

استفاده از مدل‌های سری زمانی، شبکه عصبی و ماشین بردار پشتیبان جهت پیش‌بینی دبی ورودی به سد گرگان

میثم صمدی و ابوالحسن فتح‌آبادی

دوره ۴، شماره ۴، زمستان ۱۳۹۷، صفحات ۲۹۹ – ۳۰۹

Vol. 4(4), Winter 2019, 299 – 309

DOI: 10.22034/jewe.2018.128256.1256

**Application of Time Series, ANN, and SVM  
Models in Forecasting the Gorgan Dam Inflow  
Rate**

Samadi M. and Fathabadi A.



[www.jewe.ir](http://www.jewe.ir)

OPEN ACCESS

ارجاع به این مقاله: صمدی م. و فتح‌آبادی ا. (۱۳۹۷). استفاده از مدل‌های سری زمانی، شبکه عصبی و ماشین بردار پشتیبان جهت پیش‌بینی دبی ورودی به سد گرگان. مجله محیط‌زیست و مهندسی آب، دوره ۴، شماره ۴، صفحات: ۲۹۹ – ۳۰۹.

**Citing this paper:** Samadi M. and Fathabadi A. (2019). Application of time series, ANN, and SVM models in forecasting the Gorgan Dam inflow rate. J. Environ. Water Eng., 4(4), 299– 309. DOI: 10.22034/jewe.2018.128256.1256

## استفاده از مدل‌های سری زمانی، شبکه عصبی و ماشین بردار پشتیبان جهت پیش‌بینی دبی ورودی به سد گرگان

میثم صمدی<sup>۱\*</sup> و ابوالحسن فتح‌آبادی<sup>۲</sup>

<sup>۱</sup> دانشجوی دکتری، گروه آبخیزداری و مدیریت مناطق بیابانی، دانشکده مرتع و آبخیزداری، دانشگاه علوم کشاورزی و منابع طبیعی گرگان، گرگان، ایران  
<sup>۲</sup> استادیار، گروه آبخیزداری، دانشکده کشاورزی و منابع طبیعی، دانشگاه گنبد کاووس، گنبد کاووس، ایران

\*نویسنده مسئول: meisamadi68@gmail.com

### مقاله اصلی

تاریخ پذیرش: [۱۳۹۷/۰۹/۱۸]

تاریخ بازنگری: [۱۳۹۷/۰۸/۱۹]

تاریخ دریافت: [۱۳۹۷/۰۲/۱۰]

### چکیده

پیش‌بینی مقادیر جریان ورودی به سیستم منابع آب به‌منظور آگاهی از شرایط آینده و برنامه‌ریزی برای تخصیص بهینه منابع آب به بخش‌های مختلف از قبیل شرب، کشاورزی و صنعتی امری ضروری در مدیریت منابع آب می‌باشد. هدف از پژوهش حاضر پیش‌بینی مقادیر دبی ماهانه ورودی به سد گرگان برای آینده بود. بدین منظور از داده‌های هیدرومتری ایستگاه قزاقلی با دوره آماری ۴۷ سال و سه مدل سری‌زمانی، شبکه عصبی و ماشین بردار پشتیبان جهت پیش‌بینی استفاده شد و نتایج مدل‌های مختلف مورد مقایسه قرار گرفت. با توجه به نتایج به‌دست آمده در مدل سری زمانی بر اساس معیارهای آکایک و شوارتز، مدل  $ARIMA(1,0,0)$  به‌عنوان مدل بهینه انتخاب شد. در مدل شبکه عصبی، شبکه با ورودی ۲ و ۱۰ نرون به‌عنوان شبکه برتر انتخاب شد و در مدل ماشین بردار پشتیبان شبکه با ورودی ۱ به‌عنوان شبکه برتر انتخاب شد. در نهایت با توجه به نتایج به‌دست آمده و با توجه به معیارهای ارزیابی مدل‌ها، مدل ماشین بردار پشتیبان بهترین عملکرد را نسبت به دو مدل دیگر داشت. مقادیر  $RMSE$  و  $AARE$  برای مدل ماشین بردار پشتیبان به ترتیب  $۵/۳۱$  و  $۱/۰۷$ ، برای مدل شبکه عصبی به ترتیب  $۹/۸۸$  و  $۲/۷۸$  و برای مدل سری زمانی به ترتیب  $۸/۸۴$  و  $۱/۲۰$  به‌دست آمد. بر اساس نتایج این پژوهش، بهترین مدل برای پیش‌بینی دبی ماهانه ورودی به سد گرگان مدل ماشین بردار پشتیبان می‌باشد.

واژه‌های کلیدی: آکایک؛ ایستگاه قزاقلی؛ شوارتز؛ مدل‌سازی؛ نرون.

## ۱- مقدمه

در یک سیستم منابع آب اغلب نیاز بخش‌های مختلف از قبیل کشاورزی، شرب و صنعت تغییر چندانی نکرده و به راحتی قابل پیش‌بینی می‌باشد. آنچه در این میان مهم می‌باشد، پیش‌بینی مقادیر جریان ورودی به سیستم منابع آب در ماه‌های آینده است. در این راستا استفاده از روش‌هایی که بتواند با حداقل خطا و با توجه به داده و اطلاعات موجود، جریان رودخانه را پیش‌بینی کند از اهمیت فراوانی برخوردار می‌باشد. در این میان روش سری‌های زمانی طی چهار دهه اخیر به‌طور گسترده‌ای برای پیش‌بینی جریان رودخانه مورد استفاده قرار گرفته است. امروزه تجزیه و تحلیل سری‌های زمانی به‌طور وسیعی در بسیاری از شاخه‌های مهندسی، علوم فیزیک و اقتصاد مورد استفاده واقع می‌شود و می‌توان گفت که بیشتر شاخه‌های علوم منجر به مطالعه داده‌هایی می‌شوند که به شکل سری‌های زمانی رخ می‌دهند. سری زمانی مجموعه‌ای از مشاهدات است که بر حسب زمان و یا هر کمیت دیگری مرتب شده باشند (Karamouz and Araghinejd 2005).

