



بهینه‌سازی با الگوریتم‌های فراکاوشی جهت تخمین ضریب پخش طولی رودخانه‌ها

سیما اوحدی^۱، سید آرمان هاشمی منفرد^{۲*}، محمد گیوه‌چی^۳

تاریخ ارسال: ۱۳۹۸/۴/۴

تاریخ پذیرش: ۱۳۹۸/۱۰/۱۴

مقاله پژوهشی برگرفته از رساله دکتری

چکیده

در سراسر جهان، رودخانه‌ها از اصلی‌ترین شریان‌های تأمین نیازهای شرب، صنعت و کشاورزی به شمار می‌آیند. از این رو، در سال‌های اخیر شناخت و مطالعه‌ی فرایندهای انتقال آلودگی در رودخانه‌ها اهمیت خاصی پیدا کرده است. ضریب پخش طولی مهم‌ترین پارامتر تأثیرگذار در معادله انتقال است که تاکنون روش‌های مختلفی جهت تخمین و محاسبه آن ارائه شده است. از جمله رویکردهای مطرح جهت تخمین ضریب پخش طولی رودخانه‌ها استفاده از روش‌های بهینه‌سازی فراکاوشی می‌باشد. در این تحقیق توانایی ۵ الگوریتم فراکاوشی شامل الگوریتم ژنتیک (GA)، کلونی زنبورهای عسل مصنوعی (ABC)، ازدحام ذرات (PSO)، گرگ خاکستری (GWO) و رقابت استعماری (ICA) در تخمین پارامتر مذکور مورد ارزیابی و مقایسه قرار گرفته است. رابطه محاسبه ضریب پخش طولی با استفاده از هر یک از الگوریتم‌ها ارائه و نتایج با داده‌های مشاهداتی مقایسه شده است. طبق نتایج، الگوریتم بهینه‌سازی ژنتیک در شرایط تکرار یکسان، از دقت و سرعت بالاتری نسبت به سایر الگوریتم‌های مورد بررسی برخوردار است. سپس از پارامترهای آماری شامل خطای مجذور میانگین مربعات، میانگین خطای مطلق و ضریب همبستگی جهت بررسی دقت الگوریتم‌ها استفاده شده است. که این مقادیر برای الگوریتم ژنتیک به ترتیب برابر با ۰/۱۰۵۲، ۰/۴۳۵۴ و ۰/۸۲۸۸ بدست آمد. در نتیجه می‌توان با دقت مناسبی، معادله حاصل از این الگوریتم را در محاسبه ضریب پخش طولی کارآمد معرفی نمود.

واژه‌های کلیدی: انتقال آلودگی، بهینه‌سازی، روش‌های فراکاوشی، ضریب پخش طولی.

^۱ دانشجوی دکترای مهندسی عمران- مدیریت منابع آب، دانشکده مهندسی شهید نیکبخت، دانشگاه سیستان و بلوچستان، زاهدان، ایران.
Ohadi_sima@pgs.usb.ac.ir

^۲ دانشیار گروه مهندسی عمران، دانشکده مهندسی شهید نیکبخت، دانشگاه سیستان و بلوچستان، زاهدان، ایران. Hashemi@eng.usb.ac.ir
(نویسنده مسئول)

^۳ استادیار گروه مهندسی عمران، دانشکده مهندسی شهید نیکبخت، دانشگاه سیستان و بلوچستان، زاهدان، ایران.
M.Givehchi@eng.usb.ac.ir



مقدمه

وضعیت کنونی منابع آب جهان و مشکل روزافزون آلودگی آب‌های سطحی و زیرزمینی، مسئله مدیریت کیفی منابع آب را در صدر برنامه‌های مدیران و پژوهشگران این حیطه قرار داده است. این مهم سبب شده که بررسی و کنترل انتقال آلودگی در آبهای سطحی، جایگاه ویژه و تأثیرگذاری در مباحث مدیریت منابع آب پیدا کند. گام اول در این راستا بررسی معادله انتقال-پخش به عنوان معادله حاکم بر رفتار آلاینده در جریان‌های سطحی می‌باشد. طبق این معادله، پارامتر ضریب پخش طولی به عنوان مهمترین پارامتر مطرح بوده و دقت تخمین آن، دارای اهمیت بالایی در مهندسی هیدرولیک می‌باشد (Noori et al., 2017). از این رو تاکنون، روش‌های مختلف آزمایشگاهی، تئوری و تجربی جهت تخمین دقیق پارامتر مذکور مورد استفاده و ارزیابی قرار گرفته است (Alizadeh et al., 2017a). زمان بر بودن و هزینه بالای مطالعات ردیابی در روش‌های آزمایشگاهی سبب توسعه معادلات تجربی بر پایه اندازه‌گیری‌های میدانی و آزمایشگاهی شده است (Alizadeh et al., 2017b).

ضریب پخش طولی به عنوان اندازه‌ای از فرآیند پخش طولی توسط معادله انتقال پخش برای اولین بار توسط Taylor ارائه گردید (Taylor, 1954). در ادامه Elder با گسترش روش Taylor، با فرض لگاریتمی بودن پروفیل سرعت در جهت عمودی، معادله‌ای برای تخمین ضریب پخش طولی برای کانال‌های باز عریض ارائه داد (Elder, 1959). اگرچه رابطه Elder موفقیت بزرگی در پیش‌بینی ضریب پخش طولی محسوب شد، با این وجود عدم در نظر گرفتن تغییرات پروفیل عرضی سرعت در رابطه، اختلاف زیاد بین مقادیر مشاهداتی و پیش‌بینی شده را نشان داد (Alizadeh et al., 2017b; Sahay and Dutta, 2009; Fisher (1968)). اظهار داشت برای پیش‌بینی ضریب پخش طولی در کانال عریض، پروفیل عرضی سرعت اهمیت بیشتری نسبت به پروفیل عمودی سرعت داراست. بر این اساس وی انتگرال سه گانه‌ای را برای تخمین ضریب پخش

طولی در کانال‌های عریض ارائه کرد (Fisher, 1968). Kashefipour and Falconer (2002) یک رابطه جدید برای پیش‌بینی ضریب پخش طولی در کانال‌های طبیعی با استفاده از آنالیز ابعادی و رگرسیون پارامترهای هیدرولیکی جریان ارائه دادند. آنها در تحقیق خود از ۸۱ سری داده بدست آمده از ۳۰ رودخانه در ایالات متحده آمریکا استفاده کردند. مقایسه پارامترهای آماری، حاکی از دقت بیشتر روش مذکور در مقایسه با سایر روابط موجود جهت محاسبه ضریب پخش طولی بود. آنها همچنین با ترکیب خطی معادله پیشنهادی خود با معادله ارائه شده توسط Soe and Cheong (1998)، معادله دیگری جهت تخمین ضریب پخش طولی ارائه دادند که آنالیز معادله مذکور با روش‌های آماری، کارآیی آن را به اثبات رساند (Kashefipour and Falconer, 2002).

گیوه‌چی و همکاران (۱۳۸۸)، تأثیر سرعت متوسط عمقی در تخمین ضریب پخش طولی را مورد مطالعه قرار دادند. آنها در تحقیق خود به تخمین توزیع سرعت متوسط عمقی با رابطه ارائه شده توسط مغربی (Maghrebi and Ball, 2006) پرداختند. سپس نتایج با داده‌های میدانی مقایسه و دقت این روش در تخمین ضریب پخش آلودگی اثبات گردید (گیوه‌چی و همکاران، ۱۳۸۸).

در ادامه تلاش‌های صورت گرفته در راستای دستیابی به روشی دقیق‌تر جهت محاسبه پارامتر ضریب پخشیدگی، سرعت بالای روشهای محاسبات نرم سبب شد در دهه‌های اخیر استفاده از الگوریتم‌های بهینه‌سازی مورد توجه بیشتری قرار گیرد. Sahay and Dutta (2009)، از الگوریتم ژنتیک و Etemad-Shahidi and Taghipour (2012)، از مدل درختی M5 که یک روش محاسبات نرم می‌باشد برای تخمین پارامتر ضریب پخش طولی در رودخانه‌های طبیعی استفاده کردند (Sahay and Dutta, 2009; Etemad-Shahidi and Taghipour, 2012).



از این رو در این پژوهش عملکرد پنج الگوریتم فراابتکاری ژنتیک، زنبورهای عسل مصنوعی، ازدحام ذرات، گرگ خاکستری و رقابت استعماری در تخمین ضریب پخش طولی به عنوان مهمترین پارامتر معادله انتقال-پخش مورد ارزیابی قرار گرفت. در مرحله اول، معادله مربوط به ضریب پخش طولی تحت تأثیر پارامترهای هیدرولیکی جریان شامل عمق، عرض، سرعت و سرعت برشی، توسط هر یک از الگوریتم‌های مذکور با هدف به حداقل رساندن خطای مجذور میانگین مربعات محاسبه شد. سپس داده‌های مشاهداتی با نتایج حاصل از الگوریتم‌های مذکور مقایسه و با استفاده از شاخص‌های آماری، قدرتمندترین الگوریتم در حل مسئله مورد بررسی مشخص شد.

