

پیش‌بینی تبخیر ماهانه به کمک مدل‌های خطی و غیرخطی سری زمانی (مطالعه موردی: ایستگاه تبخیرسنجی سد اکباتان)

حامد نوزری^{۱*}، سعید آزادی^۲، نادیا صدق نژاد^۲، سجاد پویان فر^۲

تاریخ دریافت: ۱۴۰۱/۰۳/۱۵

تاریخ پذیرش: ۱۴۰۱/۱۲/۱۶

چکیده

پیش‌بینی تبخیر به‌عنوان یک جزء اصلی چرخه هیدرولوژیکی، اهمیت زیادی در مطالعات هواشناسی و منابع آب دارد. در این پژوهش، کارایی مدل‌های ARIMA، SARIMA، برنامه‌ریزی بیان ژن، رگرسیون خطی چندگانه، مونت کارلو و توماس فیرینگ در پیش‌بینی مقادیر ماهانه تبخیر از تست بررسی گردید. بدین منظور، داده‌های تبخیر ماهانه ایستگاه تبخیرسنجی سد اکباتان در یک دوره ۴۷ ساله (۱۳۹۶-۱۳۵۰) مورد استفاده قرار گرفتند. از آمار دوره ۴۰ ساله ۱۳۸۹-۱۳۵۰ برای واسنجی و از داده‌های سال‌های ۱۳۹۶-۱۳۹۰ جهت اعتبارسنجی نتایج مدل‌ها استفاده گردید. معیارهای ارزیابی ضریب تبیین، ریشه میانگین مربعات خطا، خطای استاندارد، معیار اطلاعاتی آکائیک و ضریب نش-ساتکلیف برای ارزیابی و مقایسه عملکرد مدل‌ها مورد استفاده قرار گرفت. نتایج نشان داد که مدل SARIMA عملکرد دقیق‌تری در پیش‌بینی تبخیر ماهانه داشته و مدل‌های برنامه‌ریزی بیان ژن، ARIMA و رگرسیون خطی چندگانه به ترتیب در رتبه‌های دوم تا چهارم قرار دارند. با توجه به این که مدل برنامه‌ریزی بیان ژن از سهولت کاربست بیشتر و تعداد پارامتر کمتری نسبت به مدل SARIMA برخوردار است، پیش‌بینی را آسان‌تر و در زمان کمتری انجام می‌دهد و در میان روش‌های مورد استفاده قابل توصیه است.

واژه‌های کلیدی: تبخیر، سد اکباتان، سری زمانی، شبیه‌سازی

مقدمه

دارای فرآیندی غیرخطی بوده، تخمین دقیق آن‌ها بسیار مشکل می‌باشد. از این رو ایجاد مدل‌هایی برای پیش‌بینی تبخیر می‌تواند اطلاعات مهمی را در شرایط خشکسالی ارائه دهد (Ahmed et al., 2004; Farahani et al., 2019). در سال‌های اخیر استفاده از مدل‌های هوش مصنوعی و سری زمانی در مدیریت مدرن منابع آب بسیار پرکاربرد بوده و توجه محققان را برای پیش‌بینی اجزای هیدرولوژیکی به خود جلب کرده است. از آن جمله می‌توان به طراحی پروژه‌های آبی، مدیریت کیفیت آب، سیستم آبیاری و بهره‌برداری پایدار از منابع آب اشاره کرد (Wang et al., 2009; Zhang et al., 2019).

امروزه یکی از مشکلات اساسی که زندگی بشر را تحت تأثیر قرار داده است، کمبود آب و توزیع نامناسب آن می‌باشد. بحران آب در کشورهایی که در ناحیه کمربند خشک و نیمه خشک جهان قرار دارند به وضوح دیده می‌شود. در این راستا، پیش‌بینی دقیق تبخیر که یکی از اجزای اصلی چرخه هیدرولوژیکی می‌باشد، برای مدیریت منابع آب و جلوگیری از افزایش بحران آب ضروری است (Maroufpoor et al., 2004; Piri et al., 2019). با توجه به اینکه پدیده‌های هیدرولوژیکی

^۱ دانشیار گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان، ایران.
^۲ دانشجو کارشناسی ارشد مهندسی منابع آب، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان، ایران.

^۱ دانشیار گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان، ایران.
*نویسنده مسئول: (h.nozari@basu.ac.ir)

نحوه ارجاع مقاله:

نوزری، ح، آزادی، س، صدق نژاد، ن، پویان فر، س. ۱۴۰۲. پیش‌بینی تبخیر ماهانه به کمک مدل‌های خطی و غیرخطی سری زمانی (مطالعه موردی: ایستگاه تبخیرسنجی سد اکباتان). نشریه هواشناسی کشاورزی، ۱۱(۱): ۳۶-۴۷. DOI: 10.22125/agmj.2023.345670.1136
Nozari, H., Azadi, S., Sedghnejad, N., Pouyanfar, S. 2023. Predicting monthly evaporation using linear and nonlinear time series models (Case study: Ekbatan Dam station). Journal of Agricultural Meteorology, 11(1): 36-47. DOI: 10.22125/agmj.2023.345670.1136

مدل‌های MLR و ARIMA دو رویکرد متداول برای مدل‌سازی سری‌های زمانی، به ویژه در مدل‌سازی منابع آب بوده‌اند (Maier and Dandy, 1997; Choubin et al., 2016). Wang et al., (2017) در پژوهشی، شش نوع محاسبات نرم شامل مدل‌های پرسپترون چند لایه (MLP)، شبکه عصبی رگرسیون عمومی (GRNN)، ژنتیک فازی (FG)، ماشین بردار پشتیبان حداقل مربع (LSSVM)، اسپلین رگرسیون تطبیقی چند متغیره (MARS) و سیستم‌های استنباط عصبی فازی تطبیقی با پارتیشن شبکه (ANFIS-GP) و دو روش رگرسیون شامل رگرسیون خطی چندگانه (MLR) و مدل استغفس و استوارت (SS) را جهت پیش‌بینی تبخیر تشت ماهانه مورد بررسی قرار دادند. نتایج نشان داد که تکنیک‌های محاسبات نرم به طور کلی بهتر از روش‌های رگرسیون عمل می‌کنند، اما مدل‌های MLR و SS می‌توانند در برخی مناطق آب و هوایی به جای مدل‌های پیچیده غیرخطی ترجیح داده شوند. (Mattar and Alazba, 2018) از مدل‌های GEP و MLR در مدل‌سازی تبخیر تعرق ماهانه مرجع استفاده کردند و دریافتند که مدل GEP عملکرد بهتری داشته است. Wang et al., (2014) در مطالعه‌ای جهت پیش‌بینی بارش ماهانه در منطقه‌ای از چین، مدل بهبود یافته‌ای از ARIMA را ارائه دادند و نشان دادند که دقت مدل بهبود یافته نسبت به ARIMA معمولی به طور قابل توجهی بیشتر است. Bari et al., (2015) برای پیش‌بینی بارش ماهانه در شهر سیلت بنگلادش، از مدل ARIMA استفاده کردند و بیان داشتند که مدل $(1,1,1) * (0,0,1)$ ARIMA برای پیش‌بینی بارش، دقت بسیار بالایی دارد. (Nozari and Tavakoli, 2020) برای پیش‌بینی سه پارامتر بارش، رواناب و تبخیر با استفاده از مدل‌های ARIMA، SVM، ARMAX و W-SVM به این نتیجه رسیدند که به ترتیب مدل‌های SVM، ARIMA، ARMAX و W-SVM برای پیش‌بینی ماهانه تبخیر عملکرد بهتری را دارند. با بررسی مطالعات صورت گرفته می‌توان به اهمیت پیش‌بینی تبخیر با روش‌های هوش مصنوعی و کاربرد آن پی برد. لذا در این تحقیق با استفاده از روش الگوریتم ژنتیک و مقایسه آن با مدل‌های ARIMA، SARIMA،

