

پیش‌بینی سطح ایستابی با استفاده از سری‌های زمانی و سیستم استنباط فازی-عصبی تطبیقی

بهرام چوبین^۱، آرش ملکیان*^۲، فرزانه ساجدی حسینی^۳، امید رحمتی^۴

۱. دانشجوی دکتری، ساری، دانشگاه علوم کشاورزی و منابع طبیعی ساری

۲. استادیار پردیس کشاورزی و منابع طبیعی، دانشکده منابع طبیعی، دانشگاه تهران

۳. دانشجوی کارشناسی ارشد، ساری، دانشگاه علوم کشاورزی و منابع طبیعی ساری

۴. دانشجوی دکتری، لرستان، دانشگاه لرستان

(تاریخ دریافت: ۱۳۹۲/۸/۵ - تاریخ تصویب: ۱۳۹۳/۲/۱)

چکیده

مدل‌سازی در مناطق خشک برای مدیریت بهینه منابع آب اهمیت ویژه‌ای دارد. آب زیرزمینی از مهم‌ترین منابع آبی در مناطق خشک محسوب می‌شود. هدف این پژوهش ارزیابی عملکرد سیستم استنباط فازی عصبی تطبیقی (انفیس) و مدل‌های سری زمانی در پیش‌بینی سطح ایستابی است. در این پژوهش، با استفاده از مدل‌های سری زمانی و مدل انفیس با توابع عضویت مختلف اقدام به پیش‌بینی یک ماه بعد سطح آب‌های زیرزمینی دشت شیراز شد. بهترین ترکیب ورودی و طول داده‌های آموزشی و صحت‌سنجی در مدل انفیس با استفاده از آزمون گاما و M برآورد شد. عملکرد مدل‌های مختلف با پارامترهای خطا و دیاگرام تیلر مقایسه شد. نتایج مدل انفیس نشان داد که این مدل با تابع عضویت II شکل عملکرد بهتری نسبت به بقیه توابع عضویت دارد ($RMSE = 1/241$ و $MAE = 0/953$). مقایسه عملکرد مدل‌ها، حاکی از کارایی بسیار مناسب مدل خطی $ARIMA(2,1,2)$ نسبت به مدل انفیس با توابع عضویت مختلف است ($MAE = 0/325$ و $RMSE = 0/241$).

کلیدواژگان: آزمون گاما، انفیس، دیاگرام تیلر، $ARIMA$

مقدمه

ایران کشوری با متوسط ریزش‌های جوی ۲۵۲ میلی‌متر است و در زمره مناطق خشک و نیمه‌خشک جهان طبقه‌بندی می‌شود. محدودیت منابع آبی و نبود آب مناسب از عمده‌ترین تنگناها و مشکلات کشاورزی در کشور است. از سوی دیگر، افزایش روزافزون جمعیت نیاز به تولید مواد غذایی بیشتر را ایجاد می‌کند. بنابراین، یافتن راهکارهایی برای غلبه بر کمبود آب در این مناطق همواره از اولویت‌های مطالعاتی و پژوهشی است تا بتوان محصولات کشاورزی را برای این جمعیت رو به تزاید تولید کرد (Afuni et al., 2002). یکی از فاکتورهای مهم در مدیریت داشتن یک دید و نگرش مناسب از اتفاقات آینده است. در مدیریت منابع آب این امر مستثنا نیست و آگاهی از وضعیت منابع آب در یک منطقه نقش تعیین‌کننده‌ای در برنامه‌ریزی‌های کشاورزی دارد. خصوصاً اگر بتوان با استفاده از تحلیل‌های آماری، مدل‌های ریاضی و شرایط منابع آب در آینده را نیز پیش‌بینی کرد. از این‌رو، می‌توان برای مواجهه بهتر و مدیریت

بهینه و صحیح در آب‌های زیرزمینی با استفاده از تکنیک‌های مدل‌سازی به پیش‌بینی نوسانات سطح آب‌های زیرزمینی در ماه‌های آینده پرداخت.

مطالعات در زمینه آب‌های زیرزمینی نشان می‌دهد که سری زمانی و سیستم‌های هوشمند مناسب‌ترین مدل‌های استفاده‌شده برای پیش‌بینی سطح ایستابی هستند (Ahn, 2000; Daliakopoulos et al., 2005; Wong et al., 2007). در سال‌های اخیر، از مدل $ARIMA$ برای پیش‌بینی پارامترهای هیدرومتئولوژیکی استفاده شده است (Abghari, 2010; Boochabun et al., 2004; Chattopadhyay and Chattopadhyay, 2010; Chattopadhyay et al., 2011; Wong et al., 2007). در دشت همدان- بهار از اطلاعات سطح آب زیرزمینی طی سال‌های ۱۳۶۲-۱۳۸۱ استفاده شد و سپس با استفاده از مدل $ARIMA(1,1,0)$ برای پیش‌بینی سطح آب زیرزمینی در ۲۰ سال آینده استفاده شد. نتایج حاصل از مقادیر پیش‌بینی شده به‌وسیله مدل سری زمانی، نشان می‌دهد که در صورت ثابت‌ماندن الگوی مصرف و هم‌چنین عدم تغییرات در روند تغذیه سفره دشت همدان- بهار طی بیست سال آینده با کاهش حدود ۱۷/۵ متر نسبت به وضعیت فعلی سطح سفره

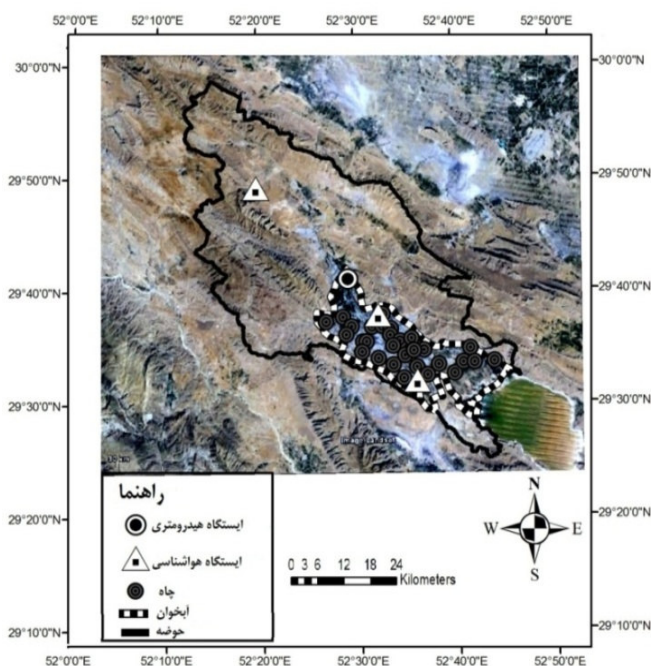
* نویسنده مسئول: malekian@ut.ac.ir

توجه به محدودیت منابع آبی در کشور، پیش‌آگاهی از وضعیت آب‌های زیرزمینی می‌تواند کمک شایانی به مدیریت منابع آبی به‌خصوص در مناطق خشک و نیمه‌خشک داشته باشد.