مواد و روش‌ها

تحلیل رابطه ضریب پخش

بر اساس مطالعات پیشین، خواص هندسی جریان (عرض مسیر جریان (W) ، ضریب شکل بستر (S_f) ، شیب کف کانال (S) ، ضریب زبری کانال (k_s) و عمق جریان (H))، خواص هیدرولیک جریان (سرعت متوسط جریان (U) و سرعت برشی جریان (U_*)) و خواص سیال جریان (چگالی سیال (ρ) و لزجت سینماتیکی جریان (μ)) پارامترهای مهم و تأثیرگذار بر روی ضریب پخشیدگی طولی در رودخانه‌ها می‌باشند که با استفاده از آنالیز ابعادی می‌توان به صورت رابطه (۱) بیان نمود (Chow, 1959):

$$\frac{K}{U_* H} = f' \left(\rho \frac{HU}{\mu}, \frac{W}{H}, \frac{U}{U_*}, \frac{k_s}{H}, S_f, S \right) \quad (1)$$

که $\frac{K}{U_* H}$ پارامتر بدون بعد ضریب پخشیدگی، $\rho \frac{HU}{\mu}$ عدد رینولدز، $\frac{W}{H}$ نسبت عرض جریان به عمق و $\frac{U}{U_*}$ مربوط به ترم مقاومت جریان است. از آوردن تأثیر مستقیم پارامترهای مربوط به شیب آبراهه، شکل بستر

(Sattar and Gharabaghi, 2015) از مدل بیان ژن^۱ برای توسعه رابطه‌ای تجربی بین ضریب پخش طولی و متغیرهای کنترل شامل عدد فرود، نسبت ابعادی و زبری بستر استفاده کردند. آنها برای این منظور از ۱۵۰ سری داده مربوط به جریان‌های طبیعی شامل پارامترهای هندسی و هیدرولیکی جریان بهره گرفتند. مقایسه عدم قطعیت مدل پیشنهادی با سایر روش‌های موجود، دقت بالاتر مدل GEP را نشان داد (Sattar and Gharabaghi, 2015) and (Alizadeh et al., 2017). از روش شبکه بیزین جهت ارائه مدلی برای پیش‌بینی ضریب پخش طولی در رودخانه‌های طبیعی استفاده کردند. آنها جهت افزایش دقت مدل پیشنهادی از کوپل آنالیز خوشه‌ای جهت پیش‌پردازش داده‌ها و شبکه بیزین بهره گرفتند. نتایج بدست‌آمده با نتایج حاصل از معادلات تجربی از لحاظ پارامترهای آماری مقایسه و دقت روش پیشنهادی نشان داده شد (Alizadeh et al., 2017a) and (Rezaie Balf et al., 2018). از مدل رگرسیون چند جمله‌ای تکاملی^۲ برای پیش‌بینی K_x به عنوان تابعی از عمق جریان، عرض کانال، سرعت متوسط و سرعت برشی استفاده کردند (Rezaie Balf et al., 2018).

تکنیک‌های بهینه‌سازی فراابتکاری زیادی در دو دهه اخیر معرفی شده‌اند. فراابتکارها به علت سادگی، انعطاف‌پذیری، نیاز نداشتن به مشتق‌گیری و فرار از بهینه محلی بسیار مورد توجه قرار گرفته‌اند. اما طبق نظریه NFL^۳ هیچ یک از روش‌های فراابتکاری برای حل تمام مسائل بهینه‌سازی مناسب نمی‌باشد (Wolpert and Macready, 1997). به عبارتی، یک فراابتکاری بخصوص ممکن است نتایج قابل‌قبولی را برای حل یک سری از مسائل خاص نشان دهد اما همین الگوریتم ممکن است کارایی ضعیفی را برای تعداد دیگری از مسائل داشته باشد (Mirjalili et al., 2014).

³ No Free Lunch

¹ GEP: Gene Expression Programming

² EPR: Evolutionary Polynomial Regression



است. تنها قید این مسئله بهینه‌سازی نیز رعایت محدوده پایین (L_b) و بالای (U_b) متغیرهای تصمیم می‌باشد.

الگوریتم‌های فراکاوشی

یکی از روش‌های انعطاف‌پذیر جهت حل مسائل پیچیده بهینه‌سازی مبتنی بر هوش جمعی می‌باشد که با الهام از سیستم‌های بیولوژیکی و فیزیکی موجود در طبیعت به وجود آمده‌اند. مکانیزم این روش‌ها به این صورت است که تعداد زیادی ذره در فضای مسأله پخش شده و به طور همزمان به دنبال جواب بهینه می‌گردند. این روش‌ها را الهام گرفته از طبیعت مبتنی بر جمعیت یا الگوریتم‌های فراکاوشی می‌نامند. الگوریتم‌های فراکاوشی نیز با در نظر گرفتن قیدها و محدودیت‌های مختلف، به مجموعه‌ای از جواب‌های غیرغالب دست می‌یابند که بهترین حالت مجموعه جواب بوده و یا می‌توان امیدوار بود که با مجموعه جواب اصلی فاصله زیادی نداشته است (Tukey, 1965; Montalvo et al., 2008). در سال‌های اخیر استفاده از روش‌های فراکاوشی در حل مسائل مهندسی آب به‌طور گسترده افزایش یافته است. در این میان برخی از آن‌ها بیش از سایر الگوریتم‌ها مورد توجه قرار گرفته‌اند. بر این اساس در این مطالعه سعی شد عملکرد الگوریتم‌های پر کاربرد در زمینه مهندسی آب نظیر GA (پیاده‌کوهسار و همکاران، ۱۳۹۸؛ حجتی و همکاران، ۱۳۹۵)، ABC (سامی کشکولی و همکاران، ۱۳۹۵)، ICA (جعفری‌اصل و همکاران، ۱۳۹۴؛ هاشمی منفرد و همکاران، ۱۳۹۶)، PSO (پیاده‌کوهسار و همکاران، ۱۳۹۸؛ Alizadeh et al., 2017a) و GWO (Yu and Lu, 2018؛ امامی و امامی، ۱۳۹۸؛ Dehghani et al., 2019) که تاکنون در حل مسائل مهندسی آب عملکرد خود را اثبات نموده‌اند، برای پیش‌بینی ضریب پخش طولی مورد ارزیابی قرار گیرند.

و مقدار سینوسی بودن مسیر رودخانه که به آسانی قابل اندازه‌گیری نیستند و تأثیر آنها در ترم مقاومت دیده می‌شود، صرف‌نظر می‌گردد. همچنین به دلیل اینکه جریان در رودخانه‌ها معمولاً آشفته است می‌توان تأثیر عدد رینولدز را نادیده گرفت و فرم نهایی رابطه (۱) را به صورت رابطه (۲) نوشت:

$$\frac{K}{U_* H} = f' \left(\frac{W}{H}, \frac{U}{U_*} \right) \quad (2)$$

در نهایت می‌توان ضریب پخشیدگی طولی در رودخانه‌ها را با در نظر گرفتن f' به عنوان یک تابع توانی به صورت رابطه (۳) بیان کرد (Sahha and Alizadeh et al., 2017a; Seo and Dutta, 2009; cheong, 1998; Deng et al., 2001):

$$\frac{K}{U_* H} = a \left(\frac{W}{H} \right)^b \left(\frac{U}{U_*} \right)^c \quad (3)$$

در این رابطه، ضریب پخش طولی (K) به چهار پارامتر U : سرعت میانگین جریان، H : عمق جریان، W : عرض کف جریان و U_* : سرعت برشی جریان وابسته می‌باشد. در رابطه فوق ضرایب a ، b و c ثابت‌های معادله (۳) می‌باشند.

مدل ریاضی مسئله بهینه‌سازی ضریب پخش طولی به صورت رابطه زیر بیان می‌شود:

$$\text{Find } X = \{x_1, x_2, x_3\} \quad (4)$$

$$\text{to minimize } RMSE = \sqrt{\frac{\sum_{i=1}^n (K_i(\text{observed}) - K_i(\text{predicted}))^2}{n}} \quad (5)$$

Subject to:

$$L_b \leq X \leq U_b$$

که بردار X متغیرهای تصمیم مسئله بهینه‌سازی شامل ضرایب a ، b و c می‌باشد و تابع هدف در این مطالعه به صورت کمینه کردن اختلاف مقادیر پیش‌بینی شده با مقادیر مشاهداتی در نظر گرفته شده



درصدی از ژنها عمل جهش صورت گیرد بستگی به انتخاب مقدار نرخ عملگر جهش دارد.

د) **ارزیابی به وسیله تابع هدف:** هر کروموزوم باید مورد ارزیابی قرار گیرد. مناسبترین کروموزومها شانس بیشتری برای زنده ماندن در طول دوره‌های دیگر و دوباره تولید شدن را دارا بوده و ضعیف‌ترها محکوم به نابودی هستند. جهت ارزیابی باید تابع هدف مشخص شود.