(2011; Machekposhti et al., 2018). در واقع این مدل‌ها از اهمیت ویژه‌ای در مطالعات زیست‌محیطی و اقتصادی برخوردار می‌باشند و یکی از مزیت‌های استفاده از آن‌ها، سرعت و دقت قابل توجه در پیش‌بینی اجزای مورد مطالعه نسبت به سایر روش‌ها است. از جمله مدل‌های سری زمانی و مدل‌های هوش مصنوعی که منجر به رفع بسیاری از مشکلات زیست‌محیطی و علوم و مهندسی آب شده است می‌توان به مدل‌های AR، MA، ARMA، ARIMA، SARIMA، ماشین بردار پشتیبان (SVM)^۱، شبکه عصبی مصنوعی (ANN)^۲ و برنامه‌نویسی بیان ژن (GEP)^۳ اشاره کرد. مدل GEP از جمله مدل‌هایی است که برای پیش‌بینی پارامترهای هیدرولوژیکی استفاده می‌گردد. در واقع مدل توسعه یافته GEP برای اولین بار توسط Ferreira, (2006) پیشنهاد گردید. Guven et al., (2008) از یک مدل مبتنی بر GEP برای پیش‌بینی تبخیر تعرق مرجع روزانه استفاده کردند و نشان دادند که مدل GEP نسبت به روش‌های معمولی کارآمدتر می‌باشد. Aytok and Alp, (2008) جهت مدل‌سازی روند بارش-رواناب کارایی مدل‌های ANN و GEP را بررسی نمودند. نتایج حاکی از توانایی خوب GEP در مدل‌سازی بارش-رواناب بود و نشان داد می‌توان GEP را به عنوان جایگزینی برای مدل‌های ANN نام برد. (Shiri and Kisi, 2011) با مقایسه نتایج مدل‌های هوش مصنوعی در برآورد تبخیر روزانه در استان خوزستان به این نتیجه رسید که مدل GEP بهتر از مدل‌های سیستم منطق-فازی (NF) و ANN عمل می‌کند. Shiri et al., (2012) برای پیش‌بینی تبخیر تعرق مرجع، مدل‌های GEP، ANFIS، Priestley-Taylor و Hargreaves-Samani را مورد بررسی قرار دادند. نتایج مطالعه آن‌ها، برتری مدل GEP را نسبت به سایر روش‌ها نشان داد. (Terzi, 2013) در پژوهشی از GEP و ANFIS برای مدل‌سازی روزانه تبخیر تشت استفاده کرد. نتایج حاکی از برتری مدل GEP نسبت به ANFIS بود. (Goodarzi and Eslamian, 2018) عملکرد مدل‌های GP، MLR و ANN را جهت پیش‌بینی تبخیر روزانه مورد بررسی قرار دادند. نتایج نشان داد که دقت مدل GP برای پیش‌بینی تبخیر بالاتر از سایر مدل‌ها می‌باشد. در دهه‌های گذشته،

³ Gene expression programming

¹ Support vector machines

² Artificial Neural Networks

۱/۷۵ کیلومتر مربع و ارتفاع سد اکباتان ۵۳/۸ متر می‌باشد. هدف از احداث این سد در درجه اول تأمین آب شرب شهر همدان و سپس تأمین بخشی از نیازهای آبی زمین‌های کشاورزی پایین دست سد بوده است. لذا به دلیل اهمیت این سد و میزان تبخیر از سطح آن، ایستگاه تبخیرسنجی سد اکباتان جهت مطالعه انتخاب گردید. داده‌های تبخیر مورد استفاده، از مهر ماه سال ۱۳۴۹ تا شهریور سال ۱۳۹۶ از این ایستگاه جمع‌آوری گردیده است. لذا طول دوره آماری در این تحقیق ۴۷ سال می‌باشد که ۴۰ سال اول به‌عنوان دوره واسنجی و ۷ سال باقیمانده نیز به‌عنوان دوره صحت‌سنجی در نظر گرفته شده است. نمودار سری زمانی تبخیر ایستگاه تبخیرسنجی سد اکباتان در شکل ۱ ارائه شده است.

مدل‌های همبستگی سری زمانی

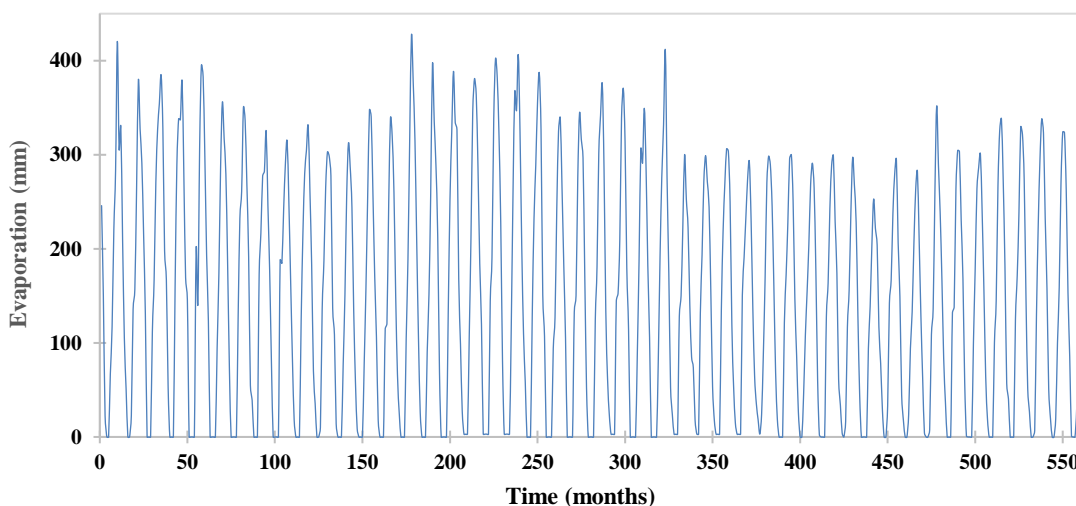
به‌طور کلی مدل‌های سری زمانی می‌توانند اشکال مختلفی داشته باشند و روندهای تصادفی مختلفی را نشان دهند. این مدل‌ها با انجام تجزیه و تحلیل سیستماتیک در هر مرحله (شناسایی، تخمین، تشخیص)، منجر به یافتن مدلی می‌شوند که بتوان به کمک آن پیش‌بینی را برای آینده انجام داد.

رگرسیون خطی و غیرخطی چند متغیره، مونت‌کارلو و توماس‌فیرینگ به پیش‌بینی تبخیر در سد اکباتان همدان پرداخته شد. با توجه به ضرورت تأمین آب شرب شهر همدان و پیش‌بینی تبخیر از این سد، و از طرف دیگر مزایای استفاده از روش GEP به دلیل سادگی ایجاد تنوع ژنتیکی، انجام عملگرهای ژنتیکی در سطح کروموزوم و همچنین طبیعت منحصر به فرد و چند ژنی آن و امکان تکمیل برنامه‌های پیچیده‌تر، تصمیم به استفاده از این روش در پیش‌بینی تبخیر در سد اکباتان گرفته شد. همچنین به‌منظور درک بهتر کارایی این روش در مقایسه با روش‌های دیگر، نتایج آن با سایر مدل‌های مذکور مقایسه و مورد بررسی قرار گرفت.

