

تعیین پارامترهای موثر و ترکیب بهینه‌ی آن‌ها به منظور مدل‌سازی دبی رودخانه‌های سیلابی

عبدالرضا ظهیری^۱ - محمد قبائی سوق^۲ - ابوالفضل مساعدی^{۳*}

تاریخ دریافت: ۹۰/۵/۳۰

تاریخ پذیرش: ۹۰/۸/۱

چکیده

محاسبه‌ی دبی رودخانه‌ها در مقاطع مرکب، به علت تغییرات آبی و قابل توجه شرایط هندسی و هیدرولیکی مقطع اصلی و دشت‌های سیلابی بسیار پیچیده است. به همین دلیل پارامترهای متعددی در برآورد آن نقش دارند. تعیین پارامترهای مهم و انتخاب ترکیب بهینه از میان ترکیب‌های مختلف، یکی از مهم‌ترین مراحل ساخت و طراحی هرگونه مدل‌سازی ریاضی و هوشمند است. در این تحقیق با محاسبه ۱۰ پارامتر بدون بعد هندسی و هیدرولیکی از ۳۰ مقطع مرکب آزمایشگاهی و صحرایی (جمع‌آوری شده)، دبی جریان به کمک شبکه‌ی عصبی مصنوعی مبتنی بر روش پیش‌پردازش گاما تست محاسبه شده است. بر اساس نتایج این روش پارامترهای عمق نسبی، کوهرنس مقطع، شیب عرضی بستر دشت‌های سیلابی و عرض نسبی دشت سیلابی موثرترین پارامترها شناخته شدند. همچنین از میان ۱۰۲۳ ترکیب حاصل از ۱۰ پارامتر ورودی، ترکیب این چهار پارامتر به عنوان مدل بهینه تعیین گردید. بررسی کفایت داده‌ها با استفاده از M تست برای مدل بهینه نشان داد که تعداد داده‌های لازم برای ایجاد یک مدل هموار برابر ۳۰۰ مجموعه می‌باشد. دبی جریان در مقاطع مرکب با استفاده از مدل بهینه (مدل ۴ پارامتری) و همچنین مدل‌های بهینه‌ی ۱، ۲ و ۳ پارامتری مدل‌سازی شد. سپس با نتایج مدل‌های حاصل از تمامی ۱۰ پارامتر ورودی و همچنین روش تجزیه‌ی قائم مقطع مرکب Divided Channel Method (DCM) براساس آماره‌های مختلف آماری مقایسه گردیدند. نتایج نشان داد که مدل بهینه با IPE برابر با ۰/۱۲۳ از بالاترین دقت و روش DCM با IPE برابر با ۰/۸۶۷ از کم‌ترین دقت برخوردارند.

واژه‌های کلیدی: برآورد دبی، شبکه عصبی مصنوعی، گاما تست، مقاطع مرکب

مقدمه

آب از مقطع اصلی رودخانه فراتر رفته و وارد دشت‌های سیلابی شود، شرایط هندسی و هیدرولیکی رودخانه (عمق جریان، ضریب زبری و عرض) با تغییر آبی و قابل توجهی مواجه می‌شود. تغییرات عمق آب و ضریب زبری باعث می‌شود که سرعت در مقطع اصلی رودخانه خیلی بیشتر از دشت‌های سیلابی شود. این تفاوت زیاد، منجر به ایجاد تنش برشی قابل ملاحظه‌ای در ناحیه‌ی اتصال مقاطع اصلی و سیلابی شده که مهم‌ترین نتیجه‌ی آن، افت انرژی و کاهش قابل توجه دبی جریان در مقطع اصلی برای یک عمق معین است (۳ و ۳۵).

مطالعات آزمایشگاهی انجام شده در زمینه مقاطع مرکب نشان داده است که روش‌های متداول تعیین دبی جریان (تجزیه‌ی قائم مقطع مرکب به مقاطع جزئی) قابل اعتماد نیستند (۲۶).

مطالعات زیادی به منظور اصلاح روش‌های محاسبه جریان در مقاطع مرکب با مسیر مستقیم انجام شده است که مهم‌ترین آن‌ها عبارتند از: روش دوبعدی تحلیلی (۳۱)، روش فاکتور اصلاح ϕ (۳۵)، روش شبه‌دوبعدی تقسیم عرضی مقطع (۳۴)، روش نیمه‌تحلیلی کوهرنس (۳)، روش شبه‌دوبعدی مبتنی بر طول اختلاط پراتل (۲۱)، روابط رگرسیونی بدون بعد (۲۶)، روش تجزیه وزنی مقطع مرکب

رودخانه‌ها و به‌ویژه رودخانه‌های آبرفتی معمولاً شامل یک مجرای عمیق اصلی و یک یا دو دشت سیلابی عریض هستند که در اغلب اوقات سال خشک بوده و تنها در صورت وقوع سیل، آب در آن‌ها جریان می‌یابد. این گونه مقاطع را مقاطع مرکب می‌نامند. دشت‌های سیلابی معمولاً ضریب زبری بیشتری نسبت به مقطع اصلی رودخانه دارند. یکی از موارد مهم و مرسوم در مهندسی رودخانه، تعیین تراز سطح آب به ازاء دبی جریان سیلاب می‌باشد. محاسبه یا برآورد دبی جریان برای شرایط جریان عادی یک رودخانه، با استفاده از روابط مقاومت جریان از قبیل رابطه‌ی شزی و مانینگ به‌سادگی امکان‌پذیر است، اما در شرایط سیلاب، محاسبه‌ی دبی جریان در مقطع اصلی و دشت‌های سیلابی بسیار پیچیده است. وقتی تراز سطح

۱-۲ استادیار گروه مهندسی آب و دانش‌آموخته‌ی کارشناسی ارشد مهندسی منابع آب، دانشگاه علوم کشاورزی و منابع طبیعی گرگان

۳- دانشیار دانشکده منابع طبیعی و محیط زیست، دانشگاه فردوسی مشهد

(Email: mosaedi@um.ac.ir)

(* نویسنده مسئول)

مشخص می‌سازد در دسترس نباشد، برای کمک به ساخت مدلی هموار از رفتار سیستمی آن پدیده بر اساس سری داده‌های اندازه‌گیری شده کمک نماید (۱۵). گاما تست نخستین بار توسط کانسار (۱۱) به‌طور خلاصه معرفی و سپس در سال‌های بعد توسط محققان مختلف با جزئیات بیشتر مورد بررسی و استفاده قرار گرفت (۱ و ۱۵). در سال‌های اخیر نیز از این آزمون برای کاهش مراحل سعی و خطا در مدل‌سازی هوشمند و غیرخطی پدیده‌های گوناگون استفاده شده است: مقدم‌نیا و همکاران (۲۴) در برآورد تبخیر، احمدی و همکاران (۵) و مقدم‌نیا و همکاران (۲۵) در برآورد تابش خورشیدی، نوری و همکاران (۳۰) در برآورد تولید زباله‌های هفتگی، نوری و همکاران (۲۸) در برآورد غلظت مونواکسید کربن، قبائی‌سوق و همکاران (۱) در برآورد تبخیر و تعرق و نوری و همکاران (۲۹) در پیش‌بینی دبی جریان از این آزمون استفاده نموده‌اند.

با توجه به پیچیدگی محاسبه‌ی دبی جریان در مقاطع مرکب و تاثیر پارامترهای مختلف بر آن، تعیین پارامترهای مهم و انتخاب ترکیب بهینه از میان ترکیب‌های مختلف، مهم‌ترین مرحله در مدل‌سازی ریاضی و هوشمند برای این مقاطع است. هدف اصلی این تحقیق، تعیین پارامترهای موثر، انتخاب ترکیب بهینه و بررسی کیفیت داده‌ها به کمک روش پیش‌پردازش گاما تست برای مدل‌سازی دبی جریان مقاطع مرکب همگن و غیرهمگن با مسیر مستقیم و در شرایط آزمایشگاهی و رودخانه‌ای است. در نهایت با استفاده از نتایج این روش، دبی جریان به کمک شبکه‌ی عصبی مصنوعی برآورد شده و نتایج آن با نتایج روش متداول تجزیه‌ی قائم مقطع مرکب^۶ (DCM) با آماره‌های مختلف آماری مقایسه می‌گردد. همچنین با تعیین تعداد پارامترها در ترکیب بهینه، مدل‌سازی دبی در شرایط مختلف با کاهش پارامترهای ورودی نیز بررسی می‌گردد.

مواد و روش‌ها

داده‌های مورد استفاده

در این تحقیق، از ۳۷۱ سری داده‌ی هندسی و هیدرولیکی مقاطع مرکب در شرایط مختلف آزمایشگاهی و صحرایی استفاده شده است. در جدول ۱ محدوده‌ی تغییرات هندسی و هیدرولیکی این مقاطع ارائه شده است.

داده‌های آزمایشگاهی از نتایج محققین مختلفی مانند بلالوک و استورم (۷)، نایت و دیمتریو (۱۶)، نایت و سلین (۱۷)، لامبرت و سلین (۲۱)، میرز و لینس (۲۶)، لامبرت و میرز (۲۰)، بوسمار و زخ (۸)، هیدرا و والتاین (۱۴)، لای و بسیح (۱۹)، آتابای و نایت (۶) و بوسمار و همکاران (۹) جمع‌آوری شده است.

(۲۰)، روش تبادل دبی (۸)، روش شبه‌دوبعدی تحلیلی (۱۳) و روش اصلاحی تجزیه وزنی مقطع مرکب (۶).