کاربرد سری‌های زمانی در هیدرولوژی از چهار دهه پیش آغاز شده و با ارائه مدل‌های باکس و جنکینز<sup>۱</sup> به اوج خود رسید. (Jalal Kamali 2002) با استفاده از سری‌های زمانی اقدام به پیش‌بینی جریان ماهانه ورودی به مخزن سد جیرفت نمود. با توجه به نتایج به‌دست آمده مدل  $ARIMA(1,1,1)(0,1,2)$  به‌عنوان مدل برتر انتخاب شد. (Valipour et al 2013) مدل‌های  $ARMA$ ،  $ARIMA$  و مدل شبکه عصبی خودهمبسته، برای مدل‌سازی جریان ماهانه ورودی به مخزن سد دز استفاده کردند. طبق نتایج حاصله مدل شبکه عصبی خودهمبسته بهتر از مدل‌های سری زمانی جریان ورودی ماهانه به سد دز را مدل‌سازی کرده است. در سال‌های اخیر استفاده از روش‌های هوش مصنوعی از جمله منطق فازی و شبکه عصبی در زمینه‌های مختلفی که در آن‌ها ارتباط بین ورودی و خروجی غیرخطی بوده، گسترش پیدا کرده است. (Karimi Masouleh et al. 2010) کاربرد شبکه عصبی در پیش‌بینی دبی ماهانه رودخانه کارون را بررسی کردند. نتایج

تحقیق آن‌ها بیانگر از دقت نسبتاً بالای شبکه عصبی مصنوعی در برآورد و تخمین دبی ماهانه رودخانه کارون بود. (Abdollah Pour Azad and Sattari 2015) جریان رودخانه اهرچای را با استفاده از روش‌های شبکه عصبی و انفیس پیش‌بینی کردند. نتایج آن‌ها نشان داد که مدل انفیس دقت بالاتری نسبت به مدل شبکه عصبی دارد. (Salarijazi et al. 2016) اقدام به پیش‌بینی دبی جریان روزانه رودخانه در حوزه آبخیز گالیکش در استان گلستان کردند. بدین منظور از سه مدل شبکه عصبی مصنوعی پرسپترون، درختی M5 و برنامه‌ریزی بیان ژن استفاده شد. نتایج آن‌ها نشان داد که هر سه مدل دارای کارایی مناسبی جهت پیش‌بینی دبی روزانه برخوردار می‌باشند. (Choubin and Malekian 2017) جهت پیش‌بینی نوسانات آب زیرزمینی از شبکه عصبی مصنوعی و  $ARIMA$  استفاده کردند. نتایج آن‌ها نشان داد که مدل  $ARIMA$  نسبت به شبکه عصبی عملکرد بهتری جهت پیش‌بینی داشته است. مدل‌های  $SVM$  که بر پایه تئوری آموزش احتمال قرار دارند، یک کلاس جدید از مدل‌هایی هستند که به‌منظور طبقه‌بندی و پیش‌بینی در علوم مختلف استفاده می‌شوند.

(Ahmadi et al. 2014) جهت پیش‌بینی جریان رودخانه باراندوزچای ارومیه از روش‌های ماشین بردار پشتیبان و برنامه‌ریزی ژنتیک استفاده کردند. نتایج آن‌ها نشان داد که اگرچه دقت برنامه‌ریزی ژنتیک اندکی نسبت به  $SVM$  بیش‌تر بود اما روش  $SVM$  به‌مراتب ساده‌تر از روش برنامه‌ریزی ژنتیک بوده و می‌تواند به‌عنوان یک روش کاربردی در پیش‌بینی جریان رودخانه به کار رود. (Choubin et al. 2018) اقدام به ارزیابی مدل  $CART$  جهت تخمین بار رسوب معلق نمودند. آن‌ها همچنین اقدام به مقایسه دقت مدل  $CART$  با مدل‌های انفیس، شبکه عصبی پرسپترون چند لایه و ماشین بردار پشتیبان نمودند. نتایج نشان داد مدل  $CART$  بهترین عملکرد را داشته است. همچنین آن‌ها اقدام به پیش‌بینی بارش با استفاده از مدل‌های  $CART$ ، سری زمانی و انفیس نمودند. نتایج نشان داد که مدل  $CART$

<sup>1</sup> Box-Jenkins

آن‌ها، زه‌کشی اراضی منطقه به‌منظور بهبود کشت، توسعه کشاورزی و قابل‌کشت‌نمودن اراضی شور و آبی‌پروری دارای اهمیت زیادی می‌باشد.

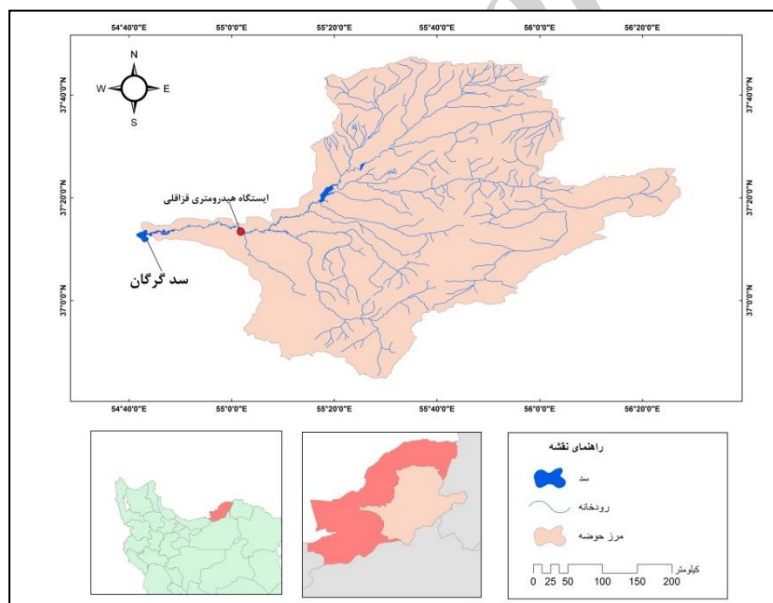
## ۲- مواد و روش‌ها

### ۱-۲- منطقه مورد مطالعه

سد مخزنی گرگان یا سد ووشمگیر در موقیت جغرافیایی  $44^{\circ}$   $55^{\circ}$  طول شرقی و  $37^{\circ} 11' 43''$  عرض شمالی و در فاصله ۶۵ کیلومتری شهرستان گرگان قرار گرفته است. ایستگاه قزاقلی که در بالادست سد گرگان و بر روی رودخانه گرگانرود واقع گردیده است، به‌عنوان ایستگاه شاخص جهت اندازه‌گیری میزان آب ورودی به مخزن سد گرگان می‌باشد. این ایستگاه در موقیت جغرافیایی  $55^{\circ} 01' 10''$  طول شرقی و  $37^{\circ} 13' 41''$  عرض شمالی قرار گرفته و ارتفاع آن از سطح دریا ۳۰m می‌باشد. شکل (۱) موقیت کلی حوضه، سد و ایستگاه مورد مطالعه را نشان می‌دهد.

نسبت به دو مدل دیگر عملکرد بهتری در پیش‌بینی بارش داشته است.