مواد و روش‌ها

منطقه مورد مطالعه

مطالعه حاضر در ایستگاه تبخیرسنجی سد اکباتان انجام شد که مهمترین سد استان همدان است و در ۱۰ کیلومتری جنوب شرقی شهر همدان بر روی رودخانه یلفان واقع شده است. این سد در عرض جغرافیایی $34^{\circ} 45'$ شمالی و طول جغرافیایی $48^{\circ} 36'$ شرقی قرار دارد. حداکثر سطح مخزن



شکل ۱- سری زمانی تبخیر ماهانه در ایستگاه تبخیرسنجی سد اکباتان همدان (۱۳۴۹-۱۳۹۶)

Figure 1- Monthly evaporation time series at the evaporation station of Ekbatan Dam in Hamadan (1970-2017)

خودهمبسته- میانگین متحرک) و SARIMA (خودهمبسته میانگین متحرک فصلی) که مهمترین این مدل‌ها می‌باشند، اشاره کرد (Box et al., 1994). ارتباط بین مقادیر زمان حال

بدین منظور مدل‌های متعددی ارائه شده است که می‌توان به مدل‌های AR (خودهمبسته)، MA (میانگین متحرک)، ARMA (خودهمبسته- میانگین متحرک)، ARIMA (تلفیقی

مقادیر $\varphi_p(B)$ ، $\varphi_p(B^S)$ ، $\theta_q(B)$ و $\theta_q(B^S)$ به ترتیب چند جمله‌ای از درجه p ، P ، q و Q می‌باشند که به کمک معادله‌های ۵ تا ۸ محاسبه می‌شوند.

$$\varphi_p(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \quad (۵)$$

$$\varphi_p(B^S) = 1 - \varphi_s B^s - \varphi_{2s} B^{2s} - \dots - \varphi_{ps} B^{ps} \quad (۶)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (۷)$$

$$\theta_q(B^S) = 1 - \theta_s B^s - \theta_{2s} B^{2s} - \dots - \theta_{qs} B^{qs} \quad (۸)$$

مدل توماس فیرینگ

مدل توماس فیرینگ یکی از مدل‌های سری زمانی جهت پیش‌بینی می‌باشد. در این مدل میزان پارامتر در ماه‌های آتی با استفاده از داده‌های آماری گذشته و نیز با استفاده از یک متغیر تصادفی بر پایه تابع توزیع احتمال نرمال به دست می‌آید. حالت کلی به صورت معادله ۹ بیان می‌شود (Elizabeth et al., 2010).

$$x_j = \bar{x}_j + b_j(x_{j-1} - \bar{x}_{j-1}) + t_j \sigma_j (1 - r_{j,j-1}^2)^{0.5} \quad (۹)$$

که x میزان تبخیر، \bar{x} میانگین تبخیر، σ انحراف معیار مقادیر تبخیر، t_j عدد نرمال استاندارد با میانگین صفر و واریانس یک و j بیانگر ماه می‌باشد. ضریب b_j به صورت معادله ۱۰ محاسبه می‌گردد.

$$b_j = r_{j,j-1} \frac{\sigma_{j-1}}{\sigma_j} \quad (۱۰)$$

که r ضریب همبستگی می‌باشد. به دلیل این که در بیشتر سال‌ها در ماه‌های آذر، دی، بهمن و اسفند مقدار تبخیر صفر گزارش شده است، لذا پیش‌بینی تبخیر در ماه‌های ذکر شده به کمک این روش انجام گرفته است.

مدل رگرسیون خطی

مدل رگرسیون خطی از انواع مدل‌های تجربی می‌باشند که به‌طور گسترده جهت بررسی داده‌های هیدرولوژیکی استفاده می‌شوند. انواع مختلف مدل رگرسیون خطی شامل رگرسیون خطی ساده، رگرسیون لجستیک، رگرسیون چند جمله‌ای و رگرسیون خطی چند متغیره می‌باشد (Ghani and Ahmad., 2010; Kis and Ozkan ., 2019). رگرسیون خطی چند متغیره به دلیل فرمول‌بندی ساده بیشتر مورد توجه و استفاده قرار گرفته است (Rencher and Schaalje., 2008; Wang et al., 2013). نحوه عملکرد رگرسیون خطی به‌صورت

و زمان‌های پیشین و سادگی در عمل و ساختار در این مدل‌ها منجر به استقبال از آن‌ها گردیده است (Nozari and Tavakoli, 2020).

مدل ARIMA(p,d,q)

مدل ARMA در کاربردهای وسیعی از مسائل هیدرومتئورولوژی مورد استفاده قرار گرفته است. اما برای استفاده از آن باید فرآیند ایستایی برقرار باشد. بنابراین Bax et al., (1994) با اضافه کردن مرتبه تفاضلی d ، مدل ARIMA را ارائه نمودند. این مدل جزء مدل‌های کلاسیک برای پیش‌بینی سری‌های زمانی است که به عنوان مدل Box و Jenkins شناخته می‌شود (Faruk et al., 2010). فرم پایه مدل‌های تلفیقی خودهمبسته- میانگین متحرک ARIMA(p,d,q) برای متغیر نرمال استاندارد شده Z_t به صورت معادله ۱ بوده که در آن $Z_t(1-B)^d$ سری تفاضلی شده Z_t می‌باشد که از یک فرآیند ایستای ARMA(p,q) پیروی می‌کند. لذا اگر $d=0$ باشد، سری اولیه Z_t ایستا می‌شود.

$$\varphi(B)(1-B)^d Z_t = \theta(B)\varepsilon_t + \theta_0 \quad (۱)$$

در این معادله $\varphi(B)$ عملگر مدل AR، $\theta(B)$ عملگر مدل MA، d مرتبه تفاضلی، θ_0 پارامتر روند قطعی، ε_t متغیر تصادفی مستقل از زمان و Z_t متغیر نرمال استاندارد شده وابسته به زمان می‌باشند. مقادیر $\varphi(B)$ و $\theta(B)$ در معادله ۱ به ترتیب چند جمله‌ای از درجه p و q بوده و به‌صورت معادله‌های ۲ و ۳ می‌باشند.

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \quad (۲)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (۳)$$

مدل SARIMA(p,d,q)(P,D,Q)s

در صورتی که اثر فصلی در سری زمانی وجود داشته باشد، از کلاس دیگر الگوهای احتمالی ARIMA که الگوی فصلی است استفاده می‌گردد. فرم پایه این مدل به صورت معادله ۴ ارائه شده که در آن p ، d ، P ، D ، Q مرتبه‌های مدل بوده و s طول فصل می‌باشد.

$$\varphi_p(B)\varphi_p(B^S)(1-B)^d(1-B^S)^D Z_t = \theta_q(B)\theta_q(B^S)\varepsilon_t + \theta_0 \quad (۴)$$

ضریب تبیین (R^2)، معیار اطلاعاتی آکائیک (AIC) و ضریب نش-ساتکلیف (NSE) استفاده شد. شکل ریاضی این آماره‌ها به صورت معادله‌های ۱۲ الی ۱۶ می‌باشد (Akaike., 1979; Azadi et al., 2020).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_m - Y_p)^2}{n}} \quad (12)$$

$$SE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (Y_m - Y_p)^2}}{\bar{Y}_m} \quad (13)$$

$$R^2 = \frac{[\sum_{i=1}^n (Y_m - \bar{Y}_m)(Y_p - \bar{Y}_p)]^2}{\sum_{i=1}^n (Y_m - \bar{Y}_m)^2 \cdot \sum_{i=1}^n (Y_p - \bar{Y}_p)^2} \quad (14)$$

$$AIC = N \cdot \ln \sigma_e^2 + 2(k) \quad (15)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (Y_m - Y_p)^2}{\sum_{i=1}^n (Y_m - \bar{Y}_m)^2} \quad (16)$$

که n تعداد ماه‌های دوره مورد مطالعه، Y_m مقادیر اندازه‌گیری شده در هر روز، Y_p مقدار پیش‌بینی شده با استفاده از مدل، \bar{Y}_m میانگین داده‌های اندازه‌گیری شده، \bar{Y}_p میانگین داده‌های پیش‌بینی شده، N تعداد نمونه، k نشان‌دهنده مجموع پارامترهای مدل و σ_e^2 واریانس باقی‌مانده‌ها می‌باشند.

نتایج و بحث

در این تحقیق عملکرد مدل‌های سری زمانی ARIMA و SARIMA و مدل‌های رگرسیون خطی و مدل GEP جهت پیش‌بینی تبخیر مورد تجزیه و تحلیل قرار گرفت.

پیش‌بینی تبخیر به کمک مدل ARIMA

در این روش، پیش‌بینی برای هر ماه به صورت جداگانه انجام گرفت. ایستایی در میانگین سری‌های زمانی با توجه به میانگین و واریانس مورد بررسی قرار گرفت و با استفاده از روش اندرسون دارلینگ نرمال بودن سری‌های زمانی در سطح اطمینان ۹۵ درصد آزمون شد. با استفاده از نمودارهای خودهمبستگی و خودهمبستگی جزئی سری‌های زمانی، مرتبه پارامترهای مدل ARIMA تعیین شد و در مرحله واسنجی، مدلی به عنوان بهترین مدل انتخاب گردید که دارای کمترین خطای استاندارد، کمترین ریشه میانگین مربعات خطا و همچنین کمترین مقدار آکائیک بود. در ادامه جهت

معادله ۱۱ بیان می‌شود که در آن X_1, X_2, \dots, X_n متغیرهای مستقل و $b_0, b_1, b_2, \dots, b_n$ ضرایب ثابت می‌باشند.

