

## برآورد ضریب پخشیدگی طولی رودخانه با استفاده از انواع روش‌های داده‌کاوی

سمیه سلطانی گردفرامرز<sup>۱\*</sup>، روح‌الله تقی‌زاده مهرجردی<sup>۲</sup>، محسن قاسمی<sup>۳</sup>

۱. استادیار دانشکده کشاورزی و منابع طبیعی دانشگاه اردکان

۲. استادیار دانشکده کشاورزی و منابع طبیعی دانشگاه اردکان

۳. دانشجوی دکتری علوم و مهندسی آب دانشگاه صنعتی اصفهان

(تاریخ دریافت: ۱۳۹۳/۲/۳۱ - تاریخ تصویب: ۱۳۹۴/۲/۲)

### چکیده

در مدل‌سازی و تعیین دقیق وضعیت آلودگی رودخانه‌ها محاسبه دقیق ضریب پراکندگی طولی آلودگی بسیار اهمیت دارد. برای محاسبه این ضریب، معادلات گوناگون با استفاده از روش‌های تجربی، تحلیلی، و ریاضی ارائه شده است. با وجود این، روش‌های تحلیلی و ریاضی به علت پیچیدگی محاسبات و روش‌های تجربی به سبب خطای زیاد تا کنون مورد توجه قرار نگرفته‌اند. این تحقیق به بررسی روش‌ها و معادلات تجربی مختلف برای تعیین ضریب پراکندگی طولی آلودگی در رودخانه‌های طبیعی و ارزیابی دقت این روش‌ها در مقایسه با داده‌های اندازه‌گیری شده واقعی پرداخت و روشی دقیق‌تر در این زمینه، با بهره‌جستن از روش‌های داده‌کاوی، همچون برنامه‌ریزی ژنتیک، شبکه عصبی، و شبکه عصبی-فازی ارائه شد. با به‌کارگیری مدل نروفازی، معیارهای ریشه مربعات خطا و ضریب تبیین به ترتیب ۷۲/۲۱ و ۰/۸۷ و ضریب جرم باقی‌مانده ۰/۱۰۳ و کارایی مدل ۰/۷۵ به دست آمد. به این ترتیب، روش نروفازی جهت پیش‌بینی ضریب پخشیدگی طولی رودخانه پیشنهاد می‌شود.

**کلیدواژگان:** آلودگی، رودخانه، روش‌های داده‌کاوی، ضریب پخشیدگی طولی.

### مقدمه

مطالعه کیفیت آب‌های سطحی از موضوعات مهم در زمینه محیط زیست است. این موضوع، در شرایطی که رودخانه‌ها منبع تهیه آب آشامیدنی انسان‌ها و آب لازم برای صنایع‌اند و در عین حال فاضلاب و پساب‌های صنعتی در رودخانه‌ها تخلیه می‌شود، به توجه بیشتری نیاز دارد (Abedi-kupaie et al., 2007). از طرف دیگر مشکلات زیست‌محیطی ناشی از خشک‌سالی اهمیت مطالعه در زمینه آلودگی آب‌های سطحی را بیشتر می‌کند. در نتیجه، به منظور حفظ سلامت عمومی، حفاظت و پاک‌سازی رودخانه‌ها همواره مورد توجه است. توانایی در تخمین و شبیه‌سازی جریان، انتقال آلودگی، و رسوب در رودخانه‌ها برای برنامه‌ریزی منابع آب اهمیت دارد؛ زیرا موجب کاهش خطر خسارت به محیط زیست می‌شود. در حال حاضر استفاده از فرایند پخشیدگی و تقویت خودپالایی رودخانه‌ها از مهم‌ترین روش‌های مدیریت زیست‌محیطی رودخانه‌ها به شمار می‌رود (Purabadehie et al., 2003). برای کنترل آلودگی رودخانه‌ها، رهاسازی آلاینده باید به صورت منطقی و تنظیم‌شده انجام شود.

این موضوع به اطلاع دقیق از توانایی حمل و پخش و پاک‌سازی آلودگی، توسط جریان آب در طول مشخصی از مسیر خود، نیاز دارد؛ که آن را طول اختلاط کامل می‌نامند. با تزریق آلاینده‌های مختلف به درون رودخانه، انتقال آن‌ها به پایین‌دست به کمک فرایند اختلاط با جریان و پخشیدگی طولی و عرضی و عمقی، تحت تأثیر فرایندهای انتقال و اختلاط، انجام می‌گیرد. قابلیت و قدرت جریان رودخانه و دیگر جریان‌های سطحی در هر سه جهت طولی و عرضی و عمقی با ضرایب پراکندگی  $K_x$ ،  $K_y$  و  $K_z$  بیان می‌شود (Tayfur and Singh, 2005). در نقاط دورتر از محل تزریق آلاینده، که اختلاط در کل مقطع عرضی کامل می‌شود، فقط پراکندگی طولی قابل توجه است. شدت پراکندگی طولی نیز با ضریب پراکندگی طولی تعیین می‌شود و سرنوشت انتقال آلاینده تابع پراکندگی طولی است. در نتیجه در مدل‌سازی و تعیین دقیق وضعیت آلودگی رودخانه‌ها محاسبه دقیق ضریب پراکندگی طولی آلودگی بسیار مهم است (Riahi Madvar and Ayyoubzadeh, 2007) و تخمین ضریب پخشیدگی طولی در رودخانه در مسائل مهندسی رودخانه، محیط زیست، ارزیابی خطر تزریق آلاینده‌های خطرناک به درون رودخانه، طراحی آبگیر، و ... اهمیتی ویژه دارد (Deng et al., 2001). بررسی دقیق تغییرات غلظت آلاینده در رودخانه‌های

\* نویسنده مسئول: [ssoltani@ardakan.ac.ir](mailto:ssoltani@ardakan.ac.ir)

روش‌های تجربی، با دقت بالا پیش‌بینی کردند. Azamathulla and Wu (2011)، با استفاده از تکنیک SVM، ضریب پخشیدگی آلودگی را در رودخانه‌ها، با استفاده از داده‌های موجود در منابع مختلف، تخمین زدند. نتایج تحقیق آن‌ها دقت بالای این روش را در تخمین ضریب پخشیدگی آلودگی، در مقایسه با روش تجربی Rajeev and Dutta (2009)، نشان داد. Toprak *et al.* (2013)، با استفاده از داده‌های موجود در منابع مختلف، ضریب پخشیدگی طولی آلودگی را در آبراهه‌های طبیعی با استفاده از چند روش هوش مصنوعی (فازی و شبکه عصبی) برآورد کردند و نشان دادند مدل شبکه عصبی و سپس مدل فازی نسبت به روش‌های تجربی کارایی بهتری دارد. Rajeev (2013) از روش الگوریتم ژنتیک برای پیش‌بینی این ضریب بهره جست و نشان داد این روش نسبت به روش‌های تجربی نتیجه بهتری ارائه می‌دهد.

