در رگرسیون فازی خطا به صورت میزان فازی بودن ساختار مدل می باشد (چنگ و ایوب 2001). برای اولین بار لطفی زاده در سال 1965 مفاهیم مجموعه های فازی را بیان کرد و کاربردهایی از داده های فازی در مدل های رگرسیونی را پیشنهاد داد (زاده 1965). در اوایل رگرسیون فازی با در نظر گرفتن مینیمم میزان شاخص فازیت (میزان عدم قطعیت) مورد توجه قرار گرفت (تاناکا و همکاران 1982). همچنین مقایسه‌ای بین مدل های مختلف رگرسیون فازی انجام و به تفاوت های رگرسیون فازی و معمولی نیز اشاره شده است (رادان و ودال، 1994). مقایسه ای بین رگرسیون فازی و رگرسیون معمولی توسط چنگ و ایوب (2001) انجام گرفت. سه روش رگرسیون فازی مورد استفاده شامل حداقل کردن میزان فازیت به عنوان یک معیار بهینه‌سازی، حداقل مربعات به عنوان معیار برازش و تحلیل رگرسیون فاصله‌ای با رگرسیون آماری بوده است (چنگ و ایوب، 2001). ویو (2003) روشی را برای تخمین پارامترهای رگرسیون خطی فازی با کمک "*Resolution Identity*" در تئوری مجموعه های فازی ارائه داد (ویو 2003). یک روش جدید با تغییراتی در قسمت بهینه سازی رگرسیون فازی جهت محاسبه ضرایب رگرسیون فازی ابتدا با در نظر گرفتن متغیرهای وابسته بصورت فازی و سپس فازی بودن متغیرهای وابسته و مستقل ارائه شده است. از محدودیت‌های روش عدم بکارگیری این روش برای توابع عضویت مثلثی نامتقارن می‌باشد (حجتی و همکاران 2005).

هدف این تحقیق بررسی عملکرد دو رهیافت مدل سری زمانی (مدل *ARIMA*) و رگرسیون فازی با تابع عضویت مثلثی در پیش‌بینی جریان سالانه رودخانه می‌باشد. در این راستا اثر روش‌های تخمین پارامترهای مدل سری زمانی و همچنین تاثیر دو تابع عضویت مثلثی

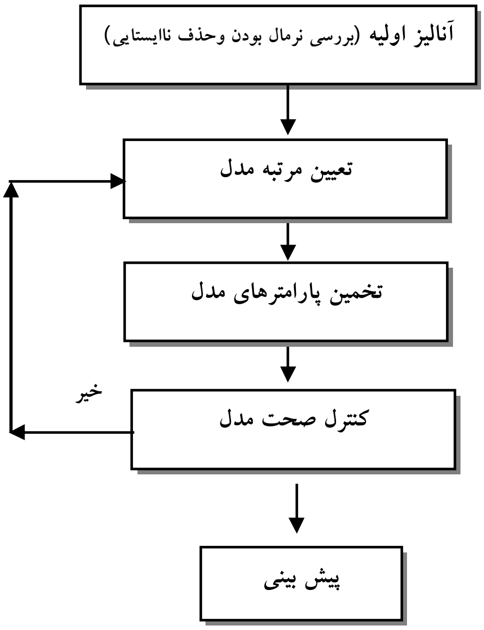
متقارن و نامتقارن رگرسیون فازی در پیش بینی جریان مورد بررسی قرار گرفته است. علت استفاده از روش رگرسیون فازی توجه کم به عدم قطعیت در ساختار مدل های کلاسیک پیش بینی جریان رودخانه می‌باشد.

مواد و روش‌ها

مدل خودهمبسته میانگین متحرک ($ARMA$) برای مدلسازی سری های زمانی ایستا و مدل $ARIMA$ برای مدلسازی سری های زمانی ناایستا کاربرد دارند. از جمله روش های تبدیل سری ناایستا به سری ایستا استفاده از روش تفاضل گیری سری ناایستا است. ساختار ریاضی مدل $ARIMA$ بصورت رابطه 1 می باشد.

$$f_P(B)(1-B)^d Z_t = q_0 + q_q(B)e_t \qquad [1]$$

$[Z_t]$: سری اصلی، $[e_t]$:سری تصادفی با میانگین صفر و واریانس s_e^2، f_i: پارامترهای مدل



شکل1- مراحل مدلسازی سری‌های زمانی

در این روند تکراری تعیین مرتبه‌های مدل بعد از آنالیز اولیه و شناسایی مدل با استفاده از برخی معیارها

خودهمبسته، q_i:پارامترهای مدل میانگین متحرک، p : مرتبه مربوط به مدل خودهمبسته، q: مرتبه مربوط به مدل میانگین متحرک، d : مرتبه مربوط به تفاضل، B: عملگرپسرو. q_i:این پارامتر برای $d=0$ و $d>0$ نقش های متفاوتی دارد. اگر $d=0$ فرآیند اولیه ایستا است و q_0 به میانگین فرآیند بستگی دارد و از رابطه 2 قابل محاسبه است. با این وجود وقتی که $d>1$ است اغلب آن را از مدل حذف می‌کنند مگراین که واقعاً لازم باشد.

$$q_0 = m(1-\sum_{i=1}^{p} f_i) \qquad [2]$$

m:میانگین ، f_i: پارامترهای مدل خودهمبسته (سالاس و همکاران، 1988). مراحل مدلسازی در سری‌های زمانی بر اساس روند تکراری باکس و جنکینز که در شکل 1 نشان داده شده است، می باشد (والنوزیولا و همکاران 2004).

انجام می‌گیرد. معیار مورد استفاده در این تحقیق تست معیارآکائیک (AIC[1]) است.

$$AIC(i) = NLn(s_{ei}^{\ 2}) + 2n_i \qquad [3]$$

i:دلالت برتعداد مدل‌های منتخب، n :تعداد پارامترها (مجموع مرتبه‌های مدل خودهمبسته و میانگین متحرک)،N: تعداد داده های مشاهداتی، s_{ei} : انحراف معیار مدل. این تست براین مبنا استوار است که هر مرتبه‌ای که معیار آکائیک کمتر داشته باشد برازش بیشتری با سری مشاهداتی خواهد داشت. دو مرحله بعدی تخمین پارامتر و کنترل صحت مدل است. یکی از مهمترین مراحل روند باکس و جنکینز تخمین پارامترهای مدل می باشد.روش‌های مختلفی به منظور تخمین پارامترهای مدل $ARIMA$ وجود دارد که برخی از آنها شامل روش‌های روابط ضرایب خود همبستگی برحسب پارامترهای مدل، درستنمایی شرطی و غیرشرطی می-

[1] Akaike Information Criterion

باشد. در بخش بعد هر یک از روش‌های تخمین مذکور، به اختصار شرح داده شده‌اند.