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n \quad (11)$$

ضرایب ثابت برای هر متغیر به گونه‌ای تخمین زده می‌شود که مجموع مربعات بین مقادیر مشاهده‌ای و شبیه‌سازی متغیر Y به حداقل برسد. در این تحقیق از نرم افزار SPSS جهت پیش‌بینی به کمک مدل رگرسیون خطی، استفاده شد.

مدل برنامه‌ریزی بیان ژن (GEP)

این مدل اولین بار توسط Ferreira در سال ۱۹۹۹ معرفی شد و در واقع شکل توسعه یافته برنامه‌ریزی ژنتیک (GP) بوده و روش مناسبی برای مدل‌سازی فرایندهای غیرخطی می‌باشد (Ferreira., 2006). این روش مانند الگوریتم ژنتیک و برنامه‌ریزی ژنتیک، جزء الگوریتم‌های تکاملی است که از جمعیتی از افراد استفاده کرده و آن‌ها را مطابق برازندگی انتخاب می‌کند و تغییرات ژنتیکی را با استفاده از یک یا چند عملگر اعمال می‌کند. در این نوع الگوریتم، کدگذاری‌ها به شکل کروموزوم‌های خطی با طول ثابت و قابل بیان به شکل ساختارهای شاخه‌ای هستند. ساختار خطی کروموزوم‌ها موجب می‌شود استفاده از عملگرهای ژنتیکی نظیر جهش و ترکیب، همیشه ساختارهای صحیح و معتبری تولید کند (Ferreira., 2001). می‌توان گفت ایجاد تنوع ژنتیکی در این روش بسیار ساده است، زیرا عملگرهای ژنتیکی در سطح کروموزوم عمل می‌کنند. یکی دیگر از نقاط قوت GEP طبیعت منحصر به فرد و چند ژنی آن است که امکان تکمیل برنامه‌های پیچیده‌تر متشکل از چندین زیر برنامه را فراهم می‌کند. در نتیجه، GEP نسبت به GP قدیمی بسیار قدرتمندتر می‌باشد (Ferreira., 2001a; Ferreira., 2001b). در این تحقیق جهت پیش‌بینی به کمک مدل رگرسیون غیرخطی چندمتغیره، نرم‌افزار GenXproTools مورد استفاده قرار گرفته است.

معیارهای ارزیابی

در پژوهش حاضر، جهت ارزیابی و صحت‌سنجی نتایج مدل‌ها در پیش‌بینی تبخیر، از شاخص‌های آماری ریشه میانگین مربعات خطا (RMSE)، خطای استاندارد (SE)،

نبوده و از روش‌های مونت کارلو و توماس فیرینگ جهت پیش‌بینی استفاده شده است. نتایج محاسبه معیارهای ارزیابی در مرحله واسنجی در جدول ۱ ارائه شده است. در این مرحله با استفاده از مدل‌های برگزیده مذکور، مقادیر تبخیر در ایستگاه مورد مطالعه پیش‌بینی، و به منظور بررسی صحت نتایج با داده‌های مشاهده‌ای مقایسه گردید (شکل ۲).

تأیید الگوهای برگزیده، نمودارهای خودهمبستگی و خودهمبستگی جزئی سری باقیمانده‌ها رسم گردید و استقلال و تصادفی بودن آن‌ها بررسی و صحت الگوهای برگزیده کنترل گردید. لازم به ذکر است، به دلیل این که در بیشتر سال‌ها در ماه‌های آذر، دی، بهمن و اسفند مقدار تبخیر صفر گزارش شده است، لذا پیش‌بینی به کمک مدل ARIMA امکانپذیر

Table 1- The values of evaluation criteria of selected ARIMA models at the calibration step
جدول ۱- مقادیر معیارهای ارزیابی مدل‌های برگزیده ARIMA در دوره واسنجی

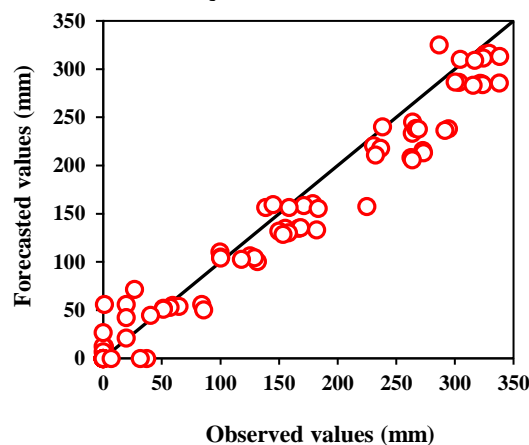
Month	Selected model	R ²	RMSE (mm.day ⁻¹)	SE	AIC	NSE
October	ARIMA(1,1,1)	0.36	22.91	0.14	10.90	-0.22
November	ARIMA(0,1,1)	0.31	24.49	0.29	10.38	-1.15
April	ARIMA(1,1,3)	0.27	25.12	0.20	14.27	-4.64
May	ARIMA(0,1,2)	0.21	35.84	0.20	10.13	-2.46
June	ARIMA(1,1,0)	0.25	37.85	0.14	7.77	-0.22
July	ARIMA(1,1,0)	0.51	31.70	0.10	8.76	0.43
August	ARIMA(1,1,0)	0.32	38.83	0.12	8.29	-0.01
September	ARIMA(0,1,1)	0.53	24.46	0.09	8.50	0.37

و نشان دادند که مدل مذکور از دقت قابل قبولی برخوردار است. لذا می‌توان نتیجه گرفت که مدل استفاده شده در این تحقیق از دقت بسیار خوبی برخوردار است.

پیش‌بینی تبخیر به کمک مدل SARIMA

به‌منظور بررسی ماهیت فصلی داده‌ها، نمودارهای خودهمبستگی و خودهمبستگی جزئی سری زمانی تهیه شد (شکل ۳). پس از مشاهده فرآیند فصلی در سری زمانی، پیش‌بینی به صورت فصلی و به کمک این مدل انجام گردید. در این حالت ایستایی در میانگین و ایستایی در واریانس سری زمانی به کمک تحلیل روند و روش Box-Cox مورد تجزیه و تحلیل قرار گرفت. شکل ۴ ایستایی در واریانس سری زمانی را پس از تبدیل Box-Cox برای پارامتر تبخیر ماهانه اصلاح شده با عامل λ را نشان می‌دهد (λ عامل تبدیل Box-Cox می‌باشد). لازم به ذکر است که در تبدیل Box-Cox جهت ایستایی داده‌ها، سری زمانی به توان λ می‌رسد، اما اگر λ برابر صفر باشد، از سری زمانی لگاریتم گرفته می‌شود. منظور از Lower CL و Upper CL، به ترتیب حد اطمینان پایین و بالا برای فاصله اطمینان ۹۵ درصد برای λ می‌باشد. هر دو حد اطمینان در شکل ۴ به صورت خطوط عمودی نشان داده شده است. اگر فاصله اطمینان برای λ برابر ۱ باشد، هیچ تغییری

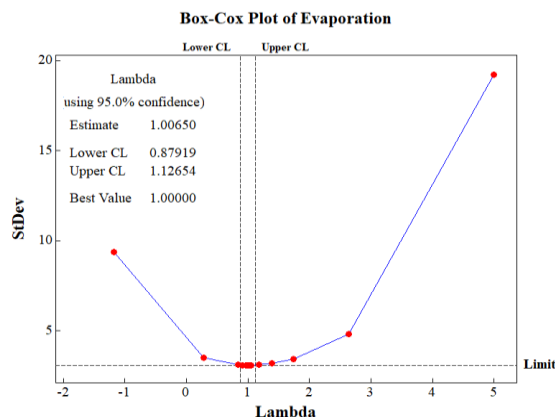
نتایج نشان داد که مقدار SE، RMSE، R² و NSE به ترتیب برابر با ۰/۱۹، ۲۷/۲ میلی‌متر بر ماه، ۰/۹۷ و ۱ می‌باشد. (Azadi et al., (2020) در تحقیقی با برآزش مدل استوکاستیکی با ضرایب ثابت بر روی سری زمانی ماهانه دبی رودخانه و استفاده از منحنی سنجه رسوب رودخانه در پیش‌بینی رسوبات معلق، از مدل فصلی و ضربی ARIMA(3,0,2)*(3,1,1) استفاده کردند.