روش پژوهش

منطقه پژوهش‌شده

حوضه شیراز با سطح برابر با ۱۴۵۰ کیلومترمربع در مختصات ۵۲ درجه و ۱۲ دقیقه تا ۵۲ درجه و ۴۵ دقیقه طول جغرافیای و ۲۹ درجه و ۲۵ دقیقه تا ۲۹ درجه و ۵۸ دقیقه عرض جغرافیای در استان فارس واقع شده است. متوسط بلندمدت بارندگی در سطح حوضه ۳۵۰ میلی‌متر است. بررسی‌های انجام‌شده نشان داده است که آبرفت دشت شیراز به‌صورت لایه لایه است و لایه‌های رسی بین لایه‌های آبدار قرار گرفته است به‌نحوی که رسوبات آبرفتی ضخامت یکنواختی ندارد و لایه‌های شنی بین لایه‌های رسی و سیلنتی قرار دارند (The Water comprehensive plan, 1991). داده‌های ورودی استفاده‌شده در این پژوهش شامل داده‌های ماهیانه از ۵ متغیر: مجموع بارش (P)، متوسط جریان ورودی به دشت (SF)، دما (T)، تبخیر (E) و سطح آب‌های زیرزمینی (Shirmohammadi *et al.*, 2013) طی دوره زمانی هجده‌ساله (۱۳۷۲-۱۳۸۹) است. شکل (۱) دشت و آبخوان شیراز و هم‌چنین موقعیت چاه‌های پیژومتری (تعداد ۲۹ حلقه چاه)، ایستگاه هیدرومتری و ایستگاه‌های هواشناسی را نشان می‌دهد.



شکل ۱. موقعیت آبخوان دشت شیراز و حوضه بالادست آن

مواجه خواهیم شد (Rahmani and Sedehi, 2005). در دشت شهرکرد با استفاده از ۱۷ مدل ARIMA نوسانات آب زیرزمینی پیش‌بینی شد و نتایج نشان داد که با ادامه روند کنونی، آبخوان شهرکرد به‌طور متوسط طی سال‌های آینده، در هر سال ۶۹ میلیون مترمکعب از حجم خود را از دست خواهد داد (Mirzaei *et al.*, 2006). در کره جنوبی نیز از سری‌های زمانی برای پیش‌بینی سطح آب زیرزمینی ماهانه استفاده شده است، این مطالعه براساس روش باکس و جنیکز (Box-Jenkins) بوده و نتایج نشان داده است که آریمای فصلی عملکرد رضایت‌بخشی در پیش‌بینی آب زیرزمینی دارد (Lee *et al.*, 2009). در ترکیه از مدل شبکه عصبی بازگشتی، ARIMA و مدل هیبرد (شبکه عصبی و ARIMA) برای پیش‌بینی کیفیت آب طی سال‌های ۱۹۹۶-۲۰۰۶ استفاده شد و نتایج نشان داد که مدل ARIMA عملکرد مناسبی دارد، درحالی‌که مدل هیبرد (شبکه عصبی و ARIMA) عملکرد بهتری از مدل شبکه عصبی بازگشتی دارد (Faruk, 2010). انفیس (ANFIS) یک مدل برآوردگر جهانی است که توانایی پیش‌بینی هر تابع پیوسته را با درجه‌ای از دقت و صحت دارد (Shirmohammadi *et al.*, 2013). نتایج پیش‌بینی مصارف آب در ترکیه با استفاده از سیستم‌های استنتاج فازی نشان داد که سیستم استنتاج فازی عصبی تطبیقی (ANFIS) نسبت به سیستم استنتاج فازی ممدانی (MFIS) با موفقیت بیشتری قادر به پیش‌بینی است (Firat *et al.*, 2009). در سال ۲۰۱۳ پژوهشی در غرب استان جیلین در کشور چین برای افزایش درک از ویژگی‌های پویای سطح آب‌های زیرزمینی استفاده شد، از داده‌های سطح آب‌های زیرزمینی ماهانه سه حلقه چاه در سال‌های ۱۹۸۶ تا ۲۰۱۱ استفاده شد. نتایج نشان داد که سطح آب زیرزمینی روند کاهشی در حدود ۷ سال دارد. علاوه بر این، هیدروگراف سه حلقه چاه، تأثیرات رفتار انسان بعد از سال ۱۹۹۵ در کاهش سطح آب‌های زیرزمینی را نشان داد. در مقایسه با مدل یکپارچه میانگین متحرک اتورگرسیو (ARIMA)، روش تجزیه (Decomposition method) برای پیش‌بینی سطح آب زیرزمینی توصیه می‌شود (Lu, *et al.*, 2013). چندین پژوهش دیگر در زمینه مدیریت منابع آب‌های زیرزمینی از مدل انفیس استفاده کرده‌اند (Chu and Chang, 2009; Firat *et al.*, 2009; Kisi and Shiri, 2012; Tutmez *et al.*, 2006). پیش‌بینی سطح ایستابی (یک ماه به جلو) با کمک سری زمانی و سیستم استنباط فازی عصبی-تطبیقی (انفیس) انجام شد. از نقاط قوت این مطالعه، بهره‌گیری از آزمون‌های گاما و M برای تعیین ترکیبات مختلف ورودی در مدل‌های غیرخطی است. با

$$O_i^2 = w_i = \mu_{A_i}(x)\mu_{B_i}(y) \quad i = 1,2 \quad (\text{رابطه ۳})$$

لایه ۳: گره i ام این لایه که با N نام گذاری شده که نسبت قوانین نرمال شده را محاسبه می کند:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1,2 \quad (\text{رابطه ۴})$$

لایه ۴: گره i در این لایه مشارکت قانون i ام را به سمت خروجی مدل با استفاده از تابع زیر محاسبه می کند.

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (\text{رابطه ۵})$$

که در آن: \bar{w}_i ، خروجی لایه ۳ و $\{p_i, q_i, r_i\}$ مجموعه پارامترها است.