الگوریتم کلونی زنبورهای عسل مصنوعی (ABC)^۲

الگوریتم ABC برگرفته از رفتار زنبورهای عسل در یافتن گلزارهای مطلوب جهت جمع آوری شهد مناسب است. در این الگوریتم، اجتماع و کلونی زنبورها شامل زنبورهای استخدام شده، زنبورهای جستجوگر و زنبورهای دیده‌بان بوده و هر منبع غذایی نشان دهنده یک جواب ممکن برای مسئله بهینه‌سازی است و مقدار شهد موجود در هر منبع بیانگر کیفیت (برازندگی) آن منبع می‌باشد (Karaboga and Basturk, 2007; Karaboga and Ozturk, 2011). مراحل این الگوریتم به شرح زیر می‌باشد:

در مرحله اول، یک جمعیت اولیه از جوابها (SN) که برابر همان موقعیت منبع غذایی است، تولید می‌شود. در واقع SN بیانگر تعداد زنبورهای استخدام شده یا جستجوگر می‌باشد. هر پاسخ X_i ($i=1, \dots, SN$) یک بردار D بعدی است که D تعداد پارامترهای بهینه‌سازی است. زنبورهای جستجوگر یک منبع غذایی را بر اساس احتمال، انتخاب می‌کنند. این انتخاب، متأثر از کیفیت آن منبع غذایی است. احتمال انتخاب هر منبع توسط رابطه (۶) محاسبه می‌گردد:

$$P_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_n} \quad (6)$$

الگوریتم ژنتیک (GA)^۱

الگوریتم بهینه‌سازی ژنتیک، تکنیک بهینه‌سازی و جستجوی تصادفی است که با الهام از طبیعت و فرآیند تکامل طبیعی برای حل مسائل بهینه‌سازی توسعه داده شده است. در این تکنیک ایده‌هایی از قبیل تکامل موازی اعضای یک جمعیت، به ارث رسیدن اطلاعات ژنتیکی از والدین از طریق زاد و ولد، تکثیر ترجیحی اعضای اصلحتر جمعیت و اعمال جهش‌های تصادفی، مورد استفاده قرار می‌گیرد. تحقیقات Halland در سال ۱۹۷۵ منجر به توسعه الگوریتم ژنتیک شد که در واقع نوع خاصی از انواع رویکردهای تکاملی به شمار می‌رود. در این الگوریتم ابتدا بطور تصادفی تعدادی از افراد به عنوان جمعیت اولیه تولید می‌شوند و میزان شایستگی هر یک از آنها بر اساس تابع ارزیابی می‌شود. سپس تعدادی بر اساس میزان شایستگی به عنوان والد انتخاب شده و فرزندان جدیدی بوجود می‌آیند. فرزندان جدید در جمعیت کپی شده و جمعیت جدید تشکیل می‌شود. این عمل تا برآورده شدن شرط خاتمه، ادامه می‌یابد (Melanie, 2008). الگوریتم ژنتیک شامل مراحل زیر خواهد بود:

الف) ساختن جمعیت اولیه: برای شروع باید از متغیرهای موجود (ژنها)، کروموزوم تشکیل شده و الگوریتم ژنتیک یک جمعیت اولیه ناهمگن از کروموزومها را ایجاد نماید.

ب) تقاطع یا جفتگیری: گام بعدی ایجاد دومین نسل از جمعیت اولیه است. تعدادی از کروموزومهای جمعیت (والدین) دو به دو با یکدیگر تقاطع پیدا کرده و هر جفت دو فرزند تشکیل می‌دهند.

ج) جهش: به کمک این عملگر می‌توان در یک بخش از یک کروموزوم تغییراتی به صورت تصادفی ایجاد کرد به این امید که کروموزوم بهتری تولید شود. اینکه در چه

² Artificial Bee Colony

¹ Genetic Algorithm



(1995) and Eberhart مطرح شد و به سرعت به عنوان یک ابزار بهینه‌سازی قدرتمند مورد توجه قرار گرفت (Kennedy and Eberhart, 1995a).

در الگوریتم PSO، می‌توان هر پاسخ را به صورت پرنده‌ای در دسته‌ای از پرندگان تصور نمود که به آن جزء می‌گویند. در چنین ساختاری پرندگان علاوه بر هوش فردی، دارای نوعی رفتار گروهی نیز می‌باشند که آنان را به سمت هدف راهنمایی می‌کند. ابتدا، فرآیند از یک دسته از اجزا، که هر کدام یک پاسخ^۲ برای مسئله می‌باشند و به صورت تصادفی انتخاب شده‌اند، آغاز می‌شود. در ادامه با حرکت دسته‌ای از اجزا، فضای جستجو در جهت یافتن بهترین پاسخ جستجو می‌شود. در این الگوریتم هر جزء دارای یک سرعت انطباقی است که بیانگر بردار حرکت جزء در فضای جستجو می‌باشد. علاوه بر آن دارای حافظه نیز می‌باشد، یعنی بهترین موقعیتی را که تا آن مرحله به آن رسیده است را به خاطر می‌سپارد. ارزش ذرات با توجه به تابع معیار مربوط به مسأله بهینه‌سازی سنجیده می‌شود. برای ذره نام، بهترین ارزش و موقعیت مربوط به آن به ترتیب در متغیرهای $pbest_i$ و $Pi=(p_{i1}, p_{i2}, \dots, p_{iD})$ ذخیره می‌شود. همچنین بهترین موقعیت ذرات در تمام تکرارها و موقعیت مربوط به آن به ترتیب در متغیرهای $gbest_i$ و $Pg=(p_{g1}, p_{g2}, \dots, p_{gD})$ ذخیره می‌شود. D بعد فضای جستجو را نشان می‌دهد. مجموعه n ذره در یک اجتماع، جمعیت نام دارد و به صورت $pop=(p_1, p_2, \dots, p_n)$ بیان می‌شود (Kennedy and Eberhart, 1995a).

اساس روش PSO اینگونه است که ذرات در هر تکرار به سمت $pbest$ و $gbest$ خود شتاب می‌گیرند و در نتیجه در هر تکرار با اضافه شدن بردار جابجایی در هر ذره، موقعیت جدیدی در فضای جستجو برای هر ذره پیدا شده و در پایان موقعیت مربوط به $gbest_i$ به عنوان جواب بهینه مسئله در نظر گرفته می‌شود. سرعت و

که مقدار fit_i برازندگی جواب i می‌باشد که نسبتی است از حجم شهد آن منبع غذا در مکان i و SN تعداد منابع غذا می‌باشد که برابر با تعداد زنبورهای استخدام شده و زنبورهای جستجوگر است (Karaboga and Basturk, 2007; Karaboga and Ozturk, 2011). ABC به منظور ایجاد یک موقعیت غذای کاندید (v_{ij}) از مکان غذای قدیمی (x_{ij})، از رابطه (۷) بهره می‌برد:

$$V_{ij} = X_{ij} + \phi_{ij}(X_{ij} - X_{kj}) \quad (7)$$

به طور تصادفی $j \in \{1, 2, \dots, D\}$ و $K \in \{1, 2, \dots, D\}$ انتخاب می‌گردند. اگرچه k به طور تصادفی تعیین می‌گردد ولی بایستی با i متفاوت باشد. ϕ_{ij} عددی تصادفی بین $[-1, +1]$ می‌باشد. فرض کنید منبع ترک شده X_i باشد و $j \in \{1, 2, \dots, D\}$ ، در آن صورت زنبورهای کاشف منبع غذایی جدیدی را که جایگزین X_i می‌شود را می‌یابند. این عمل می‌تواند طبق رابطه (۸) تعریف گردد:

$$X_i^j = X_{min}^j + \text{rand}[0, 1](X_{max}^j - X_{min}^j) \quad (8)$$

پس از اینکه هر موقعیت منبع کاندید غذا (V_{ij}) انتخاب شد توسط زنبور مصنوعی مورد ارزیابی قرار گرفته و با موقعیت قبلی مقایسه می‌شود. اگر منبع غذایی جدید دارای شهد برابر یا بهتر از منبع‌های قدیمی باشد آن را با یکی از موقعیت‌های قدیمی در حافظه جایگزین می‌کند (Karaboga and Basturk, 2007; Karaboga and Ozturk, 2011).

الگوریتم بهینه‌سازی ازدحام ذرات (PSO)^۱

الگوریتم بهینه‌سازی مبتنی بر رفتار اجزای جمعی (PSO)، اولین بار در سال ۱۹۹۵ توسط Kennedy

¹ Particle Swarm Optimization

^۲ مجموعه مقادیری که به مجموعه متغیرهای تصمیم نسبت داده شده است.



$$\vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{X}_p(t)| \quad (11)$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \quad (12)$$

که t بیانگر تکرار، X_p بردار موقعیت طعمه، X بردار موقعیت گرگ خاکستری و A ، C بردارهای ضرایب که به صورت روابط زیر محاسبه می‌شوند:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (13)$$

$$\vec{C} = 2 \cdot \vec{r}_2 \quad (14)$$

که مقادیر a بطور خطی و به ازای تکرارها از مقدار دو تا صفر کاهش می‌یابند. r_1 و r_2 مقادیر تصادفی در بازه صفر و یک می‌باشند.

در الگوریتم گرگ خاکستری ابتدا سه پاسخ اول از میان پاسخ‌های مطلوب برگزیده شده و سایر عامل‌ها ملزم به بهبود موقعیت خود نسبت به این سه پاسخ می‌شوند تا موقعیت خود را نسبت به این سه پاسخ با استفاده از روابط زیر بهبود بخشند.

$$\begin{aligned} \vec{D}_\alpha &= |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \vec{D}_\beta \\ &= |\vec{C}_2 \cdot \vec{X}_\beta \\ &\quad - \vec{X}|, \vec{D}_\delta \\ &= |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}| \end{aligned} \quad (15)$$

$$\begin{aligned} \vec{X}_1 &= \vec{X}_\alpha - \vec{A}_1 (\vec{D}_\alpha), \vec{X}_2 = \vec{X}_\beta - \\ &\vec{A}_2 (\vec{D}_\beta), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 (\vec{D}_\delta) \quad , \quad (16) \\ \vec{X}(t+1) &= \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \end{aligned}$$

جست و جو در فضای مسأله توسط بردار \vec{C} با مقادیر تصادفی بین ۱ و -۱ که عاملها را ملزم به فاصله گرفتن از طعمه میکند، تضمین می‌گردد. بردار \vec{C} که نشان دهنده اکتشاف است مقادیر وزنی تصادفی در محدوده ۰ و ۲ برای طعمه انتخاب می‌کند. در واقع در الگوریتم گرگ خاکستری پروسه بهینه‌سازی با تولید مجموعه‌ای از جواب‌های تصادفی اولیه آغاز شده و در حین پروسه

موقعیت جدید ذره‌ها در هر تکرار توسط روابط زیر به‌روز می‌شوند (Kennedy and Eberhart, 1995b).

$$v_{id}^k = C_1 \times rand_1 \times (P_{id} - x_{id}^{k-1}) + C_2 \times rand_2 \times (P_{gd} - x_{id}^{k-1}) + w \times v_{id}^{k-1} \quad (9)$$

$$X_{id}^k = X_{id}^{k-1} + v_{id}^k \quad (10)$$

که در آن c_1 و c_2 ضرایب‌های تصمیم‌گیری هستند که با توجه به اولویت بیشتر و یا اینکه بیشتر به سمت تجربیات خود حرکت کنیم یا به سمت تجربیات شخص الگو، مقدارشان تعیین می‌شود. معمولاً در پیاده‌سازی‌ها c_1+c_2 را کمتر یا مساوی ۴ در نظر می‌گیرند. $rand_1$ و $rand_2$ هم اعداد تصادفی هستند که از توزیع یکنواخت بین صفر و یک به دست می‌آیند. $w \times v_{id}^{k-1}$ به ضریب اینرسی معروف است و از آن چشم‌پوشی می‌کنند. قابل ذکر است که مقدار w در بازه ۰/۹ تا ۰/۴ به صورت خطی کاهش می‌یابد.