بنابراین، هدف از انجام این پژوهش تعیین کارایی مدل‌های سری زمانی، شبکه عصبی و SVM در پیش‌بینی دبی ورودی به سد گرگان در استان گلستان و در نهایت مقایسه نتایج مدل‌های مختلف و ارائه روش مناسب و با دقت کافی جهت پیش‌بینی جریان ورودی به مخزن سد گرگان می‌باشد. از آنجایی که در تحقیقات اندکی اقدام به استفاده از دو روش در پیش‌بینی و مدل‌سازی جریان رودخانه با استفاده از داده‌های یک‌سان نموده‌اند، لذا در تحقیق حاضر سعی شد که عملکرد هر دو روش در مدل‌سازی و پیش‌بینی جریان رودخانه با استفاده از داده‌های یک‌سان نسبت به یکدیگر مقایسه شود. همچنین در این پژوهش از مدل جدید ماشین پشتیبان تصمیم به‌منظور پیش‌بینی دبی استفاده شد. سد گرگان که بر روی گرگانرود واقع گردیده به‌دلیل بهبود اراضی ساحل راست و چپ گرگانرود، کنترل سیلاب‌ها و کاهش خسارات ناشی از



شکل ۱- موقیت کلی منطقه، موقیت سد و ایستگاه مورد مطالعه

Fig. 1 Location of study area, dam and stations

ابتدا نرمال بودن سری داده‌ها بررسی شد. با توجه به این که داده‌ها نرمال نبودند، از تبدیل کاکس و باکس<sup>۲</sup> (رابطه ۲) برای نرمال‌سازی استفاده شد. در ادامه کلوروگرام خود

### ۲-۲- پیش‌بینی دبی با استفاده از سری‌های زمانی

جهت پیش‌بینی دبی از سری داده‌های دبی ماهانه ایستگاه تمر با دوره‌ی آماری  $y$  ۴۷ (۱۳۴۸ تا ۱۳۹۴) استفاده شد.

<sup>2</sup> Cox-Box

هر یک از مدل‌ها که مقدار آکائیک و شوارتز پایین‌تری داشت به‌عنوان مدل برتر برای مراحل بعدی انتخاب شد. در نهایت با استفاده از مدلی که درستی آن تعیین و با توجه به معیارهای آکائیک و شوارتز به‌عنوان مدل بهینه‌تر انتخاب گردید اقدام به پیش‌بینی دبی و تولید داده‌های مصنوعی دبی شد.

### ۳-۲- پیش‌بینی دبی با استفاده از شبکه عصبی

در شبکه عصبی مصنوعی ابتدا لازم است که داده‌ها معیارسازی شوند، یعنی بین یک دامنه عددی که معمولاً (۰ و ۱) می‌باشند، قرار گیرند. در این ارتباط و با توجه به پژوهش‌های پیشین (Salamatian et al. 2006; Kisi) پژوهش‌های (Alborzi 2001; 2004) به دلیل این که مقادیر آستانه ۰ و ۱ در داده‌های ورودی به شبکه نباشد، داده‌ها با استفاده از رابطه (۵) که باعث می‌شود داده‌ها بین ۰/۰۵ و ۰/۹۵ قرار گیرند، معیارسازی شد.

$$y = 0.9 * \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} + 0.05 \quad (5)$$

در این رابطه  $X_{\min}$  کوچک‌ترین داده و  $X_{\max}$  بزرگ‌ترین داده سری داده‌ها می‌باشد. پس از معیارسازی الگوهای آموزشی، صحت‌سنجی و تست انتخاب شد. از آنجایی که در این پژوهش هدف مقایسه عملکرد روش‌های هوش مصنوعی با مدل‌های سری‌های زمانی می‌باشد و در مدل‌های سری‌های زمانی ترتیب داده‌ها برای پژوهش‌گران مهم بود، لذا در پژوهش حاضر از رویکرد تصادفی برای تعیین ورودی‌ها نمی‌توان استفاده نمود. بدین منظور تا حد ممکن باید در صد داده‌های آموزشی را بالا گرفت (۷۰٪ داده‌های کل). در این پژوهش، ۱۵٪ کل داده‌ها را برای تصدیق و ۱۵٪ آخر داده‌ها به‌عنوان داده‌های تست در نظر گرفته شد. در شبکه عصبی ساختار شبکه شامل نوع ورودی‌ها و خروجی‌ها، تعداد لایه‌های مخفی و تعداد نورون‌ها در هر لایه مخفی، نوع تابع فعال‌سازی و نوع الگوریتم آموزشی می‌باشد. لازم است تا روابط ما بین ورودی‌ها و خروجی‌ها تعیین و بهترین ورودی‌ها تعیین گردد. در این پژوهش، برای روش‌های هوش مصنوعی، ورودی‌های شبکه مقادیر گذشته دبی بود. در ابتدا ورودی شبکه را مقادیر

همبستگی و خودهمبستگی جزئی ترسیم شد. شکل کلوروگرام تاثیر به‌سزایی در انتخاب نوع مدل و مرتبه مدل دارد. بعد از تعیین نوع مدل و مرتبه آن، مقادیر پارامترهای مدل تخمین زده شد. در این تحقیق برای تعیین پارامترها از نرم‌افزار MINITAB استفاده شد. در این نرم‌افزار از روش حداکثر درستی برای تعیین پارامترها استفاده می‌شود. برای آزمایش درستی مدل محاسبه‌شده ابتدا باقیمانده‌ها با استفاده از رابطه (۱) محاسبه شد.

$$E_k = Z_t - Z_t(k) \quad (1)$$

در این رابطه  $E_k$  خطای  $k$ ام می‌باشد که از تفریق  $Z_t$  از مقدار پیش‌بینی شده آن بدست می‌آید. پس از آن که مقادیر باقیمانده‌ها محاسبه شد، برای بررسی استقلال باقیمانده‌ها از توابع خودهمبستگی و خودهمبستگی جزئی و هم‌چنین آماره باکس پیرسون (رابطه ۲) استفاده شد. با مقایسه  $Q$  با مقدار به‌دست آمده از جدول کای دو با درجه آزادی  $L-P$  (که  $L$  تعداد تأخیرها و  $p$  تعداد پارامترها می‌باشند) می‌توان پی به استقلال باقیمانده‌ها برد. در صورتی که  $Q < \chi^2(l-p)$  باشد، غیروابسته بودن مقادیر باقیمانده‌ها ثابت شده و مدل پذیرفته می‌شود. در صورتی که خلاف رابطه برقرار شده باشد، باید مدل دیگر با رتبه بالاتر  $(p+1)$  را بررسی کرد.

$$Q = n(n+2) \sum_{k=1}^k \frac{r_k^2}{n-k} \quad (2)$$

معمولاً برای هر سری زمانی چند مدل مختلف وجود دارد که صحت آن‌ها نیز تعیین شده در نتیجه مدلی به‌عنوان مدل نهایی انتخاب می‌شود که بهینه‌تر باشد. که برای این منظور از معیار آکائیک (رابطه ۳) استفاده شد.

$$AIC(M) = n \ln \sigma_a^2 + 2M \quad (3)$$

که در رابطه (۳)، مقدار  $\sigma_a^2$  واریانس تخمینی خطا با استفاده از روش حداکثر درستی‌نمایی،  $n$  تعداد داده‌ها و  $M$  تعداد پارامترهای مدل می‌باشد. از آنجایی که معیار آکائیک مرتبه مدل‌های اتورگرسیو را بالا نشان می‌دهد. بنابراین، در این پژوهش، برای انتخاب مدل بهینه از معیار شوارتز (رابطه ۴) نیز استفاده شد.

$$SBC(M) = n \ln \sigma_a^2 + M \ln n \quad (4)$$

<sup>3</sup> Akaike information criterion



دبی تا تاخیر ۵ گرفته و در کل برای این قسمت ۵ ورودی وجود دارد که به صورت جدول (۱) می باشد.