لایه ۵: تنها گره این لایه، خروجی کلی ANFIS به عنوان جمع بندی از تمام ورودی های مدل، که غیرفازی شده است مطابق زیر محاسبه می کند:

$$O_i^5 = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (\text{رابطه ۶})$$

آزمون گاما

آزمون گاما (Gamma Test) نوعی مدل سازی غیرخطی و ابزاری برای آنالیز است و اجازه می دهد رابطه بین ورودی ها و خروجی ها در یک مجموعه داده های عددی امتحان شود. این آزمون می تواند کمترین مقدار قابل محاسبه میانگین مربعات خطا (آماره گاما δ) را برای هر مدل ارائه دهد، در اصل تمام ترکیب های ورودی محتمل را بررسی می کند و میزان صحت و دقت مدل سازی را در ترکیب های مختلف تعیین می کند (Agalbjörn *et al.*, 1997; Končar, 1997). در این پژوهش پارامتر مقایسه شده برای ترکیب مختلف ورودی مقدار آماره گاما δ بود که ما برای اطمینان بیشتر از نسبت (V ratio) هم استفاده شد. نسبت (V ratio) از تقسیم آماره گاما بر واریانس داده های خروجی حاصل می شود. کوچک تر بودن مقادیر آماره گاما و نسبت V نشان دهنده بهترین ترکیب ورودی است (Agalbjörn *et al.*, 1997; Končar, 1997).

آزمون M

آزمون (M-Test) نوعی روش اندازه گیری برای برآورد قابلیت اعتماد آماره گاما است، که با توجه به داده های موجود با برآورد آماره گاما حاصل می شود. این آزمون تعداد داده های مورد نیاز برای یک مدل مطلوب را تعیین می کند. آزمون گاما و M هر دو در نرم افزار WinGamma قابل اجرا است.

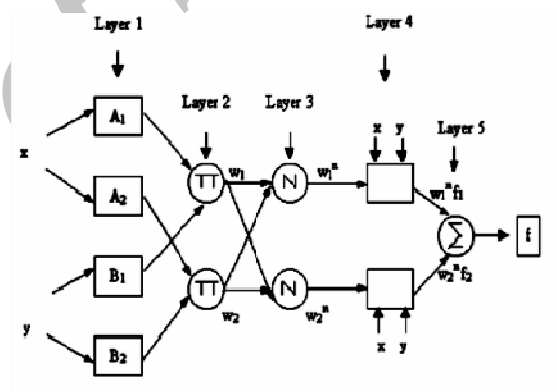
سیستم استنباط فازی-عصبی تطبیقی

انفیس اولین بار توسط جانگ معرفی شد (Jang, 1993). نکته اصلی هنگام طراحی یک مدل انفیس (ANFIS)، انتخاب سیستم استنتاج فازی (FIS)، است. برای یک مدل فازی سوگنو (Sugeno) درجه اول، با دو ورودی x و y و خروجی f مجموعه قوانین به صورت IF-Then، به صورت زیر ارائه می شود:

قانون اول: اگر x مساوی A_1 و y مساوی B_1 باشد، آن گاه $p_1 x + q_1 y + r_1 = f_1$

قانون دوم: اگر x مساوی A_2 و y مساوی B_2 باشد، آن گاه $p_2 x + q_2 y + r_2 = f_2$

که در آن: A_1, A_2, B_1, B_2 به ترتیب توابع عضویت برای ورودی های x و y هستند (برای مثال توابع گوسی، مثلثی و... که نوع تابع عضویت قابل تغییر است). $p_1, q_1, r_1, p_2, q_2, r_2$ پارامترهای تابع خروجی هستند (Jang, 1993). ساختار معمول ANFIS در شکل ۲ ارائه شده است.



شکل ۲. معماری مدل ANFIS (Jang, 1993)

کارکرد ANFIS بدین شرح است:

لایه ۱: هر گره در این لایه درجات عضویت از یک متغیر ورودی را تولید می کند:

$$O_i^1 = \mu_{A_i}(x) \quad i = 1,2 \quad (\text{رابطه ۱})$$

$$O_i^1 = \mu_{B_{i-2}}(y) \quad i = 3,4 \quad (\text{رابطه ۲})$$

که در آن: x (یا y) ورودی گره مورد نظر، A_i یا B_{i-2} مجموعه فازی مرتبط با این گره است. μ_{A_i} و μ_{B_i} توابع عضویت مناسب برای مجموعه فازی A_i یا B_i هستند. O_i خروجی مربوط به هر لایه است.

لایه ۲: هر گره در این لایه در سیگنال ورودی ضرب می شود و خروجی نتیجه ای از تمامی سیگنال های ورودی است:

مدل خودهمبسته میانگین متحرک تجمعی

یکی از فرضیات اساسی در مدل‌سازی سری‌های زمانی هیدرولوژیکی ایستابودن سری زمانی است. در این پژوهش از دو روش رایج (ADF Unit root test) یا آزمون ریشه واحد (Dickey and Fuller, 1979) و روش Phillip - Perron یا PP (Phillips and Perron, 1988) برای آزمون ایستایی سری‌های زمانی سطح ایستایی با نرم‌افزار Eviews استفاده شد. توسعه مدل‌های سری زمانی شامل سه مرحله شناسایی، برآورد پارامترها و آزمون کفایت مدل است (Box and Jenkins, 1976). مرحله شناسایی شکل عمومی مدل با توجه به رفتار سری زمانی تعیین می‌شود. در مرحله برآورد پارامتر، پارامترهای مدل انتخاب شده در مرحله قبل، محاسبه شده و در نهایت در مرحله آزمون کفایت مدل، توانایی مدل منتخب در مدل‌سازی سری زمانی سنجیده می‌شود.

در این پژوهش سری زمانی داده‌های ماهیانه سطح ایستایی مربوط به دوره آماری ۱۳۷۲ تا ۱۳۸۹ برای توسعه سری‌های زمانی استفاده شد. در مرحله شناسایی، ساختار همبستگی زمانی سری، به وسیله آزمون توابع خودهمبستگی (ACF) و خودهمبستگی جزئی (PACF) تعیین می‌شود و توجه به تعداد گام‌های معنادار در هر یک از نمودارهای ACF و PACF می‌توان محدودۀ تغییرات مرتبه مدل‌ها را به دست آورد و در نهایت با ترکیب آن‌ها به مدل‌های داوطلب دست یافت. مدلی که حداقل مقدار پارامتر آکائیکه (Akaike Information Criterion-AIC) و شوارتز (Schwarz's Bayesian Criterion-SBC) را داشته باشد، به عنوان بهترین مدل انتخاب می‌شود. فرمول ریاضی (Akaike, 1974) به صورت زیر است:

$$AIC = -2\log(L) + 2m \quad (\text{رابطه ۷})$$

که در آن: $m = (p + q + P + Q)$ تعداد پارامترهای تخمین زده شده و L تابع درست‌نمایی مدل‌های ARIMA است.