روش‌های تخمین پارامترهای مدل *ARIMA*

استفاده از فرمول‌های ضرایب خود همبستگی برحسب پارامترهای مدل

در این روش با تشخیص مرتبه‌های مدل، نمایی از شکل مدل مشخص می‌شود و سپس با تابع خود همبستگی بدست آمده (ρ_k) پارامترهای مدل تعیین می شوند. به عنوان مثال در مورد مدل *ARMA(1,1)* تابع خودهمبستگی را می توان با روابط زیر بدست آورد (سالاس و همکاران 1988).

$$r_0=1 \qquad [4]$$

$$r_1=\frac{(f_1-q_1)(1-f_1q_1)}{1+q_1^{\,2}-2f_1q_1} \qquad [5]$$

$$r_k=f_1r_{k-1} \qquad k\ \geq\ 2 \qquad [6]$$

روش حداکثر درست‌نمایی

اگر از یک جامعه تعدادی نمونه تصادفی در نظر گرفته شود، مقدار حاصلضرب تابع چگالی احتمال بازای مقادیر متغیر $x_N,\ldots,x_2,x_1$، تابع درستنمایی توزیع مربوطه نامیده می‌شود.

$$L=\overset{N}{\underset{i=1}{p}}\ f\ (x_i,a) \qquad [7]$$

α: پارامترهای مجهول، L: تابع درستنمایی و $f(x_i,a)$: تابع فراوانی.

اصل درستنمایی بیان می‌کند (با فرض اینکه مدل موردنظردرست باشد) همه اطلاعاتی را که داده‌ها درباره پارامترها در بر می گیرند در تابع درستنمایی نهفته است. روش حداکثر تابع درستنمایی بر این اصل استوار است که پارامترهای توزیع چنان برآورد شوند که بازای آمار داده شده تابع درستنمایی حداکثر باشد. از آنجائی‌که تابع درست‌نمایی بیشتر به صورت توان می‌باشد و از طرف دیگر حداکثر کردن لگاریتم یک تابع با حداکثر کردن خود تابع هم ارز است لذا برای سهولت محاسبات، حداکثر درست‌نمایی بر روی لگاریتم تابع انجام می‌گیرد. در این روش از تابع فراوانی خطاها که از توزیع نرمال با میانگین صفر و واریانس $s_e^{\,2}$ پیروی می‌کنند، استفاده شده است. تابع چگالی احتمال خطاها به صورت رابطه زیر می‌باشد.

$$f(e_t)=\frac{1}{\sqrt{2ps_e^{\,2}}}e^{\frac{-(e_t-\bar{e})^2}{2s_e^{\,2}}}=\frac{1}{\sqrt{2ps_e^{\,2}}}e^{\frac{-e_t^{\,2}}{2s_e^{\,2}}} \qquad [8]$$

بعد از اعمال تابع درستنمایی و سپس لگاریتم‌گیری از آن رابطه زیر حاصل می‌شود.

$$LL=-NLn(\sqrt{2p}\,s_e)-\frac{1}{2s_e^{\,2}}\sum_{t=1}^{N}e_t^{\,2} \qquad [9]$$

یکی از گزینه های ماکزیمم نمودن لگاریتم تابع درست‌نمایی، کاهش مقدار $-\frac{1}{2s_e^{\,2}}\sum_{t=1}^{N}e_t^{\,2}$ می باشد، به عبارتی دیگر حداقل کردن مقدار $\sum_{t=1}^{N}e_t^{\,2}$، که منجر به روش کمترین مربعات می شود. حال جهت مینیمم کردن مجموع مربعات خطاها دو روش درستنمایی شرطی و غیرشرطی بیان می‌شود.

درست‌نمایی شرطی

وقتی تعداد نمونه ها متوسط و یا بزرگ باشد می‌توان یک تقریب مناسبی از درستنمایی غیر شرطی را با استفاده از درست‌نمایی شرطی و استفاده از مقادیر اولیه خطاها و سری تفاضلی بدست آورد. در مدل *ARIMA* مقدار e_t از معادله 10 بدست می آید (مشکانی 1371).

$$e_t = \tilde{W}_t - f_1\tilde{W}_{t-1} - \ldots - f_p\tilde{W}_{t-p} + q_1e_{t-1} + \ldots + q_qe_{t-q} \ \ and$$
$$E[W_t] = m, \tilde{W}_t = W_t - m \qquad [10]$$

که W_t سری تفاضلی Z می‌باشد. بدلیل مشکل بودن شروع معادله تفاضلی 10 نمی‌توان W ها را بلافاصله در رابطه قرارداد تا مقادیر ε ها محاسبه شوند. اما با فرض اینکه در آغاز محاسبات P مقدار W_* از W ها و q مقدار e_* از ε معلوم باشند، در این‌صورت مقادیر $e_1, e_2, \ldots, e_n$ مشروط به مقادیر آغازین را می توان به ترتیب از معادله 10 محاسبه نمود.

در این حالت نیاز به انتخاب مقادیر آغازی برای محاسبه شرطی است و همان طورکه بیان شد وقتی تعداد نمونه ها متوسط و یا بزرگ باشد می توان یک تقریب کافی از درستنمایی غیر شرطی را با استفاده از درستنمایی شرطی و با قرار دادن مقادیر مناسب بجای e_*, W_* در معادله 10 بدست آورد. یک شیوه کار این است که عناصر e_*, W_* مساوی امید ریاضی غیر شرطی آنها قرار داده شوند. امید ریاضی غیر شرطی e_* صفر است و اگر امید ریاضی غیر شرطی W_*، صفر باشد، صفر قرار داده می شود و اگر مخالف صفر باشد می‌توان از میانگین سری بجای هر یک از عناصر W_* استفاده کرد. بنابراین روند کلی در این روش به منظور تخمین پارامترهای مربوطه به این ترتیب است که ابتدا با در نظر گرفتن شرایط اولیه و در نظر گرفتن محدوده‌هایی برای پارامترها، مجموع مربعات خطاها برای هر کدام از پارامترها محاسبه شده و پارامترهای مربوط به مینیمم مجموع به عنوان پارامتر منتخب در نظر گرفته می شوند.

درست‌نمایی غیرشرطی یا محاسبه مجموع توان‌های دوم غیر شرطی

در حالت شرطی، از مقادیر شرطی W وε جهت تخمین پارامترها استفاده می شود ولی در روش درست-نمایی غیرشرطی جهت محاسبه مقادیر صحیح W و ε در پریدهای قبل از دو معادله پسرو و پیشرو استفاده می‌شود.

معادله پیشرو

$$f(B)\tilde{W}_t = q(B)e_t \qquad [11]$$

معادله پسرو

$$f(F)\tilde{W}_t = q(F)e_t \qquad [12]$$

برای محاسبه پیش‌بینی‌های گذشته سری، ابتدا رابطه پسرو بکار برده می شود و بخاطر خاصیت ایستایی مدل خود همبسته برآوردهای $[\tilde{W}_t]$ در عمل بازاء مقدار محدودی ازw ها در نقطه t=-Q (Q: زمانی است که Wها تقریبا صفر می‌شوند) اساسا برابر صفر خواهند بود. پس به طورکلی، مجموعه دوگانی برای تولید امید ریاضی شرطی (به شرطθ وφ) با گرفتن امید ریاضی از معادلات پسرو و پیشرو حاصل می‌شود، ابتدا از رابطه $f(F)[\tilde{W}_t] = q(F)[e_t]$ استفاده می‌شود تا پیش‌بینی‌های گذشته سری ایجاد شوند و سپس از رابطه $f(B)[\tilde{W}_t] = q(B)[e_t]$ استفاده می‌شود تا مقادیر $[e_t]$ تولید شوند. بعداز محاسبه Wها وε ها از طریق رابطه پیشرو شروع به محاسبه e_t می‌شود و برای هر کدام از مجموعه پارامترها که مجموع مربعات حداقل بود، آن مجموعه انتخاب می‌شود. در این روش می توان محاسبات بازگشتی را با استفاده از روابط 13 و 14 پیش برد (مشکانی،1371)

$$[e_{-j} | f, q, W] = 0 \qquad j = 0,1,2,\ldots \qquad [13]$$

$$[e_{-j} | q, f, W] = 0 \qquad j > Q-1 \qquad [14]$$

رگرسیون فازی

منطق فازی ابزاری توانمند جهت حل مسایل مربوط به سیستم های پیچیده ای که درک آنها مشکل و یا