در شبکه عصبی پرسپترون چند لایه یک لایه ورودی و یک لایه خروجی و یک یا چند لایه مخفی وجود دارد که تعداد نرون‌ها در لایه‌های ورودی و خروجی برابر با تعداد بردارهای ورودی و خروجی می باشد. (Hornik et al. 1989) بیان کرده است که یک شبکه با یک لایه مخفی و تعداد کافی نرون در لایه مخفی می‌تواند با هر درجه صحتی ارتباط بین داده‌های ورودی و خروجی را تعیین کند. برای تعیین تعداد نرون‌ها در لایه‌های مخفی الگوریتم مدونی وجود ندارد طوری که اغلب به روش سعی و خطا تعیین می‌گردد. در ارتباط با نوع تابع فعال‌سازی با توجه به تحقیقات قبلی نوع تابع فعال‌سازی سیگموئید استفاده شد (Rezaei 2004). نوع الگوریتم آموزش نیز با توجه به کارایی روش پس‌انتشار خطا در مطالعات هیدرولوژیکی، از این الگوریتم آموزش استفاده گردید. در این پژوهش به دلیل جامعیت و سهولت کاربردی از نرم‌افزار MATLAB استفاده گردید. در این نرم‌افزار برای ساخت مدل‌های مختلف شبکه عصبی از جعبه ابزار Artificial Neural Network استفاده شد.

## ۲-۵- معیارهای ارزیابی

در این تحقیق در هر مرحله برای تعیین بهترین ساختار مدل و مقایسه مدل‌های مختلف از معیارهای مختلف شامل معیار RMSE<sup>f</sup> (میانگین مربعات خطا) (رابطه ۶)، معیار AARE<sup>5</sup> (متوسط مقادیر مطلق خطای نسبی) (رابطه ۷)، معیار ضریب NSE (معیار نش) (رابطه ۸) استفاده شد.

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^n (Q_i^o - Q_i^c)^2} \quad (6)$$

$$AARE = \frac{1}{N} \sum_{i=1}^n \left| \frac{Q_i^o - Q_i^c}{Q_i^o} \right| * 100 \quad (7)$$

$$NSE = 1 - \frac{\sum (Q_{obs} - Q_{sim})^2}{\sum (Q_{obs} - \bar{Q})^2} \quad (8)$$

که،  $n$  تعداد داده‌ها،  $Q_i^o$  دبی مشاهداتی و  $Q_i^c$  دبی شبیه‌سازی شده می‌باشند.

## ۳- یافته‌ها و بحث

برای سری‌های زمانی ابتدا نرمال بودن داده‌ها تعیین شد و مشخص شد که در ایستگاه مورد مطالعه داده‌ها نرمال نبودند و برای نرمال‌سازی، از روش باکس-کاکس استفاده شد. در ادامه مقادیر ضرایب خودهمبستگی و خود همبستگی جزئی در تاخیرهای مختلف تعیین و کلوروگرام خودهمبستگی و خودهمبستگی جزئی ترسیم شد (شکل‌های ۲ و ۳). با توجه به شکل‌های (۲) و (۳) ملاحظه می‌شود که آهنگ کاهش مقادیر ضریب خودهمبستگی در تأخیرهای مختلف بسیار پایین بوده و به صورت سینوسی نوسان داشته که این امر حاکی از غیرایستابودن سری زمانی مورد مطالعه دارد. همچنین ضرایب خودهمبستگی در تأخیرهای مختلف نوسان داشته و در تأخیر مشخصی بریده نشده است، این امر نشان دهنده وجود مدل‌های

جدول ۱- ترکیب ورودی‌های مختلف در مدل شبکه عصبی

Table 1 Combining of different inputs in the Neural Network model

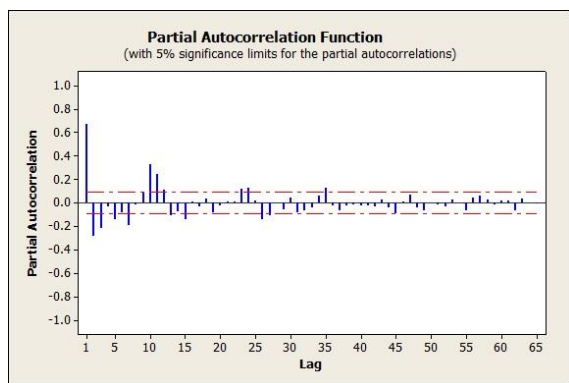
| Model  | Input type  |
|--------|---|
| Input1 | $Q_{r-1}$   |
| Input2 | $Q_{t-2} \quad Q_{r-1}$   |
| Input3 | $Q_{t-3} \quad Q_{t-2} \quad Q_{r-1}$                             |
| Input4 | $Q_{t-4} \quad Q_{t-3} \quad Q_{t-2} \quad Q_{r-1}$               |
| Input5 | $Q_{t-5} \quad Q_{t-4} \quad Q_{t-3} \quad Q_{t-2} \quad Q_{r-1}$ |

## ۲-۴- پیش‌بینی دبی با استفاده از مدل SVM

در روش ماشینی بردار پشتیبان (SVM) پارامترهای مدل شامل مقادیر  $C$ ،  $e$  و پارامترهای تابع کرنل آن می‌باشد. جهت تعیین این پارامترهای الگوریتم مشخصی وجود ندارد؛ بدین منظور از روش سعی و خطا و به صورت جستجوی شبکه‌ای مقادیر پارامترهای آن تعیین شد (Yu and Chen 2007).