طی دو دهه‌ی اخیر، استفاده از مدل‌های مبتنی بر روش‌های هوشمند به‌طور فزاینده‌ای در علوم مختلف گسترش یافته‌است. لیو و جیمز (۲۲) با استفاده از شبکه‌های عصبی، هیدرولیک جریان در مقاطع مرکب پیچانرود را در شرایط آزمایشگاهی بررسی و شبیه‌سازی نمودند. آن‌ها حداکثر خطا در برآورد دبی کل جریان را ۱۵ درصد گزارش نمودند. ظهیری و دهقانی (۳۶) با انتخاب ۱۳ پارامتر بدون بعد به‌عنوان ورودی در شبکه‌ی عصبی مصنوعی، دبی مقاطع مرکب آزمایشگاهی و رودخانه‌ای با مسیر مستقیم را شبیه‌سازی نمودند. نتایج تحلیل حساسیت پارامترهای ورودی نشان داد که پارامتر بدون بعد عمق نسبی (D_r) با $47/6\%$ درصد، مهم‌ترین پارامتر ورودی بوده و پارامترهای عرض نسبی (B_r) و زبری نسبی (n_r) در مرتبه‌ی بعدی اهمیت قرار دارند. نیک‌مهر و همکاران (۲) با استفاده از شبکه‌های عصبی و روش فازی، پروفیل عرضی سرعت در مقاطع مرکب آزمایشگاهی را شبیه‌سازی نمودند.

پیش‌پردازش پارامترهای ورودی جهت انتخاب ترکیب بهینه، یکی از فرایندهای پیچیده در مدل‌سازی سیستم‌های غیرخطی است. هم‌زمان با گسترش استفاده از روش‌های هوشمند جهت مدل‌سازی پدیده‌های گوناگون، به‌منظور سرعت بخشیدن به فرآیند مدل‌سازی و کاهش مراحل زیاد سعی و خطا برای مقایسه‌ی مدل‌های مختلف حاصل از ترکیب پارامترهای ورودی و انتخاب ترکیب بهینه، استفاده از روش‌های پیش‌پردازش پارامترهای ورودی نیز گسترش یافت. روش‌های پیش‌پردازش جهت پاسخ به سوالات مهمی (از جمله: کدام یک از پارامترهای ورودی از اهمیت بیشتری برخوردار است؟ ترکیب بهینه‌ی پارامترهای ورودی کدام است؟ چه تعداد داده برای مدل‌سازی مورد نیاز است؟) در مدل‌های هوشمند و غیرخطی به کار می‌روند. طی سال‌های اخیر، محققان به‌منظور کاهش مراحل سعی و خطا و پاسخ به برخی از پرسش‌های مطرح شده، روش‌های پیش‌پردازش مختلفی شامل رگرسیون گام‌به‌گام، گزینش پیشرو^۱، گاما تست^۲، معیارهای AIC ^۳ و BIC ^۴ و تحلیل مولفه‌های اصلی^۵ را با مدل‌های هوشمند تلفیق نموده‌اند.

گاما تست (GT) به‌عنوان یک روش پیش‌پردازش، آن بخش از واریانس داده‌های خروجی که نمی‌تواند با ایجاد یک مدل هموار از روی داده‌های ورودی محاسبه شوند را محاسبه می‌کند. این آزمون، می‌تواند در شرایطی که هیچ دانشی از معادلاتی که رفتار سیستم را

- 1- Forward Selection
- 2- Gamma Test
- 3- Akaike's information criterion
- 4- Bayesian information criterion
- 5- Principal Component Analysis

جدول ۱- محدوده تغییرات هندسی و هیدرولیکی مقاطع مورد استفاده

متغیر مورد استفاده	علامت	محدوده‌ی تغییرات
ارتفاع مقطع لبریز (m)	h	۰/۰۳۱ - ۶
عرض مقطع اصلی (m)	b	۰/۱۵۲ - ۲۱/۴
عرض دشت‌های سیلابی (m)	b _f	۰ - ۶۳
شیب جانبی مقطع اصلی	S _c	۰ - ۲
شیب جانبی دشت‌های سیلابی	S _f	۰ - ۷/۱
شیب عرضی بستر دشت‌های سیلابی	b _i	۰ - ۰/۰۶۲
ضریب زبری مانینگ مقطع اصلی	n _c	۰/۰۱ - ۰/۰۳۶
ضریب زبری دشت‌های سیلابی	n _f	۰/۰۱ - ۰/۰۵
شیب طولی	S ₀	۰/۰۰۰۱۹ - ۰/۰۰۵
عمق جریان (m)	H	۰/۰۳۶ - ۷/۸۱
دبی جریان (m ³ /s)	Q	۰/۰۰۳ - ۵۶۰

صورت $f_* = f_f / f_c$ و $P_* = P_f / P_c$ ، $A_* = A_f / A_c$ استفاده گردید. در این تحقیق با بررسی پارامترهای هندسی مقاطع مرکب، ۱۰ پارامتر $b_i, S_{fi}, S_c, S_0, S_f, n_{fi}, B_c, B_f, D_r$ و b_i و COH به عنوان پارامترهای ورودی و نسبت دبی کل مقطع به دبی مقطع پر (Q_t / Q_b) به عنوان پارامتر خروجی انتخاب گردید.

تعیین رابطه‌ی دبی - اشل مقاطع مرکب

روش متداول برای محاسبه‌ی دبی کل جریان در مقاطع مرکب، روش تجزیه‌ی قائم مقطع مرکب (DCM) است که در آن، مقطع مرکب به مقاطع جزئی (مقطع اصلی و دشت‌های سیلابی) تقسیم شده و با استفاده از رابطه‌ی مانینگ، دبی جریان در هر یک از این مقاطع جزئی محاسبه می‌شود. دبی کل جریان از مجموع این دبی‌های جزئی به‌دست می‌آید (۱۰).

$$Q = \sum_{i=1}^N Q_i = \sum_{i=1}^N \frac{A_i R_i^{2/3} S_0^{1/2}}{n_i} \quad (2)$$

در رابطه ۲، Q : دبی کل جریان، i : بیان کننده هر یک از مقاطع جزئی، N : تعداد کل مقاطع جزئی، A : سطح مقطع جریان، R : شعاع هیدرولیکی هر یک از مقاطع جزئی، S_0 : شیب طولی کانال و n : ضریب زبری مانینگ می‌باشند. تفکیک مقطع مرکب به مقاطع اصلی و دشت‌های سیلابی در شکل ۲ ارائه شده است.

روش فوق با وجود سادگی محاسبات، به دلیل عدم در نظر گرفتن تنش برشی بین مقطع اصلی و دشت‌های سیلابی، دبی کل جریان را به مقدار قابل توجهی بیشتر از مقدار واقعی محاسبه می‌کند. خطای این روش در مقاطع مرکب همگن آزمایشگاهی (زبری یکسان در مقطع اصلی و دشت‌های سیلابی) حدود ۱۰ درصد (۴) و در مقاطع مرکب غیرهمگن رودخانه‌ای (با دشت‌های سیلابی زبرتر از مقطع اصلی) حدود ۴۰ درصد گزارش شده است (۲۳).

داده‌های صحرائی مقاطع مرکب از مقادیر اندازه‌گیری شده عمق و دبی جریان رودخانه‌های سورن^۱ (۳) و مین^۲ (۲۳) در انگلستان و رودخانه‌ی ریو کلرادو^۳ در آرژانتین (۳۲) استخراج شده‌اند.

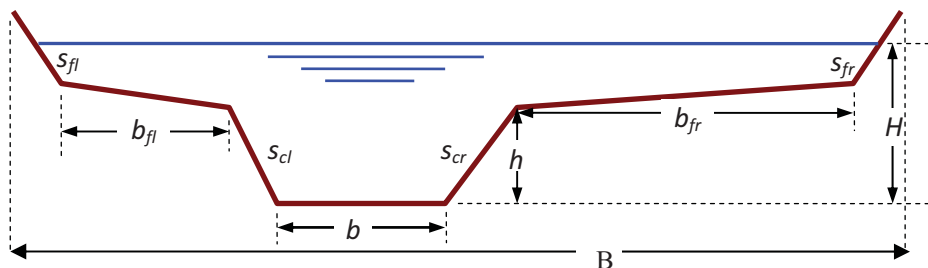
در شکل ۱ متغیرهای مورد استفاده در این تحقیق در یک مقطع مرکب نشان داده شده است. با استفاده از متغیرهای ارائه شده در جدول ۱ و شکل ۱، ۱۵ پارامتر بدون بعد به صورت عمق نسبی $D_r = \frac{H-h}{H}$ ، عرض نسبی دشت سیلابی $B_f = \frac{b_f}{B}$ ، عرض نسبی مقطع اصلی $B_c = \frac{b_c}{B}$ ، ضریب نسبی $n_f = \frac{n_f}{n_c}$ ، فاکتور شکل $S_f = \frac{b_c}{h}$ ، شیب طولی S_0 ، شیب مقطع اصلی S_c ، شیب جانبی دشت-های سیلابی S_{fi} ، شیب عرضی بستر دشت‌های سیلابی b_i و کوهپرنس مقطع COH برای هر یک از مقاطع مرکب تعریف شد. در این مقاطع ممکن است متغیرهای هندسی سمت چپ (l) و راست (r) مقاطع مرکب متفاوت باشند.