مسایلی که وابسته به استدلال، تصمیم گیری و استنباط بشری می باشد، بشمار می آید. به همین دلیل بهترین وسیله برای مدلسازی سیستم‌هایی است که دارای پیچیدگی زیاد بوده و داده‌های کافی از آنها موجود نیست و یا اطلاعات آنها مبهم و غیر صریح می‌باشد. استفاده از این ابزار در رگرسیون فازی سبب ارتقاء این مدل از یک مدل آماری محض به سمت مدلی با قابلیت و کارآیی بیشتری با تاکید بر در نظر گرفتن عدم قطعیت حاکم بر مسئله شده است (کوره پزان 1384).

هدف اصلی در مدل رگرسیون پیدا کردن مدل ریاضی مناسب و تعیین ضرایب مدل با هدف بهترین برازش نتایج مدل رگرسیون با مقادیر مشاهداتی می باشد. اختلاف بین مقادیر مشاهداتی و تخمینی ناشی از خطای مشاهداتی است ولی این تفاوت در رگرسیون فازی ناشی از ابهام در ساختار سیستم می باشد (ین و همکاران، 1999). مدلسازی خطی فازی که به صورت معادله 15 نشان داده می شود مبین وابستگی متغیرهای ورودی و خروجی می باشد.

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \dots + \tilde{A}_n x_n \qquad [15]$$

$$\tilde{A} = (\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_n) \qquad [16]$$

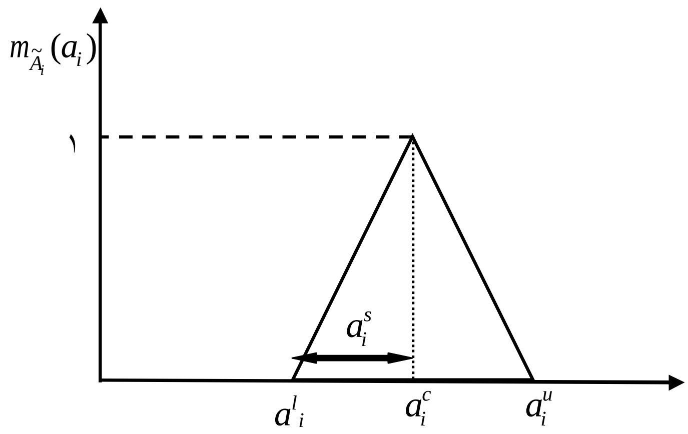
$\tilde{A}$ و $\tilde{Y}$ به ترتیب ضرایب و خروجی فازی و x مقادیر مشاهداتی می باشند. هدف این است که براساس یک مجموعه از داده های مشاهداتی ضرایب فازی معادله به گونه ای بدست آورده شود که معادله مزبور بهترین برازش بر روی مقادیر مشاهداتی را داشته باشد. در این تحقیق، اثر تابع عضویت مثلثی در دو حالت متقارن و نامتقارن به منظور درنظر گرفتن ضرایب فازی رگرسیون بررسی شده است.

رگرسیون فازی با توابع عضویت مثلثی متقارن

در این حالت توابع عضویت متغیر فازی خروجی به صورت زیر تعریف می‌شود:

$$m_{\tilde{Y}}(y) = \begin{cases} \max\limits_{\substack{(a_1,\dots,a_n) \\ (a_1,\dots,a_n)=f^{-1}(y,x)}} \left\{ \min \tilde{m}_{A_j}(a_j) \right\} & if \quad f^{-1}(y,x) \neq f \\ 0 & \text{در غیر اینصورت} \end{cases} \qquad [17]$$

ضرایب فازی با در نظر گرفتن توابع عضویت مثلثی متقارن به دو صورت $\tilde{A}_i = \{a_i^c, a_i^s\}$ و $\tilde{A}_i = \{a_i^u, a_i^l\}$ تعریف می‌شوند. که ضرایب همراه با اندیس l,u,s,c به ترتیب مرکز و پهنای تابع عضویت و حدود بالا و پایین را نشان می‌دهند. شکل 2 تابع عضویت مثلثی متقارن یک عدد فازی را نشان می‌دهد (ین و همکاران 1999).



شکل 2- تابع عضویت مثلثی متقارن ضرایب فازی

مدل رگرسیون امکانی بهترین معادله رگرسیون را با حداقل کردن میزان فازی بودن بدست می‌آورند. این کار را با حداقل کردن مجموع کل پهنای توابع عضویت ضرایب فازی معادله رگرسیون انجام می‌دهند. بنابراین برای دستیابی به بهترین برازش باید یک مدل بهینه سازی تهیه گردد. در صورتی که از توابع عضویت مثلثی برای نمایش اعداد فازی این مدل ها استفاده شود رگرسیون فازی را می توان در قالب یک مساله برنامه ریزی خطی فرموله و حل نمود. هدف مدل رگرسیون

تعیین مقادیر بهینه ضرایب می باشد به گونه ای که درجه عضویت متغیر خروجی فازی برای تمام داده ها از یک مقدار معین مانند h بزرگتر باشد.

$$m_{\tilde{y}}(y_i) \geq h \qquad j=1,\ldots,m \qquad [18]$$

$m_{\tilde{y}}$: تابع عضویت متغیر خروجی می باشد. در نهایت مدل برنامه‌ریزی خطی به صورت زیر می‌باشد:

تابع هدف:

$$Minimize \quad ma_0^s + \sum_{i=1}^{n} a_i^s \sum_{j=1}^{m} |x_{ij}| \qquad [۱۹]$$

قیود:

$$a_0^c + \sum_{i=1}^{n} a_i^c x_{ij} - (1-h)\left[a_0^s + \sum_{i=1}^{n} a_i^s x_{ij}\right] \leq y_j \qquad [۲۰]$$

$$a_0^c + \sum_{i=1}^{n} a_i^c x_{ij} + (1-h)\left[a_0^s + \sum_{i=1}^{n} a_i^s x_{ij}\right] \geq y_j \qquad [۲۱]$$

با حل مدل برنامه‌ریزی خطی فوق، ضرایب رگرسیون فازی بدست می‌آیند. با قرار دادن مقادیر بدست آمده برای ضرایب در معادله رگرسیون، مقدار متغیر خروجی به صورت فازی تعیین می‌شود. روش‌های مختلفی جهت غیر فازی سازی متغیر فازی و تبدیل آن به حالت قطعی وجود دارد که می توان به روش میانه حداکثر، روش مرکز مجموع ها و روش مرکز سطح اشاره کرد (ین و همکاران 1999 و کوره پزان 1384).