هدف اصلی این تحقیق بررسی روش‌ها و معادلات تجربی مختلف برای تعیین ضریب پراکندگی طولی آلودگی در رودخانه‌های طبیعی و ارزیابی دقت این روش‌ها در مقایسه با داده‌های اندازه‌گیری شده واقعی و بررسی روش‌های هوش مصنوعی در این زمینه است. در قدم بعدی، استفاده از برنامه‌ریزی ژنتیک، شبکه عصبی مصنوعی، شبکه عصبی-فازی، و رگرسیون چندمتغیره به منزله روش‌های داده‌کاوی در تخمین ضریب پخشیدگی طولی آلودگی ارائه و در پایان نتایج روابط تجربی و مدل‌ها با یکدیگر مقایسه می‌شود.

## مواد و روش‌ها

### مجموعه داده‌های استفاده شده

برآورد ضریب پراکندگی طولی با استفاده از روابط تجربی نیازمند مجموعه‌ای از داده‌های هیدرولیکی و هندسی رودخانه است. در این تحقیق از داده‌های یادشده در مراجع، که در آن‌ها مقادیر همه پارامترهای مورد نیاز روابط تجربی ارائه شده، به منظور ارزیابی دقت روابط تجربی و همچنین توسعه مدل و نیز ارزیابی دقت آن استفاده شد. به این منظور مجموعه داده‌های Seo and Cheong (2002) Kashefipour and Falconer (1998) به کار رفت. این داده‌ها را محققان در بیش از سی رودخانه در آمریکا در زمان‌های مختلف (از ۱۹۶۸ تا ۱۹۹۸) اندازه گرفته‌اند. مجموعه داده‌ها، شامل عمق جریان (متر)، عرض رودخانه (متر)، سرعت متوسط جریان (متر بر ثانیه)، سرعت برشی جریان (متر بر ثانیه)، و ضریب پخشیدگی طولی (متر مربع بر ثانیه) است. در جدول ۱، متغیرهای مورد نیاز روابط تجربی و پارامترهای استفاده شده و محدوده تغییرات آن‌ها

طبیعی، با استفاده از مدل‌ها، نیازمند انتخاب و تخمین درست ضریب پراکندگی طولی آلودگی است. در صورتی که داده‌های واقعی در دسترس باشد، این ضریب به سادگی قابل تعیین است. ولی در رودخانه‌هایی که ویژگی‌های پراکندگی آلودگی نامشخص باشد روابط تجربی کاربرد دارد (Kashefipour and Falconer, 2002). برای محاسبه ضرایب پخشیدگی، معادلات گوناگونی با استفاده از روش‌های تجربی، تحلیلی، و ریاضی ارائه شده است. Deng *et al.* (2001) با استفاده از روش حل عددی و به کمک معادله حاکم انتقال-انتشار، مقدار ضریب انتقال آلودگی را پیش‌بینی کردند. روش‌های تحلیلی و ریاضی، به علت پیچیدگی محاسبات، و روش‌های تجربی، به علت خطای زیاد، مورد توجه قرار نگرفته‌اند. در این زمینه محققان روابطی را ارائه کرده‌اند که به پارامترهای هیدرولیکی رودخانه-مانند سرعت جریان، سرعت برشی جریان رودخانه، عمق جریان، عرض رودخانه-وابسته است. هر یک از این روابط در دامنه‌ای از شرایط هیدرولیکی جریان، که واسنجی شده‌اند، معتبرند و برای مقادیر بزرگ‌تر یا کوچک‌تر از این دامنه نتایج درستی ارائه نمی‌کنند (Azamathulla and Wu, 2011). بنابراین، مهندسین هیدرولیک روش‌های هوش مصنوعی را پیشنهاد کردند. Riahi-Madvar and Ayyoubzadeh (2007) ضریب پخشیدگی آلودگی را به روش سیستم استنتاج فازی-عصبی انطباقی تخمین زدند. آن‌ها نشان دادند با استفاده از این روش دقت تخمین این ضریب به میزانی قابل توجه، نسبت به روش‌های تجربی، بهبود می‌یابد. Tayfur and Singh (2005) و Toprak and Cigizoglu (2008) با استفاده از روش شبکه عصبی مصنوعی ضریب پخشیدگی طولی آلودگی را تخمین زدند و اعلام کردند دقت این روش نسبت به روش‌های تجربی بهتر است. Toprak and Savci (2007) از تکنیک فازی به منظور پیش‌بینی ضریب پخشیدگی طولی آلودگی بهره جستند. Riahi-Madvar *et al.* (2009) با استفاده از ترکیب شبکه عصبی و منطق فازی ضریب آلودگی طولی را تخمین زدند و این مدل را نسبت به مدل‌های تجربی آن مناسب‌تر دانستند. Tayfur (2009) با استفاده از الگوریتم ژنتیک این ضریب را در رودخانه‌ها برآورد و دقت آن را نسبت به روش‌های تجربی ارزیابی کرد. Adarsh (2010) روش SVM و برنامه‌ریزی ژنتیک و شبکه عصبی را به منظور تخمین ضریب پخشیدگی طولی آلودگی در آبراهه‌های طبیعی به کار برد و نشان داد روش برنامه‌ریزی ژنتیک نسبت به سایر روش‌های تجربی خطای کمتری دارد. Azamathulla and Ghani (2010)، با استفاده از روش برنامه‌ریزی ژنتیک، ضریب پخشیدگی طولی آلودگی را در آبراهه‌های طبیعی، نسبت به

داد (Afzalimehr, 2011):

$$\frac{\varepsilon}{u^* h} = f_2 \left( \frac{u}{u^*}, \frac{w}{h}, Re^* \right) \quad (\text{رابطه } ۲)$$

$\frac{\varepsilon}{u^* h}$  ضریب اختلاط یا پخشیدگی طولی بدون بعد و  $Re^*$  عدد رینولدز برشی است. در صورتی که جریان آشفته زبر باشد ( $Re^* \geq 70$ )، می توان از تأثیر لزوجت سیال صرف نظر و رابطه ۲ را به صورت رابطه ۳ خلاصه کرد (Afzalimehr, 2011):

$$\frac{\varepsilon}{u^* h} = f_3 \left( \frac{u}{u^*}, \frac{w}{h} \right) \quad (\text{رابطه } ۳)$$

به همین دلیل روابط تجربی موجود از پارامترهای سرعت بدون بعد و عمق جریان بدون بعد برای تخمین این ضریب استفاده می کنند.

#### روابط تجربی

محققان، به منظور برآورد ضریب پخشیدگی طولی در رودخانه ها، روابط تجربی مختلفی ارائه کردند که مهم ترین آن ها در جدول ۲ می آید. همان طور که مشاهده می شود، همه این روابط ضریب پخشیدگی طولی را با استفاده از متغیرهای مربوط به شرایط متوسط جریان در مقطع رودخانه محاسبه می کنند. این متغیرها شامل متوسط عمق جریان در مقطع ( $h$ )، سرعت جریان ( $u$ )، سرعت برشی جریان ( $u^*$ )، و عرض سطح آب ( $w$ ) هستند.

می آید. از شصت و پنج مجموعه داده موجود ۷۰ درصد (۴۸ مورد) برای آموزش مدل و ۳۰ درصد (۱۷ مورد) برای ارزیابی مدل ها استفاده شد. شایان ذکر است تقسیم بندی داده ها به دو گروه آموزش و ارزیابی به صورت کاملاً تصادفی انجام شد.