<sup>5</sup> Average Absolute Relative Error

<sup>f</sup>Root Mean Square Error

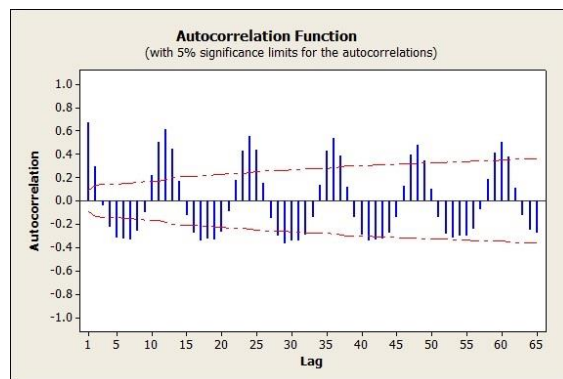


شکل ۳- خود همبستگی جزئی داده‌های ایستگاه قزاقلی

Fig. 3 Partial Autocorrelation of the data of Kazagholi Station

نتایج مدل‌های برازش داده شده‌ی ARIMA و مقادیر ضرایب مختلف به همراه مقدار معیارهای آکایک و شوارتز برای هر کدام از مدل‌ها در جدول (۲) آورده شده است. برای تعیین صحت هریک از مدل‌های تعیین شده از آزمون‌های باکس-پیرسون، ضرایب خودهمبستگی و خودهمبستگی جزئی باقی‌مانده‌ها استفاده شد.

اتورگرسیو در سری مورد مطالعه می‌باشد. بنابراین، در گام اول مدل‌های ARIMA به عنوان مدل مناسب تشخیص داده شد و در مرحله بعد مدل‌ها مورد بررسی بیشتر قرار گرفته و با استفاده از حداکثر درست‌نمایی، پارامترهای هر یک از مدل‌ها تعیین شد.



شکل ۲- خود همبستگی داده‌های ایستگاه قزاقلی

Fig. 2 autocorrelation of the data of Kazagholi Station

جدول ۲- نتایج مدل‌های ARIMA

Table 2 The results of ARIMA models

| Model               | $\phi_1$ | $\phi_2$ | $\phi_3$ | $\theta_1$ | $\theta_2$ | $\Phi_1$ | $\Theta_1$ | Akaike   | Schwarz  |
|---------------------|----------|----------|----------|------------|------------|----------|------------|----------|----------|
| ARIMA(1,0,1)(1,0,0) | 0.508    |          |          | -0.104     |            | 0.491    |            | -579.617 | -567.496 |
| ARIMA(1,0,0)(1,0,0) | 0.573    |          |          |            |            | 0.508    |            | -579.927 | -571.846 |
| ARIMA(1,0,1)(1,0,1) | 0.583    |          |          | 0.077      |            | 0.996    | 0.921      | -701.617 | -685.456 |
| ARIMA(1,0,0)(1,0,1) | 0.528    |          |          |            |            | 0.996    | 0.923      | -703.844 | -691.724 |
| ARIMA(2,0,0)(1,0,1) | 0.508    | 0.038    |          |            |            | 0.996    | 0.919      | -701.162 | -685.001 |
| ARIMA(2,0,0)(1,0,0) | 0.622    | 0.075    |          |            |            | 0.483    |            | -579.617 | -567.496 |
| ARIMA(3,0,0)(1,0,0) | 0.633    | -0.039   | -0.079   |            |            | 0.447    |            | -577.617 | -561.456 |
| ARIMA(3,0,0)(1,0,1) | 0.507    | 0.027    | 0.022    |            |            | 0.9956   | 0.918      | -698.481 | -678.28  |
| ARIMA(1,0,0)(1,1,1) | 0.527    |          |          |            |            | 0.042    | 0.942      | -697.07  | -684.949 |
| ARIMA(2,0,0)(1,1,1) | 0.511    | 0.028    |          |            |            | 0.041    | 0.942      | -695.518 | -679.357 |
| ARIMA(3,0,0)(1,1,1) | 0.509    | 0.02     | 0.018    |            |            | 0.039    | 0.937      | -691.282 | -671.081 |
| ARIMA(1,0,1)(1,1,1) | 0.569    |          |          | 0.059      |            | 0.041    | 0.9416     | -694.399 | -678.238 |
| ARIMA(1,0,2)(1,1,1) | 0.602    |          |          | 0.091      | 0.0271     | 0.042    | 0.941      | -692.399 | -672.197 |
| ARIMA(2,0,1)(1,1,1) | 0.362    | 0.105    |          | -0.151     |            | 0.041    | 0.94       | -692.175 | -671.974 |
| ARIMA(3,0,1)(1,1,1) | 0.707    | -0.083   | 0.017    | 0.197      |            | 0.039    | 0.936      | -690.175 | -665.934 |

$\phi_1$ : ضریب خودهمبستگی غیرفصلی مرتبه ۱  
 $\theta_1$ : ضریب میانگین متحرک غیرفصلی مرتبه ۱  
 $\Theta_1$ : ضریب میانگین متحرک فصلی مرتبه ۱  
 $\phi_2$ : ضریب خودهمبستگی غیرفصلی مرتبه ۲  
 $\theta_2$ : ضریب میانگین متحرک غیرفصلی مرتبه ۲  
 $\Phi_1$ : ضریب خودهمبستگی فصلی مرتبه ۱

شناخته می‌شود. بنابراین، مدل (1,0,1) (1,0,0) ARIMA با معیار آکایک -703.844 و معیار شوارتز -691.724 به عنوان

با توجه به جدول (۲) مدلی که معیارهای آکایک و شوارتز کم‌تری داشته باشد، به عنوان مدل بهینه‌تر در سری زمانی

سپس یک گام زمانی آینده پیش‌بینی گردید. در مرحله بعد برای پیش‌بینی یک گام زمانی بعد مقدار مشاهداتی دبی در گام زمانی قبل را به داده‌ها اضافه کرده و پارامترهای مدل برای یک گام زمانی آینده تعیین گردید. در جدول (۳) مقادیر دبی مشاهداتی و پیش‌بینی شده با سه مدل مختلف را در داده‌های تست آورده شده است.

مدل بهینه‌تر انتخاب شد و در گام بعدی با این مدل اقدام به پیش‌بینی دبی شد. با توجه به اینکه مدل‌های سری زمانی در پیش‌بینی‌های بیش‌تر از یک گام زمانی بسیار بد عمل می‌کنند، بنابراین، فقط یک گام زمانی آینده پیش‌بینی گردید و اصطلاحاً در هر گام زمانی پیش‌بینی به هنگام شد. بدین صورت که در هر گام زمانی برای پیش‌بینی مقدار دبی آینده در ابتدا با استفاده از مقادیر قبلی دبی پارامترهای مدل منتخب تعیین و