رگرسیون فازی با توابع عضویت مثلثی نامتقارن

در این حالت به علت آن که پهنای تابع عضویت در دو طرف مرکز تابع عضویت متقارن نمی باشد پس متغیر جدیدی در معرفی ضریب فازی بوجود خواهد آمد. در این حالت ضرایب فازی به صورت $\tilde{A}_i = \{a_i^l, a_i^p, a_i^R\}$ و یا $\tilde{A}_i = \{s_i^l, a_i^p, s_i^R\}$ که $S_i^R = k_i S_i^l$ تعریف می شود. S_i^R و S_i^l به ترتیب پراکندگی به سمت راست و چپ را نشان می‌دهند k_i فاکتورهای چولگی می‌باشد (شکل 3).

همچنین مدل برنامه ریزی خطی تعیین ضرایب معادله رگرسیون در ادامه آورده شده است.

تابع هدف:

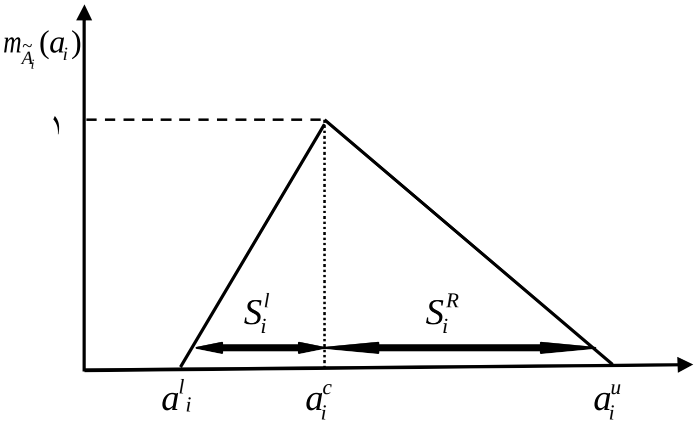
$$Minimize \quad (1+k_0)S_i^l + \sum_{i=1}^{n}\left[(1+k_i)S_i^l \sum_{j=1}^{m}|x_{ji}|\right] \qquad [22]$$

قیود:

$$(1-h)S_0^l + (1-h)\sum_{i=1} S_i^l|x_i| + \sum_i a_i^p x_i + a_0^p \leq y_j \qquad [23]$$

$$(1-h)k_0 S_0^l + (1-h)\sum_{i=1} k_i S_i^l|x_i| - \sum_i a_i^p x_i - a_0^p \leq -y_j \qquad [24]$$

در دو مدل برنامه‌ریزی مذکور، مقدار دقیق پارامتر سطح اعتماد تابع عضویت متغیرهای فازی (h) و همچنین فاکتور چولگی معلوم نیست و باید با آنالیزحساسیت بدست آید (ین و همکاران 1999 و کوره پزان 1384).



شکل 3- تابع عضویت مثلثی نامتقارن ضرایب فازی

ارزیابی مدل ها

به منظور ارزیابی عملکرد مدل می توان از معیارهای مربوطه استفاده کرد. در این تحقیق از معیارهای متوسط خطای نسبی (MRE[1]) و خطای ریشه متوسط مربعات ($RMSE$[2]) جهت این امر استفاده شده است و ارزیابی

[1]Mean relative error
[2]Root mean square error

عملکرد مدل براساس کمترین مقدار معیارها خواهد بود (ویو و همکاران 2005).

$$RMSE = \left[\frac{\sum_{i=1}^{N} (\hat{Q}_i - Q_i)^2}{N} \right]^{\frac{1}{2}} \quad [25]$$

$$MRE = \frac{\sum_{i=1}^{N} \left| \frac{Q_i - \hat{Q}_i}{Q_i} \right|}{N} \quad [26]$$

Q_i و $\hat{Q}_t$: به ترتیب جریان مشاهداتی و پیش‌بینی شده در زمان t، N: تعداد سال‌های آماری می‌باشند.

منطقه مورد مطالعه

حوضه آبریز دریاچه ارومیه براساس یکی ازتقسیم‌بندی‌های حوضه‌های کشور، جزء یکی از شش حوضه آبریز اصلی کشور است. حوضه دریاچه ارومیه با مساحت 51866 کیلومتر مربع در مختصات جغرافیایی'33°44 تا '53° 47 طول شرقی و'39 °35 تا '30 38 عرض شمالی قرار دارد. در تحقیق حاضر، به منظور ارزیابی عملکرد دو مدل *ARIMA* و رگرسیون خطی فازی، از آمار آبدهی سالانه ایستگاه های هیدرومتری ونیار بر روی رودخانه آجی‌چای با مختصات '24°46 طول جغرافیایی و'07 °38 عرض جغرافیایی (ایستگاه درجه 1) و ایستگاه ورودی سد نهند با مختصات'29°46 طول جغرافیایی و'08 °37 عرض جغرافیایی (ایستگاه درجه 3) بر روی رودخانه نهند چای استفاده شده است. آمار ایستگاه ونیار 33 سال آمار به ترتیب از سال های (42-1341) تا (74-1373) و ایستگاه ورودی سد نهند 20 سال آمار از سال 1350تا 1369 می باشد. شکل4 موقعیت منطقه مورد مطالعاتی را نشان می‌دهد.