پارامتر COH، اغلب متغیرهای هندسی و هیدرولیکی یک مقطع مرکب را در بردارد. هرچه مقدار این پارامتر به ۱ نزدیکتر باشد بیانگر ساده بودن کانال و هرچه به صفر نزدیکتر باشد بیانگر مرکب بودن کانال می‌باشد و به صورت رابطه‌ی ۱ تعریف می‌شود (۳).

$$COH = \frac{(1 + A_*) \sqrt{[(1 + A_*) / (1 + f_* P_*)]}}{1 + A_* \sqrt{(A_* / f_* P_*)}} \quad (1)$$

در این رابطه، A : سطح مقطع، P : محیط مرطوب و f : ضریب اصطکاک دارسی - ویسباخ است. هم‌چنین از اندیس * برای بیان نسبت متغیرها در دشت سیلابی به مقطع اصلی به-

- 1- River Severn
- 2- River Main
- 3- Rio Colorado



شکل ۱- نمایش متغیرهای مورد استفاده در یک مقطع مرکب طبیعی



شکل ۲- مقطع مرکب و تجزیه آن به مقطع اصلی و دشت‌های سیلابی

رگرسیون خطی به صورت رابطه‌ی ۵ رسم می‌شود که ترسیم گرافیکی آن برای ترکیبی با کد رشته‌ی "۱۱۰۰۰۰۱۰۱" در شکل ۳ ارائه شده است.

$$\gamma = A\delta + \Gamma \quad (5)$$

عرض از مبدا خط رگرسیون ترسیم شده در شکل ۳، برابر آماره‌ی گاما (Γ) می‌باشد و مقدار آن برابر آن بخش از واریانس داده‌های خروجی است که نمی‌توانند توسط مدل برآورد گردند. همچنین شیب این خط نشان دهنده‌ی پیچیدگی مدلی است که از روی مجموعه داده‌های ورودی و خروجی ساخته می‌شود. هرچه این شیب تندتر باشد، پیچیدگی مدل نیز بیشتر است (۱۵ و ۳۳). علاوه بر دو آماره‌ی Γ و گرادیان (شیب خط رگرسیون، A)، دو معیار مهم دیگر را می‌توان با استفاده از گاما تست به دست آورد. نخست معیار بدون بعد $VRatio$ است که از نسبت Γ بر واریانس داده‌های خروجی‌های y ($\Gamma/\sigma^2(y)$) به دست می‌آید که هرچه مقدار آن به صفر نزدیک‌تر باشد مطلوب‌تر است. همچنین مقدار خطای استاندارد^۱ (SE) هرچه به عدد صفر نزدیک‌تر باشد با اطمینان بیشتری می‌توان مقدار Γ را نشان-دهنده واریانس اغتشاش در داده‌های خروجی دانست (۲۴). مقدار SE از رابطه ۶ به دست می‌آید.

$$SE(\Gamma) = \sqrt{\frac{1}{M-2} \sum_{k=1}^{P_{max}} (\Gamma(k) - \bar{\Gamma})^2} \quad (6)$$

روش پیش‌پردازش گاما تست

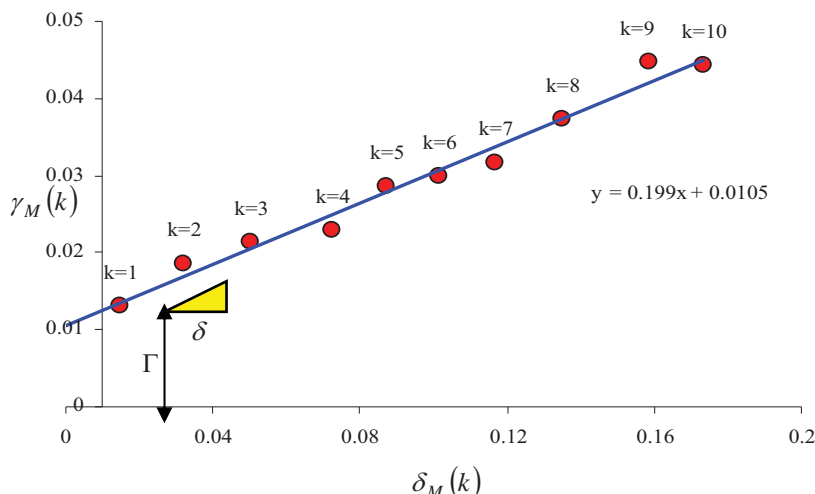
با فرض آن که مجموعه‌ای از داده‌های ورودی x_i و خروجی y_i مشاهده شده از پدیده‌ای به صورت رابطه‌ی $\{(x_i, y_i), 1 \leq i \leq M\}$ در اختیار باشد که با توجه به مجموعه ورودی x_i خروجی y_i به دست آید و M نشان‌دهنده‌ی تعداد سری مجموعه‌ی مشاهده شده از پدیده‌ی مورد بررسی باشد، آن‌گاه گاما تست رابطه‌ی حاکم بر سیستم را به صورت $y = f(x_1, x_2, \dots, x_M) + r$ در نظر می‌گیرد. در این رابطه با پذیرش خطا در به دست آوردن خروجی از روی ورودی‌ها، ارتباط بین مجموعه داده‌های ورودی و خروجی با تابع هموار f و متغیر تصادفی خطا r نشان داده می‌شود. با فرض آن که میانگین توزیعی که r از آن پیروی می‌کند، برابر صفر باشد و فرض این که واریانس خطا کران‌دار باشد، مدل فوق به مدلی با مشتقات جزئی درجه اول محدود می‌شود. گاما تست برابر آن بخش از واریانس داده‌های خروجی است که توسط مدل هموار برآورد نمی‌گردند. گاما تست بر اساس $N[i, k]$ می‌باشد که در آن: $1 \leq i \leq M$ ، $1 \leq k \leq p$ و p بیانگر تعداد نزدیک-ترین همسایگی‌ها می‌باشد. مقادیر $\delta_M(k)$ و $\gamma_M(k)$ به ترتیب برای مجموعه داده‌های ورودی و خروجی بر اساس روابط ۳ و ۴ به دست می‌آیند که $| \dots |$ بیانگر فاصله‌ی اقلیدسی است (۱۵ و ۳۳). مقدار $y_{N(i, k)}$ برابر مقدار y_i به ازای نزدیک‌ترین مقدار همسایگی x_i در رابطه‌ی ۳ می‌باشد.

$$\delta_M(k) = \frac{1}{M} \sum_{i=1}^M |x_{N(i, k)} - x_i| \quad 1 \leq k \leq p \quad (3)$$

$$\gamma_M(k) = \frac{1}{2M} \sum_{i=1}^M |y_{N(i, k)} - y_i| \quad 1 \leq k \leq p \quad (4)$$

سپس بین p مجموعه‌ی $(\gamma_M(k), \delta_M(k))$ ، یک رابطه‌ی

1- Standard Error



شکل ۳- ترسیم گرافیکی خط رگرسیون برای ترکیبی با کد رشته‌ی "۱۱۰۰۰۰۰۰۱۰۱"

تعیین گردید. سپس به منظور در نظر گرفتن تمامی آماره‌های مذکور در گزینش ترکیب بهینه، با استفاده از مفاهیم آماره IPE^۱ (۱۲)، آماره‌ی ترکیبی MIPE^۲ توسعه داده شد (رابطه‌ی ۷). آماره‌ی MIPE با وزن دادن و ترکیب آماره‌های Γ ، گرادیان، خطای استاندارد و VRatio امکان دخالت همه‌ی آن‌ها در انتخاب ترکیب بهینه را فراهم می‌نماید. در ادامه مقادیر MIPE برای تمامی ترکیب‌های ممکن محاسبه و بر اساس آن ترکیب بهینه نیز مشخص می‌گردد. مقدار آماره MIPE بین صفر و یک تغییر می‌کند و مقادیر آن هرچه به صفر نزدیکتر باشد، مدل مورد نظر از دقت بالاتری برخوردار است.

$$MIPE = \left\{ 0.25 \left[\left(\frac{SE_j}{\max SE_j} \right)^2 + \left(\frac{\Gamma_j}{\max \Gamma_j} \right)^2 + \left(\frac{Grad_j}{\max |Grad_j|} \right)^2 + \left(\frac{VRatio_j}{\max VRatio_j} \right)^2 \right] \right\}^{1/2} \quad (7)$$

در رابطه ۷ z بین یک تا n تغییر می‌کند که n برابر تعداد تمام ترکیب‌های ممکن حاصل از N پارامتر ورودی است (۱ - ۲^N).

تعیین تعداد الگوهای مورد نیاز با استفاده از M تست

یکی از چالش‌های اصلی در مدل‌سازی پدیده‌ها، تعیین تعداد الگوهای لازم برای ایجاد یک مدل هموار از پدیده‌ی مورد بررسی است. اگر چه به‌طور کلی با افزودن اطلاعات بیشتر عملکرد مدل در مرحله‌ی واسنجی بهبود می‌یابد، ولی باید عدم کفایت داده‌ها برای ایجاد یک مدل هموار نیز بررسی گردد. در روش گاما تست مقدار آماره‌های Γ ، گرادیان، خطای استاندارد و VRatio برای حداقل سه

با محاسبه‌ی مقدار گاما برای ترکیبات مختلف می‌توان ترکیب بهینه را از میان تمام ترکیب‌های ممکن استخراج، ترتیب اهمیت پارامترهای ورودی را تعیین و در ترکیب بهینه تعداد الگوهای لازم که منحنی‌های مجانب گاما و خطای استاندارد را به حالت پایدار برای ایجاد مدل هموار می‌رساند، تعیین نمود.