جدول ۳- مقادیر دبی مشاهداتی و پیش‌بینی شده با مدل‌های سری زمانی، شبکه عصبی و SVM

Table 3 Observed and predicted discharge values with time series models, neural network and SVM

| SVM   | ANN   | Time series | Observed | Row | SVM   | ANN   | Time series | Observed | Row |
|-------|-------|-------------|----------|-----|-------|-------|-------------|----------|-----|
| 3.06  | 6.47  | 6.11        | 1.00     | 37  | 4.47  | 2.03  | 2.87        | 2.73     | 1   |
| 3.34  | 4.48  | 2.96        | 1.34     | 38  | 3.62  | 2.97  | 3.88        | 1.68     | 2   |
| 3.24  | 3.15  | 3.51        | 1.22     | 39  | 4.80  | 4.53  | 3.84        | 3.14     | 3   |
| 3.57  | 3.40  | 2.93        | 1.61     | 40  | 3.93  | 4.55  | 4.84        | 2.06     | 4   |
| 9.91  | 3.44  | 6.44        | 9.44     | 41  | 8.69  | 4.92  | 6.87        | 7.92     | 5   |
| 11.06 | 10.64 | 18.57       | 10.87    | 42  | 8.41  | 9.26  | 15.25       | 7.58     | 6   |
| 6.76  | 15.35 | 23.63       | 5.55     | 43  | 16.54 | 10.73 | 19.77       | 17.87    | 7   |
| 4.76  | 8.06  | 6.24        | 3.09     | 44  | 16.98 | 24.46 | 10.17       | 18.45    | 8   |
| 4.77  | 6.32  | 1.71        | 3.10     | 45  | 3.87  | 21.32 | 7.53        | 1.99     | 9   |
| 3.11  | 5.48  | 1.38        | 1.05     | 46  | 7.60  | 4.00  | 0.73        | 6.58     | 10  |
| 2.33  | 4.46  | 1.47        | 0.09     | 47  | 3.81  | 7.66  | 3.61        | 1.92     | 11  |
| 4.53  | 2.84  | 0.32        | 2.81     | 48  | 2.85  | 5.79  | 1.67        | 0.73     | 12  |
| 8.72  | 2.77  | 4.37        | 7.96     | 49  | 4.05  | 3.67  | 3.87        | 2.21     | 13  |
| 17.58 | 9.94  | 6.11        | 19.24    | 50  | 5.09  | 3.22  | 3.44        | 3.49     | 14  |
| 10.39 | 25.22 | 12.00       | 10.03    | 51  | 7.67  | 5.18  | 5.12        | 6.67     | 15  |
| 12.48 | 8.91  | 8.21        | 12.66    | 52  | 5.73  | 8.82  | 6.49        | 4.28     | 16  |
| 33.62 | 18.20 | 16.23       | 45.50    | 53  | 11.05 | 7.32  | 9.27        | 10.86    | 17  |
| 36.25 | 61.56 | 38.82       | 52.15    | 54  | 28.28 | 14.92 | 16.79       | 35.12    | 18  |
| 24.23 | 37.98 | 46.84       | 28.61    | 55  | 17.78 | 39.58 | 38.94       | 19.51    | 19  |
| 21.00 | 3.94  | 16.41       | 23.91    | 56  | 19.66 | 7.25  | 12.46       | 22.06    | 20  |
| 4.93  | 15.21 | 8.47        | 3.29     | 57  | 4.70  | 22.15 | 7.46        | 3.02     | 21  |
| 3.59  | 2.88  | 1.61        | 1.64     | 58  | 4.02  | 3.22  | 1.46        | 2.17     | 22  |
| 7.45  | 4.80  | 1.66        | 6.39     | 59  | 11.16 | 4.92  | 1.95        | 10.99    | 23  |
| 3.46  | 7.14  | 3.36        | 1.49     | 60  | 3.95  | 13.22 | 4.25        | 2.08     | 24  |
| 4.24  | 5.55  | 4.11        | 2.44     | 61  | 3.76  | 5.75  | 5.44        | 1.85     | 25  |
| 4.49  | 4.03  | 4.34        | 2.76     | 62  | 6.17  | 4.21  | 3.53        | 4.82     | 26  |
| 5.98  | 4.90  | 4.26        | 4.59     | 63  | 5.64  | 5.88  | 6.27        | 4.17     | 27  |
| 5.00  | 6.35  | 5.27        | 3.38     | 64  | 5.91  | 6.93  | 5.14        | 4.50     | 28  |
| 9.89  | 6.27  | 9.85        | 9.41     | 65  | 12.16 | 6.99  | 9.49        | 12.25    | 29  |
| 9.91  | 12.29 | 17.99       | 9.43     | 66  | 33.58 | 17.03 | 19.49       | 45.41    | 30  |
| 24.71 | 13.07 | 18.84       | 29.35    | 67  | 40.73 | 59.87 | 41.42       | 75.74    | 31  |
| 3.27  | 29.63 | 15.42       | 1.25     | 68  | 25.18 | 45.46 | 31.06       | 30.08    | 32  |
| 5.24  | 4.29  | 1.08        | 3.68     | 69  | 9.28  | 3.15  | 9.43        | 8.65     | 33  |
| 3.26  | 4.50  | 1.52        | 1.24     | 70  | 11.75 | 1.70  | 3.06        | 11.74    | 34  |
| 2.40  | 4.79  | 1.59        | 0.17     | 71  | 7.21  | 17.00 | 6.64        | 6.10     | 35  |
| 0.42  | 3.02  | 0.48        | 0.32     | 72  | 4.85  | 8.36  | 3.09        | 3.19     | 36  |