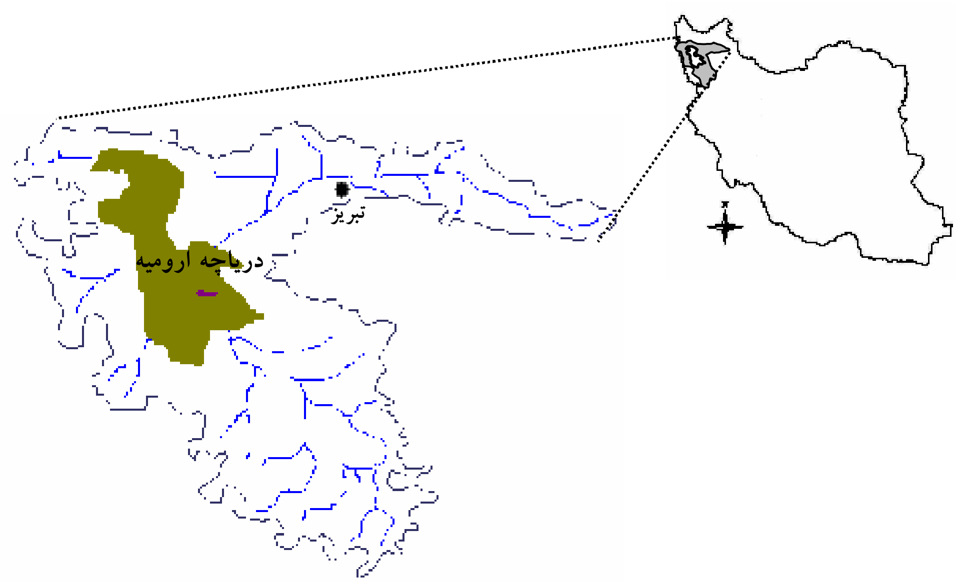
نتایج و بحث

جهت گسترش مدل های سری زمانی سه مرحله شناسایی، تخمین و برازش نکویی (شکل1) لازم می‌باشد. در آنالیز اولیه با رسم نمودار سری زمانی مشاهده می‌شود که یک بار تفاضل گیری جهت تبدیل سری زمانی ناایستا به سری زمانی ایستا کافی می‌باشد. مرحله بعدی مربوط به شناسایی و یا تعیین مر تبه‌های مدل است. یکی از روش‌های شناسایی مدل استفاده از نمودارهای تابع خودهمبستگی (همبستگی پیاپی) (*ACF*[1]) و تابع خود همبستگی جزئی (*PACF*[2]) می‌باشد. این روش پیچیده است و با افزایش مرتبه های مدل تعیین تاخیرهای نمودار آسان نمی باشد. روش دیگر برای شناسایی مدل استفاده از برخی معیارهای اطلاعاتی مانند تست معیارآکائیک (*AIC*) می‌باشد. با توجه به برازش بیشتر با سری مشاهداتی براساس حداقل مقدار*AIC* به محاسبه معیار *AIC* سری‌های زمانی جریان رودخانه ایستگاه ونیار و ایستگاه ورودی سد نهند پرداخته شد. مدل *ARIMA(1,1,1)* برای سری جریان ایستگاه ونیار و ورودی سد نهند به ترتیب کمترین مقدار *AIC* (به ترتیب374/13 و 136/75) را از بین مدل ها با مرتبه‌های گوناگون داشته است. بعد از مرحله شناسایی مدل نوبت به مرحله تخمین پارامترهای مدل می رسد. در این تحقیق جهت تخمین پارامترهای مدل*ARIMA* از روش های 1): استفاده از فرمول های ضرایب خودهمبستگی برحسب پارامترهای مدل، 2): درستنمایی شرطی و 3): درست-نمایی غیر شرطی استفاده شد.

برای هر کدام از روش‌های تخمین پارامتر مذکور، در محیط برنامه نویسی *MATLAB* برنامه‌هایی توسعه داده شد. در روش‌های درست‌نمایی شرطی و درست-نمایی غیر شرطی هدف مینیمم کردن مقدار خطای مدل (ε) می‌باشد. نمودارهای مجموه مربعات خطا در برابر پارامترهای مدل در شکل5 به عنوان نمونه برای ایستگاه

[1]Autocorrelation function
[2]Partial autocorrelation function

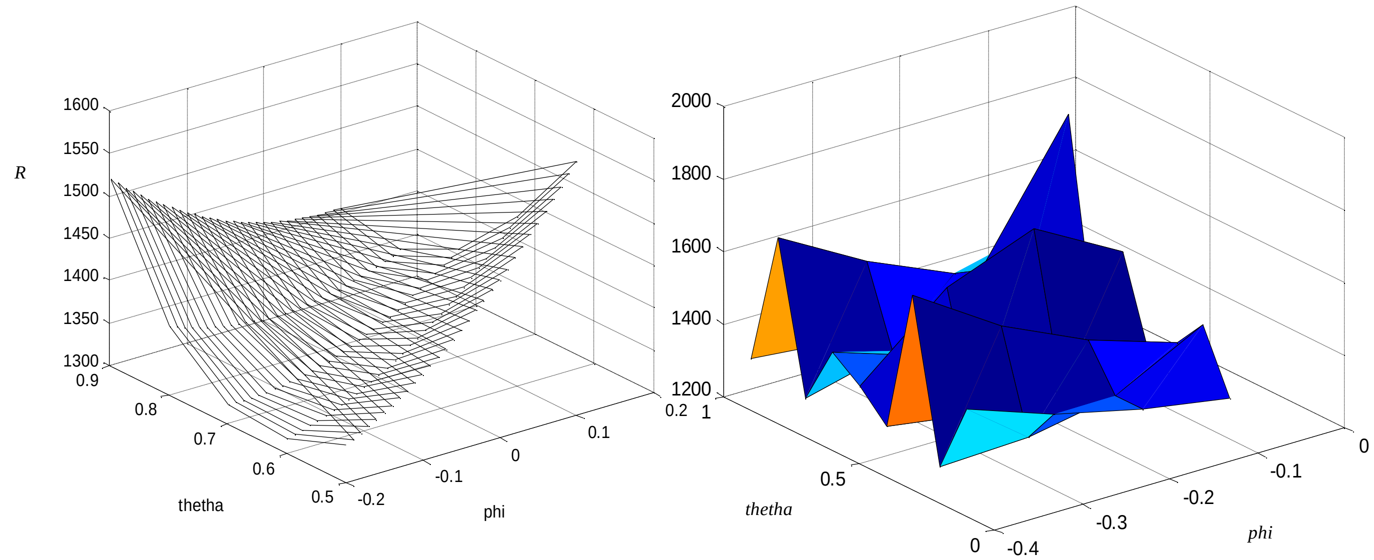
ورودی سد نهند و پارامترهای تخمینی با روش‌های بیـان شده در جدول 1 آورده شده اند.



شکل4- موقعیت منطقه مورد مطالعه

جدول1- پارامترهای تخمینی مدل *ARIMA* با روش‌های مختلف

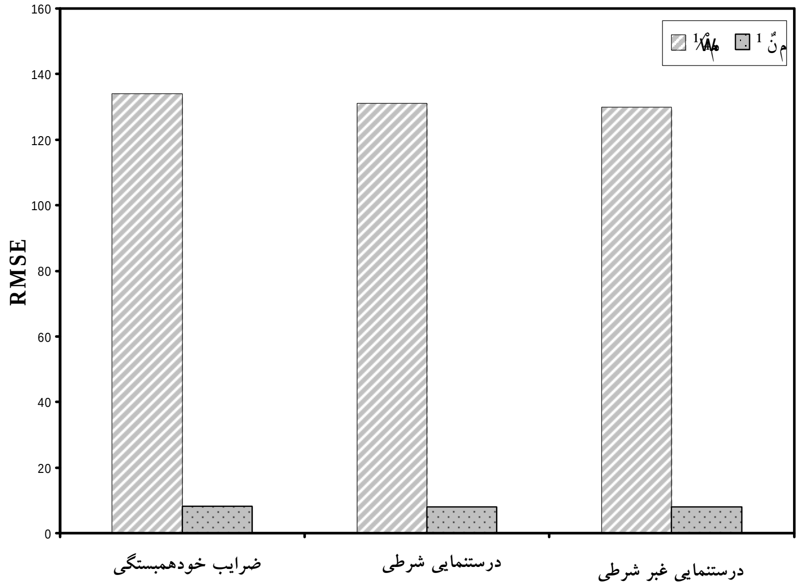
روشهای تخمین پارامتر	ایستگاه ونیار		ایستگاه نهند	
	پارامتر Φ	پارامتر θ	پارامتر Φ	پارامتر θ
فرمول‌های ضرایب خود همبستگی برحسب پارامترها	-0/062	0/9	-0/15	0/1
حداکثر درست‌نمایی شرطی	- ./1	0/9	-0/2	0/6
حداکثر درست‌نمایی غیر شرطی	-0/12	0/9	-0/15	0/58