الگوریتم بهینه‌سازی گرگ خاکستری (GWO)

الگوریتم بهینه‌سازی گرگ خاکستری با الهام گیری از زندگی و مکانیسم شکار گرگ‌های خاکستری در طبیعت توسط (Mirjalili et al., 2014) معرفی شد. چهار نوع گرگ خاکستری نظیر الف، بتا، دلتا و امگا به منظور شبیه‌سازی سلسله مراتب رهبری به کار گرفته می‌شوند. همچنین این الگوریتم از سه بخش جست‌جوی شکار، محاصره شکار و حمله به شکار تشکیل شده است. در این الگوریتم مناسب‌ترین راه‌حل را الف، دومین بتا، سومین دلتا و بقیه‌ی راه حل‌ها به صورت امگا در نظر گرفته می‌شود. و همین‌طور شکار توسط گرگ‌های الف، بتا و دلتا رهبری می‌شود و گرگ‌های امگا از این گرگ‌ها تبعیت می‌کنند. شبیه‌سازی رفتار گرگ‌ها در هنگام محاصره و شکار به صورت روابط زیر بیان می‌شود (Mirjalili et al., 2014):



قدرت نسبی نرمالیزه هر امپریالیست، به صورت رابطه (۱۸) محاسبه شده و بر مبنای آن کشورهای مستعمره بین این امپریالیست‌ها تقسیم می‌شوند:

$$P_n = \left| \frac{c_n}{\sum_{i=1}^{N_{imp}} C_i} \right| \quad (18)$$

قدرت نرمالیزه شده یک امپریالیست، نسبت مستعمراتی است که توسط آن امپریالیست اداره می‌شود. تعداد اولیه مستعمرات یک امپریالیست برابر خواهد بود با:

$$N \cdot C_n = \text{round}\{P_{in}, N_{col}\} \quad (19)$$

$N \cdot C_n$: تعداد اولیه مستعمرات یک امپراتوری.
 N_{col} : تعداد کل کشورهای مستعمره موجود در جمعیت کشورهای اولیه.

round : تابعی است که نزدیک‌ترین عدد صحیح به یک عدد اعشاری را می‌دهد.

بعد از تقسیم تمام مستعمره‌ها بین امپراتورها و تشکیل امپراتوری‌های اولیه، مستعمره‌ها شروع به حرکت به سمت امپراتورها می‌کنند. این حرکت مدل‌کننده سیاست جذب است. شکل (۱) نشان‌دهنده حرکت یک مستعمره به سمت امپراتور است.

بهینه‌سازی، سه پاسخ از میان بهترین پاسخها ذخیره شده و موقعیت گرگ‌ها آپدیت می‌گردد. در این میان، با افزایش تکرارها، مقادیر بطور خطی افزایش می‌یابند. از این‌رو، گرگها تمایل به فاصله گرفتن از طعمه و نزدیک شدن به طعمه خواهند داشت و در انتها، موقعیت و مقدار پاسخ آلفا به عنوان بهترین پاسخ یافته می‌شود.

الگوریتم رقابت استعماری (ICA)^۱

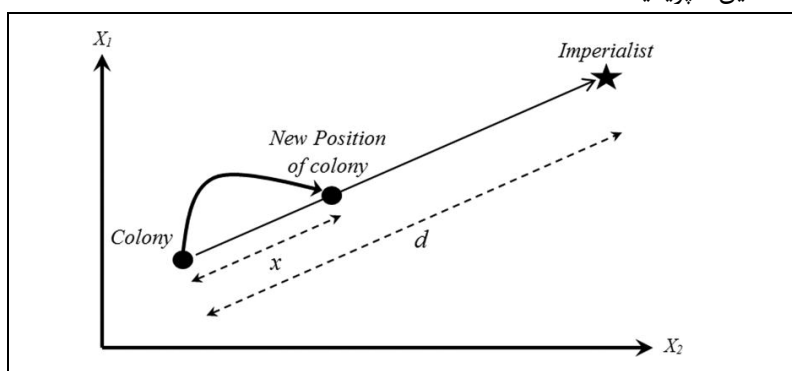
الگوریتم رقابت استعماری برگرفته از یک پدیده اجتماعی انسانی است که با یک جمعیت اولیه شروع به کار می‌کند. هر عضو از این جمعیت کشور نام داشته و دارای مشخصاتی است که مکان آن را در فضای جستجو مشخص می‌کند. از بین این نقاط، نقاطی که کمترین هزینه را طبق تابع بهینه سازی داشته باشند، به عنوان استعمارگر و مابقی به عنوان مستعمره در نظر گرفته می‌شوند (Atashpaz-Gargari et al., 2008). برای هر استعمارگر در ابتدا هزینه نرمالیزه به صورت زیر محاسبه می‌شود:

$$C_n = c_n - \max_i\{C_i\} \quad (17)$$

C_n : هزینه امپریالیست n ام.

$\max_i\{C_i\}$: بیشترین هزینه میان امپریالیست.

c_n : هزینه نرمالیزه شده این امپریالیست.



شکل (۱): نحوه حرکت کشورها در فضای جستجو بر اساس الگوریتم رقابت استعماری

¹ Imperialist Competitive Algorithm



امپراطوری از لیست امپراطوری‌ها خارج شده و در یک رقابت استعماری به‌عنوان مستعمره به امپراطوری‌های دیگر داده می‌شود. روند تکامل در یک حلقه قرار دارد که تا برآورده شدن یک شرط توقف ادامه می‌یابد.

نتایج و بحث

تخمین ضریب پخش طولی نیازمند داده‌های هیدرولیکی و هندسی جریان است. در این مطالعه مجموعه وسیعی از داده‌های مطالعات پیشین که در آنها مقادیر پارامترهای مورد نیاز ارائه شده بود، مورد استفاده قرار گرفت. این اطلاعات شامل ۱۰۷ سری داده مربوط به مشخصات جریان از قبیل عرض، عمق، سرعت، سرعت برشی و ضریب پخش طولی اندازه‌گیری شده در رودخانه‌های مختلف است که شرط $(\frac{W}{H} > 28)$ برای تمامی داده‌ها برقرار می‌باشد (Alizadeh et al., 2017a). لازم به ذکر است که در مقاله مرجع هدف از جدا نمودن داده‌ها با شرط $(\frac{W}{H} > 28)$ ، افزایش دقت مدل توسعه داده شده برای پیش‌بینی دقیق‌تر ضریب پخش طولی می‌باشد. محدوده تغییرات داده‌های مورد استفاده در جدول (۱) ارائه شده است.

در شکل (۱) d، فاصله مستقیم بین فاصله استعمارگر و مستعمرات، θ زاویه حرکت مستعمره به‌اندازه x واحد به سمت استعمارگر می‌باشد.

$$X = \sim U(0, \beta \times d) \quad (20)$$

$$\theta = \sim U(-\gamma, \gamma) \quad (21)$$

β و γ اعداد دلخواهی بوده که فضای جستجوی کلونی اطراف امپراتوری را تعیین می‌کنند. اگر حین حرکت کشورهای مستعمره به سمت کشورهای استعمارگر، بعضی این مستعمرات به موقعیت بهتری از استعمار برسند. در این حالت، کشور استعمارگر و کشور مستعمره، جای خود را باهم عوض می‌کنند. در مرحله بعد ضعیف‌ترین مستعمره از ضعیف‌ترین امپراطوری انتخاب شده و به یک امپراطوری قوی (نه لزوماً قویترین امپراطوری) داده می‌شود. البته نحوه انتخاب امپراطوری مورد نظر به صورتی است که هرچه یک امپراطوری قوی تر باشد، احتمال انتخاب آن بالاتر می‌رود. در نهایت وقتی یک امپراطوری همه مستعمره‌های خود را از دست داد، آن

جدول (۱): محدوده تغییرات داده‌های مشاهداتی (Alizadeh et al., 2017a)

پارامترها	کمینه	بیشینه	میانگین	انحراف از معیار
W	۱/۴	۲۵۳/۶	۴۶/۱۴	۴۷/۴۷
H	۰/۱۴	۸/۲	۱/۲۵	۱/۳۰
U	۰/۰۳	۱/۷۳	۰/۴۶۵	۰/۳۱۴
U_*	۰/۰۰۲	۰/۵۵۳	۰/۰۸۶	۰/۰۶۹
$\frac{W}{H}$	۲/۲۰	۴۰۳/۷۵	۴۶/۵۲	۴۵/۳۹
$\frac{U}{U_*}$	۰/۷۷	۲۰/۲۵	۶/۶۸	۴/۶۳
$K(\frac{m^2}{s})$	۰/۲۰	۱۴۸۰/۵۰	۷۶/۹۷	۱۷۵/۴۰
$\frac{K}{U_*H}$	۳/۰۸	۳۷۱۴۰/۷۴	۱۱۲۷/۲۵	۳۵۰۹/۶۴

الگوریتم‌ها به روش سعی و خطا به گونه‌ای مشخص گردد که تابع هدف مورد نظر کمینه یا بیشینه گردد. از این رو در این تحقیق پس از چندین بار اجرای مدل

در استفاده از الگوریتم‌های فراکاوشی به منظور دست‌یابی به بهترین عملکرد الگوریتم در مسئله مورد بررسی، ضروری است ابتدا پارامترهای کنترلی

برای هر یک از الگوریتم‌ها، مقادیر بهینه پارامترهای کنترلی بدست آمد. این مقادیر برای هر الگوریتم در جدول (۲) ارائه شده است.