انتخاب ترکیب بهینه

اگر N پارامتر اولیه بر وقوع پدیده‌ی موثر باشند، تعداد ۱ - ۲^N ترکیب معنی‌دار از آن‌ها به‌وجود می‌آید که در مدل‌سازی، بررسی تک‌تک ترکیبات ایجاد شده برای یافتن بهترین ترکیب، کاری وقت‌گیر و دشوار است. گاما تست برای ترکیبات مختلف حاصل از پارامترهای ورودی مقادیر آماره‌های Γ ، گرادیان، خطای استاندارد و VRatio را محاسبه می‌نماید که هرچه مقادیر این آماره‌ها در مدلی مشخص کمتر باشد، نشانگر دقت بالای آن برای یافتن خروجی‌های مطلوب از روی ورودی‌ها است. یکی از چالش‌های مهم در این روش پیش‌پردازش، انتخاب آماره‌ی مناسب از میان آماره‌های ذکر شده برای تعیین ترکیب بهینه است. اگر چه ترکیبی که دارای کمترین مقدار این آماره‌ها باشد به‌عنوان ترکیب مناسب انتخاب خواهد شد، ولی در بیشتر مواقع ممکن است امکان انتخاب چنین ترکیب ایده‌آلی محقق نگردد. بررسی مقالات منتشر شده (۱، ۵، ۲۴، ۲۵، ۲۸، ۲۹ و ۳۰) نشان می‌دهد که محققان مقدار Γ را ملاک انتخاب ترکیب بهینه قرار دادند. مقدم‌نیا و همکاران (۲۴) در تحقیق خود با تاکید بر جدید بودن این تکنیک و اهمیت هر یک از این آماره‌ها در انتخاب ترکیب بهینه، تاکید نمودند که تحقیقات بیشتری پیرامون این که کدام آماره ملاک ارزیابی قرار گیرد، انجام شود. در این تحقیق ابتدا ترکیب بهینه بر اساس هر یک از آماره‌های Γ ، گرادیان، خطای استاندارد و VRatio

1- Ideal Point Error

2- Modified Ideal Point Error

در الگوریتم پس‌انتشار خطا^۱ (BP) توابع آموزشی مختلفی جهت بهینه نمودن وزن‌ها و اربها به کار می‌رود که در این تحقیق از توابع آموزشی گرادیان ادغامی (CG) استفاده شده‌است. الگوریتم معروف BP، وزن‌ها را در جهت منفی گرادیان تعدیل می‌کند. اگرچه تابع عملکرد به سرعت در جهت خلاف گرادیان کاهش می‌یابد، اما الزامی را برای همگرایی سریع ایجاد نمی‌کند. در الگوریتم‌های CG جستجو در خلال جهت‌های آمیخته و توأم که به‌طور کلی همگرایی سریعتری را نسبت به جهت‌های نزولی ایجاد می‌کنند، صورت می‌گیرد. تمامی الگوریتم‌های گرادیان ادغامی در اولین تکرار، با جستجو در جهت منفی گرادیان شروع می‌شوند.

$$P_0 = -g_0 \quad (۸)$$

که P_0 و g_0 بردارهای ورودی و گرادیان اولیه می‌باشند. آن‌گاه جستجو برای تعیین فاصله بهینه برای حرکت در امتداد جهت جستجوی جدید، شروع می‌شود.

$$x_{k+1} = x_k + a_k P_k \quad (۹)$$

که x_k و x_{k+1} بردار وزن‌های جاری و اربها در تکرارهای k و $k+1$ هم‌چنین، a_k و P_k نرخ یادگیری و بردار ورودی جاری می‌باشند. سپس جهت جستجوی بعدی برای ادغام با روش جستجوی قبلی تعیین می‌شود. روش کلی برای تعیین جهت جستجوی جدید، ترکیب جهت نزولی جدید با جهت جستجوی قبلی است.

$$P_k = -g_k + \beta_k P_{k-1} \quad (۱۰)$$

تفاوت الگوریتم‌های مختلف گرادیان ادغامی در شیوه تعیین و محاسبه مقدار β_k است که نسبت نرم حسابی گرادیان جدید به نرم حسابی گرادیان قبلی است (۳۰). هر یک از الگوریتم‌های گرادیان ادغامی نیازمند یک رشته جستجو در هر تکرار می‌باشند. الگوریتم کاهش شیب مقیاس شده (SCG)^۲ برای پرهیز از جستجوهای زمان-بر طراحی شده است که در این تحقیق از آن استفاده می‌شود.

آماره‌های ارزیابی نتایج

برای ارزیابی نتایج مدل‌های مختلف هوشمند و روش DCM از آماره‌ی IPE^۳ استفاده شده‌است. این آماره‌ی نسبتاً جدید با وزن دادن به آماره‌های میانگین خطای مطلق (MBE)^۴، ریشه میانگین مربعات خطا (RMSE)^۵، میانگین قدرمطلق خطای نسبی (MARE)^۶ و ضریب همبستگی R، از ترکیب آن‌ها مطابق رابطه‌ی ۱۵ محاسبه می‌شود و بهترین عملکرد بر اساس این آماره برابر صفر است (۱۲).

الگوی ورودی در ترکیب بهینه محاسبه می‌گردد سپس در هر مرحله یک الگو به مجموعه الگوهای ورودی اضافه و هر مرتبه مقدار آماره-های مذکور محاسبه می‌گردد. این عمل برای M الگوی ورودی در ترکیب بهینه انجام می‌گیرد. در نهایت مقادیر آماره‌های Γ و خطای استاندارد را در برابر تعداد الگوهای ورودی در یک نمودار مشخص می‌شوند که برای بررسی بهتر چگونگی تغییرات آماره‌های مذکور با افزایش الگوهای ورودی، نقاط ایجاد شده را با منحنی‌هایی به هم متصل می‌نمایند که این منحنی‌ها را منحنی‌های مجانب و نمودار حاصل را نمودار M تست می‌نامند. به عبارتی دیگر، این نمودار با ترسیم تغییرات منحنی‌های مجانب آماره‌های Γ و خطای استاندارد با افزایش الگوهای ورودی، این امکان را فراهم می‌نماید که تعداد الگوهای لازم که مقدار Γ و خطای استاندارد را تقریباً به حالت ثابت می‌رسانند تعیین و از آن‌ها برای آموزش مدل استفاده شود (۱۱، ۱۵ و ۳۳).

اهمیت پارامترهای ورودی

برای مشخص نمودن ترتیب اهمیت پارامترهای ورودی، ابتدا برای ترکیبی که همه‌ی پارامترهای ورودی در آن حضور دارند، مقدار Γ محاسبه می‌شود. سپس یکی از پارامترهای ورودی را به دلخواه از مجموعه اولیه حذف و مجدداً مقدار Γ محاسبه می‌شود. در ادامه، متغیر حذف شده را به مجموعه پارامترهای ورودی برگردانده و پارامتر دیگری را حذف و مقدار Γ محاسبه می‌شود. این فرآیند برای تمام پارامترهای ورودی به ترتیب انجام می‌گردد و هر بار مقدار Γ محاسبه می‌گردد. در این فرآیند با حذف پارامتر تاثیرگذار مقدار Γ افزایش می‌یابد و با حذف پارامتری که از اهمیت کمتری برخوردار است، مقدار آن کاهش می‌یابد (۲۹).

شبکه‌های عصبی مصنوعی

از دیدگاه ریاضی شبکه عصبی به‌عنوان یک تقریب‌گر است. توانایی آن در تقریب بین الگوهای یک مساله باعث می‌شود تا بتوان مسائل با پیچیدگی زیاد را انجام داد. به‌طور کلی هر شبکه عصبی مصنوعی از سه لایه‌ی ورودی، پنهان و خروجی تشکیل شده که در هر لایه تعدادی نرون در نظر گرفته می‌شود که به‌وسیله‌ی اتصالات به نرون‌های لایه‌های مجاور وصل می‌شوند. تعداد نرون‌های موجود در لایه‌های ورودی و خروجی با توجه به ماهیت مساله مشخص می‌شود و تعداد نرون‌های موجود در لایه‌های پنهان و هم‌چنین تعداد این لایه‌ها با سعی و خطا در جهت کاهش مقدار خطا توسط طراح مشخص می‌گردد (۲۴).

- 1- Back Propagation
- 2- scaled conjugate gradient
- 3- Ideal Point Error
- 4- Mean Bias Error
- 5- Root Mean Square Error
- 6- Mean Absolute Relative Error

نتایج انتخاب ترکیب بهینه

در این تحقیق با در نظر گرفتن ۱۰ عامل موثر بر نسبت (Q_i/Q_b) ، تعداد ۱۰۲۳ $(2^{10}-1)$ ترکیب مختلف ایجاد گردید. به منظور انتخاب ترکیب بهینه از میان آن‌ها، ابتدا با ملاک قرار دادن هر یک از آماره‌های Γ ، گرادیان، خطای استاندارد و VRatio، ترکیب بهینه انتخاب گردید. سپس با استفاده از آماره‌ی ترکیبی MIPE، ترکیبی را که دارای کمترین مقدار MIPE بود مشخص گردید. در این تحقیق برای نمایش ترکیب پارامترهای مختلف از کد رشته‌های اعداد شامل: "۱" و "۰" استفاده شده است که به ترتیب نشانگر حضور و عدم حضور یک پارامتر در ترکیب کد رشته‌ای است. برای مثال ترکیبی با کد رشته "۱۱۱۱۱۱۱۱۰" نشان می‌دهد که از میان ۱۰ پارامتر اولیه، همه پارامترهای ورودی به‌جز پارامتر اول مورد استفاده قرار گرفتند. اعداد کد رشته‌ها به ترتیب از راست به چپ نشانگر پارامترهای $D_f, B_f, D_c, B_c, n_f, S_f, S_0, S_c, S_{fl}, b_i$ و COH می‌باشند. نتایج انتخاب ترکیب بهینه بر اساس آماره Γ با نتایج انتخاب ترکیب بهینه بر اساس آماره‌های VRatio و خطای استاندارد یکسان و با آماره گرادیان متفاوت می‌باشد بنابراین نتایج حاصل از انتخاب ترکیب بهینه بر اساس آماره‌های Γ و گرادیان در جدول ۲ ارائه شده است. بر اساس آماره‌های Γ ، VRatio و خطای استاندارد ترکیب کد رشته‌ای ۱۱۰۰۰۰۰۱۰۱ با پارامترهای D_f, B_c, b_i و COH و بر اساس آماره گرادیان، ترکیب کد رشته‌ای ۰۱۰۱۰۱۱۰ با پارامترهای B_f, B_c, S_c, b_i به‌عنوان ترکیب‌های بهینه انتخاب می‌شوند. با توجه به متفاوت بودن نتایج آماره‌های ذکر شده و هم‌چنین اهمیت هر یک در انتخاب ترکیب بهینه، در ادامه ترکیب بهینه با استفاده از آماره‌ی ترکیبی MIPE نیز تعیین گردید که بر اساس آن، ترکیب کد رشته‌ای ۱۱۰۰۰۰۰۱۰۱ با MIPE برابر ۰/۰۲۴۱ به‌عنوان ترکیب بهینه برای مدل‌سازی نسبت Q_i/Q_b تعیین گردید که با نتایج انتخاب ترکیب بهینه بر اساس آماره‌های Γ ، VRatio و خطای استاندارد مشابه می‌باشند.