شکل ۲- مقایسه میزان تبخیر مشاهده‌ای و پیش‌بینی شده با مدل ARIMA در مرحله صحت‌سنجی

Figure 2- The comparison of observed and predicted evaporation with ARIMA model in validation step

این محققین شاخص‌های SE و R² را در شبیه‌سازی جریان متوسط ماهانه به ترتیب ۰/۳۴ و ۰/۸۶ به دست آوردند



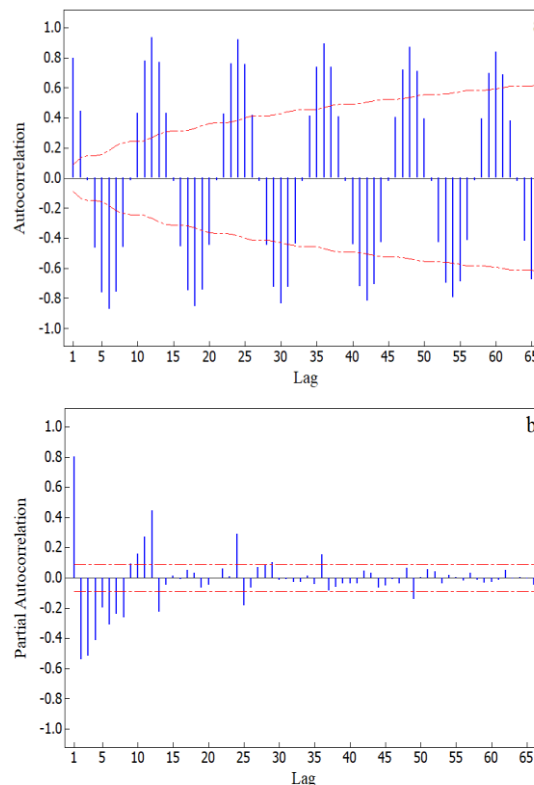
شکل ۴- تغییرات انحراف معیار تبخیر ماهانه در ایستگاه سد

اکباتان در برابر Lambda

Figure 4- Changes of monthly evaporation standard deviation at Ekbatan Dam station versus Lambda

در ادامه جهت تأیید الگوی برگزیده، نمودارهای خودهمبستگی و خودهمبستگی جزئی این الگو رسم گردید و استقلال و تصادفی بودن باقیمانده‌ها و در واقع صحت الگوی برگزیده اثبات شد. همچنین مقادیر تبخیر در ایستگاه مورد مطالعه با استفاده از الگوی برگزیده برای دوره صحت‌سنجی پیش‌بینی گردید. نمودار مقادیر مشاهده شده و پیش‌بینی شده توسط این مدل در شکل ۵ ارائه شده است. نتایج پیش‌بینی به کمک الگوی منتخب مدل SARIMA نشان داد که مقدار SE، RMSE، R^2 و NSE به ترتیب برابر با ۰/۱۷، ۲۳/۶ میلی‌متر بر ماه و ۰/۹۸ و ۰/۹۵ می‌باشد. با مقایسه نتایج مدل‌های ARIMA و SARIMA می‌توان گفت مدل SARIMA از دقت بالاتری در پیش‌بینی تبخیر برخوردار است. Shirvani and Honar, (2011) در ایستگاه باجگاه از سری زمانی SARIMA جهت تبخیر تفرق استفاده کردند. این محققین مدل‌های SARIMA(3,0,2)(0,1,4) و SARIMA(3,0,2)(0,1,4) را پیشنهاد کردند. در نهایت نتایج نشان داد که مدل‌های سری زمانی، ابزار دقیق و مناسبی برای پیش‌بینی مقادیر تبخیر تفرق می‌باشند. همچنین Babamiri and et al., (2017) به برآورد و پیش‌بینی تبخیر تفرق پتانسیل براساس مدل‌های تصادفی سری زمانی در ایستگاه تبریز پرداختند. مدل‌سازی و پیش‌بینی سری‌های زمانی تبخیر تفرق با استفاده از مدل‌های ARIMA و SARIMA مورد مطالعه قرار گرفت. در این تحقیق از داده‌های سال‌های ۱۹۸۶ الی ۲۰۱۰

لازم نبوده و سری زمانی، در واریانس ایستا است. در این شکل، فاصله اطمینان بین ۰/۸۸ و ۱/۱۳ قرار گرفته که شامل ۱ بوده و لذا هیچ تغییری برای سری زمانی نیاز نیست (Box and Cox., 1964; He et al., 2013). لذا با توجه به شکل ۴ می‌توان گفت واریانس سری زمانی ایستا می‌باشد. مرتبه پارامترهای مدل SARIMA در ابتدا به کمک نمودارهای خودهمبستگی و خودهمبستگی جزئی سری زمانی انتخاب گردید و سپس جهت دستیابی به بهترین مدل، مرتبه‌های دیگر نیز آزمون شدند. در جدول ۲ به طور مختصر نتایج تعدادی از مدل‌های SARIMA در مرحله واسنجی ارائه شده است. در این مرحله، مدل SARIMA(1,0,0)(2,1,1) به دلیل داشتن کمترین خطای استاندارد و همچنین کمترین مقدار آکائیک به عنوان الگوی برگزیده انتخاب شد.



شکل ۳- خودهمبستگی (a) و خودهمبستگی جزئی (b) تبخیر در ایستگاه سد اکباتان با حدود اطمینان ۵ درصد
Figure 3- Self-correlation (a) and partial self-correlation (b) of evaporation at Ekbatan Dam station with 5% confidence interval

استفاده شد. محاسبات مربوطه در دو حالت متمایز انجام گردید. در حالت اول همه داده‌ها به عنوان یک سری سالانه و در حالت دوم سری‌های ماهانه به صورت مجزا در طول دوره آماری انتخاب شد. نتایج نشان داد حالت اول با مدل

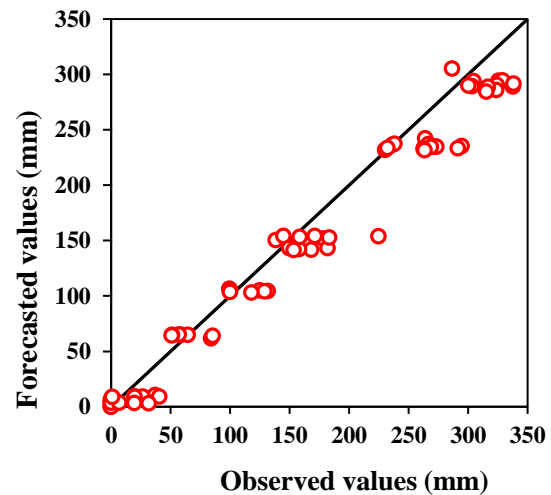
مطابق با جدول ۴ الگوی شماره ۱۰ مقدار SE و RMSE کمتری نسبت به سایر الگوها داشته، در نتیجه به عنوان الگوی مناسب جهت پیش‌بینی انتخاب می‌شود. در معادله ۱۶ رابطه رگرسیونی منتخب جهت پیش‌بینی تبخیر آورده شده است.