جدول ۱. محدوده داده های استفاده شده

پارامتر	محدوده تغییرات	متوسط
سرعت جریان (متر بر ثانیه)	۰٫۱۳-۱٫۷۴	۰٫۵۴۷
عمق جریان (متر)	۰٫۲-۱۹٫۹۴	۱٫۶۱
عرض جریان (متر)	۱۱٫۸۹-۷۱۱٫۲	۷۷٫۸۸
سرعت برشی جریان (متر بر ثانیه)	۰٫۰۲-۰٫۵۵۳	۰٫۰۸۴
ضریب پخشیدگی طولی (مترمربع بر ثانیه)	۱٫۹-۱۴۸۶٫۵	۱۱۸٫۰۱

#### مدل های استفاده شده در تحقیق

استفاده از روش آنالیز ابعادی و ارائه متغیرهای بدون بعد امکان بررسی دقیق تر عوامل مؤثر بر روابط حاکم بر ضریب پخشیدگی طولی را فراهم می کند. مهم ترین پارامترهایی که در پدیده پخشیدگی ( $\varepsilon$ ) مؤثرند سه دسته مشخصات سیال (چگالی سیال،  $\rho$ ، ضریب لزوجت سیال،  $\mu$ )، مشخصات هیدرولیکی جریان (سرعت متوسط جریان،  $u$ ، عمق جریان،  $h$ ، سرعت برشی جریان،  $u^*$ )، و پارامترهای هندسی سطح مقطع جریان (عرض رودخانه،  $w$ ، شکل مسیر جریان، و ...) هستند. به عبارت دیگر:

$$\varepsilon = f_1(u, h, w, u^*, \rho, \mu) \quad (\text{رابطه } ۱)$$

با به کارگیری قضیه باکینگهام می توان رابطه ۲ را به دست

جدول ۲. برخی روابط تجربی مهم ارائه شده برای تخمین ضریب پخشیدگی طولی

شماره رابطه	معادله	محقق (سال)
(۴)	$K_x = 0.58 \left( \frac{h}{u^*} \right)^2 u w$	McQuivey and Keefer (1974)
(۵)	$K_x = 0.11 \frac{u^2 w^2}{h u^*}$	Fisher et al. (1979)
(۶)	$K_x = 5.92 \left( \frac{u}{u^*} \right)^{1.42} \left( \frac{w}{h} \right)^{-0.62} h u^*$	Seo and Cheong (1998)
(۷)	$K_x = 0.2 \left( \frac{w}{h} \right)^{1.2} \left( \frac{u}{u^*} \right)^{1.2} h u^*$	Li et al. (1998)
(۸)	$K_x = 10.612 \left( \frac{u}{u^*} \right) h u$	Kashefipour and Falconer (2002)
(۹)	$K_x = 7.428 + 1.775 \left( \frac{w}{h} \right)^{-0.62} \left( \frac{u}{u^*} \right)^{1.572} h u$	Tavakollizadeh and Kashefipour (2007)
(۱۰)	$K_x = 2 \left( \frac{w}{h} \right)^{-0.96} \left( \frac{u}{u^*} \right)^{1.25} h u^*$	Rajeev and Dutta (2009)

## برنامه‌ریزی ژنتیک

برنامه‌ریزی ژنتیک، که اولین بار Koza (1992) آن را ارائه کرد، جزء روش‌های الگوریتم گردشی محسوب می‌شود که مبنای همه آن‌ها نظریه تکاملی داروین است. الگوریتم‌های یادشده یک تابع هدف را در قالب معیارهای کیفی تعریف و سپس تابع یادشده را برای اندازه‌گیری و مقایسه روش‌های مختلف حل می‌کنند و در فرایندی گام‌به‌گام ساختار داده‌ها را تصحیح و در نهایت روش حل مناسب را ارائه می‌کنند. برنامه‌ریزی ژنتیک جدیدترین شیوه از بین روش‌های الگوریتم گردشی است که به دلیل داشتن دقت کافی به مثابه یک روش کاربردی مطرح می‌شود (Alvisi et al., 2005). در برنامه‌ریزی ژنتیک، ابتدا بلوک‌های موجود، که شامل متغیرهای ورودی و هدف و نیز تابع ارتباط‌دهنده آن‌هاست، تعریف و سپس ساختار مناسب مدل و ضرایب آن تعیین می‌شود. این روش شامل یک معادله ارتباط‌دهنده بین متغیرهای ورودی و خروجی است. بنابراین، قادر به انتخاب خودکار متغیرهای مناسب مدل و حذف متغیرهای غیر مرتبط است؛ که سبب کاهش ابعاد متغیرهای ورودی خواهد شد. این موضوع در شرایطی که از داده‌های ورودی ثانویه نیز بهره برده می‌شود اهمیتی مضاعف می‌یابد. زیرا ارائه داده‌های ورودی غیر مرتبط سبب کاهش دقت مدل و ایجاد مدل‌های پیچیده‌تری می‌شود که تفسیر آن‌ها با دشواری بیشتری مواجه است. در کاربردهای مهندسی، از برنامه‌ریزی ژنتیک به طور وسیع در مدل‌سازی تعیین ساختار پدیده‌ها استفاده می‌شود.

فرایند گام‌به‌گام برنامه‌ریزی ژنتیک عبارت است از: ۱. در نظر گرفتن یک جمعیت اولیه از توابع مرکب نشان‌دهنده مدل‌های پیش‌بینی، به صورت تصادفی (ایجاد کروموزوم‌ها)؛ ۲. معرفی جمعیت اولیه (کروموزوم‌ها) به رایانه و ارزیابی هر یک از افراد (ژن) جمعیت مذکور با استفاده از توابع برازش (شناسایی مؤثرترین افراد در ماهیت پدیده)؛ ۳. انتخاب ژن‌های مؤثر به منظور تکثیر، جهش، جفت‌گیری، و تولید مثل افراد جدید با صفات اصلاح‌شده (فرزندان)؛ ۴. اعمال فرایند توسعه‌ای تکراری بر فرزندان در هر تولید. گام چهارم به تعداد معین یا تا حصول بهترین پاسخ تکرار خواهد شد (Borelli et al., 2006; Aytek and Kisi, 2008)

## شبکه عصبی مصنوعی

شبکه عصبی مصنوعی یک روش شبیه‌سازی است که از مطالعه سیستم مغز و شبکه عصبی موجودات زنده الهام گرفته است. یکی از شبکه‌های عصبی متداول شبکه عصبی پرسپترون

چندلایه است. این شبکه از یک لایه ورودی، یک لایه خروجی، و یک یا چند لایه بین آن‌ها، که مستقیم به داده‌های ورودی و نتایج خروجی متصل نیستند، تشکیل می‌شود. واحدهای لایه ورودی وظیفه توزیع مقادیر ورودی‌ها به لایه بعد و لایه خروجی نیز وظیفه توزیع مقادیر را به لایه بعد دارد. در لایه ورودی و خروجی شبکه عصبی مصنوعی، تعداد نورون‌ها برابر با تعداد ورودی‌ها و خروجی‌هاست و لایه یا لایه‌های پنهان وظیفه ارتباط لایه‌های ورودی به لایه‌های خروجی را دارند. در حقیقت، لایه پنهان ارتباط بین دو لایه ورودی و خروجی را بر عهده دارد (Minasny et al., 1999).