جدول (۲): مقادیر بهینه پارامترهای کنترلی الگوریتم‌ها

الگوریتم	پارامتر کنترلی	مقدار	علامت اختصاری
ABC	Number of food sources	۲۰	SN
	Limit of abandonment	۶۰	Limit
	Maximum cycle number	۴۰۰	MFE
PSO	The size of the population	۲۰	npop
	The learning factors	۲	c_2 و c_1
	The maximum number of iterations	۴۰۰	MaxIt
GA	The replication rate	۵۰	N
	The initial population	۱۰۰	npop
	The mutation rate	۰/۳۵	μ
GWO	The selection rate	۲۰	
	The grid inflation parameter	۰/۱	α
	The leader selection pressure parameter	۴	β
	The number of grid spreach dimension	۱۰	nGrid
	The initial countries	۱۰۰	Npop
ICA	The initial imperialism	۶	$N.C_n$
	The number of colony	۹۴	N_{col}
	Other parameters	۲	β
		$\frac{\pi}{۴}$	γ
		۰/۰۱	ξ

نظر گرفته شد. معادله پیشنهادی جهت محاسبه ضریب پخش طولی در هر الگوریتم با حداقل سازی مجذور میانگین مربعات خطا به عنوان تابع هدف در جدول (۳) نشان داده شده است. همچنین مقدار بهینه تابع هدف و سایر شاخص‌های آماری مربوط به اجرای نهایی هر الگوریتم در جدول (۳) ارائه شده است

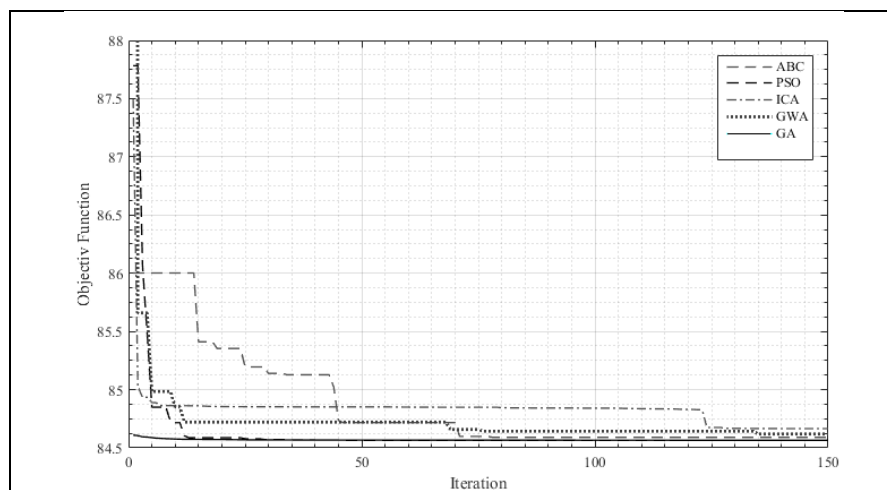
پس از تعیین مقادیر پارامترهای کنترلی برای هر الگوریتم، تابع هدف مسئله (رابطه ۳) با هر یک از الگوریتم‌های مورد بررسی بهینه گردید. به منظور دستیابی به جوابی با صحت بالاتر، هر الگوریتم ۶ نوبت اجرا شد و از این میان، جوابی که بهترین نتیجه را برای تابع هدف نشان دهد به عنوان جواب نهایی الگوریتم در

جدول (۳): مقایسه نتایج بدست آمده از الگوریتم‌ها

Algorithm	Optimal	Maximum	Mean	Standard Deviation	Predicted Equation
GA	۸۴/۵۶۴۹	۸۴/۶۱۰۱	۸۴/۵۶۵۹	۰/۰۰۳۹	$\frac{K}{U_*H} = 5.7762 \left(\frac{W}{H}\right)^{0.2728} \left(\frac{U}{U_*}\right)^{1.8635}$
PSO	۸۴/۵۶۵۰	۸۷/۷۸۲۳	۸۴/۵۸۶۷	۰/۲۱۹۲	$\frac{K}{U_*H} = 5.7650 \left(\frac{W}{H}\right)^{0.2732} \left(\frac{U}{U_*}\right)^{1.8635}$
ABC	۸۴/۵۸۳۱	۸۶/۰۰۲۱	۸۴/۵۹۸۶	۰/۱۶۵۸	$\frac{K}{U_*H} = 5.9670 \left(\frac{W}{H}\right)^{0.2659} \left(\frac{U}{U_*}\right)^{1.8617}$
GWO	۸۴/۵۶۵۲	۱۲۲/۲۱۷۳	۸۴/۶۸۴۲	۱/۶۸۵۲	$\frac{K}{U_*H} = 5.8567 \left(\frac{W}{H}\right)^{0.2680} \left(\frac{U}{U_*}\right)^{1.8654}$
ICA	۸۴/۵۷۳۲	۸۷/۴۹۵۵	۸۴/۶۳۱۸	۰/۱۲۶۳	$\frac{K}{U_*H} = 6.2479 \left(\frac{W}{H}\right)^{0.2421} \left(\frac{U}{U_*}\right)^{1.8794}$

از این رو نتایج حاصل از همگرایی الگوریتم‌های مورد بحث در رسیدن به پاسخ بهینه بصورت نمودارهایی در شکل (۷) ارائه شده است. طبق این نمودارها می‌توان اظهار داشت، الگوریتم GA نسبت به چهار الگوریتم دیگر علاوه بر دقت، سرعت بالاتری نیز در همگرایی به پاسخ مسئله دارد

طبق جدول (۳)، کمترین مقدار برای تابع هدف از روش GA به دست آمد که برابر ۸۴/۵۶۴۹ است که در میان سایر الگوریتم‌ها بهینه‌ترین پاسخ را ارائه داده است. در بررسی توانایی الگوریتم‌ها در حل مسائل بهینه‌سازی این نکته حائز اهمیت است که هر چه الگوریتم با تعداد تکرار کمتری به پاسخ بهینه دست یابد نشان دهنده عملکرد بهتر و سرعت بالاتر آن در حل مسئله می‌باشد.



شکل (۲): نمودار همگرایی الگوریتم‌ها به پاسخ بهینه

روش‌های آماری که بر پایه مقادیر مشاهداتی و پیش‌بینی شده استوار هستند نظیر خطای استاندارد، نمی‌توانند حساسیت مسئله را نشان دهند. در نتیجه استفاده از روش‌های آماری مبتنی بر فرم لگاریتمی داده‌ها برای بررسی نتایج، برتری دارد (Kashefipour

ضریب پخش طولی تحت تأثیر پارامترهای مختلف جریان و پارامترهای هندسی کانال قرار دارد و برای کانال‌ها با اندازه و انواع مختلف در محدوده وسیعی تغییر می‌کند. از این رو ممکن است خطای نسبی زیادی در پیش‌بینی این پارامتر رخ دهد. بنابراین استفاده از

(DR) ، میانگین خطای مطلق (ME) و خطای مجذور میانگین مربعات (RMS) استفاده شده که روابط هر یک در زیر ارائه شده است.

and Falconer, 2002). در این تحقیق، جهت مقایسه کارایی مدل‌های پیشنهادی و میزان دقت و صحت نتایج الگوریتم‌ها، از پارامترهای آماری شامل نسبت توافق

$$DR = \log_{(10)} \frac{K_{(predicted)}}{K_{(observed)}} \quad (22)$$

$$ME = \frac{1}{N} \sum_{i=1}^N |DR_i| \quad (23)$$

$$RMS = \frac{1}{N} \sqrt{\sum_{i=1}^N (DR_i)^2} \quad (24)$$

N

الگوریتم محاسبه شده، که این مقادیر در جدول (۴) نشان داده شده است.

: تعداد کل داده‌ها می‌باشد.

