پیش‌بینی خطر وقوع جریان‌های گردابی در آبگیر قائم با استفاده از مدل k نزدیک‌ترین همسایگی

احسان فدائی کرمانی[۱]، غلامعباس بارانی[۲]، مهناز قائینی حصاروئیه [۳]

مقاله برگرفته از طرح پژوهشی مطالعاتی درون دانشگاهی توسط نویسندگان

تاریخ دریافت: ۹۱/۰۵/۰۴
تاریخ پذیرش: ۹۱/۰۸//۱۹

چکیده

استفاده از آبگیرهای قائم یکی از روش‌های معمول برداشت آب از رودخانه‌ها و یا مخازن به منظور تامین آب شرب و مورد نیاز صنعت و کشاورزی می‌باشد. یکی از مشکلات اساسی که معمولا آبگیرهای قائم با آن مواجه می‌باشند، وقوع جریان‌های گردابی در در دهانه آبگیرها می‌باشد. وقوع جریان‌های گردابی باعث کاهش بازدهی سیستم آبگیری می‌شود و عمر مفید سازه را نیز کاهش می‌دهد. در این مطالعه با استفاده از مدل k نزدیک‌ترین همسایگی و با توجه به شرایط هیدرولیکی و هندسی آبگیر قائم، خطر وقوع گرداب‌های مخرب برای سه نسبت مختلف طول به عرض کانال ورودی آبگیر تخمین زده شد. سپس با استفاده از ضرایب آماری مختلف میزان کارایی و دقت مدل مورد ارزیابی قرار گرفت. مقادیر مناسب ضریب همبستگی ۰/۹۱۲، ضریب کارایی مدل ۰/۸۲۷ و جذر میانگین مربعات خطا ۰/۱۰۵، نشان داد این تکنیک قابل قبول و توانمند می‌باشد.

واژه‌های کلیدی: آبگیر قائم، پیش‌بینی خطر، جریان‌های گردابی، مدل k نزدیک‌ترین همسایگی.

Archive of SID

[۱] دانشجوی کارشناسی ارشد ، بخش مهندسی عمران، دانشگاه شهید باهنر، کرمان، ایران ۰۹۱۳۲۹۵۱۴۸۰
ehsanhard@gmail.com (نویسنده مسئول)
[۲] استاد ، بخش مهندسی عمران، دانشگاه شهید باهنر، کرمان ، ایران ۰۳۴۱-۳۲۲۰۰۵۴
gab@mail.uk.ac.ir
[۳] استاد یار ، بخش مهندسی عمران، دانشگاه شهید باهنر، کرمان ، ایران ۰۳۴۱-۳۲۲۰۰۵۴
mghaeini@mail.uk.ac.ir

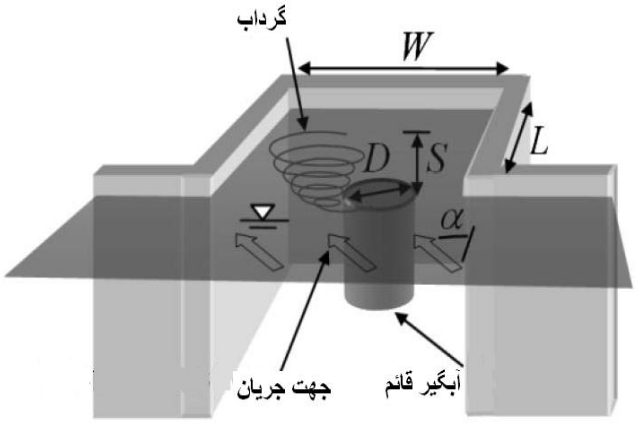
www.SID.ir

مقدمه

آبگیرهای قائم از جمله سازه‌هایی به شمار می‌آیند که به منظور آبگیری مستقیم از رودخانه‌ها و یا مخازن مورد استفاده قرار می‌گیرند. این سازه‌ها متشکل از لوله‌ای قائم بوده که در یک مخزن قرار می‌گیرد (شکل۱). آبگیرهای قائم سازه‌هایی اقتصادی می‌باشند و در صورت طراحی صحیح و اصولی، بهره‌برداری از آن‌ها با مشکلات کمتری همراه است. علاوه بر این از این جهت که این سازه‌ها در نزدیکی سطح آب کار گذاشته می‌شوند، از ورود رسوبات درشت دانه به داخل سیستم جلوگیری می‌نمایند (زمردیان و شجاعیان ، ۱۳۸۳).