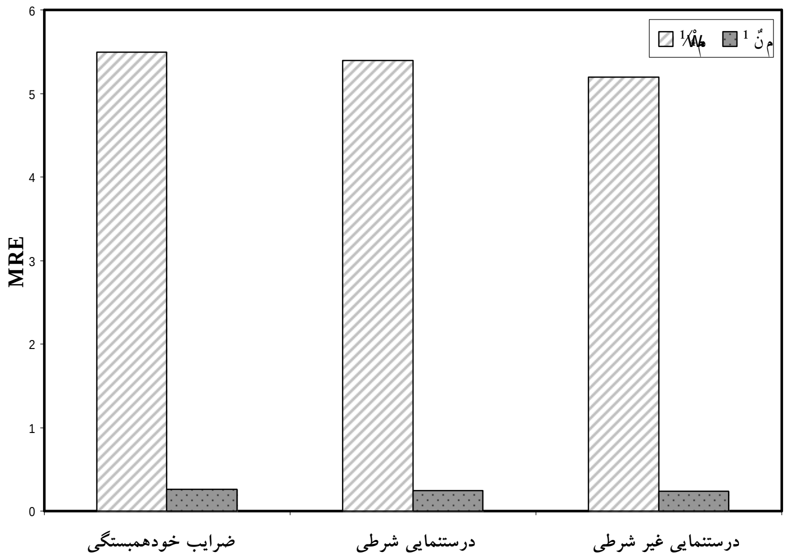
الف　　　ب

شکل 5- نمودار مجموع مربعات خطا (R) در برابر پارامترها، الف) روش شرطی، ب) روش غیرشرطی

بـه منظـور مقایـسه عملکـرد روش‌هـای تخمـین پارامترهای مـورد اسـتفاده، از پـیش بینـی مقـادیر آینـده سری اسـتفاده شـد. بـه ایـن صـورت کـه بـا اسـتفاده از مجموعـه پارامترهـای تخمـین زده شـده بـر اسـاس روش‌های مذکور، مقـادیر آینـده سـری پـیش‌بینـی شـده است. برای سری جریان ایستگاه ونیار پـیش بینـی بـرای سال های 74- 1373 تا 81-1381 و بـرای ایـستگاه نهنـد در سال های 70 -1373 انجام گرفته است. برای ارزیـابی جریان‌های پیش بینی شده بـا جریـان هـای مـشاهداتی از معیارهـای *RMSE* و *MRE* اسـتفاده شـده اسـت. مقادیر *RMSE* و *MRE* به طور شماتیک در شـکل هـای 6 و 7 نشان داده شده است. هر دو شـکل بیـانگر آن هـستند که مقادیر دو معیار از روش یک به سه کاهش می یابـد و مبین بهبود روش های تخمینی می باشد. ولی اخـتلاف دو معیار در روش دو و سه کم می باشد. معیار دیگـر جهـت ارزیابی ضریب همبستگی بوده است که مبـین ایـن نتـایج بودند. جهت مقایسه دو روش درستنمایی شـرطی و غیـر شرطی از مجموع مربعات خطـا اسـتفاده شـده اسـت کـه مقادیر مربوطه در جدول 2 آورده شـده اسـت و کمتـرین مقدار مجموع مربوط به روش تخمینـی درسـتنمایی غیـر شرطی می‌باشد.



شکل 6- ارزیابی عملکرد روش‌های تخمین پارامتر براساس معیار **RMSE**



شکل 7- ارزیابی عملکرد روش‌های تخمین پارامتر براساس معیار **MRE**

جدول 2- مقادیر مجموع مربعات خطا

روشهای تخمین پارامترهای مدل موردنظر	ایستگاه ونیار	ایستگاه نهند
حداکثر درست‌نمایی شرطی	183850	1317/3
حداکثر درست‌نمایی غیر شرطی	182000	1292/6

رگرسیون فازی

دراین حالت رگرسیون فازی تک و دو متغیره مورد بررسی قرار گرفت به گونه‌ای که در حالت تک متغیره مقدار جریان سالانه در زمان حاضر تنها به یک زمان قبل مربوط می‌شود و در حالت دو متغیره جریان در حال حاضر حداکثر به دو گام زمانی قبل وابسته است که می توان به صورت روابط (27) و (28) بیان کرد.

$$(y_t^L, y_t^P, y_t^R)\times Q_t = (a_t^L, a_t^P, a_t^R)\times Q_{t-1} + (c^L, c^P, c^R) \qquad [27]$$

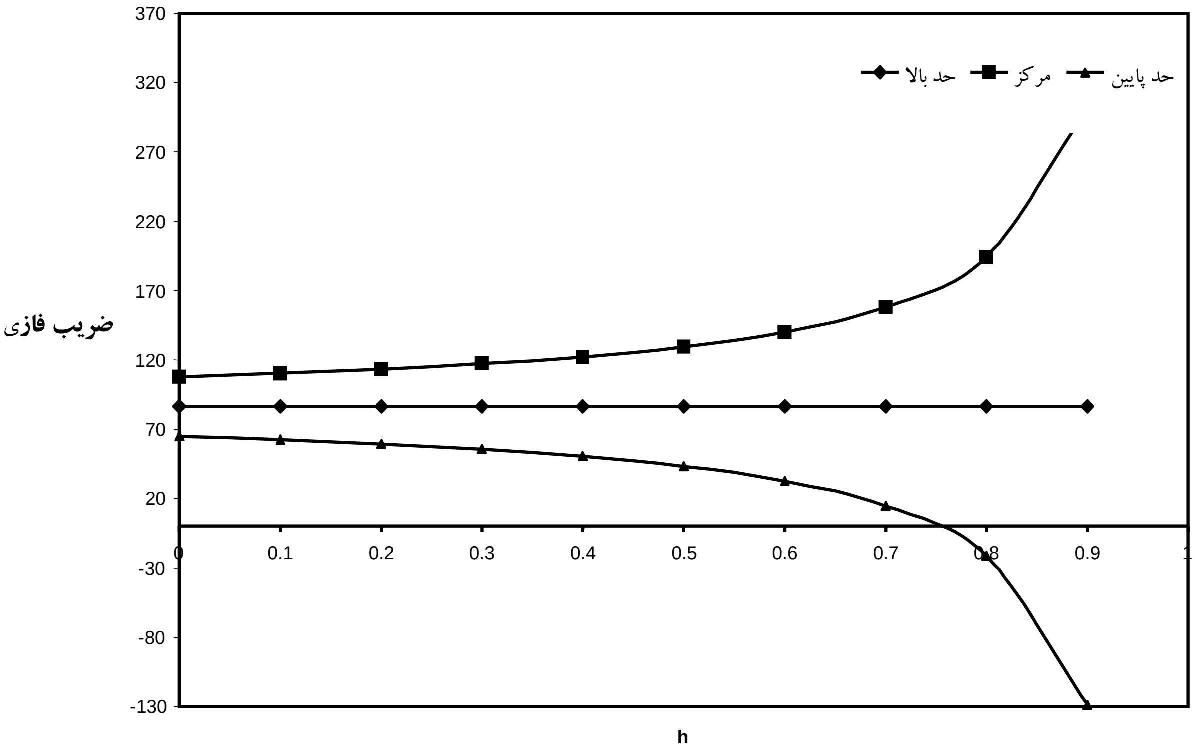
$$(y_t^L, y_t^P, y_t^R)\times Q_t = (a_t^L, a_t^P, a_t^R)\times Q_{t-1} + (b_t^L, b_t^P, b_t^R)\times Q_{t-2} + (d^L, d^P, d^R) \qquad [۲۸]$$

رگرسیون فازی با توابع عضویت متقارن

نتایج حاصل از بهینه سازی پارامترهای فازی که بیانگر ضرایب رگرسیون می باشند بازاء مقادیر مختلف h برای حالت‌های تک و دو متغیره مورد بررسی قرار گرفته اند. ضرایب رگرسیون ایستگاه ونیار در حالت تک متغیره در جدول 3، مرکز و پهنای تابع عضویت حالت دو متغیره در شکل 8 آورده شده است. پارامتر مربوط به پهنای تابع عضویت میزان فازی بودن هر متغیر را نشان می دهد. انجام آنالیز حساسیت بر روی h بیانگر این است که تغییری در مراکز فازی متغیرهای رگرسیون ایجاد نشده ولی بر میزان گستردگی این متغیرها تاثیر می گذارد به طوری که با افزایش سطح اعتماد (h) بر میزان گستردگی متغیرهای رگرسیون (پهنای اعداد فازی) افزوده می‌شود. در تعیین ضرایب رگرسیون فازی با روش بهینه سازی با استفاده از برنامه ریزی خطی از نرم افزار *Lingo4* استفاده شده است.