مسئله اصلی در این شبکه‌ها تعیین تعداد لایه‌های پنهان و تعداد نورون‌های آن‌هاست. در این زمینه نظریات متفاوتی وجود دارد. شبکه عصبی یک لایه مخفی با تابع سیگموئید در لایه میانی و تابع خطی در لایه خروجی دارد که قادر به تقریب همه توابع مورد نظر با هر درجه تقریب است؛ مشروط بر اینکه نورون به اندازه کافی در لایه مخفی وجود داشته باشد. این قضیه به تقریب‌ساز جهانی معروف است (Menhaj, 2009). در این تحقیق شبکه‌ها با یک لایه مخفی، که دارای تابع فعال‌سازی سیگموئید در لایه مخفی و تابع فعال‌سازی خطی در لایه خروجی‌اند و تعداد نورون‌های آن از دو تا ده متغیر است و بهترین تعداد نورون به صورت سعی و خطا تعیین می‌شود، در نظر گرفته شد. همچنین، به علت کارایی، سادگی، و سرعت بالا در این تحقیق الگوریتم آموزشی لونیگ مارکوارت به کار رفت (Amini et al., 2005).

## شبکه عصبی-فازی

مدل ANFIS (سیستم فازی تطبیق‌پذیر مبتنی بر شبکه عصبی) را اولین بار ژانگ معرفی کرد (Jang, 1993). این مدل یک سیستم فازی سوگنو را در یک ساختار عصبی اجرا می‌کند و برای فرایند آموزش ترکیب روش‌های آموزش پسانتشار خطا و حداقل مربعات خطا را به کار می‌برد. مدل ANFIS یک شبکه چندلایه پیش‌رونده است که از الگوریتم‌های یادگیری شبکه‌های عصبی مصنوعی و استدلال‌های فازی برای توصیف یک فضا یا مکان ورودی به یک فضا یا مکان خروجی استفاده می‌کند (Firat and Gungor, 2007). تجربیات اخیر بیان‌کننده قدرت و توانایی ANFIS در مدل‌سازی فرایندهای متعدد و مختلف است. ANFIS دارای خصوصیتی از قبیل قدرت یادگیری، نوشتن و تدوین کردن، هزینه کردن، و طبقه‌بندی است. همچنین این حسن را دارد که اجازه استخراج قوانین فازی را از داده‌های متعدد و تخصصی می‌دهد و به گونه‌ای متناسب قوانین پایه و اساسی را می‌سازد. مشکل اصلی این

ضریب پخشیدگی طولی، دقت مناسبی ندارند و خطای آن‌ها قابل توجه است. رابطه Kashefipour and Falconer (2002) بهترین رابطه، دارای ضریب همبستگی ۰/۷۸، مجذور مربعات خطای ۹۸/۳۲ متر مربع در ثانیه، مجموع مربعات خطای  $۰/۴۶ \times ۱۰^۶$  و ضریب جرم باقی‌مانده ۰/۱۰۳، است. کارآیی مدل برای این روش تجربی ۰/۷۵ است که مقدار برابر با ۱ این پارامتر بر کارآمدی این مدل در تخمین ضریب پخشیدگی دلالت دارد. نکته شایان توجه در این روابط آن است که مطابق نتایج جدول ۳ و روابط جدول ۲ کمترین دقت را رابطه Quivey and Keefer (1974) (رابطه ۴) دارد؛ طوری که دارای ضریب همبستگی ۰/۲۴، مجذور مربعات خطای ۲۵۰۴۰۶ متر مربع در ثانیه، مجموع مربعات خطای  $۳ \times ۱۰^{۱۲}$  و ضریب جرم باقی‌مانده ۵۸۹۷- است. کارآیی مدل برای این روش تجربی ۱۵۵۶۷۹۸- است. مقدار زیاد و منفی این پارامتر بر ناکارآمدی این مدل در تخمین این ضریب دلالت دارد. همان‌طور که شکل ۱ و جدول ۲ نشان می‌دهد، از بین مدل‌های تجربی سه مدل اول (روابط ۴ تا ۶) مقداری بیشتر از مقدار اندازه‌گیری شده ارائه می‌دهد. به همین دلیل مقدار ضریب جرم باقی‌مانده (CRM) منفی است. مطابق جدول ۲، در این رابطه ضریب پخشیدگی طولی با توان دوم عمق جریان رابطه مستقیم دارد؛ در صورتی که این موضوع با فیزیک پدیده همخوانی ندارد. زیرا ضریب پخشیدگی طولی تابعی از توزیع عرضی متوسط سرعت عمقی است و هر چه عمق جریان کمتر باشد توزیع عرضی اهمیت بیشتری می‌یابد و پخشیدگی طولی، که حاصل اختلاف سرعت متوسط عمقی در عرض رودخانه با متوسط سرعت در کل مقطع است، بیشتر می‌شود (Riahi-Madvar and Ayyoubzadeh, 2007). همچنین ضریب پخشیدگی طولی با عمق رابطه معکوس دارد، نه مستقیم. نکته شایان توجه دیگر اینکه در روابطی که یکی از پارامترهای عرض یا سرعت متوسط جریان حذف می‌شود، به علت حذف مهم‌ترین پارامترهای مؤثر بر شدت تلاطم عرضی رودخانه، دقت روابط به میزان چشمگیری کاهش می‌یابد. نیز، نتایج این جدول نشان می‌دهد اثر سرعت جریان بر ضریب پخشیدگی طولی بیشتر از عرض جریان است. زیرا در رابطه Kashefipour and Falconer (2002) (رابطه ۸)، به‌رغم اینکه عرض جریان حذف شده، دقت آن در مقایسه با سایر روابط، که مقدار عرض جریان در آن‌ها لحاظ شده، بهتر است.

#### شبکه عصبی مصنوعی

نتایج شبکه عصبی نشان داد مدل با یک لایه مخفی و شش نورون در لایه مخفی بهترین ساختار را دارد. مقادیر میانگین

سیستم مدت زمان لازم جهت تهیه ساختار و نحوه آموزش سیستم و تعیین پارامترهای مؤثر است.