نتایج اهمیت پارامترهای ورودی

برای تعیین ترتیب اهمیت پارامترهای ورودی اثر حذف هر یک از ۱۰ پارامتر ورودی از ترکیبی که همه‌ی پارامترهای ورودی در آن حضور دارند (۱۱۱۱۱۱۱۱۱) طی ۱۰ مرحله انجام گردید و در هر مرحله مقدار Γ محاسبه و نسبت به مقدار Γ ترکیب بدون حذف پارامتر مقایسه گردید. نتایج اهمیت پارامترهای ورودی در جدول ۳ ارائه شده است. مطابق نتایج آن، با حذف پارامتر D_f مقدار Γ به ۰/۰۷۶۶ افزایش می‌یابد که نسبت به ترکیب بدون حذف پارامتر با Γ برابر ۰/۰۳۹۹ مقدار ۰/۰۳۶۷ افزایش داشته است و این افزایش نشان دهنده‌ی اهمیت زیاد این پارامتر است.

بنابراین از میان آماره‌های مختلف، ملاک ارزیابی آماره‌ی ترکیبی IPE می‌باشد. همچنین جهت اطلاع از چگونگی توزیع خطا در مدل‌های مختلف از نمودار تحلیل آستانه‌های خطا (TS) (رابطه‌ی ۱۶) استفاده شد. شاخص TS_x برای مقدار x درصد از پیش‌بینی‌ها، نشان‌دهنده‌ی توزیع خطا در مقادیر برآوردی هر مدل می‌باشد. این شاخص که بر حسب درصد تعریف می‌شود، برای مقادیر مختلف قدرمطلق خطای نسبی $\left(\frac{O_i - P_i}{O_i}\right)$ ارائه می‌شود. مقدار TS_x برای x درصد از پیش‌بینی‌ها مطابق رابطه‌ی ۱۶ به‌دست می‌آید، که در آن، Y_x تعداد دبی‌های سیلابی برآوردی (از کل تعداد M) برای هر مقدار مطلق خطای نسبی کمتر از x درصد باشد (۲۷).

$$MBE = \frac{1}{M} \sum_{i=1}^M (O_i - P_i) \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^M (O_i - P_i)^2}{M}} \quad (12)$$

$$MARE = \frac{1}{M} \sum_{i=1}^M \left| \frac{O_i - P_i}{O_i} \right| \quad (13)$$

$$R = \frac{\sum_{i=1}^M (O_i - \bar{O}_i)(P_i - \bar{P}_i)}{\sqrt{\sum_{i=1}^M (O_i - \bar{O}_i)^2 \sum_{i=1}^M (P_i - \bar{P}_i)^2}} \quad (14)$$

$$IPE = \left\{ 0.25 \left[\left(\frac{RMSE_{ij}}{\max RMSE_{ij}} \right)^2 + \left(\frac{MARE_{ij}}{\max MARE_{ij}} \right)^2 + \left(\frac{MBE_{ij}}{\max |MBE_{ij}|} \right)^2 + \left(\frac{R_{ij} - 1.0}{1/\max R_{ij}} \right)^2 \right] \right\}^{1/2} \quad (15)$$

$$TS_x = \frac{Y_x}{M} \cdot 100 \quad (16)$$

در روابط ۱۱ تا ۱۴ پارامترهای O_i, P_i به ترتیب مقادیر مشاهداتی و برآوردی و \bar{O}_i و \bar{P}_i نیز به ترتیب میانگین مقادیر مشاهداتی و برآوردی می‌باشند. در رابطه‌ی ۱۵ برای مدل‌های مختلف تعریف شده (۶ مدل)، به کار می‌رود که در ادامه معرفی شده‌اند.

نتایج و بحث

همان‌طور که بیان گردید با استفاده از روش پیش‌پردازش GT می‌توان بهترین ترکیب از میان تمام ترکیب‌های ممکن، ترتیب اهمیت پارامترهای ورودی و تعداد داده‌های مورد نیاز برای ایجاد یک مدل هموار از پدیده‌ی مورد بررسی را به‌دست آورد.

جدول ۲- نتایج انتخاب ترکیب بهینه بر اساس آماره‌های Γ و گرادیان

مدل بهینه	کد رشته	مقدار گاما	گرادیان	خطای استاندارد	VRatio	MIPE
بر اساس مقدار Γ	۱۱۰۰۰۰۰۱۰۱	۰/۰۱۰۵۴	۰/۱۹۸۷	۰/۰۰۱۵	۰/۰۴۲۱	۰/۰۲۴۱
بر اساس مقدار گرادیان	۰۱۰۱۰۱۰۱۱۰	۰/۲۰۳۲۸	۰/۰۰۱۲	۰/۰۲۴۲	۰/۸۱۳۱	۰/۷۴۷۹

جدول ۳- نتایج اهمیت پارامترهای ورودی بر اساس مقادیر Γ و MIPE

پارامتر حذف شده	کد رشته	مقادیر Γ	مقادیر MIPE
D_r	۱۱۱۱۱۱۱۱۱۰	۰/۰۷۶۶	۰/۲۰۵۸
B_f	۱۱۱۱۱۱۱۱۰۱	۰/۰۳۷۵	۰/۰۸۸۱
B_c	۱۱۱۱۱۱۱۰۱۱	۰/۰۳۷۸	۰/۰۸۹۰
n_r	۱۱۱۱۱۱۰۱۱۱	۰/۰۳۶۴	۰/۰۸۵۳
S_r	۱۱۱۱۱۰۱۱۱۱	۰/۰۳۶۳	۰/۰۸۴۵
S_0	۱۱۱۱۰۱۱۱۱۱	۰/۰۳۰۱	۰/۰۶۸۹
S_c	۱۱۱۰۱۱۱۱۱۱	۰/۰۳۱۵	۰/۰۷۲۹
S_{fl}	۱۱۰۱۱۱۱۱۱۱	۰/۰۳۹۶	۰/۰۹۳۷
b_i	۱۰۱۱۱۱۱۱۱۱	۰/۰۴۰۵	۰/۰۹۶۱
COH	۰۱۱۱۱۱۱۱۱۱	۰/۰۴۲۲	۰/۱۰۰۷
بدون حذف پارامتر	۱۱۱۱۱۱۱۱۱۱	۰/۰۳۹۹	۰/۰۹۴۶

شیب عرضی دشت‌های سیلابی تاکنون به صورت جدی مورد تحقیق و بررسی قرار نگرفته است. اما با توجه به مکانیسم جریان در مقاطع مرکب، هرچه این شیب به صفر نزدیکتر باشد (یعنی دشت سیلاب به صورت افقی باشد) مقطع مرکب دارای دبی کمتری خواهد بود یعنی اثر متقابل شدیدتر است. البته در رودخانه‌های طبیعی دشت‌های سیلابی به صورت کاملاً افقی نیستند و یک شیب جانبی بسیار کمی دارند.

نتایج تعیین الگوهای لازم برای آموزش مدل با M تست

مطابق آنچه در بخش مواد و روش‌ها بیان گردید منحنی‌های مجانب آماره‌های Γ و خطای استاندارد برای ترکیب بهینه (۱۱۰۰۰۰۰۱۰۱) با افزایش الگوهای ورودی در شکل ۴ رسم شده است. بر اساس این شکل، مقدار آماره‌های Γ و خطای استاندارد از نقطه‌ی ۳۰۰ به بعد تغییرات چندانی نمی‌کنند و منحنی‌های مجانب آن‌ها تقریباً به حالت ثابت می‌رسند. بنابراین الگوهای ورودی برای آموزش و صحت‌سنجی مدل بهینه‌ی مورد بررسی کفایت می‌کنند و از ۳۰۰ الگوی آن برای آموزش مدل و از ۷۱ الگوی باقی‌مانده برای ارزیابی نتایج در مرحله‌ی آزمون استفاده می‌شود.