جدول 3- ضرایب رگرسیون فازی بازای مقادیر مختلف سطح اعتماد و عملکرد مدل

RMSE	a_0^c	a_0^s	a_1^c	a_1^s	h
207/52	71/92	2/81	0/92	./88	0
196/87	71/92	3/12	0/92	./98	0/1
184/89	71/92	3/51	0/92	1/1	0/2
170/17	71/92	4/01	0/92	1/26	0/3
152/83	71/92	4/68	0/92	1/47	0/4
131/42	71/92	5/62	0/92	1/77	0/5
105/96	71/92	7/03	0/92	2/21	0/6
75/38	71/92	9/37	0/92	2/95	0/7
46/5	71/92	14/06	0/92	4/42	0/8
68/6	71/92	28/12	0/92	8/85	0/9



شکل 8- تغییرات مرکز و گستردگی حالت دو متغیره بازای مقادیر مختلف سطح اعتماد (**h**)

جهت پیدا کردن مقدار بهینه پارامتر سطح اعتماد بایستی عملکرد سیستم تحت تاثیر مقادیر مختلف این پارامتر تعیین گردد به این صورت که بازاء مقادیر مختلف h مقادیر خروجی محاسبه و سپس از طریق معیارهایی مانند *RMSE* گزینه برتر معین می شود. قابل ذکر است که روش مورد استفاده در این تحقیق جهت تبدیل متغیر خروجی از حالت فازی به حالت قطعی، روش مرکز سطح می باشد. جدول4 مقادیر بهینه پارامتر سطح اعتماد را در دو ایستگاه مورد بررسی در دو حالت تک متغیره و دو متغیره نشان می‌دهد.

جدول 4- تعیین سطوح بهینه براساس معیار *RMSE*

H بهینه	ایستگاه	دو متغیره	تک متغیره
0/4	نهند	-	6/69
0/1	نهند	6/5	-
0/8	ونیار	-	46/5
0/8	ونیار	46/1	-

به عنوان مثال ساختار رگرسیونی حالت تک و دو متغیره ایستگاه ونیار به ترتیب براساس روابط 29 و 30 می باشد.

$$Q_t = (0.92, 4.42) \times Q_{t-1} + (71.92, 14.06) \quad [29]$$

$$Q_t = (0.022, 0) \times Q_{t-1} + (0.71, 0.63) \times Q_{t-2} + (86.37, 21.52) \quad [30]$$

مقادیر *RMSE* در حالت دو متغیره نسبت به تک متغیره کمتر می باشد چرا که وابستگی به دو زمان قبل در

عملکرد مدل اهمیت بسزایی دارد ولی اختلاف بین این مقادیر کم می‌باشد و دلیل آن در صفر بودن پارامترهای گستردگی در ضریب مربوط به حالت دو متغیره می باشد.

رگرسیون فازی با توابع عضویت نامتقارن

در این حالت علاوه بر تغییرات h، تغییرات ضریب چولگی نیز باید در نظر گرفته شود. به اینصورت که ابتدا بازاء فاکتور چولگی مختلف، فاکتور چولگی که بهترین عملکرد سیستم را نشان دهد انتخاب می شود و سپس آنالیز بر روی فاکتور سطح اعتماد انجام می گیرد. در این تحقیق برای برخی از فاکتورهای چولگی متوسط $RMSE$ در جدول 5 محاسبه شده و با مقایسه این مقادیر فاکتور چولگی با مقادیر 0/1 مورد قبول واقع شد. همچنین مقادیر $RMSE$ در حالت دو متغیره نسبت به تک متغیره کاهش یافته است و مقادیر h بهینه برای هر دو حالت در ایستگاه ونیار 0/9 و در ایستگاه نهند برای حالت تک متغیره 0/1 و برای حالت دو متغیره 0/5 می باشد.

با توجه به ضرایب محاسبه شده مشاهده می گردد که با افزایش فاکتور چولگی مقادیر a_0^s کاهش می یابد و در بین مقادیر ضریب چولگی در جدول 6 k_0 را می توان به عنوان ضریب چولگی غالب دانست. نکاتی که در مقایسه بین نتایج می توان بدست آورد به این صورت می باشد که عملکرد رگرسیون فازی با توابع عضویت متقارن بهتر از رگرسیون فازی با توابع عضویت نا متقارن می باشد که البته حالت دو متغیره هر کدام گزینه منتخب است.

جدول 5- مقادیر $RMSE$ با فاکتورهای چولگی مختلف ونیار

متوسط $RMSE$	فاکتور چولگی
241/12	$k_0 = 1.4, k_1 = 1.6$
183/61	$k_0 = 1.2, k_1 = 1.1$
160/75	$k_0 = 1.1, k_1 = 1.1$
237/04	$k_0 = 1.4, k_1 = 1.6, k_2 = 1.9$
178/18	$k_0 = 1.2, k_1 = 1.1, k_2 = 1$
150/005	$k_0 = 1.1, k_1 = 1.1, k_2 = 1.1$

جدول 6- تغییرات a_0^s نسبت به فاکتورهای چولگی

$k_0 = 1.1\ ,\ k_1 = 1.1$	$k_0 = 1.2\ ,\ k_1 = 1.1$	$k_0 = 1.4, k_1 = 1.6$	***H***
171/86	164/03	150/37	0
190/15	182/27	167/088	0/1
214/82	205/64	187/97	0/2
245/51	234/35	214/82	0/3
286/43	273/41	250/63	0/4
343/72	328/1	300/75	0/5
429/65	410/12	375/94	0/6
572/87	546/83	501/26	0/7
859/31	820/25	751/89	0/8
1718/622	1640/5	1503/79	0/9

در این تحقیق استفاده از روش حداکثر درستنمایی در تخمین پارامترهای مدل *ARIMA* منجر به روش کمترین مربعات شده است. جهت مینیمم کردن مجموع مربعات خطا از دو روش درستنمایی شرطی و غیر شرطی استفاده شد. در بین روش های تخمین پارامترهای مدل *ARIMA* روش مربوط به استفاده از فرمول های ضرایب خود همبستگی برحسب پارامترهای مدل نتایج قابل قبولی داشت. ولی به بعلت این که امکان بدست آوردن روابط بین پارامترهای مدل و ضرایب خودهمبستگی و حل آنها با افزایش مرتبه های مدل مشکل می باشد، بهمین جهت کمتر استفاده می شود. در روش مربوط به درستنمایی شرطی با مقادیر اولیه میزان محاسبات و برنامه نویسی راحت تر می باشد ولی عدم استفاده از مقادیر گذشته سری در این روش در نتایج تاثیر می گذارد. روش درستنمایی غیر شرطی با استفاده از دو معادله پسرو و پیشرو منجر به تعیین مقادیر صحیح سری زمانی(در پریدهای قبلی) در محاسبه مجموع مربعات خطا می شود.