Table 2- The values of evaluation criteria of SARIMA models in the calibration step

جدول ۲- مقادیر معیارهای ارزیابی مدل‌های SARIMA در دوره واسنجی

Selected model	R ²	RMSE (mm.d ⁻¹)	SE	AIC	NSE
SARIMA(1,0,0) (2,1,1)	0.96	25.50	0.17	8.82	0.96
SARIMA(2,0,0) (2,2,1)	0.95	27.97	0.19	10.47	0.95
SARIMA(0,0,1) (2,1,2)	0.96	25.85	0.17	10.78	0.96
SARIMA(0,0,2) (2,1,2)	0.96	25.67	0.17	12.81	0.96
SARIMA(1,0,1) (1,1,2)	0.96	25.34	0.17	10.79	0.96
SARIMA(2,0,1) (1,2,1)	0.95	29.50	0.20	10.29	0.95
SARIMA(1,0,2) (0,2,1)	0.94	32.25	0.22	7.57	0.94
SARIMA(2,0,2) (1,2,1)	0.95	29.32	0.20	12.30	0.95
SARIMA(1,1,1) (1,2,2)	0.95	29.85	0.20	10.26	0.94
SARIMA(2,1,1) (0,2,2)	0.95	27.65	0.19	10.42	0.95
SARIMA(0,1,2) (2,1,1)	0.96	62.16	0.18	10.75	0.96
SARIMA(2,1,2) (1,1,0)	0.95	28.39	0.19	10.39	0.95

همچنین نمودار مقادیر مشاهده‌ای و پیش‌بینی شده توسط الگوی منتخب در دوره صحت‌سنجی نیز در شکل ۶ ارائه شده است. نتایج پیش‌بینی تبخیر به کمک این مدل نشان داد که مقدار خطای استاندارد ۰/۲۶، مقدار ریشه میانگین مربعات خطا ۳۶/۷۵ میلی‌متر بر ماه، ضریب تبیین برابر با ۰/۹۳ و ضریب نش-ساتکلیف ۰/۹۱ می‌باشد.



شکل ۵- مقایسه میزان تبخیر مشاهده‌ای و پیش‌بینی شده با مدل SARIMA در مرحله صحت‌سنجی

Figure 5- The comparison of observed and predicted evaporation with SARIMA model in validation step

جدول ۳- الگوهای ورودی سری زمانی جهت پیش‌بینی تبخیر ماهانه

Table 3- Time series input patterns for predicting monthly evaporation

Input pattern number	Input pattern structure
1	$E(t) = f\{E(t-1), E(t-2)\}$
2	$E(t) = f\{E(t-1), \dots, E(t-3)\}$
3	$E(t) = f\{E(t-1), \dots, E(t-4)\}$
4	$E(t) = f\{E(t-1), \dots, E(t-5)\}$
5	$E(t) = f\{E(t-1), \dots, E(t-6)\}$
6	$E(t) = f\{E(t-1), \dots, E(t-7)\}$
7	$E(t) = f\{E(t-1), \dots, E(t-8)\}$
8	$E(t) = f\{E(t-1), \dots, E(t-9)\}$
9	$E(t) = f\{E(t-1), \dots, E(t-10)\}$
10	$E(t) = f\{E(t-1), \dots, E(t-11)\}$

پیش‌بینی تبخیر با مدل رگرسیونی خطی

جهت پیش‌بینی تبخیر ماهانه به کمک مدل رگرسیونی خطی از نرم‌افزار IBM SPSS Statistics 25 استفاده شد. برای شبیه‌سازی، داده‌های ماهانه تبخیر با توالی بازگشت (تأخیر زمانی) تا ۱۲ ماه قبل، به عنوان داده‌های آموزش و در ترکیبات مختلف مطابق جدول ۳ استفاده گردید. همچنین نتایج مربوط به هر یک از الگوهای مذکور در جدول ۴ ارائه شده است.

مدل‌سازی تبخیر ماهانه ابتدا پارامترهای مدل برنامه‌ریزی بیان ژن جهت آموزش مدل انتخاب شدند (جدول ۵). سپس داده‌های ماهانه تبخیر با توالی بازگشت (تأخیر زمانی) تا ۱۲ ماه قبل، به‌عنوان داده‌های آموزش و در ترکیبات مختلف مطابق جدول ۳ استفاده گردید و شبیه‌سازی تبخیر با استفاده از پارامترهای مدل GEP انجام گرفت. نتایج مربوط به الگوهای استفاده شده در جدول ۶ ارائه شده است. طبق نتایج جدول ۶ مشاهده می‌شود که الگوی شماره ۱۰ مقدار SE و RMSE کمتری نسبت به سایر الگوها دارد، در نتیجه به عنوان الگوی نهایی انتخاب گردید. معادله ۱۷ مدل منتخب جهت پیش‌بینی تبخیر می‌باشد.

$$E(t) = \frac{0.249 + \cos\left(\frac{E(t-1) + (-66.99 - (E(t-4) - 0.882))}{2}\right)}{2} E(t-1) + \frac{\tan^{-1}\left(\cos(E(t-5)) \times \left(-3.318 + \frac{E(t-11)}{2}\right)\right) + E(t-11)}{2} + \tan^{-1}\left(\frac{1}{1 - (E(t-2) \times (E(t-6) - 9.163))}\right) \quad (17)$$

به‌منظور بررسی اعتبار مدل مذکور، مقدار تبخیر با استفاده از این الگو و توسط مدل GEP پیش‌بینی شد. نمودار مقادیر مشاهده‌ای و مقادیر پیش‌بینی شده در دوره صحت‌سنجی توسط الگوی منتخب در شکل ۷ ارائه شده است. نتایج پیش‌بینی به کمک این مدل نشان داد که مقدار خطای استاندارد، ریشه میانگین مربعات خطا، ضریب تبیین و ضریب نش-ساتکلیف به ترتیب برابر با ۰/۱۸، ۰/۱۷، ۲۶/۱۷ میلی‌متر بر ماه، ۰/۹۵ و ۰/۹۵ می‌باشد.

جدول ۵- پارامترهای مورد استفاده در اجرای مدل GEP

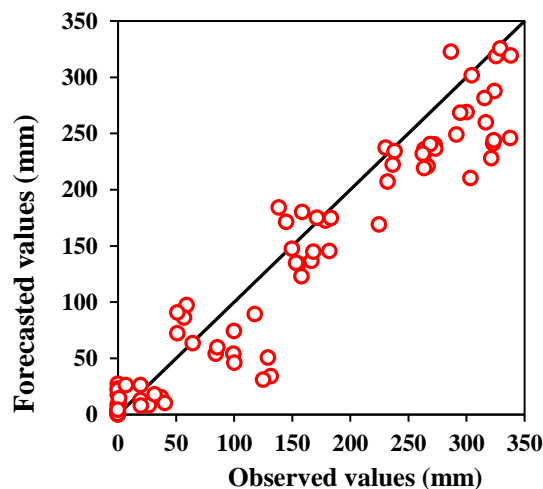
Parameters	Values
Number of chromosomes	30
Head size	8
Number of genes	3
Leap rates	0.044
Inversion rate	0.1
Single-point recombination rate	0.1
Two-point recombination rate	0.3
Gene recombination rate	0.1
Gene transfer rate	0.1
Consecutive song rates	0.1
Root song rate of consecutive insertion	0.1
Two-point recombination rate	0.3
Gene recombination rate	0.1

جدول ۴- نتایج پیش‌بینی تبخیر در مراحل واسنجی و

صحت‌سنجی با مدل رگرسیون خطی

Table 4- Evaporate prediction results in calibration and validation steps with the linear regression model

Input pattern number	stage	R ²	RMSE (mm.d ⁻¹)	SE	NSE
1	Calibration	0.75	64.98	0.44	0.74
	Validation	0.79	55.93	0.39	0.78
2	Calibration	0.79	59.01	0.4	0.86
	Validation	0.79	66.19	0.47	-2.38
3	Calibration	0.83	54.95	0.39	0.81
	Validation	0.87	45.16	0.32	0.86
4	Calibration	0.85	51.59	0.34	0.83
	Validation	0.92	48.04	0.34	0.84
5	Calibration	0.94	47.55	0.32	0.86
	Validation	0.91	38.13	0.27	0.9
6	Calibration	0.88	46.07	0.31	1
	Validation	0.93	33.51	0.24	1
7	Calibration	0.88	45.96	0.31	0.87
	Validation	0.93	38.95	0.27	0.9
8	Calibration	0.88	45.21	0.3	0.87
	Validation	0.92	35.76	0.25	0.91
9	Calibration	0.9	42.31	0.28	0.89
	Validation	0.89	42.05	0.3	0.88
10	Calibration	0.92	37.27	0.25	0.91
	Validation	0.93	36.75	0.26	0.91



شکل ۶- مقایسه میزان تبخیر مشاهده‌ای و پیش‌بینی شده با مدل رگرسیون خطی در مرحله صحت‌سنجی

Figure 6- The comparison of observed and predicted evaporation with linear regression model in validation step