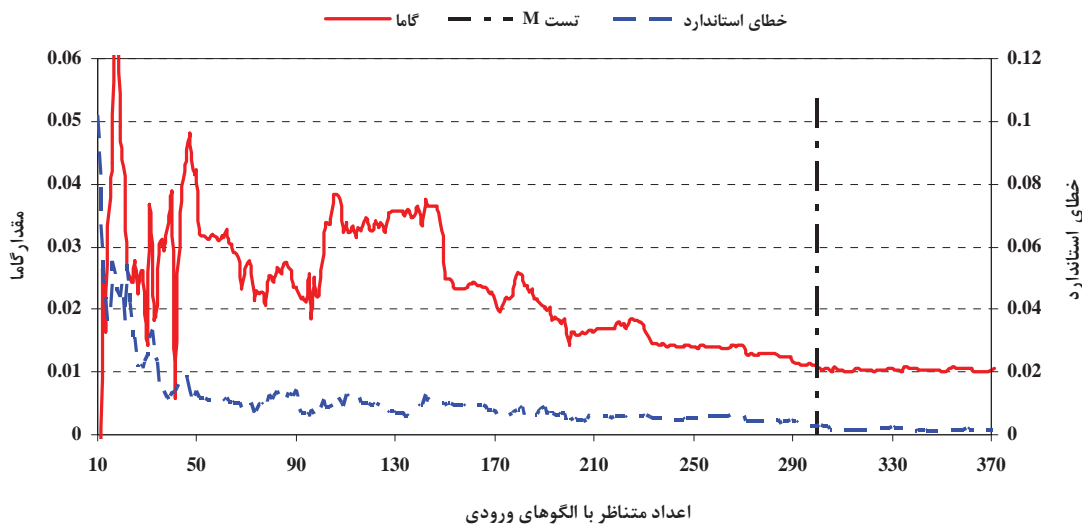
پارامترهای مهم دیگری که با حذف آن‌ها مقدار Γ از ۰/۰۳۹۹ به ۰/۰۴۲۲ و ۰/۰۴۰۵ افزایش می‌یابد، پارامترهای COH و b_i هستند که در مرتبه‌های دوم و سوم اهمیت قرار دارند. بر اساس مقدار آماره‌ی Γ پارامترهای S_{fl} ، B_f ، B_c ، n_r ، S_r ، S_c و S_0 به ترتیب در مرتبه‌های بعدی اهمیت قرار دارند. همچنین مطابق آنچه برای Γ ذکر گردید، اهمیت پارامترهای ورودی با آماره‌ی MIPE نیز ارزیابی گردید که مشابه آزمون Γ به دست آمد.

نتایج مدل‌های بهینه با افزایش تعداد پارامترهای ورودی در جدول ۴ ارائه شده است. مطابق نتایج آن مشاهده می‌شود که در تمامی مدل‌های بهینه پارامترهای D_r و COH حضور دارند.

پارامتر عمق نسبی در مقاطع مرکب دارای اهمیت زیادی است و دقیقاً نقش عمق جریان در مقاطع ساده را ایفا می‌کند و متغیر مستقل اصلی در رابطه دبی- اشل است که هم در مقاطع ساده و هم مقاطع مرکب ظاهر می‌شود. کوهپرنس مقطع یک پارامتر بدون بعد است که میزان اثر متقابل بین مقطع اصلی و دشت سیلابی را بیان می‌کند. این پارامتر مشخصات هندسی و هیدرولیکی مقطع اصلی و دشت‌های سیلابی را به صورت توأم در بر می‌گیرد و نشان می‌دهد که کانال مورد بررسی چه مقدار به یک مقطع ساده نزدیک است. ظاهر شدن آن به عنوان دومین پارامتر موثر نشان می‌دهد که باید از تفکیک مقطع اصلی و دشت‌های سیلابی برای تحلیل دبی جریان استفاده نمود.

جدول ۴- نتایج مدل‌های بهینه با افزایش تعداد پارامترهای ورودی

تعداد پارامتر	کد رشته	مقادیر Γ	مقادیر MIPE
۱	۰۰۰۰۰۰۰۰۰۱	۰/۰۸۵۲	۰/۲۳۱۱
۲	۱۰۰۰۰۰۰۰۰۱	۰/۰۲۲۱	۰/۱۱۷۰
۳	۱۰۰۰۰۰۱۰۱	۰/۰۱۰۷	۰/۰۲۵۸
۴	۱۱۰۰۰۰۱۰۱	۰/۰۱۰۵۴	۰/۰۲۴۱
۵	۱۱۱۰۰۰۱۰۱	۰/۰۱۲۰	۰/۰۲۷۰
۶	۱۱۰۱۰۱۰۰۱۱	۰/۰۱۲۸	۰/۰۲۹۱
۷	۱۱۱۱۰۱۰۰۱۱	۰/۰۱۴۰	۰/۰۳۱۶
۸	۱۱۱۱۰۱۱۰۱۱	۰/۰۲۱۵	۰/۰۴۸۳
۹	۱۱۱۱۰۱۱۱۱۱	۰/۰۳۰۱	۰/۰۶۸۹
تمامی پارامترها	۱۱۱۱۱۱۱۱۱۱	۰/۰۳۹۹	۰/۰۹۴۶



شکل ۴ - نمودار M تست مقادیر Γ و خطای استاندارد برای ترکیب بهینه (۱۰۰۰۰۰۱۰۱)

مدل‌های بهینه بر اساس آماره‌های MIPE و گرادیان برای دوره‌ی آزمون در جدول ۵ ارائه شده است. مطابق آن، مدل بهینه بر اساس آماره‌ی MIPE، با RMSE برابر با ۰/۲۰۶ و ضریب همبستگی برابر با ۰/۹۸۹ از دقت بالاتری نسبت به مدل بهینه بر اساس آماره‌ی گرادیان (RMSE برابر با ۰/۴۲۶) برخوردار است. بنابراین آماره‌ی MIPE برای انتخاب ترکیب بهینه و تعیین اهمیت پارامترهای ورودی در روش گاما تست انتخاب می‌شود.

نتایج انتخاب آماره مناسب در روش GT

نتایج انتخاب ترکیب بهینه بر اساس آماره‌های Γ ، خطای استاندارد، VRatio و MIPE یکسان می‌باشد، بنابراین دو ترکیب بهینه بر اساس آماره‌های MIPE و گرادیان در قالب ۲ مدل با استفاده از شبکه‌ی عصبی مصنوعی با الگوریتم BP و تابع آموزشی CGS مدل‌سازی شدند. بر اساس نتایج M تست از ۳۰۰ الگو برای آموزش و از ۷۱ الگوی باقی‌مانده برای ارزیابی استفاده گردید. نتایج ارزیابی

جدول ۵- نتایج ارزیابی مدل‌های بهینه بر اساس آماره‌های MIPE، گرادیان در دوره‌ی آزمون

مدل بهینه	کد رشته	ساختار*	RMSE	MBE	MARE	R
بر اساس مقدار MIPE	۱۱۰۰۰۰۰۱۰۱	۱-۵-۴	۰/۲۰۶	-۰/۰۴۳	۰/۰۷۰	۰/۹۸۹
بر اساس مقدار گرادیان	۰۱۰۱۰۱۰۱۱۰	۱-۴-۵	۰/۴۲۶	-۰/۴۲۳	۰/۵۱۱	۰/۴۲۶

*: تعداد پارامترهای ورودی - تعداد نرون‌های لایه میانی اول - تعداد نرون‌های لایه میانی دوم - پارامتر خروجی (از چپ به راست)

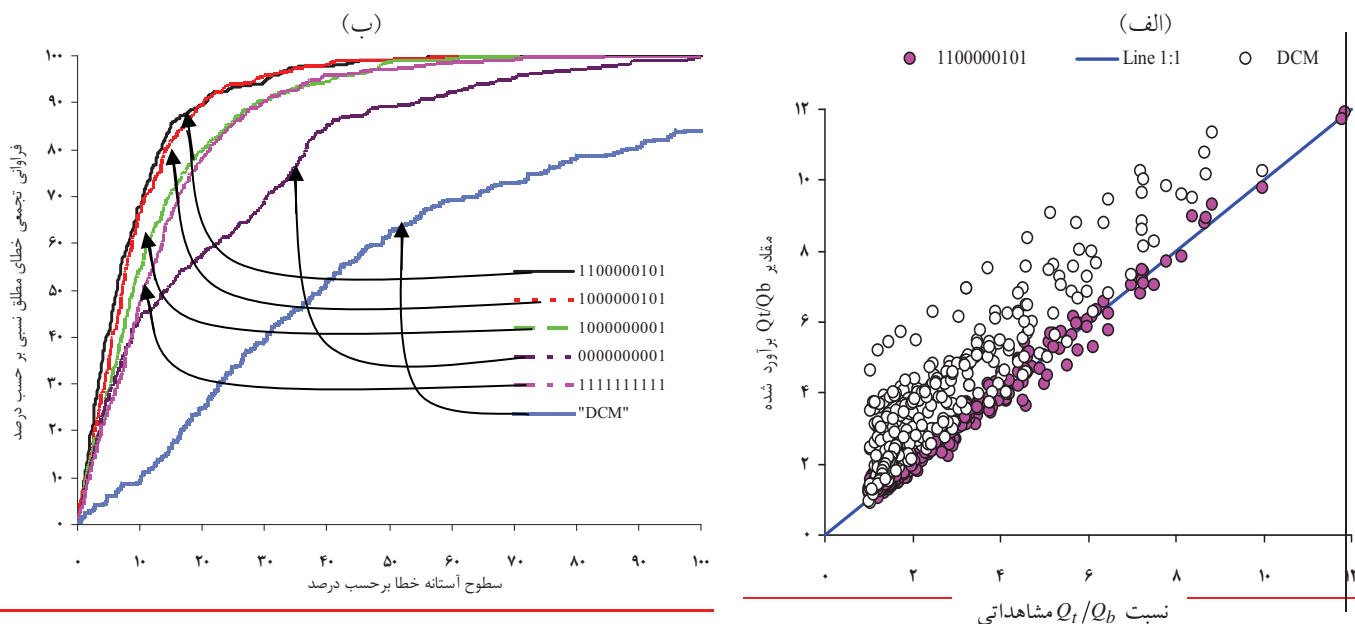
برای اطلاع از چگونگی توزیع خطا، نمودار پراکنش درصد فراوانی تجمعی مقادیر قدرمطلق خطای نسبی برای هر یک از مدل‌های ساخته شده بر حسب درصد سطوح آستانه‌ی خطا در شکل ۵ (ب) ترسیم شده است. مطابق آن، مدل ۱۱۰۰۰۰۰۱۰۱، در ۹۰ درصد موارد سطح آستانه‌ی خطای قدرمطلق نسبی کمتر از ۲۰/۵ درصد می‌باشد و این مقدار به ازای ۷۰ درصد موارد، کمتر از ۱۰/۵ درصد است. نمودار توزیع خطا برای مدل ۱۱۰۰۰۰۰۱۰۱ نیز تقریباً مشابه مدل ۱۱۰۰۰۰۰۱۰۱ است و به ازای مقادیر خطای قدرمطلق نسبی کمتر از ۲۰ درصد در پائین آن قرار دارد یا به عبارتی در ۹۰ درصد موارد دارای خطای بیشتری است و تنها در چند درصد ناچیز خطای کمتری دارد. نمودار توزیع خطا برای دو مدل ۱۰۰۰۰۰۰۰۰۱ و ۱۱۱۱۱۱۱۱۱۱ نیز تقریباً مشابه هستند با این تفاوت که مدل ۱۰۰۰۰۰۰۰۰۱ با قرار گرفتن در بالای مدل ۱۱۱۱۱۱۱۱۱۱ در بیش از ۹۵ درصد موارد دارای خطای کمتری است. همان‌طور که گفته شد مدل DCM بر اساس آماره R دارای دقت بالاتری نسبت به مدل ۱۰۰۰۰۰۰۰۰۱ می‌باشد ولی بررسی نمودار توزیع خطا برای این دو مدل نشان می‌دهد که مدل DCM در ۸۰ درصد موارد، دارای سطح آستانه‌ی خطایی برابر ۱۰۰ درصد بوده و در ۲۰ درصد باقی‌مانده دارای سطح آستانه‌ی خطایی بیش از ۱۰۰ درصد می‌باشد و نمودار آن با اختلاف زیادی در زیر نمودار توزیع خطا برای مدل ۱۰۰۰۰۰۰۰۰۱ قرار دارد. آماره‌هایی مانند R با بیان متوسط خطا نمی‌توانند چگونگی توزیع خطا را نشان دهند و گاهی استفاده از یک آماره آماری برای انتخاب مدل بهینه می‌تواند منجر به انتخابی نادرست گردد این در حالی است که آماره ترکیبی IPE با ترکیب آماره‌های مختلف و ارزش‌دهی به آن‌ها نشان داد که مدل ۱۰۰۰۰۰۰۰۰۱ با IPE برابر با ۰/۳۹۴ از مدل DCM با IPE برابر با ۰/۸۶۷ دارای دقت بیشتری است که این امر در نمودار توزیع خطای این دو مدل مشهود است.