فرمول ریاضی (Schwarz, 1978) (SBC) به صورت زیر است:

$$SBC = -2\log(L) + m \ln(n) \quad (\text{رابطه ۸})$$

که در آن: n تعداد مشاهدات است.

برای برآورد پارامترها از نرم‌افزار MINITAB16 استفاده شد. در مدل ARIMA تفاضل اصلی با هر مرتبه و تناوبی که لازم باشد، انجام می‌گیرد تا سری حاصل ایستا شود.

در مرحله آزمون کفایت مدل، ابتدا سری باقیمانده‌ها را از تفاضل سری زمانی مشاهده شده و سری زمانی حاصل از برآزش مدل منتخب بر سری زمانی سطح ایستایی ارائه شد، سپس آزمون‌های صحت سنجی روی باقیمانده‌های دوره واسنجی

(۸۰ درصد داده‌ها)، برای بررسی کفایت مدل منتخب انجام شد که در این پژوهش با استفاده از نمودار PP برای بررسی نرمال بودن باقیمانده و آزمون پورتمنتو (Portmanteau) برای تعیین استقلال باقیمانده استفاده شد. برای ارزیابی کارایی مدل‌ها از معیارهای ریشه میانگین مربعات خطا (RMSE) و میانگین قدر مطلق خطا (MAE) استفاده شد. همچنین برای مقایسه دقت پیش‌بینی مدل‌ها از نمودار تیلر (Taylor diagrams) استفاده شد. نمودار تیلر یک روش گرافیکی است که چگونگی نزدیکی یک الگو یا مجموعه‌ای از الگوها را نسبت به مشاهدات نشان می‌دهد. این نمودار در سال ۲۰۰۱ توسط تیلر ارائه شده است (Taylor, 2001). نمودار تیلر با استفاده از سه آماره: ضریب همبستگی، ریشه میانگین مربعات خطا و انحراف معیار دو سری زمانی (پیش‌بینی شده و مشاهداتی) ترسیم می‌شود. این نمودار به خصوص در ارزیابی مدل‌های متعدد یا در اندازه‌گیری مهارت نسبی بسیاری از مدل‌های مختلف مفید است (IPCC, 2001). نمودار تیلر در نرم‌افزار مطلب قابل ترسیم است.

نتایج و بحث

در این پژوهش ابتدا داده‌ها بررسی و دوره مشترک آماری انتخاب شد (۱۳۷۲-۱۳۸۹). سپس با استفاده از نرم‌افزار WinGamma و با در نظر گرفتن دو پارامتر گاما و نسبت V از روش الگوریتم ژنتیک اقدام به تعیین بهترین ترکیب ورودی از بین ترکیب‌های ورودی مختلف شد و همچنین بهترین تعداد داده‌های آموزشی با استفاده از روش M-Test تعیین شد. تعداد $1 + 2^n$ ترکیب ورودی وجود دارد که n نشان‌دهنده تعداد ورودی‌ها است. در این پژوهش از داده‌های مجموع بارش ماهانه (P)، متوسط ماهانه جریان ورودی به دشت (SF)، دما (T)، تبخیر (E) و سطح آب زیرزمینی (GL) به عنوان ورودی برای تعیین سطح آب زیرزمینی (GL) ماه بعد (Shirmohammadi et al., 2013) استفاده شد. با توجه به اینکه تعداد ورودی‌ها پنج تا بود، براساس فرمول یادشده تعداد ۳۱ نوع ترکیب مختلف دارند. جدول ۱ بهترین ترکیب را از بین ورودی‌های مختلف نشان می‌دهد. به علت محدودیت فضا فقط تعدادی از ترکیب‌ها در این جدول ارائه شده است. در این جدول نمونه‌ای از پارامترهای ورودی با علائم اختصاری نشان داده شده است. آماره شیب، همان شیب رگرسیون خطی است که برای محاسبه آماره گاما استفاده می‌شود. خطای استاندارد هم مقدار صحت رگرسیون خطی را ارائه می‌کند. کوچک‌تر بودن آماره نشان‌دهنده صحت بیشتر آماره گاما است. Mask الگوی ورودی را نشان

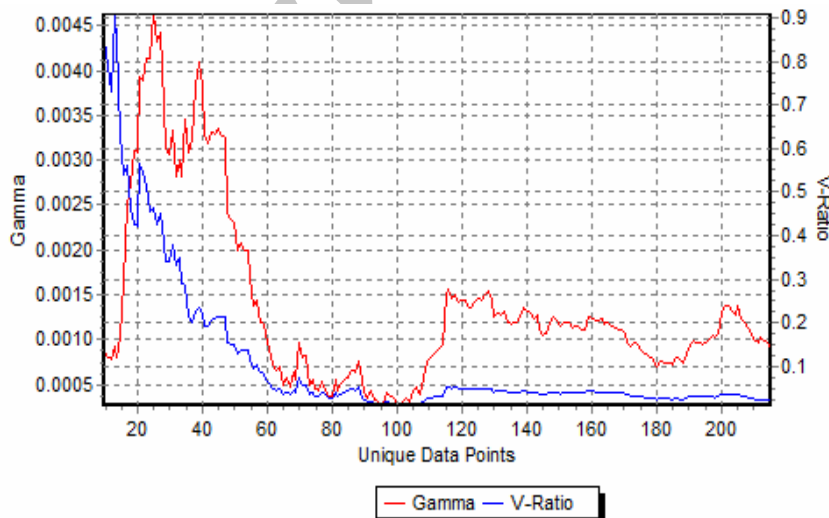
پایین ترین مقدار آماره گاما (میانگین مربعات خطا) و نسبت V را دارد، تعداد بهینه داده های آموزشی را مشخص می سازد (Agalbjörn et al., 1997; Končar, 1997). همان طور که در شکل ۳ مشخص است، در قسمت انتهایی و جایی که نمودار در حال افقی شدن است (توجه شود آماره های گاما بسیار کوچک هستند) کمترین مقدار آماره گاما و نسبت V مربوط به تعداد داده های ۱۸۰ است (به عبارت دیگر مطابق با پژوهش های Agalbjörn et al., 1997; Končar, 1997، بعد از افقی شدن نمودار در تعداد داده برابر با ۱۸۳ آماره گاما و نسبت V کمترین مقدار را دارد که ۸۳ درصد کل داده های ما را شامل می شوند). بنابراین، ۸۳ درصد داده ها برای سری آموزش و ۱۷ درصد برای ارزیابی کارایی مدل انفیس استفاده می شود.