یکی از معضلات و مشکلات اساسی که معمولا آبگیرهای قائم با ان مواجه هستند، تشکیل جریان‌های گردابی در دهانه آبگیرها می‌باشد. وقوع جریان‌های گردابی افت بازدهی سیستم آبگیری را در پی دارد. علاوه بر کاهش ظرفیت آبگیری، گرداب‌ها باعث خطرات دیگری نظیر ایجاد سر و صدا، ورود هوا و نیز لغزش و خوردگی در آبگیرها می‌شوند (کناس، ۱۹۸۷). از جمله عوامل مهم در وقوع جریان‌های گردابی در دهانه آبگیرها می‌توان به شرایط نامتوازن سازه آبگیر، هندسه سازه، کافی نبودن عمق استغراق، جدایش جریان و نیز تغییرات ناگهانی در مسیر جریان اشاره کرد (ریندلز و گالیور، ۱۹۸۳).

تاکنون مطالعات و تحقیقات گوناگونی در زمینه چگونگی تشکیل گرداب‌ها و نیز شناخت و تعیین ابعاد ناحیه گردابی در ورودی کانال آبگیر انجام شده است. هکر (۱۹۸۱) طی یک تقسیم‌بندی گرداب‌ها را بر مبنای وضعیت گرداب و قدرت آن‌ها از ضعیف به قوی به شش گروه طبقه‌بندی کرد. تحقیقات بعدی نشان دادکه عمق استغراق بحرانی، عمق شروع گرداب نوع پنج می‌باشد و مشکلات ناشی از وقوع جریان‌های گردابی در گرداب‌های نوع پنج و شش به وقوع می‌پیوندد (ورمین، ۱۹۹۹).



شکل(۱): مشخصات یک آبگیر قائم

استفاده از مدل‌های فیزیکی و هیدرولیکی از دیگر ابزار محققین برای بررسی این پدیده به شمار می‌آید به عنوان مثال کانستنتیسکو و پاتل (۱۹۹۸). استفاده از مدل‌های هیدرولیکی نقش مهم و موثری در پیش‌بینی شکل‌گیری جریان‌های گردابی تحت شرایط ویژه و کنترل شده آزمایشگاه دارد. اما متاسفانه به دلیل شرایط نامشخص و یا شرایط ویژه حاکم بر این نوع مدل‌ها، شرایط استفاده از آن‌ها را به منظور تعیین ضوابط طراحی محدود می‌کند و معمولا از روابط تجربی به جای آن‌ها در کارهای طراحی استفاده می شود (تراویس و میز، ۲۰۱۱).

صورت کلی روابط تجربی حاکم بر طراحی آبگیرها به منظور جلوگیری از وقوع جریان‌های گردابی معادله(۱) می‌باشد:

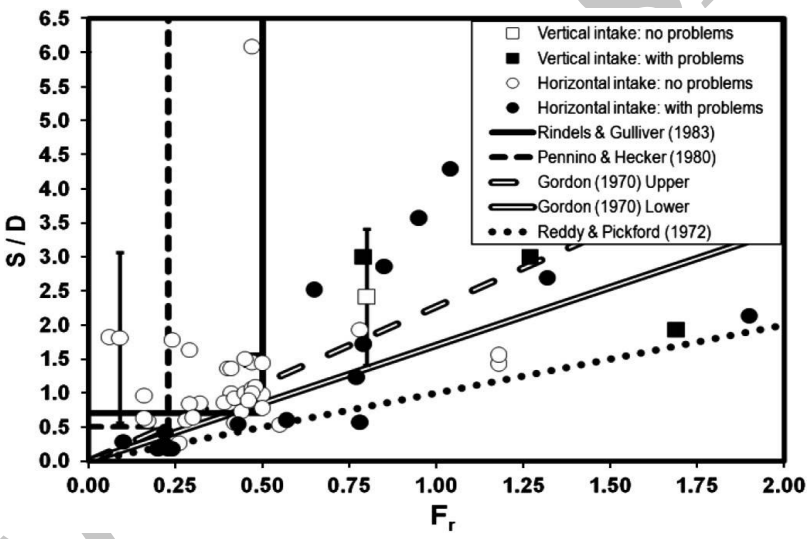
$$S / D = aFr^b \qquad (۱)$$

که در معادله فوق، S عمق استغراق، D قطر دهانه آبگیر، a و b اعداد ثابت و Fr عدد فرود مربوط به آبگیر می‌باشد. با توجه به تحقیقات و آزمایشات صورت گرفته، مقادیر مختلفی برای ثابت های a و b توصیه شده است. به عنوان مثال اگر b=0 و یا بینهایت باشد یک ناحیه چهار

www.SID.ir

گوش تشکیل می‌گردد و در صورتی که b=1 فرض شود یک رابطه خطی بین متغیرها برقرار می‌گردد. در شکل(۲) وضعیت برخی از این روابط تجربی در پیش‌بینی وقوع جریان‌های گردابی برای آبگیرهای قائم و افقی نمایش داده شده است. تراویس و میز (۲۰۱۱) ، میزان کارایی و صحت این روابط را مورد بررسی قرار دادند و نشان دادند که این روابط به طور کلی از قابلیت مناسبی برخوردار نمی‌باشند. همچنین آن‌ها نقش جهت جریان ورودی را در پیش‌بینی خطر وقوع جریان‌های گردابی در آبگیرهای قائم را مورد بررسی قرار دادند.