با توجه به آن‌که در ساختار ترکیب بهینه بر اساس آماره‌ی MIPE (۱۱۰۰۰۰۰۱۰۱)، ۴ پارامتر b_1, B_c, D_r و COH ظاهر شدند، در ادامه مدل‌های بهینه‌ی دیگری با پارامترهای ورودی کمتر از ترکیب بهینه (۳، ۲ و ۱ پارامتر ورودی) بر اساس آماره‌ی MIPE تعیین گردید (جدول ۴) و مطابق آن‌چه که برای ترکیب بهینه‌ی ۱۱۰۰۰۰۰۱۰۱ بیان گردید با استفاده از شبکه عصبی مدل‌سازی گردیدند. هم‌چنین برای ترکیب بدون پیش‌پردازش (شامل همه‌ی پارامترهای ورودی) نیز این مدل‌سازی انجام گردید. نتایج این مدل‌ها در کل دوره (آموزش و آزمون) با یکدیگر و با مدل DCM با آماره‌های مختلف آماری مقایسه گردیدند که نتایج آن در جدول ۶ ارائه شده است. مطابق نتایج آن، مدل ۱۱۰۰۰۰۰۱۰۱ با IPE برابر با ۰/۱۲۳ نسبت به ۴ مدل دیگر از دقت بالاتری برخوردار است. مدل‌های بهینه با ۳ و ۲ پارامتر ورودی نسبت به مدل بدون پیش‌پردازش (۱۱۱۱۱۱۱۱۱۱) دارای مقدار IPE کمتر و دقت بالاتر هستند. هم‌چنین مدل ۱۱۱۱۱۱۱۱۱۱ با IPE برابر با ۰/۲۲۹ نسبت به مدل بهینه با پارامتر ورودی عمق نسبی دقت بیشتری دارد. از میان مدل‌های جدول ۶، مدل DCM با IPE برابر با ۰/۸۶۷ ضعیف‌ترین نتایج را داشت اگر چه مدل DCM بر اساس آماره‌ی R دارای دقت بالاتری نسبت به مدل ۱۰۰۰۰۰۰۰۰۱ می‌باشد.

نمودار پراکنش مقادیر برآوردی نسبت Q_t/Q_b توسط دو مدل بهینه بر اساس آماره‌ی MIPE (۱۱۰۰۰۰۰۱۰۱) و روش معمول تجزیه‌ی قائم مقطع مرکب (DCM) در کل دوره (آموزش و آزمون) روی خط ۱:۱ در شکل ۵ (الف) مقایسه شده‌اند. بر اساس آن، مدل ۱۱۰۰۰۰۰۱۰۱ به‌خوبی توانسته است مقادیر Q_t/Q_b را شبیه‌سازی نماید، به‌طوری که مقادیر برآوردی و مشاهداتی تقریباً روی خط ۱:۱ قرار دارند در حالی که نتایج روش DCM با قرار گرفتن در بالای خط مذکور، دارای بیش‌برآوردی زیادی می‌باشند.

جدول ۶- نتایج ارزیابی مدل‌های بهینه با ۴، ۳، ۲ و ۱ پارامتر ورودی به همراه مدل بدون پیش‌پردازش و DCM در کل دوره

مدل	کد رشته	RMSE	MBE	MARE	R	IPE
با ۴ پارامتر ورودی	۱۱۰۰۰۰۰۱۰۱	۰/۲۵۸	-۰/۰۷۸	۰/۰۹۱	۰/۹۹۱	۰/۱۲۳
با ۳ پارامتر ورودی	۱۰۰۰۰۰۰۱۰۱	۰/۳۰۹	-۰/۰۱۵	۰/۰۹۷	۰/۹۸۸	۰/۱۳۵
با ۲ پارامتر ورودی	۱۰۰۰۰۰۰۰۰۱	۰/۵۰۱	-۰/۰۰۹	۰/۱۲۹	۰/۹۶۲	۰/۲۰۴
با ۱ پارامتر ورودی	۰۰۰۰۰۰۰۰۰۱	۰/۹۹۸	-۰/۰۱۵	۰/۲۱۹	۰/۸۳۷	۰/۳۹۴
بدون پیش‌پردازش	۱۱۱۱۱۱۱۱۱۱	۰/۵۶۶	-۰/۰۴۸	۰/۱۴۱	۰/۹۵۱	۰/۲۲۹
روش تجزیه‌ی قائم مقطع مرکب (DCM)	(DCM)	۱/۵۰۱	-۱/۱۸۰	۰/۵۶۱	۰/۹۱۵	۰/۸۶۷



شکل ۵- نمودارهای پراکنش نقاط روی خط ۴۵ درجه برای مدل بهینه (۱۱۰۰۰۰۰۱۰۱) و مدل DCM (الف) و توزیع آستانه‌ی خطا برای مدل‌های بهینه و مدل DCM (ب) در کل دوره

نتیجه‌گیری

در این تحقیق با استفاده از روش پیش‌پردازش GT، پارامترهای مهم ورودی و ترکیب بهینه‌ی آن‌ها برای ایجاد یک مدل هموار از دبی جریان مقاطع مرکب تعیین و دبی جریان به کمک شبکه‌ی عصبی محاسبه شده است. نتایج حاصل از این تحقیق به شرح زیر می‌باشد:

در روش GT، انتخاب آماره مناسب از میان آماره‌های Γ ، گرادیان، خطای استاندارد و VRatio از اهمیت زیادی در انتخاب ترکیب بهینه و تعیین اهمیت پارامترهای ورودی برخوردار است. در این تحقیق ضمن بررسی آماره‌های فوق در گزینش ترکیب بهینه و تعیین اهمیت پارامترهای ورودی، با دخالت دادن آماره‌های ذکرشده در قالب آماره MIPE امکان دخالت همه‌ی آن‌ها در این گزینش فراهم گردید. هرچند در این تحقیق نتایج انتخاب ترکیب بهینه بر اساس آماره‌های Γ ، VRatio و خطای استاندارد با آماره MIPE یکسان بود اما این آماره ترکیبی می‌تواند به استفاده بهتر از نتایج این

روش کمک زیادی نماید.

بر اساس نتایج GT، از میان ۱۰۲۳ ترکیب حاصل از ۱۰ پارامتر ورودی، ترکیب پارامترهای عمق نسبی، کوهپرنس مقطع، شیب جانبی دشت‌های سیلابی و عرض نسبی دشت سیلابی با کد رشته‌ی ۱۱۰۰۰۰۰۱۰۱ به‌عنوان ترکیب بهینه و پارامترهای موثر در تعیین دبی رودخانه‌های سیلابی شناخته شدند.

مدل‌سازی دبی جریان در مقاطع مرکب برای ترکیب بهینه (۴ پارامتر ورودی) و مدل‌های بهینه با ۲، ۳ و ۱ پارامتر ورودی و مقایسه نتایج آن‌ها با نتایج مدل بدون پیش‌پردازش پارامترهای ورودی و روش DCM با آماره‌های مختلف خطا نشان داد مدل ۱۱۰۰۰۰۰۱۰۱ با آماره IPE برابر با ۰/۱۲۳ از دقت بالایی نسبت به سایر مدل‌ها در کل دوره برخوردار است و روش DCM با آماره IPE برابر ۰/۸۶۷ دبی رودخانه‌های سیلابی را بسیار بیشتر از مقدار واقعی برآورد می‌نماید به طوری که نتایج آن نسبت به مدل ۱۱۰۰۰۰۰۰۰۱ با تنها یک پارامتر ورودی (D_f) دقت کمتری دارد.