می دهد که اعداد یک نشان دهنده دخالت دادن متغیر ورودی مرتبط در ترکیب ورودی است و عدد صفر نشان دهنده استفاده نکردن از آن است (اعداد براساس ترتیبی است که ورودی ها وارد نرم افزار می شوند). به عنوان مثال، الگوی ۱۱۰۰۱ نشان دهنده ترکیبی با ورودی های اول، دوم و پنجم (بارندگی، دما و سطح آب زیرزمینی) است. با توجه به کمترین مقدار پارامترها گاما (۰/۰۰۰۹) و نسبت (021/0) بهترین ترکیب ورودی با الگوی ۱۰۱۱۱ و با ورودی های بارندگی، دما، تبخیر و سطح آب زیرزمینی انتخاب می شود.

بهترین طول داده های آموزشی با استفاده از M-Test تعیین شد. در این آزمون خروجی به صورت نمودار است که در قسمت انتهایی که نمودار در حال افقی شده است، نقطه ای که

جدول ۱. تعیین بهترین ترکیب ورودی با استفاده از آماره گاما و نسبت v

P, T, E, GL	P, SF, T, E, GL	P, SF, E, GL	P, T, GL	SF, T, E, GL	SF, E, GL	SF, T, GL	P, SF, GL	E, GL	T, GL	P, GL	ترکیب ورودی
۰/۰۰۰۹	۰/۰۰۱۳	۰/۰۰۱۱	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱	۰/۰۰۲	۰/۰۰۲	۰/۰۰۲	۰/۰۰۲	گاما
۰/۲۲۵	۰/۱۴۵	۰/۱۵۵	۰/۲۰۳	۰/۱۶۴	۰/۱۸۵	۰/۱۵۰	۰/۱۱۷	۰/۲۵۸	۰/۲۴۱	۰/۲۱۷	شیب
۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	خطای استاندارد
۰/۰۲۱	۰/۰۲۸	۰/۰۲۶	۰/۰۲۹	۰/۰۲۶	۰/۰۳۳	۰/۰۴۲	۰/۰۴۵	۰/۰۴۵	۰/۰۴۸	۰/۰۵۰	نسبت V
۱۰۱۱۱	۱۱۱۱۱	۱۱۰۱۱	۱۰۱۰۱	۰۱۱۱۱	۰۱۰۱۱	۰۱۱۰۱	۱۱۰۰۱	۰۰۰۱۱	۰۰۱۰۱	۱۰۰۰۱	الگو (Mask) ورودی



شکل ۳. خروجی M-Test برای تعیین بهترین طول داده های آموزشی با استفاده از نرم افزار WinGamma

انتشار است. در این پژوهش، از تمامی توابع عضویت داده های ورودی ارائه شده در جدول ۳ استفاده و تابع عضویت داده های خروجی خطی انتخاب شد. با توجه به اینکه در روش جداسازی تورانه ای تعداد قوانین فازی زیاد است و برای تعداد توابع عضویت زیاد محاسبه پارامترهای این مدل سخت است

برای طبقه بندی داده های ورودی و ایجاد قوانین در مدل انفیس از روش Grid partition (جداسازی تورانه ای) (Jang, 1993) که رایج ترین روش است استفاده شد و الگوریتم یادگیری استفاده شده در این پژوهش هیبرید (hybrid) بوده است که ترکیبی از روش حداقل مربعات خطا و روش کاهش شیب پس

جدول ۳ نتایج ارزیابی پیش‌بینی مدل انفیس با توابع عضویت مختلف را نشان می‌دهد. همان‌طور که مشخص است تمامی توابع با تعداد تابع عضویت برابر با دو بهترین عملکرد را داشته‌اند و از بین مدل‌های مختلف، ریشه میانگین مربعات خطا (RMSE= ۱/۲۴۱) و میانگین مطلق خطای (MAE= ۰/۹۵۳) تابع عضویت II شکل در داده‌های آزمون کمتر از بقیه توابع عضویت بوده است که نشان‌دهنده عملکرد بهتر مدل انفیس با این تابع عضویت است.

(Farokhnia et al. 2011)، تعداد توابع عضویت (از دو تا پنج تابع عضویت) براساس سعی و خطا تعیین شد. انفیس برای ایجاد قوانین فازی از پارمترهای خطی و غیرخطی استفاده می‌کند، که با تغییر تعداد توابع عضویت این پارامترها نیز تغییر می‌کند. جدول ۲ تعداد پارامترهای استفاده‌شده به وسیله مدل انفیس براساس تعداد توابع عضویت را در مدل مطلب ۲۰۱۱ نشان می‌دهد که پارامترها به تعداد توابع عضویت استفاده‌شده بستگی دارد.

جدول ۲. تعداد پارامترهای استفاده‌شده به وسیله مدل ANFIS

تعداد قوانین فازی	تعداد کل پارامترها	تعداد پارامترهای غیرخطی	تعداد پارامترهای خطی	تعداد گره‌ها	تعداد توابع عضویت ورودی
۲	۹۶	۱۶	۸۰	۵۵	۲
۳	۴۲۹	۲۴	۴۰۵	۱۹۳	۳
۴	۱۳۱۲	۳۲	۱۲۸۰	۵۵۱	۴
۵	۳۱۶۵	۴۰	۳۱۲۵	۱۲۹۷	۵

جدول ۳. نتایج حاصل از توابع عضویت مختلف در روش انفیس

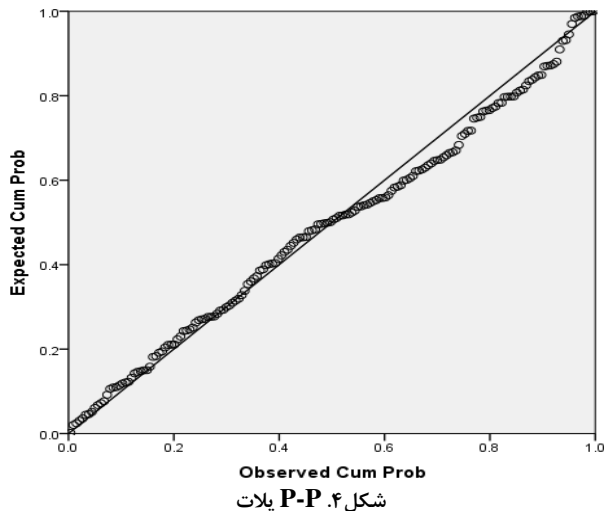
نوع توابع عضویت	تعداد توابع عضویت	آموزش		آزمون	
		MAE	RMSE	MAE	RMSE
گوسی	۲	۰/۲۰۶	۰/۱۵۳	۲/۰۴۸	۱/۶۲۲
گوسی ۲	۲	۰/۲۰۴	۰/۱۵۲	۱/۹۵۲	۱/۳۸۸
زنگوله‌ای	۲	۰/۲۰۴	۰/۱۵۱	۲/۷۷۴	۱/۱۰۴
دوزنقه‌ای	۲	۰/۲۱۶	۰/۱۵۱	۱/۵۸۹	۱/۱۷
مثلثی	۲	۰/۲۳۵	۰/۱۶۶	۱/۵۶۴	۱/۸۶۵
II شکل	۲	۰/۲۲۵	۰/۱۶۰	۱/۲۴۱	۰/۹۵۳
دوسیگموئیدی	۲	۰/۱۹۸	۰/۱۴۵	۱/۷۹۳	۱/۳۹۸