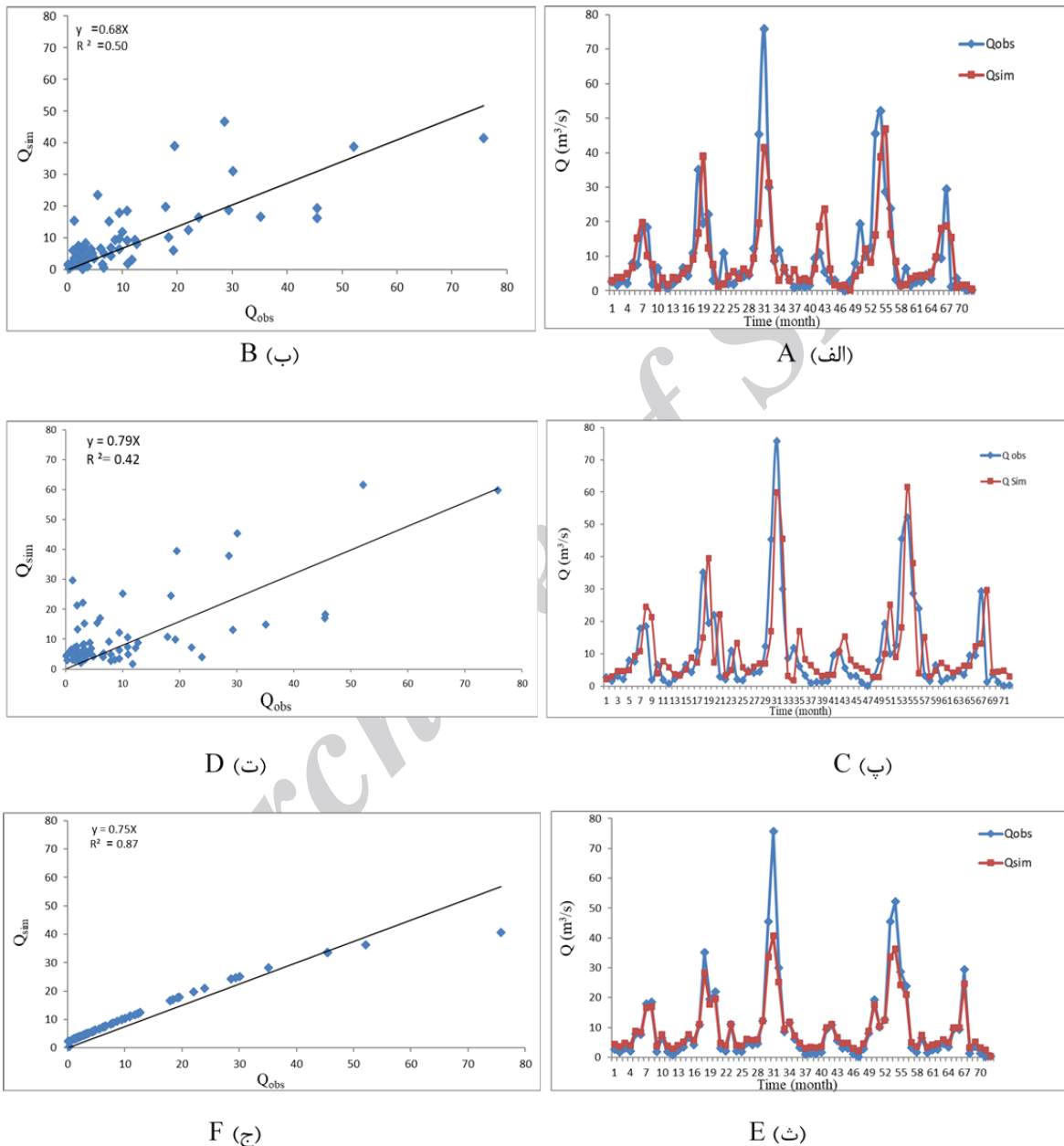
با یک لایه مخفی و پنج ورودی با تعداد ۱۰ نرون مورد بررسی قرار گرفت و در نهایت در این قسمت نیز با استفاده از معیارهای ضریب همبستگی، RMSE و MAE بهترین ورودی‌ها و

در شبکه عصبی ساختار شبکه شامل تعداد لایه‌های مخفی، تعداد نرون‌های هر یک از لایه‌های مخفی، نوع تابع فعال‌سازی و ورودی‌های مدل که بر روی خروجی مدل تاثیر گذاشته، باید تعیین گردد. برای این منظور در ایستگاه مورد نظر شبکه‌هایی



انتخاب شد که مقادیر ضریب همبستگی، RMSE و MAE برای این شبکه به ترتیب ۲۷/۹۳، ۵/۳۱ و ۲/۵۹ به دست آمد. شکل (۴) هیدروگراف و دیاگرام پراکنش دبی‌های پیش‌بینی-شده با سه مدل سری زمانی، شبکه عصبی و SVM را نشان می‌دهد.

ساختار مدل شبکه عصبی انتخاب گردید. با توجه به نتایج به-دست آمده شبکه با ورودی ۲ و تعداد ۱۰ نرون به عنوان شبکه برتر انتخاب گردید که مقادیر ضریب همبستگی، RMSE و MAE برای این شبکه به ترتیب ۸۱/۶۴، ۹/۸۸ و ۶/۹۵ به دست آمد. در مدل SVM شبکه با ورودی ۱ به عنوان شبکه برتر



شکل ۴- الف- مقادیر دبی مشاهده‌ای و پیش‌بینی شده با استفاده از مدل‌های سری زمانی، ب- شبکه عصبی، پ- SVM، ت- دیاگرام پراکنش مقادیر دبی مشاهده‌ای و پیش‌بینی شده با استفاده از مدل‌های سری زمانی، ث- شبکه عصبی و ج- SVM

Fig. 4 a) Observed and predicted discharge values with time series models, b) ANN, c) SVM, d) distribution diagram of observed and predicted discharge values with time series models, e) ANN, and f) SVM

نسبی) و معیار ضریب NSE (ضریب نش) می‌باشند. مقادیر معیارهای ارزیابی برای مدل‌های مختلف در داده‌های تست در جدول (۴) آورده شده است. با توجه به جدول (۴) ملاحظه می‌شود که مقدار RMSE برای مدل SVM برابر با ۵/۳۱ می‌باشد که پایین‌تر از دو مدل دیگر است. هم‌چنین معیار AARE نیز برای مدل SVM کم‌تر از مدل‌های دیگر می‌باشد. با توجه به اینکه هر چقدر این دو معیار (RMSE و AARE) کم‌تر باشد، مدل بهتر می‌باشد، بنابراین مدل SVM با توجه به این دو معیار بهترین عملکرد را داشته است.

#### ۴- نتیجه‌گیری

۱- در مدل سری زمانی بر اساس معیارهای آکایک و شوارتز مدل  $(1,0,1)$   $(1,0,0)$  ARIMA به عنوان مدل برتر انتخاب شد. در مدل شبکه عصبی، شبکه با ورودی ۲ و ۱۰ نرون و در مدل ماشین بردار پشتیبان شبکه با ورودی ۱ به عنوان شبکه‌های برتر انتخاب شدند.

۲- اکثر مدل‌ها مقادیر دبی کمینه را بهتر از مقادیر دبی بیشینه پیش‌بینی کردند. به عبارت دیگر عملکرد مدل در پیش‌بینی دبی‌های کم در مقایسه با دبی‌های زیاد، بهتر بود.

۳- بر خلاف مدل‌های SVM و سری زمانی که عملکرد آن‌ها در دبی‌های پایین مناسب بوده، عملکرد مدل شبکه عصبی در دبی‌های بالا بهتر بوده و جهت پیش‌بینی دبی‌های بالا مناسب‌تر بود.

۴- از بین مدل‌های مورد ارزیابی در این تحقیق مدل SVM بهترین مدل برای پیش‌بینی دبی ماهانه ورودی به سد گرگان انتخاب شد.

با توجه به شکل (۴) ملاحظه می‌شود که نمودارهای مقادیر دبی مشاهداتی و دبی پیش‌بینی شده برای مدل SVM تا حد بسیار زیادی برهم منطبق هستند و در واقع مقادیر دبی پیش‌بینی شده نزدیک به مقادیر مشاهداتی دبی می‌باشند. هم‌چنین با توجه با دیگرام پراکنش ترسیم شده برای مدل‌های مختلف، ملاحظه می‌شود که مدل SVM بهترین تطابق بین دبی‌های مشاهداتی و دبی‌های پیش‌بینی شده را نشان می‌دهد. هم‌چنین مقادیر  $R^2$  برای مدل‌های سری زمانی، شبکه عصبی و SVM به ترتیب برابر با ۰/۵۰، ۰/۴۲ و ۰/۸۷ می‌باشد. همانگونه که مشخص است مقدار  $R^2$  برای مدل SVM از مدل‌های دیگر بالاتر است که بیانگر کارایی بالای این مدل در پیش‌بینی دبی در ایستگاه مورد مطالعه می‌باشد. هم‌چنین با توجه به شکل نمودارها و هم‌چنین مقادیر دبی مشاهداتی و پیش‌بینی شده ملاحظه می‌شود که هر سه مدل مقادیر دبی‌های پایین را بهتر از دبی‌های بالا پیش‌بینی کرده‌اند.