پیش‌بینی غیرخطی تبخیر به کمک مدل برنامه‌ریزی بیان ژن (GEP)

جهت پیش‌بینی تبخیر به کمک مدل GEP از نرم‌افزار GeneXproTools 5.0 استفاده گردید. برای

جدول ۷- خلاصه نتایج پیش‌بینی تبخیر ماهانه

Table 7- Summary results of monthly evaporate prediction

Model	Evaluation criteria			
	R2	RMSE (mm.d ⁻¹)	SE	NSE
ARIMA	0.97	27.22	0.19	1
SARIMA	0.98	23.59	0.17	0.95
GENE	0.95	26.17	0.18	0.95
Linear regression	0.93	36.75	0.26	0.91

نتیجه‌گیری

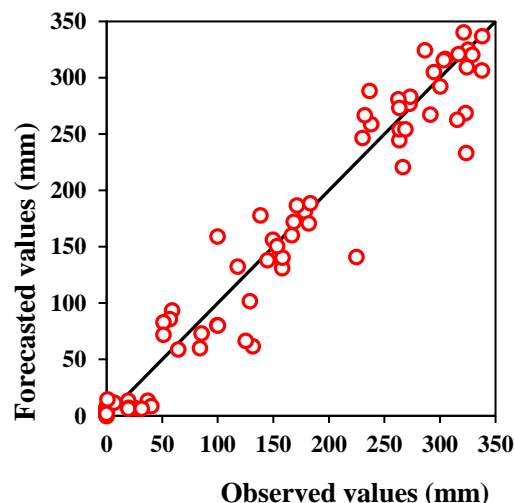
تبخیر از پدیده‌های مهم چرخه آب‌شناسی بوده و تعیین و پیش‌بینی آن در مدیریت و برنامه‌ریزی اصولی علوم آب ضروری می‌باشد. به همین دلیل به پیش‌بینی این پدیده در شهر همدان و در ایستگاه تبخیرسنجی سد اکباتان که بخش مهمی از آب شرب این شهر را تأمین می‌کند، پرداخته شد. در این مطالعه دقت و کارایی مدل‌های ARIMA، SARIMA، GEP و رگرسیون خطی جهت پیش‌بینی مقادیر تبخیر ماهانه مورد بررسی قرار گرفت. به‌منظور صحت‌سنجی نتایج حاصل از مدل‌های مذکور از شاخص‌های آماری R^2 ، SE، RMSE و NSE جهت تعیین میزان برازش میان مقادیر مشاهده‌ای و شبیه‌سازی شده، استفاده شد. در شبیه‌سازی مقادیر تبخیر با استفاده از مدل $ARIMA(0,1,1)$ ، مقدار این شاخص‌ها به ترتیب 0.97 ، 27.22 ، 0.19 و 1 به دست آمد. همچنین مدل منتخب $SARIMA(1,0,0)(2,1,1)$ به دلیل داشتن کمترین خطای استاندارد و همچنین کمترین مقدار آکائیک به عنوان الگوی برگزیده انتخاب گردید. مقادیر شاخص‌های آماری به ترتیب برابر 0.98 ، 23.59 ، 0.17 و 0.95 ، 0.95 و 0.95 محاسبه شد. به‌منظور بررسی اعتبار مدل GEP، مقدار SE برابر با 0.17 ، RMSE برابر با 23.59 میلی‌متر بر روز، R^2 برابر با 0.98 و NSE برابر با 0.95 به دست آمد. در نهایت شاخص‌های آماری R^2 ، SE، RMSE و NSE در مدل رگرسیون خطی به ترتیب 0.93 ، 36.75 ، 0.26 و 0.91 به دست آمد. با بررسی نتایج به دست آمده و دقت قابل قبول بین مقادیر شبیه‌سازی شده نسبت به شرایط واقعی، نشان داده شد که مدل SARIMA عملکرد دقیق‌تری در پیش‌بینی تبخیر ماهانه داشته و مدل‌های GEP، ARIMA و رگرسیون خطی به ترتیب در رتبه‌های دوم تا چهارم قرار

نتایج پیش‌بینی تبخیر در ایستگاه سد اکباتان توسط تمامی مدل‌های استفاده شده در این تحقیق، در جدول ۷ ارائه شده است. نتایج نشان می‌دهد که مدل SARIMA در پیش‌بینی تبخیر ماهانه نسبت به سایر مدل‌ها از دقت بالاتری برخوردار است.

Table 6- Evaporation prediction results in calibration and validation steps with GEP model

جدول ۶- نتایج پیش‌بینی تبخیر در مراحل واسنجی و صحت‌سنجی با مدل GEP

Input pattern number	stage	R ²	RMSE (mm.d ⁻¹)	SE	NSE
1	Calibration	0.83	53.45	0.36	0.82
	Validation	0.88	42.68	0.30	0.87
2	Calibration	0.91	39.67	0.27	1
	Validation	0.23	205.34	1.45	-1.91
3	Calibration	0.91	39.07	0.26	0.90
	Validation	0.95	27.36	0.19	0.95
4	Calibration	0.88	44.61	0.30	0.88
	Validation	0.91	37.41	0.26	0.90
5	Calibration	0.95	40.49	0.27	0.90
	Validation	0.94	31.03	0.22	0.93
6	Calibration	0.88	52.68	0.35	0.83
	Validation	0.21	221.40	1.56	-2.38
7	Calibration	0.90	40.71	0.27	0.90
	Validation	0.94	28.88	0.20	0.94
8	Calibration	0.90	40.58	0.27	0.90
	Validation	0.94	30.60	0.22	0.94
9	Calibration	0.93	35.24	0.24	0.92
	Validation	0.77	59.04	0.42	0.76
10	Calibration	0.93	34.06	0.23	0.93
	Validation	0.95	26.17	0.18	0.95



شکل ۷- مقایسه میزان تبخیر مشاهده‌ای و پیش‌بینی شده با مدل GEP در مرحله صحت‌سنجی

Figure 7- The comparison of observed and predicted evaporation with GEP model in validation step

- based on large-scale climate signals. *Hydrological Sciences Journal*, 61(6), 1001–1009.
- Elizabeth, MS., Keith, JB., Nick, A. 2010. *Hydrology in Practice*. 4th ed, 546 p. Amazon.Co.Uk.
- Farahani, N., Karami, H., Farzin, S., Ehteram, M., Kisi, O., El Shafie, A. 2019. A new method for flood routing utilizing four-parameter nonlinear Muskingum and shark algorithm. *water resources management*, 31, 1–15.
- Faruk, DO. 2010. A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence Journal*, 23 (4), 586–594.
- Ferreira, C. 2001a. Gene expression programming in problem solving. 6th Online World Conf. on Soft Computing in Industrial Applications.
- Ferreira, C. 2001b. Gene expression programming, A new adaptive algorithm for solving problems. *Complex Systems Journal*, 13(2), 87–129.
- Ferreira, C. 2006. Gene expression programming, Mathematical modeling by an artificial intelligence, Springer, Berlin, Heidelberg, New York, 478.
- Ghani, IMM., Ahmad, S. 2010. Stepwise multiple regression method to forecast fish landing. *Procedia-Social and Behavioral Sciences*, 8, 549–554.
- Goodarzi, M., Eslamian, S. 2018. Performance evaluation of linear and nonlinear models for the estimation of reference evapotranspiration. *International Journal of Hydrology Science and Technology*, 8(1), 1–15.
- Güven, A., Aytekin, A., Yuce, M., Aksoy, H. 2008. Genetic programming-based empirical model for daily reference evapotranspiration estimation. *Clean–Soil, Air, Water*. 36(10-11), 905–912. <https://doi.org/10.1002/clen.200800009>.
- He, W., Wan, S., Jiang, Y., Jin, H., Zhang, W., Wu, Q., He, T. 2013. Detecting abrupt change on basis of skewness, Numerical test and applications. *International Journal of Climatology*, 33(12), 2713–2727. <https://doi.org/10.1002/joc.3624>.
- Kisi, O., Ozkan, C. 2017. A New Approach for Modeling Sediment-Discharge Relationship, Local Weighted Linear Regression. *Water Resources Management*, 31(1), 1-23.
- Machekposhti, KH., Sedghi, H., Telvari, A., Babazadeh, H. 2018. Modeling climate variables of rivers basin using time series analysis (case study, Karkheh River basin at Iran). *Civil Engineering Journal*, 4, 478. <https://doi.org/10.28991/cej-030970>.
- Maier, HR., Dandy, GC. 1997. Determining inputs for neural network models of multivariate time series. دارند. همچنین مقایسه نتایج مدل‌های SARIMA و GEP نشان داد که دقت این دو مدل در پیش‌بینی تبخیر، اختلاف زیادی با یکدیگر ندارند. ولی با توجه به این که مدل GEP از سهولت بیشتری نسبت به مدل SARIMA برخوردار است و به پارامترهای کمتری نسبت به مدل SARIMA نیاز دارد، از این رو پیش‌بینی را به آسانی و در زمان کمتری انجام می‌دهد. بنابراین مدل GEP را نسبت به سایر مدل‌ها می‌توان ارجح دانست.