فقدان برازش Portmanteau نشان داد که داده‌ها به صورت یکسان و مستقل توزیع شده‌اند، چراکه مقدار Q (r) کمتر از χ^2 است ($\chi^2 = ۳۳/۴ < Q = ۳۲/۳۹$) که نشان‌دهنده این است که انتخاب مدل، صحیح است. P-P پلات نشان داد که داده‌ها توزیع نرمال دارند، چراکه نقاط بر روی شکل ۴ از الگوی خط ۴۵ درجه تبعیت می‌کنند که نشان‌دهنده این است که توزیع نرمال برای این مدل مناسب است. ضرایب مدل ARIMA(2,1,2) در جدول ۶ ارائه شده است. بنابراین، معادله پیش‌بینی سطح آب‌های زیرزمینی برای یک ماه بعد (t+1) به صورت زیر خواهد بود:

(رابطه ۱)

$$Y_{t+1} = Y_t + \phi_1 Y_t - \phi_1 Y_{t-1} + \phi_2 Y_{t-1} - \phi_2 Y_{t-2} - \theta_1 e_t - \theta_2 e_{t-1} + c$$

جدول ۴ نتایج بررسی ایستایی بودن سری زمانی سطح ایستایی را قبل و بعد از تفاضلی کردن با دو روش رایج ADF و PP با نرم‌افزار Eviews نشان می‌دهد. همان‌طور که مشخص است، بعد از تفاضلی کردن فرض صفر ($H_0: \theta=0$)، یعنی داده‌ها نایستا هستند و نیاز به تفاضلی کردن است (چون مقادیر سطح معناداری کوچک‌تر از سطح اطمینان ۰/۰۱ درصد بوده است) و فرض مخالف مبنی بر ایستابودن داده‌ها پذیرفته می‌شود. مطابق با روش باکس-جینکس و براساس خواص گرافیکی توابع خودهمبستگی (ACF) و خودهمبستگی جزئی (PACF) مرتبه p و q ($p \leq ۲$ و $q \leq ۴$) تعیین شد. سپس، براساس حداقل مقدار معیارهای AIC و SBC، ARIMA (2, 1, 2) (به عنوان بهترین مدل انتخاب شد (جدول ۵)). نتایج آزمون



در مدل $ARIMA(2,1,2)$ مقادیر ریشه میانگین مربعات خطا و میانگین مطلق خطای داده های آموزشی به ترتیب 0.325 و 0.241 و برای داده های صحت سنجی به ترتیب 0.209 و 0.171 به دست آمد. مقایسه دقت نتایج عملکرد مدل $ARIMA(2,1,2)$ و انفیس با توابع عضویت متفاوت به وسیله دیاگرام تیلر انجام شد (شکل ۵). همان طور که در شکل ۵ مشخص است مدل هشت به مشاهدات ما نزدیک تر است که نشان دهنده عملکرد بهتر مدل $ARIMA$ نسبت به سایر مدل هاست. از بین پیش بینی های مدل انفیس مطابق با مقادیر ریشه میانگین مربعات خطا و میانگین مطلق خطای (جدول ۳) نمودار تیلر نیز نشان داد که عملکرد پیش بینی با تابع عضویت Π شکل عملکرد بهتری نسبت به بقیه توابع عضویت دارد. شکل های ۶ و ۷ موفقیت پیش بینی مدل $ARIMA(2,1,2)$ در برابر داده های مشاهداتی برای دوره آزمون یا صحت سنجی مدل را نشان می دهد. پژوهش های گذشته نیز موفقیت مدل $ARIMA$ در پیش بینی سطح آب های زیرزمینی را نشان داده اند (Adhikary *et al.*, 2012; Aflatooni and Mardaneh, 2011; Lee *et al.*, 2009; Lu, *et al.*, 2013; Voudouris, 2002).

(Rahmani and Sedehi (2005) در دشت همدان- بهار از مدل $ARIMA(1,1,0)$ برای پیش بینی سطح آب زیرزمینی در ۲۰ سال آینده استفاده کردند که نتایج نشان دهنده عملکرد خوب سری های زمانی است. Mirzaei *et al.* (2006) در دشت شهرکرد نیز نشان دادند که با مدل $ARIMA$ به خوبی می توان نوسانات آب زیرزمینی آبخوان شهرکرد را پیش بینی کرد. Lee *et al.* (2009) در کره جنوبی نیز از سری های زمانی برای پیش بینی سطح آب زیرزمینی ماهانه استفاده کرد، که نتایج این مطالعه نتایج نشان داده است آریمای فصلی عملکرد رضایت بخشی در پیش بینی آب زیرزمینی دارد.

$$Y_{t+1} = Y_t + 1.7281 Y_t - 1.7281 Y_{t-1} - 0.9988 Y_{t-1} + 0.9988 Y_{t-2} - 1.7073 e_t + 0.958 e_{t-1} - 0.006$$

که در آن: Y سطح آب زیرزمینی و e اختلاف بین سطح آب زیرزمینی مشاهده و پیش بینی شده است.

جدول ۴. نتایج بررسی ایستایی بودن سری زمانی سطح ایستابی قبل و بعد از تفاضلی کردن

مرحله بررسی	نوع آزمون	آماره	سطح معناداری**
قبل از تفاضلی کردن	-Augmented Dickey Fuller test -Phillips	۱/۱۹	۰/۵۵
بعد از تفاضلی کردن	-Augmented Dickey Fuller test -Phillips Perron test	۱۱/۳	**۰/۰۰۳۴
		۵۴/۱	**۰/۰۰۰