با توجه به مقادیر ضریب پخش طولی بدست آمده از هر یک از الگوریتم‌ها، میزان پارامترهای آماری در هر

جدول (۴): عملکرد الگوریتم‌های مورد بررسی بر اساس پارامترهای آماری

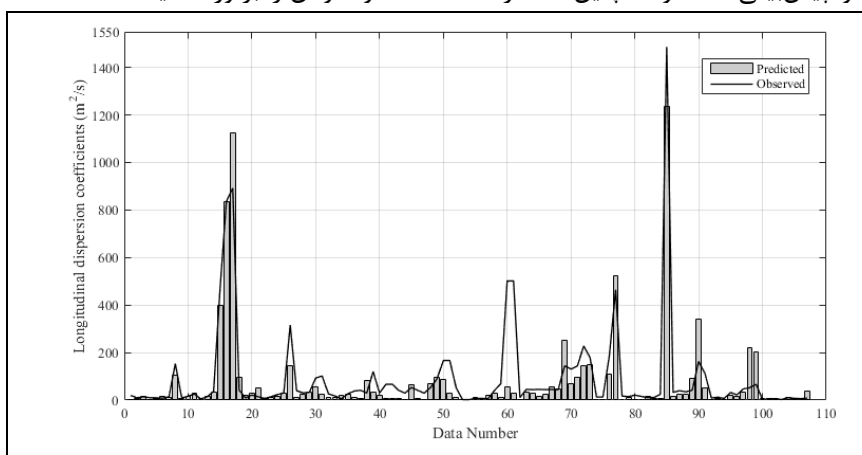
الگوریتم	DR	ME	RMS
GA	-۲/۸۹۳ ، ۱/۴۹۳	۰/۴۳۵۴	۰/۱۰۵۲
PSO	-۲/۸۵۷ ، ۱/۴۹۹	۰/۴۵۱۹	۰/۱۰۶۲
ABC	-۲/۸۶۰ ، ۱/۴۹۸	۰/۴۴۹۶	۰/۱۰۶۱
GWO	-۲/۸۶۶ ، ۱/۴۹۷	۰/۴۵۳۷	۰/۱۰۶۳
ICA	-۲/۹۲۱ ، ۱/۴۸۳	۰/۴۶۸۸	۰/۱۰۴۷

مقدار DR به صفر نزدیک‌تر باشد، مقادیر ضریب پخش پیش‌بینی شده به مقادیر واقعی نزدیک‌تر بوده و به عبارت دیگر الگوریتم عملکرد بهتری در تخمین این پارامتر داشته است. طبق جدول (۴)، پارامتر ME در الگوریتم ژنتیک در مقایسه با سایر الگوریتم‌ها کمتر بوده و به صفر نزدیک‌تر است. اگرچه الگوریتم ژنتیک کمترین مقدار RMS را دارا نمی‌باشد ولی با توجه به معیارهای همگرایی و مقدار تابع هدف که به تفصیل بیان شد، می‌توان نتیجه گرفت با وجود قابل قبول بودن جواب تمام الگوریتم‌های مذکور در تخمین ضریب پخش طولی، الگوریتم ژنتیک موفق‌تر عمل کرده است. همان‌طور که گفته شد الگوریتم GA بهینه‌ترین پاسخ را برای تابع هدف مسئله ارائه می‌کند و با سرعت

بطور کلی یک مدل پیشنهادی در صورتی بهینه خواهد بود که میزان DR، ME و RMS آن به مقدار صفر نزدیک باشد. به عبارت دیگر فاصله کمتر این پارامترهای آماری با مقدار صفر نشان دهنده صحت و کارایی بالاتر روش پیشنهادی است. در جدول فوق نسبت توافق (DR)، تطابق بین مقادیر ضریب پخش طولی اندازه‌گیری شده و مقادیر تخمین زده شده توسط هر الگوریتم را نشان می‌دهد. اگر مقادیر پیش‌بینی شده با مقادیر مشاهداتی یکسان باشند، مقدار DR برابر صفر است. در صورتیکه مقدار DR مثبت باشد مقادیر پیش‌بینی شده بیش از مقادیر اندازه‌گیری شده است و در صورتیکه مقدار DR منفی باشد نشان دهنده کمتر تخمین زدن مقادیر توسط الگوریتم‌ها است. هر چقدر

کلی روش GA قابل رؤیت است. با ملاحظه نمودار، همانگونه که مشخص است پیروی مدل پیش‌بینی ضریب پخش از داده‌های مشاهداتی که به صورت میدانی جمع‌آوری شده‌اند، به خوبی مشهود است. به‌گونه‌ای که مقادیر ضریب پخش طولی پیش‌بینی شده با دقت بالایی به مقادیر واقعی نزدیک است به این معنی که مدل به خوبی و با خطای کمی توانسته است الگوی حاکم بر انتقال آلودگی و ضریب پخش طولی را درک نموده و آن را برآورد نماید.

بالاتری نسبت به سایر الگوریتم‌ها به جواب نهایی همگرا می‌گردد. در شکل (۳) مقادیر ضریب پخش پیش‌بینی شده توسط الگوریتم بهینه‌سازی GA در کنار مقادیر مشاهداتی در مرحله صحت‌سنجی نشان داده شده است. برای محسوس بودن نتایج، مقادیر ضریب پخش پیش‌بینی شده به صورت نمودار میله‌ای و مقدار ضریب پخش اندازه‌گیری شده به صورت نمودار خط ممتد نمایش داده شده‌اند. در این نمودار اختلاف مقادیر اندازه‌گیری شده و پیش‌بینی شده و همپنین عملکرد

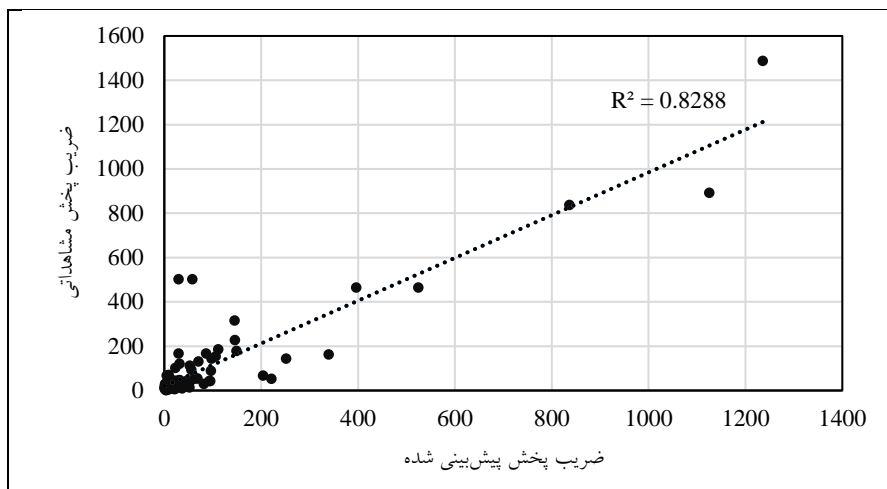


شکل (۳): عملکرد الگوریتم GA در صحت‌سنجی مدل

ارائه شده را نشان می‌دهد که این مقدار هرچه به عدد ۱ نزدیک‌تر باشد نشان دهنده‌ی همبستگی بیشتر داده‌های تخمین زده شده با داده‌های مشاهداتی است که مقدار این پارامتر در شکل (۴) نشان داده شده است

علاوه بر پارامترهای آماری بررسی شده، طبق رابطه (۲۵) از ضریب همبستگی (R^2) نیز جهت بررسی دقت ضریب پخش طولی پیش‌بینی شده با استفاده از الگوریتم GA استفاده شد. R^2 درصد عدم قطعیت مدل

$$R^2 = \frac{[\sum (K_{i(observed)} - \bar{K}_{(observed)})(K_{i(predicted)} - \bar{K}_{(predicted)})]^2}{\sum (K_{i(observed)} - \bar{K}_{(observed)})^2 \sum (K_{i(predicted)} - \bar{K}_{(predicted)})^2} \quad (25)$$



شکل (۴): مقدار R2 برای ضریب پخش طولی مشاهده‌ای و محاسباتی به روش GA

پیش‌بینی ضریب پخش طولی استفاده شد. نتایج بدست آمده نشان داد الگوریتم بهینه سازی GA در شرایط تکرار یکسان نسبت به الگوریتم‌های ABC, PSO, ICA, GWO از دقت و سرعت بالاتری در تخمین ضریب پخش طولی در رودخانه‌ها برخوردار می‌باشد. به عبارتی می‌توان اظهار داشت اگرچه تمامی الگوریتم‌های فراکاوشی مورد بررسی عملکرد قابل قبولی در تخمین ضریب پخش طولی داشته‌اند با این وجود در این میان الگوریتم بهینه‌سازی ژنتیک نتایج بهتر و قابل قبول‌تری ارائه داده است. هم‌چنین مقایسه نتایج الگوریتم GA با داده‌های واقعی حاکی از قابل قبول بودن الگوریتم مذکور در مسئله مورد بررسی است. در نتیجه معادله ارائه شده توسط الگوریتم ژنتیک به عنوان معادله منتخب در میان معادلات پیشنهادی، در نظر گرفته شد. به منظور بررسی صحت مدل‌های پیشنهادی، از سه معیار خطای مجذر میانگین مربعات، میانگین خطای مطلق و ضریب همبستگی استفاده شد، که این مقادیر برای رابطه به دست آمده از الگوریتم ژنتیک به ترتیب برابر ۰/۱۰۵۲، ۰/۴۳۵۴ و ۰/۸۲۸۸ می‌باشد که عملکرد مطلوب این الگوریتم را نشان می‌دهد.