منابع

- ۱- قبائی سوق م، مساعدی ا، حسام م، و هزارجریبی ا. ۱۳۸۹. ارزیابی تاثیر پیش‌پردازش پارامترهای ورودی به شبکه عصبی مصنوعی (ANNs) با استفاده از روش‌های رگرسیون گام‌به‌گام و گاما تست به منظور تخمین سریع‌تر تبخیر و تعرق روزانه. مجله آب و خاک، ۲۴ (۳)، ۶۲۴-۶۱۰.
- ۲- نیک‌مهر س، فرهودی ج، و صمدیان فرد س. ۱۳۸۸. تخمین پروفیل سرعت در عرض کانال با استفاده از دو روش استنتاج فازی- عصبی و

- 3- Ackers P. 1992. Hydraulic design of two- stage channels. *Journal of Water and Maritime Engineering*, 96, 247-257.
- 4- Ackers P. 1993. Stage-Discharge functions for two-stage channels. *Water and Environmental Management*, 7, 52-61.
- 5- Ahmadi A., Han D., Karamouz M., and Remesan R. 2009. Input data selection for solar radiation estimation. *Hydrological Processes*, 23: 2754–2764.
- 6- Atabay S., and Knight D.W. 2004. 1-D modeling of conveyance, boundary shear and sediment transport in overbank flow. *Journal of Hydraulic Research. IAHR*,
- 7- Blalock M.E., and Sturm T.W. 1981. Minimum specific energy in compound channel. *Journal of the Hydraulics Division, ASCE*, 107, 699–717.
- 8- Bousmar D., and Zech Y. 1999. Momentum transfer for practical flow computation in compound channels. *Journal of Hydraulic Engineering, ASCE*, 125(7), 696-70.
- 9- Bousmar D., Wilkin N., Jacquemart H., and Zech Y. 2004. Overbank flow in symmetrically narrowing floodplains. *Journal of Hydraulic Engineering, ASCE*, 130(4), 305-312.
- 10- Chow V.T. 1959. *Open channel hydraulics*, McGraw-Hill, London.
- 11- Concar N. 1997. *Optimisation methodologies for direct inverse neuro control*. Ph.D. thesis, Department of Computing, Imperial College, University of London, U.K.
- 12- Elshorbagy A., Corzo G., Srinivasulu S., and Solomatine D. 2009. Experimental investigation of the predictive capabilities of soft computing techniques in hydrology. Centre for Advanced Numerical Simulation (CANSIM), Department of Civil & Geological Engineering, University of Saskatchewan, Saskatoon, SK, CANADA. 49 p.
- 13- Ervine D.A., Babaeyan-Koopaei K., and Sellin R.H.J. 2000. Two-dimensional solution for straight and meandering overbank flows. *Journal of Hydraulic Engineering, ASCE*, 126(9), 653-669.
- 14- Haidera M.A., and Valentine E.M. 2002. A practical method for predicting the total discharge in mobile and rigid boundary compound channels. *International Conference on Fluvial Hydraulics, Belgium*, 153-160.
- 15- Jones A.J., Evans D., Margetts S., and Durrant P. 2002. The Gamma Test. Chapter IX in *Heuristic and Optimization for Knowledge Discovery*. Edited by Ruhul Sarker, Hussein Abbass and Charles Newton. Idea Group Publishing, Hershey, PA. 27 pp.
- 16- Knight D.W., and Demetriou J.D. 1983. Flood plain and main channel flow interaction. *Journal of Hydraulic Division, ASCE*, 109(8), 1073-1092.
- 17- Knight D.W., and Sellin R.H.J. 1987. The SERC flood channel facility. *Journal of Institution of Water and Environment Management*, 1(2), 198-204.
- 18- Knight D.W., Shiono K., and Pirt J. 1989. Predictions of depth mean velocity and discharge in natural rivers with overbank flow. *International Conference on Hydraulics and Environmental Modeling of Coastal, Estuarine and River Waters. England*, 419-428.
- 19- Lai S.H., and Bessaih N. 2004. Flow in compound channels. *1st International Conference on Managing Rivers in the 21st Century, Malaysia*, 275-280.
- 20- Lambert M.F., and Myers R.C. 1998. Estimating the discharge capacity in straight compound channels. *Water, Maritime and Energy*, 130, 84-94.
- 21- Lambert M.F., and Sellin R.H.J. 1996. Discharge prediction in straight compound channels using the mixing length concept. *Journal of Hydraulic Research, IAHR*, 34, 381-394.
- 22- Liu W., and James C.S. 2000. Estimating of discharge capacity in meandering compound channels using artificial neural networks. *Canadian J. of Civil Engineering*, 27(2): 297-308.
- 23- Martin L.A., and Myers R.C. 1991. Measurement of overbank flow in a compound river channel. *Journal of Institution of Water and Environment Management*, 645-657.
- 24- Moghaddamnia A., Ghafari Gousheh M., Piri J., Amin S., and Han D. 2008. Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Advances in Water Resources*. 32 : 88–97.
- 25- Moghaddamnia A., Remesan R., Hassanpour Kashani M., Mohammadi M., Han D., and Piri J. 2009. Comparison of LLR, MLP, Elman, NNARX and ANFIS Models with a case study in solar radiation estimation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 71: 975–982.
- 26- Myers R.C., and Lyness. J.F. 1997. Discharge ratios in smooth and rough compound channels. *Journal of Hydraulics Engineering, ASCE*, 123(3), 182-188.
- 27- Nayak P.C., Sudheer K.P., Rangan D.M., and Ramasastry K.S. 2004. A neuro-fuzzy computing technique for modeling hydrological time series. *Journal of Hydrology*, 291, 52-66.
- 28- Noori R., Hoshyaripour G.H., Ashrafi K.H., and Najdar Araabi B. 2009. Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration, *Atmospheric Environment*, 44:476-482.
- 29- Noori R., Karbassi A.R., Moghaddamnia A., Han D., Zokaei-Ashtiani M.H., Farokhnia A., and Ghafari Gousheh, M. 2011. Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction. *Journal of Hydrology*, 401, 177–189.

- 30- Noori R., Karbassi A., and Sabahi M.S. 2009. Evaluation of PCA and gamma test techniques on AAN operation for weekly solid waste prediction, *Journal of Environmental Management*.91: 767-771.
- 31- Shiono K., and Knight D.W. 1988. Two-dimensional analytical solution for a compound channel. 3rd International Symposium on Refined Flow Modeling and Turbulence Measurements, Japan, 503-510.
- 32- Tarrab L., and Weber J.F. 2004. Predicción del coeficiente de mezcla transversal en cauces aturales. *Mecánica Computacional*, XXIII, Asociación Argentina de Mecanica Computacional, San Carlos de Bariloche, 1343-1355.
- 33- Tsui A.P.M. 1999. Smooth data modelling and stimulus-response via stabilisation of neural chaos. Ph.D. thesis, Department of Computing, Imperial College, University of London, U.K. 199 P.
- 34- Wark J.B., Samuels P.G., and Ervine D.A. 1990. A practical method of estimating velocity and discharge in compound channels. *International Conference on River Flood Hydraulics*, London, 163-172.
- 35- Wormleaton P.R., and Merrett D.J. 1990. An improved method of calculation for steady uniform flow in prismatic main channel/floodplain sections. *Journal of Hydraulic Research*, IAHR, 28, 157-174.
- 36- Zahiri A., and Dehghani A.A. 2009. Flow Discharge Determination in Straight Compound Channels Using ANNs. *World Academy of Science, Engineering and Technology*, 58. 1-4.

Determination of Influence Parameters and its Optimal Combination in Order to Model Flood Rivers Discharge

A. Zahiri¹- M. Ghabaei Sough²- A. Mosaedi^{3*}

Received:21-8-2011

Accepted:23-10-2011

Abstract

Flood rivers discharge in compound channels is extremely complicated because of rapidly changes in geometric and hydraulic conditions of main channel and flood plains. Different parameters play important roles in determination of discharge in compound channel. Determination of influence parameters and selection of optimal combination between different combinations is one of the major stages in design of intelligence and mathematical modeling. In this study, flow discharge were calculated by using artificial neural network based on gamma test (GT) preprocessing method from 10 geometric and hydraulic dimensionless parameters of 30 laboratory and river compound channels. Based on result of GT, 4 parameters of relative depth, coherence, berm inclination and relative flood plains width, were recognized as influence variables. Moreover, among 1023 different combinations that established by 10 input parameters, combination of these 4 parameters was selected as optimal combination. Analysis of data adequacy by using M test for best combination showed that among of data needed for establishment of a smooth model is equal to 300 data points. Flood discharge of compound channel was modeled for optimal combination and optimal models of 1, 2, and 3 input parameters. The results of these models were compared with models of all 10 input parameters and divided channel method (DCM) model based on different statistical criteria. The results showed that the optimal model with IPE=0.123 and DCM with IPE=0.867 have the highest and lowest accuracy respectively

Keywords: Discharge estimation, Compound Channel, Optimal Combination, Gamma Test, Artificial Neural Network

1,2- Assistant Professor and Former MSc Student, Department of Water Engineering, Gorgan University of Agricultural Sciences and Natural Resources

3- Associate Professor, Faculty of Natural Resources and Environment, Ferdowsi University of Mashhad

(*-Corresponding Author Email: mosaedi@um.ac.ir)