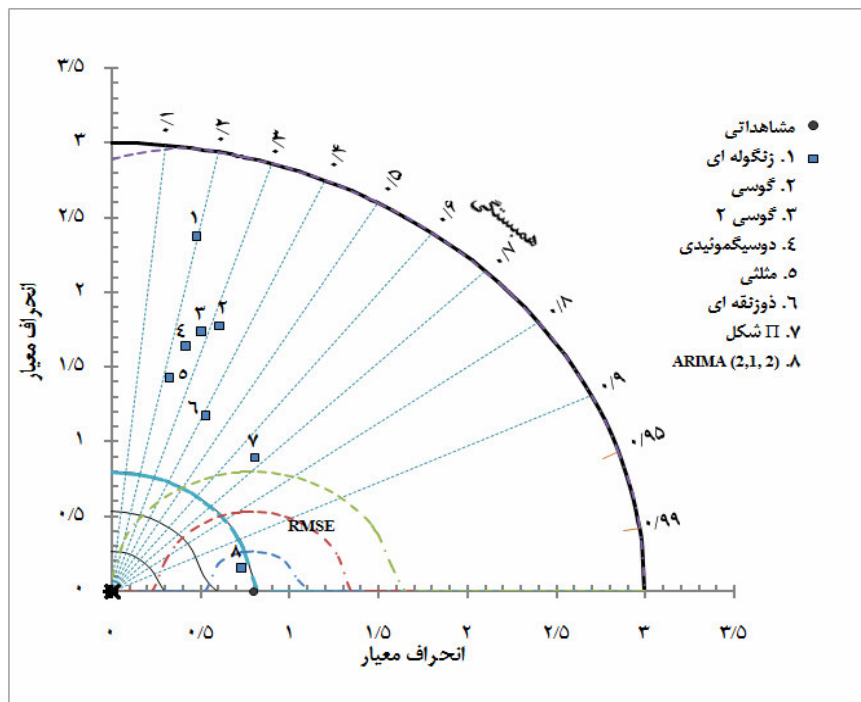
** در سطح ۹۹ درصد معنادار

جدول ۵. معیارهای AIC و SBC برای مدل های مختلف

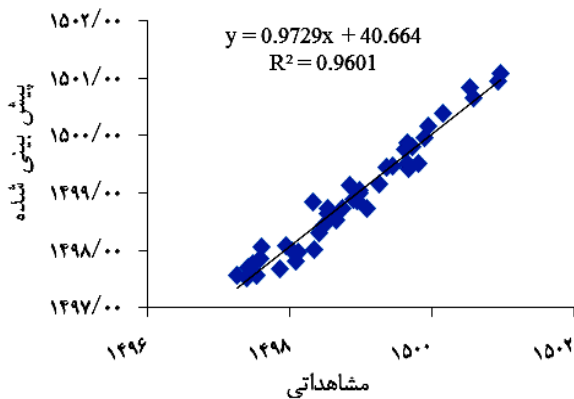
Model	AIC	SBC
ARIMA (2, 1, 4)	۱۹۴/۳	۲۱۶/۳
ARIMA (2, 1, 3)	۱۹۶/۰	۲۱۴/۹
ARIMA (2, 1, 2)	۱۱۴/۲	۱۲۹/۹
ARIMA (2, 1, 1)	۲۰۴/۴	۲۱۷/۰
ARIMA (1, 1, 4)	۱۹۹/۵	۲۱۸/۴
ARIMA (1, 1, 3)	۲۰۰/۵	۲۱۶/۲
ARIMA (1, 1, 2)	۲۰۲/۴	۲۱۵/۰
ARIMA (1, 1, 1)	۲۰۳/۰	۲۱۲/۴
ARIMA (1, 1, 0)	۲۰۱/۰	۲۰۷/۳
ARIMA (0, 1, 4)	۱۹۸/۱	۲۱۳/۸
ARIMA (0, 1, 3)	۲۰۰/۲	۲۱۲/۸
ARIMA (0, 1, 2)	۲۰۳/۶	۲۱۳/۰
ARIMA (0, 1, 1)	۲۱۲/۴	۲۱۸/۷

جدول ۶. ضرایب مدل $ARIMA(2,1,2)$

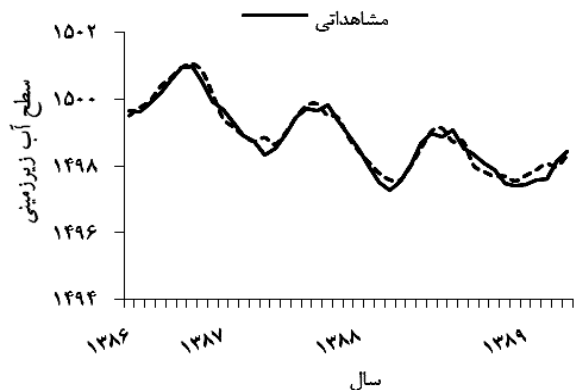
پارامترهای مدل	ضرایب	سطح معنی داری	T	خطای استاندارد
ϕ_1	۱/۷۲۸۱	۰/۰۰۰	۴۶۸/۸۵	۰/۰۰۴
ϕ_2	-۰/۹۹۸۸	۰/۰۰۰	-۲۵۴/۷۵	۰/۰۰۴
θ_1	۱/۷۰۷۳	۰/۰۰۰	۴۳/۵۴	۰/۰۳۹
θ_2	-۰/۹۵۸	۰/۰۰۰	-۲۱/۱۵	۰/۰۴۵
Constant	-۰/۰۰۶	۰/۸۰۷	-۰/۲۴	۰/۰۲۳



شکل ۵. دیاگرام تیلر برای مقایسه دقت مدل های پیش بینی



شکل ۷. Scatter پلات مربوط به داده های مشاهداتی و پیش بینی مدل ARIMA



شکل ۶. پیش بینی مدل ARIMA (2,1,2) در برابر داده های مشاهداتی

مقایسه نتایج انفیس مشخص شد که پیش بینی با تابع عضویت شکل II عملکرد بهتری نسبت به بقیه توابع عضویت دارد. امروزه به رغم استفاده کمتر از مدل های خطی این پژوهش نشان داد که مدل خطی ARIMA (2,1,2) عملکرد بهتری نسبت مدل غیرخطی انفیس در پیش بینی سطح ایستابی دارد. مشخص بودن روابط و همچنین استفاده از داده های پیشینه سطح ایستابی (استفاده نکردن از متغیرهای ورودی مختلف با خطای مختلف) در مدل ARIMA موجب موفقیت بهتر این مدل شده است.