#### معیارهای ارزیابی مدل

به منظور ارزیابی کارایی مدل‌های استفاده‌شده (هم مدل‌های تجربی و هم مدل‌های هوش مصنوعی)، داده‌ها به دو بخش آموزش (۷۰٪) و اعتبارسنجی (۳۰٪) تقسیم شدند و از شاخص‌های آماری جذر میانگین مربعات خطا (RMSE)، کارآیی مدل (EF)، ضریب جرم باقی‌مانده (CRM)، مجموع مربعات خطا (SSE)، و ضریب تبیین ( $R^2$ ) مطابق روابط ۱۱ تا ۱۵ استفاده شد. کارآیی مدل مقادیر پیش‌بینی‌شده و متوسط مقادیر اندازه‌گیری‌شده را مقایسه می‌کند. مقدار منفی این پارامتر نشان‌دهنده آن است که متوسط مقادیر اندازه‌گیری‌شده نسبت به مقادیر پیش‌بینی‌شده تخمین مناسب‌تری دارد و بهترین حالت وقتی اتفاق می‌افتد که این پارامتر برابر با ۱ باشد.

$$RMSE = \sqrt{\frac{1}{n} \sum (Y_i - X_i)^2} \quad \text{(رابطه ۱۱)}$$

$$EF = 1 - \frac{\sum_{i=1}^n (Y_i - X_i)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad \text{(رابطه ۱۲)}$$

$$CRM = \frac{\sum_{i=1}^n (X_i - Y_i)}{n\bar{X}} \quad \text{(رابطه ۱۳)}$$

$$SSE = \sum_{i=1}^n (X_i - Y_i)^2 \quad \text{(رابطه ۱۴)}$$

$$R^2 = \frac{\sum (X_i - \bar{X})^2 (Y_i - \bar{Y})^2}{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2} \quad \text{(رابطه ۱۵)}$$

$Y_i$  و  $X_i$  به ترتیب مربوط به مقادیر اندازه‌گیری‌شده و برآوردشده است.  $\bar{Y}$  و  $\bar{X}$  به ترتیب میانگین مقادیر مشاهداتی و میانگین مقادیر برآوردی‌اند و  $n$  تعداد داده‌ها را نشان می‌دهد.

#### یافته‌ها و بحث

##### مدل‌های تجربی

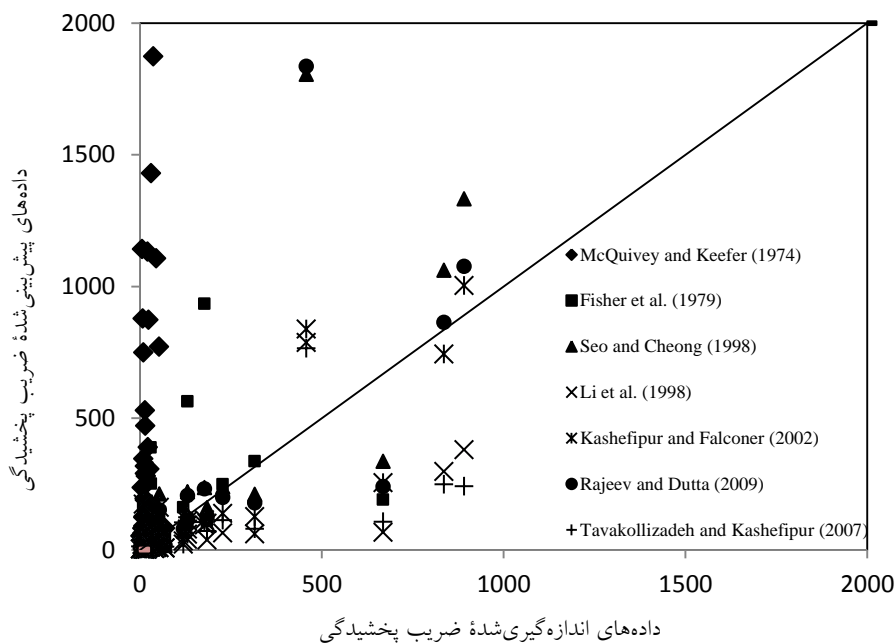
با استفاده از ۳۰ درصد مجموعه داده‌های گردآوری‌شده در این تحقیق، که به منظور آزمایش مدل‌ها به کار رفت، ضریب پخشیدگی طولی با استفاده از هر یک از روابط جدول ۲ محاسبه و دقت آن‌ها به کمک داده‌های اندازه‌گیری‌شده بررسی شد. در جدول ۳ نتایج روابط محققان می‌آید. مطابق نتایج، هیچ‌یک از روابط تجربی ارائه‌شده از سوی محققان، به منظور تخمین

معلوم نیست. علت نامشخص بودن روند تغییرات میانگین ریشه مربعات خطا تصادفی بودن اوزان اولیه است؛ به همین علت باید با سعی و خطا بهترین تعداد نورون لایه مخفی را به دست آورد.

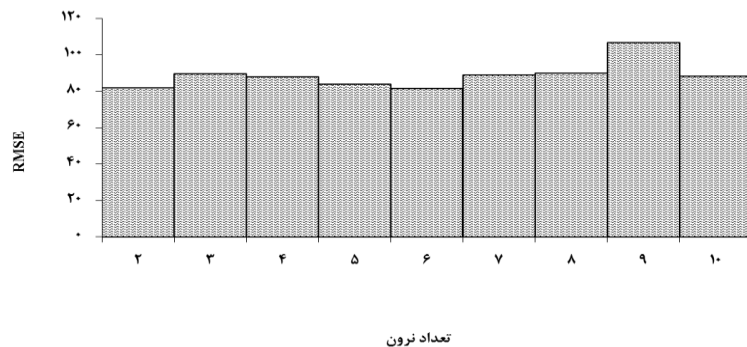
ریشه مربعات خطا و ضریب تبیین برآورد این پارامتر در بهترین ساختار شبکه عصبی مصنوعی، به ترتیب، برابر ۹۲/۱۱ و ۰/۷۱ به دست آمد. شکل ۲ نشان می‌دهد که تغییرات میانگین ریشه مربعات خطا در تعداد نورون مختلف دارای روندی مشخص و

جدول ۳. مشخصات آماری معادلات محققان برای تخمین ضریب پخشیدگی طولی

EF	SSE	CRM	RMSE	R <sup>2</sup>	معادله برآورد ضریب پخشیدگی طولی
-۱۵۵۶۷۹۸	$3 \times 10^{12}$	-۵۸۹,۷	۲۵۰۴۰۶	۰,۲۴	رابطه ۴. Quivey and Keefer (1974)
-۵۴,۸۷	$1.08 \times 10^6$	-۳,۳۶	۱۵۰۰,۱	۰,۳۴	رابطه ۵. Fischer <i>et al.</i> (1979)
-۰,۱۷	$2,26 \times 10^6$	-۰,۴۷	۲۱۷,۲۹	۰,۶۶	رابطه ۶. Seo and Cheong (1998)
۰,۳۹	$1,17 \times 10^6$	۰,۵۲۱	۱۵۶,۵۹	۰,۴۶	رابطه ۷. Li <i>et al.</i> (1998)
۰,۷۵	$۰,۴۶ \times 10^6$	۰,۱۰۳	۹۸,۳۲	۰,۷۸	رابطه ۸. Kashefipour and Falconer (2002)
-۰,۱	$2,18 \times 10^6$	۰,۲۹۸	۱۶۳,۵۱	۰,۳۷	رابطه ۹. Tavakollizadeh and Kashefipour (2007)
۰,۳۴	$1,28 \times 10^6$	۰,۳۸۷	۲۱۳,۳۹	۰,۵۶	رابطه ۱۰. Rajeev and Dutta (2009)