منابع

- Ahmed, AN., Othman, FB., Afan, HA., Ibrahim, RK., Fai, CM., Hossain, MS., Ehteram, M., Elshafie, A. 2019. Machine learning methods for better water quality prediction. *Journal of Hydrology*. 578, 124084.
- Akaike, HA. 1979. Bayesian Extension of the Minimum AIC Procedure of Autoregressive Model Fitting. *Biometrika*. 66, 237-242. Doi:10.1093/biomet/66.2.237.
- Aytek, A., Alp, M. 2008. An application of artificial intelligence for rainfall-runoff modeling. *Journal of Earth System Science*, 117(2), 145-155.
- Azadi, S., Nozari, H., Godarzi, E. 2020. Predicting Sediment Load Using Stochastic Model and Rating Curves in a Hydrological Station. *Journal of Hydrologic Engineering*. DOI:10.1061/(ASCE)HE.1943-5584.0001967.
- Babamiri, A., Nozari, H., Marofi, S. 2017. Estimation of potential evapotranspiration based on random time series models (Case study of Tabriz station). *Journal of Watershed Management*, 8(15), 146-137. (In Farsi)
- Bari, SH., Rahman, MT., Hussain, MM., Ray, S. 2015. Forecasting Monthly Precipitation in Sylhet City Using ARIMA Model. *Civil and Environmental Research*, 7(1), 69-77.
- Box, GEP., Cox, DR. 1964. An analysis of transformations. *Journal of the royal statistical society series b-methodological*, 26(2), 211–252. <https://doi.org/10.1111/j.2517-6161.1964.tb0553.x>.
- Box, GEP., Jenkins, GM., Reinsel, GC. 1994. *Time Series Analysis; Forecasting and Control*. 3rd Edition, Prentice Hall, Englewood Cliff, New Jersey.
- Choubini, B., Khalighi-Sigaroodi, S., Malekian, A., Kisi, O. 2016. Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation

- programming approach in the Basque Country (Northern Spain). *Journal of Hydrology*, 414–415, 302–316.
- Shirvani, A., Honar, T. 2011. Application of time series models to predict evapotranspiration at Bajgah station. *Iranian Water Research Journal*, 0,142-135.(In Farsi)
- Terzi, O. 2013. Daily pan evaporation estimation using gene expression programming and adaptive neural-based fuzzy inference system. *Neural Computing and Applications*, 23(3–4), 1035–1044. <https://doi.org/10.1007/s00521-012-1027-x>.
- Wang, HR., Wang, C., Lin, X., Kang, J. 2014. An improved ARIMA model for precipitation simulations, *Nonlin. Processes Geophysics*, 21, 1159–1168. doi,10.5194/npg-21-1159-2014.
- Wang, H., Shangguan, L., Wu, J., Guan, R. 2013. Multiple linear regression modeling for compositional data. *Neurocomputing*, 122, 490-500.
- Wang, L., Kisi, O., Zounemat-Kermani, M., Li, H. 2017. Pan evaporation modeling using six different heuristic computing methods in different climates of China. *Journal of Hydrology*, 544, 407-427.
- Wang, WC., Chau, KW., Cheng, CT., Qiu, L. 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*, 374(3–4), 294–306.
- Zhang, Q., Wang, BD., He, B., Peng, Y., Ren, ML. 2011. Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting. *Water Resources Management*, 25(11), 2683–2703.
- Computer-Aided Civil and Infrastructure Engineering, 12(5), 353–368.
- Maroufpoor, E., Sanikhani, H., Emamgholizadeh, S., Kişi, O. 2018. Estimation of wind drift and evaporation losses from sprinkler irrigation systems by different data-driven methods. *Journal of Irrigation and Drainage*, 67(2), 222–232.
- Mattar, MA., Alazba A. A. 2018. GEP and MLR approaches for the prediction of reference evapotranspiration. *Neural Computing and Applications*, doi:10.1007/s00521-018-3410-8.
- Nozari, H., Tavakoli, F. 2020. Forecasting hydrologic parameters using linear and nonlinear stochastic models. *Journal of Water and Climate Change*, (11)4, 1284–1301.
- Piri, J., Mohammadi, K., Shamshirband, S., Akib, S. 2019. Expression of concern, assessing the suitability of hybridizing the Cuckoo optimization algorithm with ANN and ANFIS techniques to predict daily evaporation. *Environmental Earth Sciences*, 78(1), 8-16.
- Rencher, AC., Schaalje, GB. 2008. *Linear models in statistics*. John Wiley & Sons.
- Shiri, J., Kisi, O. 2011. Application of artificial intelligence to estimate daily pan evaporation using available and estimated climatic data in the Khuzestan Province (South Western Iran). *Journal of Irrigation and Drainage Engineering*, 137 (7), 412–425. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000315](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000315).
- Shiri, J., Kisi, O., Landaras, G., Lopez, JJ., Nazemi, AH., Stuyt, L. 2012. Daily reference evapotranspiration modeling by using genetic



Predicting monthly evaporation using linear and nonlinear time series models (Case study: Ekbatan Dam station)

H. Nozari^{1*}, S. Azadi², N. Sedghnejad³, S. Pouyanfar³

Received: 05/06/2022

Accepted: 07/03/2023

Abstract

Prediction of evaporation as a key component of the hydrological cycle is one of the most important issues in water resources management and meteorology studies. In this study, the performance of ARIMA, SARIMA, gene expression programming, multiple linear regression, Monte Carlo and Thomas Fairing models in prediction of monthly evaporation values of Ekbatan Dam station, west of Iran in a 47 years period (1971-2017) were evaluated. For calibration of these models, 40 years data (1971-2010), and for validation, data from 2011-2017 (7-year) were used. The statistical metrics of the correlation coefficient, root mean square error, standard error, the Akaike information criterion, and NSE were selected for evaluation and comparison of models. The results showed that the SARIMA model has more accurate performance in predicting monthly evaporation. The GEP model, ARIMA, and MLR are ranked second to fourth. However, since the GEP model is easier to use than the SARIMA model and requires fewer variables than the SARIMA model, it shows promise to generate faster results, therefore, the GEP models can be the preferred option compared to others.

Keywords: Ekbatan Dam, Evaporation, Simulation, Time Series



¹ Associate Professor, Department of Water Science Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

(*Corresponding Author Email Address: h.nozari@basu.ac.ir)

² Ph. D. Graduate in Water Resources Engineering, Department of Water Science Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

³ M. Sc. Student in Water Resources Engineering, Department of Water Science Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

نحوه ارجاع مقاله:

نوذری، ح، آزادی، س، صدق نژاد، ن، پویان فر، س. ۱۴۰۲. پیش‌بینی تبخیر ماهانه به کمک مدل‌های خطی و غیرخطی سری زمانی (مطالعه موردی):

ایستگاه تبخیرسنجی سد اکباتان). نشریه هواشناسی کشاورزی، ۱۱(۱): ۳۶-۴۷. DOI: 10.22125/agmj.2023.345670.1136

Nozari, H., Azadi, S., Sedghnejad, N., Pouyanfar, S. 2023. Predicting monthly evaporation using linear and nonlinear time series models (Case study: Ekbatan Dam station). Journal of Agricultural Meteorology, 11(1): 36-47. DOI: 10.22125/agmj.2023.345670.1136