جدول ۴- نتایج معیارهای ارزیابی برای مدل‌های مختلف در ایستگاه قزاقلی

Fig. 4 Results of evaluation criteria for different models at Ghazaghi Station

| NSE  | AARE | RMSE | Model |
|------|------|------|-------|
| 0.62 | 1.20 | 8.84 | ARIMA |
| 0.54 | 1.07 | 5.31 | SVM   |
| 0.81 | 2.78 | 9.88 | ANN   |

برای مقایسه عملکرد مدل‌های مختلف از سه معیار ارزیابی استفاده شد. این معیارها شامل معیار RMSE (میانگین مربعات خطا)، معیار AARE (متوسط مقادیر مطلق خطای

## References

- Abdollah Pour Azad M. R. and Sattari M. T. (2015). Forecasting daily river flow of Ahar Chay River using artificial neural networks (ANN) and comparison with adaptive neuro fuzzy inference system (ANFIS). *J. Water Soil Conserv.*, 22(1), 287-298 [In Persian].
- Ahmadi F., Radmanesh F. and Mirabbasi Najaf Abadi R. (2014). Comparison between genetic programming and support vector machine

methods for daily river flow forecasting (case study: Barandoozchay River). *J. Water Soil.*, 28(6), 1162-1171 [In Persian].

Alborzi M. (2001). Introduction to neural network. Sharif scientific publications, Tehran, Iran. 138pp. [In Persian].

Choubin B., Darabi H., Rahmati O., Sajedi-Hosseini F. and Kløve B. (2018). River suspended sediment modelling using the CART model: a comparative study of machine

- learning techniques. *Sci. Total Environ.*, 615, 272-281.
- Choubin B. and Malekian A. (2017). Combined gamma and M-test-based ANN and ARIMA models for groundwater fluctuation forecasting in semiarid regions. *Environ. Earth Sci.*, 76, 538.
- Choubin B., Zehtabian G., Azareh A., Rafiei-Sardooi E. and Sajedi-Hosseini F. (2018). Precipitation forecasting using classification and regression trees (CART) model: a comparative study of different approaches. *Environ. Earth Sci.*, 77, 314.
- Hornik K., Stinchcombe M., White H. (1989). Multilayer feedforward networks are universal approximators. *Neural Net.*, 2(5), 359-366.
- Jalal Kamali N. (2002). Monthly flow prediction into the Jiroft reservoir Dam by the theory of time series. *Proc. 2004 2<sup>nd</sup> Congress of International River Engineering*, Shahid Chamran University, Ahvaz, Iran [In Persian].
- Karamouz M. and Araghinejd Sh. (2005). *Advanced hydrology*, Amirkabir University of Technology Press, Tehran, Iran. 464pp. [In Persian].
- Karimi Masouleh, H., Ahmadvand, M., Moazed, H., 2010. Neural networks application in Karun river flow prediction based on upstream stations rainfall data during the past six months. *Proc. 2010 8<sup>th</sup> Congress of International River Engineering*, Shahid Chamran University, Ahvaz, Iran, [In Persian].
- Kisi O. (2004). River flow modeling using artificial neural networks. *J. Hydrol. Eng.*, 9(1), 26-38.
- Nanduri U. V. and Swain P. C. (2005). Streamflow forecasting using neuro-fuzzy inference system. *AGU Fall Meeting*.
- Rezaei A. (2004). Regional modeling of peak discharges in sub-watershed of Sefidroud Dam using artificial neural network. PhD Thesis, Tehran University, Tehran, Iran [In Persian].
- Salamatian S. deghani A. and Aghdasian M. (2006). Smart estimating of river discharge using neural network. *Proc. 2006 1<sup>ed</sup> regional conference of optimum utilization of Karoon and Zayandehrud Basins water resources*. University of Shahrekord, Shahrekord, Iran [In Persian].
- Salarijazi M., Ghorbani K., Sohrabian E. and Abdolhosseini M. (2016). Prediction of daily stream-flow using data driven models. *Iran. J. Irrig. Drain.*, 4(10), 479-488 [In Persian].
- Chen S. T., Yu P. S. (2007). Real-time probabilistic forecasting of flood stages. *J. Hydrol.*, 340, 63-77
- Valipour M., Banihabib M. E. and Behbahani S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez Dam River. *J. Hydrol.*, 476, 433-441.

# Application of Time Series, ANN, and SVM Models in Forecasting the Gorgan Dam Inflow Rate

Meisam Samadi<sup>1\*</sup> and Abolhasan Fathabadi<sup>2</sup>

<sup>1</sup>PhD Scholar, Department of Watershed Management Science and Engineering, Faculty of Rangeland and Watershed Management, Gorgan University of Agricultural and Natural Resources, Gorgan, Iran

<sup>2</sup> Assist. Professor, Department of Rangeland and Watershed, Faculty of Natural Resources, Gonbadekavous University, Gonbadekavous, Iran

\*Corresponding author: meisamadi68@gmail.com

## Original Paper

Received: December 30, 2017

Revised: September 7, 2018

Accepted: October 10, 2018

## Abstract

In water resources management, there is a critical need to the prediction of the amount of inflow into the water supply system in order to be aware of future conditions and planning for optimal allocation of water resources to different sectors such as drinking, agriculture and. The aim of this study is to forecasting the monthly inflow to the Gorgan dam for future. To this aim, the data of the Qazaghli station with a 47-years history period and three Time series, neural network and Support vector machine models used for prediction. According to the obtained results, the ARIMA (1, 0, 0) (1, 0, 1) was found to be the premier parsimonious time series model based on the Akaike and Schwarz criteria. Moreover, The ANN model with 2 input and 10 neurons tuning and the SVM model with one input were the best performing models. Finally, according to the obtained results and evaluation criteria, the SVM model has the best efficacy in comparison with two other methods. The RMSE and AARE was 5.31 and 1.07 for SVM model, respectively; 9.88 and 2.78 for neural network, respectively and 8.84 and 1.07 has been obtained for Time Series model, respectively. Based on the results of this research, the best model to predict the monthly discharge input to the Gorgan dam was SVM.

**Keywords:** Akaike; Ghazaghli Station; Schwarz; Modeling; Neuron.