با توجه به شکل (۴) و معیار R² می‌توان دریافت که نتایج مدل GA تخمین مناسبی از داده‌های مشاهده‌ای داده است به عبارتی می‌توان اینگونه اظهار نظر کرد که، مقادیر ضریب پخش طولی محاسبه شده بسیار نزدیک به مقادیر ضریب پخش طولی اندازه‌گیری شده است. در نهایت رابطه پیشنهادی برای تخمین ضریب پخش طولی بدست آمده از روش GA به عنوان دقیق‌ترین رابطه به صورت زیر پیشنهاد می‌شود:

$$\frac{K}{U_*H} = 5.7762 \left(\frac{W}{H}\right)^{0.2728} \left(\frac{U}{U_*}\right)^{1.8635} \quad (26)$$

نتیجه‌گیری

استفاده از الگوریتم‌های بهینه‌سازی فراکاوشی در مدیریت کیفی سیستم‌های منابع آب در سال‌های اخیر به صورت گسترده‌ای مورد توجه قرار گرفته است. در این مطالعه چند مورد از مطرح‌ترین الگوریتم‌های بهینه‌سازی فراکاوشی برای حل مسئله تخمین ضریب پخش طولی در رودخانه مورد ارزیابی قرار گرفتند. بدین منظور از ۱۰۷ سری داده مربوط به پارامترهای هیدرولیکی و هندسی جریان جهت ارائه معادله‌ای برای



منابع

- . امامی، ح. و س. امامی، ۱۳۹۸. ارائه روشی جدید برای ارزیابی کیفیت هیدروژنوشیمیایی آب‌های زیرزمینی با استفاده از الگوریتم‌های هوش جمعی، نشریه اکوهیدرولوژی، دوره ۶، شماره ۱، ص ۱۹۰-۱۷۷.
- پیاده کوهسار، ج.، ح. مازتدرانی‌زاده و س. م. ک. صدر، ۱۳۹۸. ارزیابی الگوریتم‌های بهینه‌سازی GA و PSO در بهره‌برداری از سیستم‌های چندمخزنه، مطالعه موردی: سدهای حوضه گرگان‌رود، نشریه پژوهش‌های حفاظت آب و خاک، جلد ۲۶، شماره ۲، ص ۲۵۰-۲۳۹.
- جعفری‌اصل، ج.، ب. سامی‌کشکولی و م. بهرامی، ۱۳۹۴. مدیریت بهینه فشار در جهت حداقل سازی نشت در سیستم‌های توزیع آب با استفاده از الگوریتم بهینه‌سازی رقابت استعماری، دو فصلنامه تخصصی علوم و مهندسی آب، سال ۵، شماره ۱۲، ص ۸۲-۶۹.
- حجتی، س. ح.، س. ه. حجتی و س. ع. ا. صالحی نیشابوری، ۱۳۹۵. بهینه‌سازی زاویه جام پرتابی مثلثی و عرض سرریز شوت با استفاده از الگوریتم ژنتیک، مجله عمران مدرس، دوره ۱۶، شماره ۲، ص ۹۰-۸۱.
- سامی‌کشکولی، ب.، م. بهرامی و م. انصاری‌جبری، ۱۳۹۵. بهره‌برداری بهینه از ایستگاههای پمپاژ سیستم‌های تأمین آب با استفاده از الگوریتم جامعه زنبورهای عسل مصنوعی (ABC)، نشریه پژوهش‌های حفاظت آب و خاک، جلد ۲۳، شماره ۵، ص ۱۸۹-۱۷۵.
- گیوه چی، م.، م. فغفور مغربی و ج. ابریشمی، ۱۳۸۸. کاربرد سرعت متوسط عمقی در تخمین ضریب پخش طولی آلودگی در رودخانه‌ها، مجله آب و فاضلاب، دوره ۲۰، شماره ۴، ص ۹۶-۹۱.
- هاشمی‌منفرد، س. آ.، ف. حسین‌زاده و ب. پیرزاده، ۱۳۹۶. کاربرد الگوریتم رقابت استعماری در بهینه‌سازی بهره‌برداری مخزن با هدف حداکثر سازی تأمین نیاز (مطالعه موردی: سد پیشین)، مجله هیدرولیک، دوره ۱۲، شماره ۲، ص ۶۷-۵۹.
- Alizadeh, M.J., Ahmadyar, D. and Afghantoloe, A., 2017. Improvement on the existing equations for predicting longitudinal dispersion coefficient. *Water resources management*, 31(6), pp.1777-1794.
- Alizadeh, M.J., Shahheydari, H., Kavianpour, M.R., Shamloo, H. and Barati, R., 2017. Prediction of longitudinal dispersion coefficient in natural rivers using a cluster-based Bayesian network. *Environmental Earth Sciences*, 76(2), p.86.
- Atashpaz-Gargari, E., Hashemzadeh, F., Rajabioun, R. and Lucas, C., 2008. Colonial Competitive Algorithm, a novel approach for PID controller design in MIMO distillation column process. *International Journal of Intelligent Computing and Cybernetics*, Vol. 1 No. 3, 2008, pp. 337-355.
- Balf, M.R., Noori, R., Berndtsson, R., Ghaemi, A. and Ghiasi, B., 2018. Evolutionary polynomial regression approach to predict longitudinal dispersion coefficient in rivers. *Journal of Water Supply: Research and Technology-Aqua*, p.jws 2018021.
- Chow VT, 1959. *Open-channel hydraulics*. McGraw-Hill, New York.
- Dehghani, M., Seifi, A. and Riahi-Madvar, H., 2019. Novel forecasting models for immediate-short-term to long-term influent flow prediction by combining ANFIS and Grey Wolf optimization. *Journal of Hydrology*.



- Deng, Z. Q., V. P. Singh, L. Bengtsson. 2001. Longitudinal dispersion coefficient in straight rivers, *Journal of Hydraulic Engineering*, 127, 919-927.
- Eberhart, R. and Kennedy, J., 1995, October. A new optimizer using particle swarm theory. In *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on* (pp. 39-43). IEEE.
- Elder, J.W., 1959. The dispersion of marked fluid in turbulent shear flow. *Journal of fluid mechanics*, 5(4), pp.544-560.
- Etemad-Shahidi, A. and Taghipour, M., 2012. Predicting longitudinal dispersion coefficient in natural streams using M5' model tree. *Journal of hydraulic engineering*, 138(6), pp.542-554.
- Fisher, H.B., 1968. Dispersion predictions in natural streams. *Journal of the Sanitary Engineering Division*, 94(5), pp.927-944.
- Karaboga, D. and Basturk, B., 2007. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of global optimization*, 39(3), pp.459-471
- Karaboga, D. and Ozturk, C., 2011. A novel clustering approach: Artificial Bee Colony (ABC) algorithm. *Applied soft computing*, 11(1), pp.652-657.
- Kashefipour, S.M. and Falconer, R.A., 2002. Longitudinal dispersion coefficients in natural channels. *Water Research*, 36(6), pp.1596-1608.
- Kennedy J., Eberhart R.C. (1995). Particle swarm optimization, in: *IEEE International Conference on Neural Networks*, Perth, Australia, IEEE Service Center, Piscataway, NJ.
- Liu, H. 1977. Predictiong dispersion coefficient of stream, *Journal of Enviroment Engineering Division*, 103(1). 59-69.
- Mitchell, M., 1999. An introduction to genetic algorithms, *Library of Congress Cataloging-in-Publication Data*.
- Mirjalili, S.A., Mirjalili, S.M. and Lewis, A., 2014. Grey wolf optimizer. *Advances in engineering software*, 69, pp.46-61.
- Montalvo I., Izquierdoa J., Perez R. and Tungb M.M. 2008. Particle Swarm Optimization applied to the design of water supply systems. *Computers and Mathematics with Applications*, (p. 769–776).
- Noori, R., Ghiasi, B., Sheikhan, H. and Adamowski, J.F., 2017. Estimation of the dispersion coefficient in natural rivers using a granular computing model. *Journal of Hydraulic Engineering*, 143(5), p.04017001.
- Roy, R., Dehuri, S. and Cho, S.B., 2011. A novel particle swarm optimization algorithm for multi-objective combinatorial optimization problem. *International Journal of Applied Metaheuristic Computing (IJAMC)*, 2(4), pp.41-57.
- Sahay, R.R. and Dutta, S., 2009. Prediction of longitudinal dispersion coefficients in natural rivers using genetic algorithm. *Hydrology Research*, 40(6), pp.544-552.
- Sattar, A. M., B. Gaharabaghi. 2015. Gene expression models for prediction of longitudinal dispersion coefficient in streams, *Journal of hydrology*, 524, pp. 587-596.



Seo, I., T. cheong. 1998. Predicting longitudinal dispersion coefficient in natural streams., Journal of Hydraulic engineering, 124, pp. 25-32.

Taylor, G.I., 1954. The dispersion of matter in turbulent flow through a pipe. Proc. R. Soc. Lond. A, 223(1155), pp.446-468.

Tukey J. 1965. The future of data analysis. Ann. Math. Stat., 33(1):13.

Wolpert, D.H. and Macready, W.G., 1997. No free lunch theorems for optimization. IEEE transactions on evolutionary computation, 1(1), pp.67-82.

Yu, S. and Lu, H., 2018. An integrated model of water resources optimization allocation based on projection pursuit model–Grey wolf optimization method in a transboundary river basin. Journal of hydrology, 559, pp.156-165.



Estimation of Longitudinal Dispersion Coefficient of Rivers by Using Meta-Heuristic Algorithms

Sima Ohadi¹, Seyed Arman Hashemi Monfared^{2*}, Mohammad Givehchi³

Abstract

Across the world, rivers are the main supply arches for drinking, industry and agriculture. Hence, in recent years, the identification and study of the processes of transport of pollution in rivers has become very important. Longitudinal dispersion coefficient is the most important parameter in the transport equation, and various methods for its estimating and calculating have been presented so far. Considering the nature of the problem, Meta-Heuristic Optimization methods are the important approaches for estimating the longitudinal dispersion coefficient. In this research, the ability of five Meta-Heuristic Algorithms to estimate the parameter was evaluated and compared. The equation of calculating the longitudinal dispersion coefficient using each of the algorithms was presented and the results were compared with the observational data. According to the results, the genetic optimization algorithm has the best accuracy and speed compared to the other algorithms under the same iteration conditions. Then statistical parameters including the mean of the absolute error, root mean squared error and correlation coefficient were used to check the accuracy of the algorithms. These values for the genetic algorithm were 0.4354, 0.1052 and 0.8288, respectively. As a result, it can be safely introduced the equation derived from this algorithm effective in calculating the longitudinal dispersion coefficient.