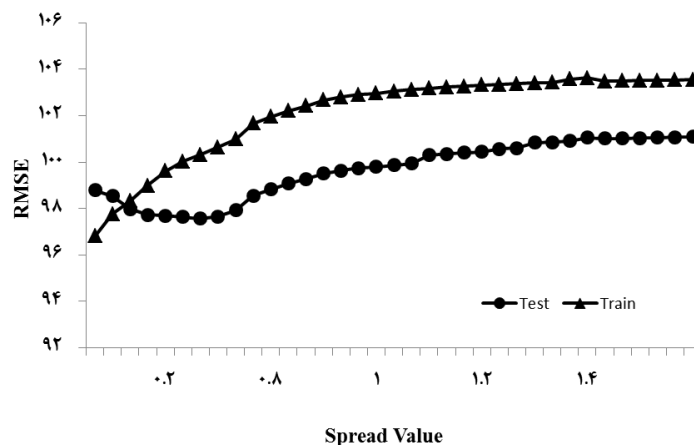


شکل ۱. داده‌های اندازه‌گیری شده در مقابل پیش‌بینی شده از روش‌های تجربی



شکل ۲. مقادیر RMSE برای تعداد نورون متفاوت در پیش‌بینی پارامتر ضریب پخشیدگی طولی رودخانه

همان‌طور که در شکل ۳ مشاهده می‌شود روند تغییرات RMSE در مقادیر متفاوت این شاخص برای داده‌های آموزش و آزمون متفاوت است. بنابراین، دو مقدار شاخص گستردگی برای ارائه انتخاب شد (Amini et al., 2005)؛ زمانی که این شاخص سبب ایجاد خطای حداقل در داده‌های آزمون می‌شود و نقطه‌ای که گراف‌های آموزش و آزمون همدیگر را قطع می‌کنند. با توجه به این دو رویکرد دو مدل پایه شعاعی توسعه داده شد. در مدل پایه شعاعی ۱ از شاخص گستردگی استفاده شد که در آن RMSE داده‌های آزمون حداقل است و در مدل پایه شعاعی ۲ از شاخص گستردگی استفاده شد که در آن گراف‌های آموزش و آزمون همدیگر را قطع می‌کنند. نتایج نشان داد آماره‌های RMSE و  $R^2$  برای مدل پایه شعاعی ۱ به ترتیب ۱۰۱/۱۱ و ۰/۶۵ و برای مدل پایه شعاعی ۲ به ترتیب ۱۰۴/۷۹ و ۰/۶۴ است. بنابراین، اختلاف معناداری بین این دو روش در برآورد ضریب پخشیدگی طولی رودخانه وجود ندارد.



شکل ۳. نمودار مقادیر RMSE برای داده‌های آموزش و آزمون

پس از بهینه کردن پارامترهای مورد نیاز و به دست آوردن معادله مورد نظر، اقدام به پیش‌بینی پارامتر مجهول در داده‌های آزمون شد. نتایج نشان داد آماره‌های RMSE و  $R^2$  به ترتیب ۸۴/۴۴ و ۰/۸۱ است.

(رابطه ۱۶)

$$\frac{\varepsilon}{u \cdot h} = 17 / 42 \frac{w}{h} + 116 / 8 \sin(17 / 42 \frac{w}{h}) + 35 / 61 \frac{u}{u} \sin(342 / 1 \frac{w}{h}) + \\ - / 88 \frac{w}{h} \sin(192 / 1 \frac{w}{h}) + 342 / 1 \sin(342 / 1 \frac{w}{h}) \sin(17 / 42 \frac{w}{h}) \cos(17 / 42 \frac{w}{h}) + \\ 116 / 8 \sin(17 / 42 \frac{w}{h}) + 35 / 61 \frac{u}{u} \sin(342 / 1 \frac{w}{h}) + 5 / 88 \frac{w}{h} \sin(192 / 1 \frac{w}{h}) - 0.7825 (\frac{w}{h})^2$$

سیستم فازی تطبیق‌پذیر مبتنی بر شبکه عصبی

در این ساختار مجموعه‌های فازی را می‌توان به عنوان وزن ارتباطات و متغیرهای ورودی و خروجی را به عنوان نورون‌ها

### شبکه عصبی مصنوعی آبخاری

برای اجرای این شبکه نیازی به بهینه‌سازی ساختار اولیه (تعداد نورون در لایه مخفی) نیست. این نکته از محاسن این شبکه است. نتایج خطای این روش بر اساس معیارهای RMSE و  $R^2$  به ترتیب ۹۸/۴۹ و ۰/۷۰ به دست آمد. نتایج نشان داد اختلاف معناداری بین روش شبکه عصبی مصنوعی آبخاری و شبکه عصبی مصنوعی پرسپترون وجود ندارد. البته یکی از برتری‌های استفاده از روش شبکه عصبی مصنوعی آبخاری این است که این روش به تعداد بهینه نورون ورودی نیاز ندارد (Shaban et al., 2009).

### شبکه عصبی پایه شعاعی

حسن این شبکه الگوریتم سریع یادگیری آن است؛ یعنی شبکه با یک دوره آموزش پیدا می‌کند. در این شبکه یگانه پارامتری که به بهینه شدن نیاز دارد شاخص گستردگی است. مقدار این شاخص حساسیت تابع به ورودی را نشان می‌دهد. در این مطالعه مقدار بهینه این شاخص به کمک شکل ۳ به دست آمد.

### برنامه‌ریزی ژنتیک

نتایج مدل برنامه‌ریزی ژنتیک در جدول ۵ می‌آید. همان‌طور که ملاحظه می‌شود، معیار توقف اجرای برنامه در مرحله ۱۹۰ است. میزان تولید جمعیت در هر یک از مراحل اجرای برنامه ۴۰۰، تعداد اجرای برنامه برای رسیدن به ضریب همبستگی مطلوب ۲۵۰، و حداکثر برنامه‌هایی که باید ارزیابی شود ۴۳۵ به دست آمد. همچنین، جمعیت اولیه مدل ۴۰۰، نرخ جهش ۹۵ درصد، و نرخ تلاقی ۲۰ درصد انتخاب شد و رابطه ۱۵ به دست آمد. انتخاب ورودی‌های مدل برنامه‌ریزی ژنتیک بر اساس آنالیز حساسیت انتخاب شد. در این مطالعه شش عملگر ریاضی شامل جمع، تفریق، ضرب، تقسیم، توان، رادیکال، و توابع مثلثاتی جهت پیش‌بینی ضریب پخشیدگی طولی رودخانه استفاده شد.

به‌رغم اینکه عرض جریان حذف شده، یگانه پارامتر ورودی آن سرعت بدون بعد جریان است. ولی دقت آن در مقایسه با سایر روابط تجربی، که مقدار عرض جریان در آن‌ها لحاظ شده، بهتر است. همچنین Rajeev (2013) نشان داد، بعد از پارامتر سینوسی بودن در رودخانه‌های مثاندر، سرعت جریان نسبت به عمق و عرض جریان اهمیت بیشتری دارد.