رهیافت دیگر استفاده از روش رگرسیون فازی در برآورد جریان می باشد. در این تحقیق از وابستگی جریان به یک و دو گام زمانی استفاده شده است. ارزیابی نتایج با معیارهای مربوطه مبین مینیمم مقدار این معیارها در وابستگی دو متغیره می باشد. نتایج نشان دادند که با نزدیکی مقدار سطح اعتماد به عدد یک، میزان فازی بودن مدل افزایش می یابد (چنگ و ایوب 2001).

مقایسه نتایج حاصل از مدل *ARIMA* با روش تخمینی درستنمایی غیر شرطی با رگرسیون فازی حاکی از دقت بالا در رگرسیون فازی می باشد (تسنگ و همکاران، 2001 و تولی و ونگ، 1999) چنانچه مقایسه معیارهایی مانند *RMSE* نیز مبین این مطلب می باشد. دلیل این مطلب را شاید این گونه بیان کرد که مدل های استوکاستیکی عدم قطعیت ناشی از تصادفی بودن پدیده را در نظر می گیرند اما در مدلسازی عدم قطعیت هایی ظاهر می شود که ماهیت تصادفی ندارند. تئوری مجموعه های فازی یک ابزار نیرومند در بیان پیچدگی های موجود در جهان واقعی می‌باشد که قادر است خلاء ناشی از عدم لحاظ کردن واقعیت حاکم بر بهره برداری در فرمول بندی مدل را پر نماید.

نتیجه گیری

در تحقیق حاضر از مدل سری زمانی *ARIMA* همراه با روش های مختلف تخمین پارامتر و رگرسیون فازی در برآورد جریان های سالانه استفاده شده است. تخمین پارامترهای مدل یکی از مراحل مهم مدلسازی می باشد. به همین جهت تعیین روش های پارامترهای مدل دارای اهمیت می باشد. همان طور که نتایج نشان دادند روش درستنمایی غیر شرطی جواب های قابل قبولی را در مقایسه با سایر روش ها براساس معیارهای مورد بررسی داشته است (میونگ 2003). رگرسیون فازی با توابع عضویت متقارن نسبت به نامتقارن جواب های متناسب با معیارها را داشت. در حالت دو متغیره نیز اختلاف بین جریان مشاهداتی و محاسباتی کم بود و k_0 به عنوان فاکتور چولگی غالب شناسایی شد. مقایسه مدل سری زمانی *ARIMA* و رگرسیون فازی با توجه به معیارهایی مانند *RMSE* بیانگر برازش بهتر مقادیر جریان پیش بینی شده توسط رگرسیون فازی می باشد.

منابع مورد استفاده

جلال کمالی ا، محمودیان شوشتری م و جلال کمالی ن، 1385. پیش بینی جریان ماهانه ورودی به مخزن سد شهید عباسپور با استفاده از مدل های سری زمانی *Box-Jenkins* . هفتمین سمینار بین المللی مهندسی رودخانه. دانشگاه شهید چمران اهواز.1-7.

کوره پزان ا،1384. اصول تئوری مجموعه های فازی و کاربردهای آن در مدلسازی مسایل مهندسی منابع آب. انتشارات جهاد دانشگاهی واحد صنعتی امیر کبیر.

مشکانی م، 1371. تحلیل سری زمانی :پیش بینی و کنترل (ترجمه). انتشارات دانشگاه شهید بهشتی .

Box EP and Jenkins GM, 1976. Time Series Analysis: Forecasting and Control. Prentice-Hall, Englewood Cliffs,NJ.

Brockwell PJ and Davis RA, 1987. Time series: Theory and Method. Springer-Verlag, New York.

Carlson RF, MacCormick AJA and Watts DG, 1970. Applications of linear models to four annual streamflow series. Water Resoure Research 6: 1070-1078.

Chang Y and Ayyub B, 2001. Fuzzy regression methods- a comparative assessment. Fuzzy Sets and Systems 119: 187-203.

Delleur JW, Tao PC and Kavvar ML, 1976. An evaluation of the practically and complexity of some rainfall and runoff time series models. Water Resoure Research 12 (5), 953-970.

Hojati M, Bector CR and Smimou K, 2005. A simple method for computation of fuzzy linear regression. European Journal of Operational Research 166:172-184.

Kurunc A, Yurekli K and Cevik O, 2005. Performance of two stochastic approaches for forecasting water quality and streamflow data Yesilrmak River, Turkey. Environmental Modeling & Software. 20: 1195-1200.

Montanari A, Rosso R and Taqqu S, 1997 .Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation. Water Resources Research 3(5): 1035-1044.

Myung I, 2003. Tutorial on maximum likelihood estimation. Journal of Mathematical Psychology 47: 90-100.

Obadage AS and Harnpornchai N, 2006. Determination of point maximum likelihood in failure domain using genetic algorithm. International Journal of Pressure Vessels and Piping 83: 276-282.

Radden DT and Woodall WH, 1994. Properties of certain fuzzy linear regression methods. Fuzzy Sets and Systems 64: 361-375.

Salas JD, Delleur JW, Yevjevich V and Lane WL, 1988. Applied Modeling of Hydrologic Time Series, Water Resource Publication (WRP) 192-194.

Shine DW and Lee JH, 2000. Consistency of the maximum likelihood estimators for nonstationary ARMA regressions with time trends. Journal of Statistical Planning and Inference 87: 55-68.

Tanaka H and Uejima S, 1982. Linear regression analysis with fuzzy model. IEEE Trans. Systems, Man, Cybernet 12:903-907.

Toly C and Wang MJ, 1999. Forecasting methods using fuzzy concepts. Fuzzy sets and systems. 105 (3): 339-352 .

Tseng YH, Durbin P and Tzeng GH, 2001. Using fuzzy piecewise regression analysis to predict the nonlinear time series of turbulent flows with automatic chang- point detection. Flow, Turbulence and Combustion. 67: 81-106.

Valenzuela O, Marquez L, Pasadas M and Rojas I, 2004. Automatic identification of ARIMA time series by expert systems using paradigms of artificial intelligence. Mongrafias del Seminaro Garcia de Galdeeano 31:425-435.

Wu, HCH, 2003. Fuzzy estimates of regression parameters in linear regression models for imprecise input and output data. Computational Statistics & Data Analysis 42: 203-217.

Wu JS, ASCE PEM, Han J, Annambhotla S and Bryant S, 2005. Artificial neural networks for forecasting watershed runoff and streamflows. Journal of Hydrology Engineering. 10 (3):216-222.

Yen KK, Ghoshary S and Roig G, 1999.A linear model using triangular fuzzy number coefficients. Fuzzy sets and systems 106:167-177.

Zadeh LA, 1965. Fuzzy sets and information. Control 8, 338-353.