نتیجه گیری

پیش آگهی از منابع آب می تواند کمک شایانی به مدیریت منابع آب کشور کند. در این پژوهش از دو روش خطی و غیرخطی برای پیش بینی سطح آب های زیرزمینی دشت شیراز تا یک ماه بعد استفاده شد. برای تعیین ترکیب ورودی مناسب و مقدار مناسب داده های آموزشی در روش غیرخطی به ترتیب از آزمون گاما و M استفاده کردیم که در هر دو روش مقدار بهینه براساس میانگین مربعات خطا یا همان آماره گاما تعیین شد. با

REFERENCES

Abghari, H., Rezaeianzadeh, M., Singh, V. P. and Moradzadeh Azar, F. (2010). Comparison of

Monthly Rainfall Prediction using Linear Stochastic-Base Models in Gharalar Rain Gauge

- Station Iran. *Geophysical Research Abstracts*, Vol. 12, EGU2010-2652-1.
- Adhikary, S.K., Rahman, M.D.M. and Gupta, A.D. (2012). A stochastic Modelling Technique for Predicting Groundwater table Fluctuations with Time Series Analysis. *Applied Sciences and Engineering Research*, 1(2), 238-249.
- Aflatooni, M. and Mardaneh, M. (2011). Time series analysis of ground water table fluctuations due to temperature and rainfall change in Shiraz plain. *Water Resources and Environmental Engineering*, 3(9), 176-188.
- Afuni, D.M., Malhojy, K. and Rashmeh, K. (2002). The use of saline water for irrigation of several wheat varieties and response. In *proceedings of the First National Conference of solutions to deal with the water crisis*, Iran, pp. 295-304. (In Farsi)
- Agalbjörn, S., Končar, N. and Jones, A. J. (1997). A note on the gamma test. *Neural Computing and Applications*, 53, 131-133.
- Ahn, H. (2000). Modeling of groundwater heads based on second-order difference time series models. *Hydrology*, 234(1-2), 82-94.
- Akaike, H. (1974). A look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723.
- Boochabun, K., Tych, W., Chappell, N. A., Carling, P. A., Lorsirirat, K. and PaObsaeng, S. (2004). Statistical modelling of rainfall and river flow in Thailand. *Geological Society of India*, 64, 503-515.
- Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control Holden-Day*, San Francisco, CA. USA.
- Chattopadhyay, S. and Chattopadhyay, G. (2010). Univariate modelling of summer-monsoon rainfall time series: comparison between ARIMA and ARNN. *Comptes Rendus Geoscience*, 342, 100-107.
- Chattopadhyay, S. J., hajharia, D. and Chattopadhyay, G. (2011). Univariate modelling of monthly maximum temperature time series over northeast India: neural network versus Yule-Walker equation based approach. *Meteorological Applications*, 18, 70-82.
- Chu, H. J. and Chang, L. C. (2009). Application of optimal control and fuzzy theory for dynamic groundwater remediation design. *Water Resource Management*, 23 (4), 647-660.
- Daliakopoulos, I. N., Coulibaly, P. and Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. *Hydrology*, 309(1-4), 229-240.
- Dickey, D. A. and Fuller, W. A. (1979). Estimators for autoregressive time series with a unit root. *American Statistical Association*, 74, 427-431.
- Farokhnia, A., Morid, S. and Byun, H. R. (2011). Application of global SST and SLP data for drought forecasting on Tehran plain using data mining and ANFIS techniques. *Theor Appl Climatol*, 104, 71-81.
- Faruk, D. (2010). A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence*, 23 (4), 586-594.
- Firat, M., Turan, M. E. and Yurdusev, M. A. (2009). Comparative analysis of fuzzy inference systems for water consumption time series prediction. *Hydrology*, 374(3-4), 235-241.
- Hasmida, H. (2009). *Water quality trend at the upper part of Johor River in relation to rainfall and runoff pattern*. Faculty of Civil Engineering. MSc. Dissertation, University Teknologi Malaysia.
- IPCC, Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K. and Johnson, C. A. (eds) (2001). *Climate change 2001: the scientific basis*, Cambridge University Press. Cambridge, United Kingdom and New York. NY. USA, 881 pp.
- Jang, J. S. R. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems Man and Cybernetics*, 23(3), 665-685.
- Kisi, O. and Shiri, J. (2012). Wavelet and neuro-fuzzy conjunction model for predicting water table depth fluctuations. *Hydrology Research*, 43 (3), 286-300.
- Končar, N. (1997). *Optimization methodologies for direct inverse neuro control*. Dept. of Computing. Imperial College of Science. Ph. D. dissertation, University of London, London.
- Lee, S. I., Lee, S. K. and Hamm, S. Y. (2009). A model for groundwater time-series from the well field of riverbank filtration. *Korea Water Resources Association*, 42, 673-680.
- Lu, W. X., Zhao, Y., Chu, H. B. and Yang, L. L. (2013). The analysis of groundwater levels influences by dual factors in western Jilin Province by using time series analysis method. *Applied Water Science*. DOI 10.1007/s13201-013-0111-4.
- Mirzaei, S.y., Chitsazan M., Chinipardaz, R. and Samadi, H. (2006). Groundwater forecasting of Shahrekord plains using time series models and discuss strategies to improve. *First Regional Conference on Optimum Utilization of Water Resources (Opportunities and Challenges)*, 5-6 SEP, Shahrekord University, Iran. (In Farsi)
- Phillips, P. C. B. and Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika*, 75, 335-346.
- Rahmani, A. and Sedehi, M. (2005). Prediction of changes in groundwater levels Hamedan - spring plain with time series model. *Water and Wastewater*, 15 (51), 42-49. (In Farsi)
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461-464.
- Shirmohammadi, B., Vafakhah, M., Moosavi, V. and Moghaddamnia, A. (2013). Application of

- Several Data-Driven Techniques for Predicting Groundwater Level. *Water Resources Management*. DOI 10.1007/s11269-012-0194-y.
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Geophysical Research*, 106, 7183-7192. (Also see PCMDI Report 55, <http://www.pcmdi.llnl.gov/publications/ab55.html>).
- The Water comprehensive plan (1991). *Groundwater resources of Shapur and Dalaki, Maharlou-Bakhtegan watersheds*. First edition, published by the Ministry of Energy. Jamab Consultant Engineering Company. Tehran. 400 pages. (In Farsi)
- Tutmez, B., Hatipoglu, Z. and Kaymak, U. (2006). Modelling electrical conductivity of groundwater using an adaptive neuro-fuzzy inference system. *Computers & Geosciences*, 32, 421-433.
- Voudouris, K. (2002). Time series analysis using ARIMA models of the groundwater table in Patras industrial area aquifer system NW Peloponnese. *Greece Terra Nostra (Bonn)*, 3, 335-340.
- Wong, H., Wc, I., Zhang, R. and Xia, J. (2007). Non-parametric time series models for hydrological forecasting. *Hydrology*, 332(3-4), 337-347.
- Zakaria, S., Al-Ansari, N., Knutsson, S. and Al-Badrany, T. (2012). ARIMA Models for weekly rainfall in the semi-arid Sinjar District at Iraq. *Earth Sciences and Geotechnical Engineering*, 2(3), 25-55.

Archive of SID