#### نتیجه‌گیری و ارزیابی مدل‌ها

تخمین درست ضریب پخشیدگی طولی آلودگی با توجه به اهمیت آن برای کنترل و مدیریت آلودگی در رودخانه‌ها اهمیتی به‌سزا دارد و به کاهش اختلاف نتایج حل تحلیلی و داده‌های آزمایشگاهی می‌انجامد. در این مطالعه سعی شد، با استفاده از مجموعه داده‌های گردآوری شده در منابع، ضریب پخشیدگی طولی از طریق چندین رابطه تجربی و روش‌های داده‌کاوی همچون برنامه‌ریزی ژنتیک، شبکه عصبی مصنوعی، و شبکه عصبی- فازی محاسبه و دقت آن‌ها با استفاده از داده‌های اندازه‌گیری شده بررسی شود. بر اساس نتایج، هیچ‌یک از مدل‌های تجربی برای تخمین ضریب پخشیدگی دقت مناسب ندارند. بین روش‌های تجربی، رابطه Kashefipour and Falconer (2002) بهترین رابطه با ضریب همبستگی ۰٫۷۸ و مجذور مربعات خطای ۹۸٫۳۲ متر مربع است. با توجه به اینکه اثر سرعت جریان بر ضریب پخشیدگی طولی بیشتر از عرض جریان است، در رابطه Kashefipour and Falconer (2002)، به‌رغم اینکه عرض جریان حذف شده، دقت آن در مقایسه با سایر روابط بهتر است. نتایج خطای مدل شبکه عصبی بر اساس معیارهای RMSE و  $R^2$ ، به ترتیب، ۹۲٫۱۱ و ۰٫۷۱ برای شبکه عصبی مصنوعی، ۹۸٫۴۹ و ۰٫۷ برای شبکه عصبی مصنوعی آبخاری، ۱۰۱٫۱۱ و ۰٫۶۵ برای مدل پایه شعاعی ۱، و ۱۰۴٫۷۹ و ۰٫۶۴، به ترتیب، برای مدل پایه شعاعی ۲ بود که اختلاف اندک بین مدل‌های شبکه عصبی مصنوعی را نشان می‌دهد. نتایج اعمال مدل برنامه‌ریزی ژنتیک بر داده‌های این تحقیق نشان داد آماره‌های RMSE و  $R^2$  به ترتیب ۸۴٫۴۴ و ۰٫۸۱ است. از محاسن این مدل نسبت به سایر روش‌ها ارائه رابطه‌ای ریاضی برای پیش‌بینی ضریب پخشیدگی طولی آلودگی است. Adarsh (2010) و Azamathulla and Ghani (2010) نیز استفاده از این مدل را برای پیش‌بینی ضریب پخشیدگی آلودگی در رودخانه‌ها پیشنهاد می‌کنند. با به‌کارگیری مدل ANFIS معیارهای ریشه مربعات خطا و ضریب تبیین به ترتیب ۷۲٫۲۱ و ۰٫۸۷ به دست آمد. نتایج نشان داد استفاده از روش نروفازی و سپس روش برنامه‌ریزی ژنتیک سبب بهبود نسبی پیش‌بینی برآورد ضریب

تفسیر کرد. الگوریتم یادگیری ساختارها، پارامترها، یا هر دوی آن‌ها را اصلاح می‌کند. برای ایجاد این شبکه باید چند پارامتر بهینه شوند که شامل نوع تابع عضویت، تعداد تابع، روش یادگیری، و تعداد تکرار است. برای این منظور همه حالات با سعی و خطا به دست آمد. نتایج خصوصیات مدل ANFIS به‌کاررفته برای پیش‌بینی ضریب پخشیدگی طولی رودخانه نشان داد سه عضویت و هفت قانون بهترین ساختار را می‌دهد. همچنین این ساختار در تکرار هزار بار روش یادگیری پس‌خور و روش دفازی کردن میانگین وزنی به دست آمد. نتایج معیارهای ریشه مربعات خطا و ضریب تبیین برای مدل ANFIS به ترتیب ۷۲٫۲۱ و ۰٫۸۷ به دست آمد. این نتایج نشان می‌دهد استفاده از روش نروفازی سبب بهبود نسبی پیش‌بینی برآورد ضریب پخشیدگی طولی رودخانه می‌شود.

جدول ۵. پارامترهای بهینه‌شده مدل برنامه‌ریزی ژنتیک

پارامتر	معیار مورد نظر	مقدار عددی
P <sub>1</sub>	تولید بدون بهبود	۱۹۰
P <sub>2</sub>	تولید از زمان شروع	۴۰۰
P <sub>3</sub>	حداکثر مقدار اجرا	۲۵۰
P <sub>4</sub>	حداکثر اندازه برنامه	۴۳۵
P <sub>5</sub>	اندازه جمعیت	۴۰۰
P <sub>6</sub>	سرعت جهش	۹۵
P <sub>7</sub>	سرعت تلاقی	۲۰

#### آنالیز حساسیت

آنالیز حساسیت بیانگر تأثیر پارامترهای مختلف ورودی مدل است و در آن مقادیر هر پارامتر شامل بازه‌ای از آن پارامتر است که بین حد پایین و بالای دامنه تغییرات آن پارامتر قرار می‌گیرد؛ به این صورت که یک پارامتر بین حد پایین و بالای خود تغییر می‌کند، در صورتی که پارامتر دیگر در میانگین خود ثابت باقی می‌ماند. نتایج آنالیز حساسیت نشان داد اهمیت پارامترهای بی‌بعد ورودی به صورت ۶۸٫۱۲٪ برای سرعت متوسط بی‌بعد ( $um^*$ ) و ۲۵٫۶۴٪ برای نسبت عرض جریان به عمق ( $w/h$ ) است. در نتیجه، پارامتر ( $um^*$ ) در تخمین ضریب پخشیدگی نسبت به عرض جریان بالاترین اهمیت را دارد. زیرا ضریب پخشیدگی طولی تابعی از توزیع عرضی متوسط سرعت عمقی است و هر چه عمق جریان کمتر باشد توزیع عرضی سرعت جریان اهمیت بیشتری می‌یابد و پخشیدگی طولی، که حاصل اختلاف سرعت متوسط عمقی در عرض رودخانه با متوسط سرعت در کل مقطع است، بیشتر می‌شود. در رابطه تجربی Kashefipour and Falconer (2002) (رابطه ۸) نیز،



خطای کمتر و کارایی بیشتر مدل‌های حاصل از روش‌های داده‌کاوی نسبت به مدل‌های تجربی و مدل‌های نروفازی و برنامه‌ریزی ژنتیک نسبت به مدل‌های شبکه عصبی است.

پخشیدگی طولی رودخانه می‌شود. جدول ۶ خلاصه نتایج به‌دست‌آمده از ارزیابی مدل‌های استفاده‌شده در این تحقیق را نشان می‌دهد. مقایسه نتایج جدول ۶ با جدول ۳ نشان‌دهنده

جدول ۶. ارزیابی مدل‌های استفاده‌شده جهت پیش‌بینی ضریب پخشیدگی طولی

EF	SSE	CRM	RMSE	R <sup>2</sup>	مدل برآورد ضریب پخشیدگی طولی
۰٫۳۲	۰٫۳۹×۱۰ <sup>۶</sup>	۰٫۰۷	۹۲٫۱۱	۰٫۷۱	شبکه عصبی مصنوعی
۰٫۳۶	۰٫۴۸×۱۰ <sup>۶</sup>	۰٫۰۹	۹۸٫۴۹	۰٫۷۰	شبکه عصبی آبخاری
۰٫۲۳	۰٫۵۱×۱۰ <sup>۶</sup>	۰٫۱۴	۱۰۱٫۱۱	۰٫۶۵	شبکه عصبی پایه شعاعی ۱
۰٫۲۵	۰٫۵۰×۱۰ <sup>۶</sup>	۰٫۱۴	۱۰۴٫۷۹	۰٫۶۴	شبکه عصبی پایه شعاعی ۲
۰٫۸۲	۰٫۳۱×۱۰ <sup>۶</sup>	۰٫۰۶	۸۴٫۴۴	۰٫۸۱	برنامه‌ریزی ژنتیک
۰٫۸۸	۰٫۲۲×۱۰ <sup>۶</sup>	۰٫۰۵	۷۲٫۲۱	۰٫۸۷	نروفازی (ANFIS)