Keywords: Contaminant transport, longitudinal dispersion coefficient, timization, Meta-Heuristic methods.

¹ PhD student of civil Engineering, Faculty of Engineering, University of Sistan and Baluchestan, Zahedan, Iran; Ohadi_Sima@pgs.usb.ac.ir.

² Associate Professor of Civil Engineering, Faculty of Engineering, University of Sistan and Baluchestan, Zahedan, Iran. Corresponding Author; Hashemi@eng.usb.ac.ir.

³ Assistant Professor of Civil Engineering, Faculty of Engineering, University of Sistan and Baluchestan, Zahedan, Iran; M.Givehchi@eng.usb.ac.ir.



Estimation of Longitudinal Dispersion Coefficient of Rivers by Using Meta-Heuristic Algorithms

Sima Ohadi¹, Seyed Arman Hashemi Monfared^{2*}, Mohammad Givehchi³

Introduction

Across the world, rivers are the main supply arches for drinking, industry and agriculture. Hence, in recent years, the identification and study of the processes of transport of pollution in rivers has become very important. Longitudinal dispersion coefficient is the most important parameter in the transport equation, and various methods for its estimating and calculating have been presented so far. Recently some of efficient soft computing approaches such as genetic algorithm, M5 tree model, Bayesian network, gene expression programming, evolutionary polynomial regression and so on were used to estimate the longitudinal dispersion coefficient parameter in natural rivers (Sahay and Dutta, 2009; Etemad-Shahidi and Taghipour, 2012; Sattar and Gharabaghi, 2015; Alizadeh et al., 2017a; Rezaie Balf et al., 2018). In this research, the ability of five Meta-Heuristic Algorithms to estimate the parameter was evaluated and compared. The equation of calculating the longitudinal dispersion coefficient using each of the algorithms was presented and the results were compared with the observational data. According to the results, the genetic optimization algorithm has the best accuracy and speed compared to the other algorithms under the same iteration conditions. Then statistical parameters including the mean of the absolute error, root mean squared error and correlation coefficient were used to check the accuracy of the algorithms. These values for the genetic algorithm were 0.4354, 0.1052 and 0.8288, respectively. As a result, it can be safely introduced the equation derived from this algorithm effective in calculating the longitudinal dispersion coefficient.

Methodology

1- Longitudinal Dispersion Coefficient

Based on previous studies, important and effective parameters on the longitudinal dispersion coefficient in rivers can be expressed as equation (1) using dimensional analysis.

$$\frac{K}{U_*H} = f'\left(\rho \frac{HU}{\mu}, \frac{W}{H}, \frac{U}{U_*}, \frac{k_s}{H}, S_f, S\right) \quad (1)$$

Where $\frac{K}{U_*H}$ is the dimensionless parameter of the dispersion coefficient, $\rho \frac{HU}{\mu}$ is the Reynolds number, $\frac{W}{H}$ is the ratio of width to depth of flow and $\frac{U}{U_*}$ is related to the term of flow resistance. The direct influence of the parameters related to the slope of the canal, the shape of the bed and

¹ PhD student of civil Engineering, Faculty of Engineering, University of Sistan and Baluchestan, Zahedan, Iran; Ohadi_Sima@pgs.usb.ac.ir.

² Associate Professor of Civil Engineering, Faculty of Engineering, University of Sistan and Baluchestan, Zahedan, Iran. Corresponding Author; Hashemi@eng.usb.ac.ir.

³ Assistant Professor of Civil Engineering, Faculty of Engineering, University of Sistan and Baluchestan, Zahedan, Iran; M.Givehchi@eng.usb.ac.ir.

the sinusoidal value of the river path, which are not easily measurable, is neglected. Also, because the flow in rivers is usually turbulent, the effect of the Reynolds number can be ignored and the final form of Equation (1) can be written as Equation (2):

$$\frac{K}{U_*H} = f' \left(\frac{W}{H}, \frac{U}{U_*} \right) \quad (2)$$

Finally, the longitudinal dispersion coefficient in rivers can be expressed as equation (3) by considering f' as an exponential function.

$$\frac{K}{U_*H} = a \left(\frac{W}{H} \right)^b \left(\frac{U}{U_*} \right)^c \quad (3)$$

In this equation, the longitudinal dispersion coefficient (K) depends on four parameters including U: average flow velocity, H: flow depth, W: width of flow and U_* : flow shear velocity. Also, a, b and c are the constant coefficients of equation (3).

2- Proposed method

The mathematical model of the longitudinal dispersion coefficient optimization problem is expressed as follows:

$$\text{Find } X = \{x_1, x_2, x_3\} \quad (4)$$

$$\text{to minimize } RMSE = \sqrt{\frac{\sum_{i=1}^n (K_{i(\text{observed})} - K_{i(\text{predicted})})^2}{n}} \quad (5)$$

Subject to:

$$L_b \leq X \leq U_b$$

The decision variables vector of the optimization problem includes the coefficients a, b and c of equation (3). Also the objective function in this study is considered as minimizing the difference between the predicted values and the observed values. The only constraint of this optimization problem is satisfying the lower (L_b) and upper (U_b) range of decision variables.

In this study, the performance of widely used algorithms in the field of water engineering including GA, ABC, ICA, PSO and GWO were evaluated to predict the longitudinal dispersion coefficient.

Discussion and Conclusion

Each algorithm was executed 6 times and among them, the answer that shows the best result for the objective function was considered as the final answer of the algorithm. The statistical indicators and the proposed equation for calculating the longitudinal dispersion coefficient in each algorithm by minimizing the mean squared error as the objective function is shown in Table (1).



Table (1): Comparison of results obtained from algorithms

Algorithm	Optimal	Maximum	Mean	Standard Deviation	Predicted Equation
GA	۸۴/۵۶۴۹	۸۴/۶۱۰۱	۸۴/۵۶۵۹	۰/۰۰۳۹	$\frac{K}{U_*H} = 5.7762 \left(\frac{W}{H}\right)^{0.2728} \left(\frac{U}{U_*}\right)^{1.8635}$
PSO	۸۴/۵۶۵۰	۸۷/۷۸۲۳	۸۴/۵۸۶۷	۰/۲۱۹۲	$\frac{K}{U_*H} = 5.7650 \left(\frac{W}{H}\right)^{0.2732} \left(\frac{U}{U_*}\right)^{1.8635}$
ABC	۸۴/۵۸۳۱	۸۶/۰۰۲۱	۸۴/۵۹۸۶	۰/۱۶۵۸	$\frac{K}{U_*H} = 5.9670 \left(\frac{W}{H}\right)^{0.2659} \left(\frac{U}{U_*}\right)^{1.8617}$
GWO	۸۴/۵۶۵۲	۱۲۲/۲۱۷۳	۸۴/۶۸۴۲	۱/۶۸۵۲	$\frac{K}{U_*H} = 5.8567 \left(\frac{W}{H}\right)^{0.2680} \left(\frac{U}{U_*}\right)^{1.8654}$
ICA	۸۴/۵۷۳۲	۸۷/۴۹۵۵	۸۴/۶۳۱۸	۰/۱۲۶۳	$\frac{K}{U_*H} = 6.2479 \left(\frac{W}{H}\right)^{0.2421} \left(\frac{U}{U_*}\right)^{1.8794}$

According to Table (1), the minimum value for the objective function was obtained from the GA method, which is equal to 84/5649, which provides the most optimal answer among other algorithms.

In this study, to compare the efficiency of the proposed models and the accuracy of the algorithm results, statistical parameters including discrepancy ratio (DR), mean absolute error (ME) and root mean square (RMS) were used. These values are shown in Table (2)

Table (2): Performance of the studied algorithms

Algorithm	DR	ME	RMS
GA	-۲/۸۹۳ ، ۱/۴۹۳	۰/۴۳۵۴	۰/۱۰۵۲
PSO	-۲/۸۵۷ ، ۱/۴۹۹	۰/۴۵۱۹	۰/۱۰۶۲
ABC	-۲/۸۶۰ ، ۱/۴۹۸	۰/۴۴۹۶	۰/۱۰۶۱
GWO	-۲/۸۶۶ ، ۱/۴۹۷	۰/۴۵۳۷	۰/۱۰۶۳
ICA	-۲/۹۲۱ ، ۱/۴۸۳	۰/۴۶۸۸	۰/۱۰۴۷

According to table (2) and the value of the objective function, it can be concluded that despite the acceptable answer of all the mentioned algorithms in estimating the longitudinal dispersion coefficient, the genetic algorithm has been more successful. As a result, the equation presented by the genetic algorithm was considered as the selected equation among the proposed equations for estimating the longitudinal dispersion coefficient.

The most important references

Alizadeh, M.J., Ahmadyar, D. and Afghantoloe, A., 2017. Improvement on the existing equations for predicting longitudinal dispersion coefficient. *Water resources management*, 31(6), pp.1777-1794.

Balf, M.R., Noori, R., Berndtsson, R., Ghaemi, A. and Ghiasi, B., 2018. Evolutionary polynomial regression approach to predict longitudinal dispersion coefficient in rivers. *Journal of Water Supply: Research and Technology-Aqua*, p.jws 2018021.

Etemad-Shahidi, A. and Taghipour, M., 2012. Predicting longitudinal dispersion coefficient in natural streams using M5' model tree. *Journal of hydraulic engineering*, 138(6), pp.542-554.

Sahay, R.R. and Dutta, S., 2009. Prediction of longitudinal dispersion coefficients in natural rivers using genetic algorithm. *Hydrology Research*, 40(6), pp.544-552.

Sattar, A. M., B. Gaharabaghi. 2015. Gene expression models for prediction of longitudinal dispersion coefficient in streams, *Journal of hydrology*, 524, pp. 587-596.