## REFERENCES

- Abedie-Kupaie, J., Nasri, Z., and Maamanpoosh, A. (2007). Investigation of Chemical quality of Zayandehrood river. 6<sup>th</sup> Iran Hydraulic conference, Shahrekord University, 131-142. (In Farsi).
- Adarsh, S. (2010). Prediction of Longitudinal Dispersion Coefficient in Natural Channels Using Soft Computing Techniques, *Transaction A: Civil Engineering*, 17(5), 363-371.
- Afzalimehr, H. (2011). *Floud Mechanic Education*, Esfahan, Arkan. (In Farsi).
- Alvisi, S., Mascellani, G., Franchini, M., and Bardossy, A. (2005). Water level forecasting through fuzzy logic and artificial neural network approaches, *Hydrology and Earth System Sciences Discussion*, 2, 1107-1145.
- Amini, M., Abbaspour, K. C., Khademi, H., Fathianpour, N., Afyuni, M., and Schulin, R. (2005). Neural network models to predict cation exchange capacity in arid regions of Iran, *European Journal of Soil Science*, 53, 748-757.
- Aytek, A. and Kisi, O. (2008). A genetic programming approach to suspended sediment modeling. *Journal of Hydrology*, 351, 288-298.
- Azamathulla, H. M. and Ghani, A. A. (2010). Genetic Programming for Predicting Longitudinal Dispersion Coefficients in Streams. *Water Resource Management*, 1-8.
- Azamathulla, H. M. and Wu, F. C. (2011). Support vector machine approach for longitudinal dispersion coefficients in natural streams. *Applied Soft Computing*, 11(2), 2902-2905.
- Borelli, A., De Falco, I., Della, C. A., Nicodemi, M., and Trautteur, G. (2006). Performance of genetic programming to extract the trend in noisy data series. *Physica A*, 370, 104-108.
- Deng, Z. Q., Singh, V. P., and Bengtsson, L. (2001). Longitudinal dispersion coefficient in single channel streams. *Journal of Hydraulic Engineering*, 128(10), 901-916.
- Firat, M. and Gungor, M. (2007). River flow estimation using adaptive neuro fuzzy inference system. *Math & Comp in Simulation*, 75, 87-96.
- Fischer, H. B., List, E. J., Koh, R. C. Y., Imberger, J., and Brooks, N. H. (1979). *Mixing in inland and coastal waters*, Academic Press, Inc., San Diego, 483.
- Kashefipour, S. M. and Falconer, A. (2002). Longitudinal dispersion coefficients in Natural channels. *Water Research*, 36, 1596-1608.
- Koza, J. R. (1992). *Genetic Programming: on the programming of computers by means of natural selection*. The MIT Press, Cambridge, MA.
- Jang, J. S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference systems, *IEEE Trans Systems Man Cybernet* 23, 665-685.
- Li, Z. H., Huang, J., and Li, J. (1998). Preliminary study on longitudinal dispersion coefficient for the Gorges reservoir. *Proc. of the 7th International Symposium Environmental Hydraulics*, 16-18. December, Hong Kong, China.
- McQuivey, R. S. and Keefer, T. N. (1974). Simple method for predicting dispersion in streams, *Journal of Environmental Engineering, ASCE*, 100(4): 997-1011.
- Menhaj, M. (2009). *Principle of Neuron network and Artificial Intelligence*, Amir Kabir University. (In Farsi).
- Minasny, B., McBratney, A. B., and Bristow, K. L. (1999). Comparison of different approaches to the development of pedotransfer functions for waterretention curves. *Geoderma*, 93, 225-253.
- Purabadehie, M., Tokeldani, M., and Liyaghat, A. (2003). Investigation of flow parameters effect on Dispersion Coefficient of Pollutants in rectangular Chanel. 6<sup>th</sup> Iran Hydraulic conference, Shahrekord University, 29-38. (In Farsi).

- Rajeev, R. S. and Dutta, S. (2009). Prediction of longitudinal dispersion coefficients in natural rivers using genetic algorithm. *Journal of Hydrolic Research*, 40(6):544–552.
- Rajeev, R. S. (2013). Predicting longitudinal dispersion coefficients in sinuous rivers by genetic algorithm, *Journal of Hydrology Hydromechanics*, 61, 3, 214–221.
- Riahi-Madvar, H. and Ayyoubzadeh, S. A. (2007). Estimating Longitudinal Dispersion Coefficient of Pollutants Using daptive Neuro-Fuzzy Inference System, *Isfahan Journal of Water and Wastewater*, 64, 15–27. (In Farsi).
- Riahi-Madvar, H., Ayyoubzadeh, S. A., Khadangi, E., and Ebadzadeh, M. M. (2009). An expert system for predicting longitudinal dispersion coefficient in natural streams by using ANFIS. *Expert System with Applications*, 36(4): 8589–8596.
- Seo, I. W. and Cheong, T. S. (1998). Predicting longitudinal dispersion coefficient in natural stream. *Journal of Hydraulics Engineering*, 124 (1), 25-32.
- Shaban, K., El-Hag, A., and Matveev, A. (2009) A Cascade of Artificial Neural Networks to Predict Transformers Oil Parameters. *IEEE Transactions on Dielectrics and Electrical Insulation*. 16(2): 516-523.
- Tavakollizadeh, A. and Kashefipour, S. M. (2007). Effects of dispersion coefficient on quality modeling of surface waters. In: *Proceedings of the sixth international symposium river engineering*, 16–18.October, Ahwaz, Iran, pp 67–78. (In Farsi).
- Tayfur, G. (2009). Optimized model predicts longitudinal dispersion coefficient in natural channels, *Hydrology Research*, 40(1), 60-78.
- Tayfur, G. and Singh, V. P.(2005). Predicting longitudinal dispersion coefficient in natural streams by artificial neural network. *Journal of Hydraulic Engineering*, 131 (11), 991-1000.
- Toprak, Z. F. and Savci, M. E. (2007). Longitudinal Dispersion Coefficient Modeling in Natural Channels using Fuzzy Logic. *Clean* 35(6): 626–637.
- Toprak, Z. F. and Cigizoglu, H. K. (2008). Predicting longitudinal dispersion coefficient in natural streams by artificial intelligence methods. *Hydrology Process*, 22: 4106–4129.
- Toprak, Z. F., Hamidi, N., Kisi, O., and Gerger, R. (2013). Modeling Dimensionless Longitudinal Dispersion Coefficient in Natural Streams using Artificial Intelligence Methods. *KSCE Journal of Civil Engineering* . In press. DOI 10.1007/s12205-014-0089-y.