در هر صورت گرداب ایجاد شده در دهانه آبگیر پدیده‌ای نامطلوب برای آن‌ها به شمار می‌آید که توجه مهندسین و طراحان را به خود معطوف ساخته است. به دلیل ماهیت پیچیده این پدیده و نیز شرایط متفاوت در طراحی‌ها و معیارهای زیاد موثر در وقوع گرداب‌ها، هنوز آیین نامه مدونی برای طراحی آبگیرها ارائه نشده است. از طرف دیگر محدودیت‌های آزمایشگاهی مدل‌های هیدرولیکی و نیز عدم کفایت کامل روابط تجربی بر دشواری کار افزوده است. در این مطالعه با استفاده از مدل غیرپارامتریک k- نزدیک‌ترین همسایگی و با بهره‌گیری از داده‌های تجربی و آزمایشگاهی موجود به بررسی پدیده وقوع جریان‌های گردابی در آبگیرهای قائم پرداخته شده است. با توجه به مدل k- نزدیک‌ترین همسایگی پایگاه داده موجود توسعه داده شده و وضعیت وقوع گرداب‌های مخرب تخمین زده شده است. برای سهولت استفاده از نتایج به دست آمده در کارهای طراحی، نتایج براساس نسبت‌های مختلف طول به عرض کانال حاوی آبگیر ارائه گردیده است.



شکل(۲): وضعیت تخمین خطر وقوع جریان‌های گردابی مخرب بر مبنای روابط تجربی (تراویس و میز، ۲۰۱۱)

مدل k- نزدیک‌ترین همسایگی

استفاده از الگوریتم‌های غیر پارامتریک در مواردی که نحوه ارتباط بین ورودی و خروجی از قبل به طور کامل مشخص نباشد، سودمند و موثر خواهد بود. الگوریتم k- نزدیک‌ترین همسایگی، یکی از مهم‌ترین و توسعه یافته‌ترین رویکردهای غیر پارامتریک می‌باشد که در بسیاری از پژوهش‌های نوین جهت تشخیص الگو و کلاسه‌بندی‌های آماری به کار گرفته شده است (جلالی و همایی، ۱۳۸۹).

در حالت کلی از این الگوریتم به دو منظور استفاده می‌شود: برای تخمین تابع چگالی توزیع داده‌های تعلیم و برای طبقه‌بندی داده‌های تست براساس الگوهای تعلیم. بر خلاف توابع انتقالی کلاسیک مدل نزدیک‌ترین همسایگی از هیچ تابع ریاضی از پیش تعریف شده‌ای برای تخمین متغیرهای مختلف استفاده نمی‌کند. به طور کلی می‌توان گفت که مدل نزدیک‌ترین همسایگی یکی از روش‌های داده کاوی می‌باشد که هدف کلی آن طبقه‌بندی و تخمین ویژگی‌های یک سری داده‌های مجهول با توجه به بیشترین شباهت این داده‌ها با داده‌های معلوم که در همسایگی (نزدیکی) آن‌ها قرار دارند، می‌باشد (ژیندانگ و کومار، ۲۰۰۹). در این مدل داده‌های هدف مورد جستجو (test data) با توجه به نزدیک‌ترین فاصله نسبت به

www.SID.ir

داده‌های آموزش (بانک داده مرجع)، (training data) طبقه‌بندی می‌شوند.

نخستین گام در استفاده از این مدل یافتن روش و رابطه‌ای برای محاسبه فاصله بین داده‌های مورد آزمایش و داده‌های تعلیم می‌باشد. معمولا برای تعیین فاصله بین داده‌های تعلیم و داده‌های آزمایش از فاصله اقلیدسی استفاده می‌شود.

$$d(X,Y)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \qquad (۲)$$

که در آن X نماینده داده‌های تعلیم با پارامترهای مشخص(X_1 تا X_n) و Y نماینده داده‌های آموزش با همان تعداد پارامترهای مشخص (y_1 تا y_n) می‌باشد.

$$X=(x_1,x_2,...,x_n)$$
$$Y=(y_1,y_2,...,y_n)$$

پس از تعیین فاصله اقلیدسی بین داده‌ها، نمونه‌های بانک داده به ترتیب صعودی از کمترین فاصله (حداکثر تشابه) تا بیشترین فاصله (حداقل تشابه) از نمونه مورد نظر طبقه‌بندی و ارزش‌گذاری می‌شوند.

قدم بعدی در این مدل یافتن تعداد نقاطی (K) از داده‌های آزمایش برای تخمین ویژگی‌های بانک داده مورد نظر می‌باشد. تعیین تعداد همسایه‌ها (k) ، یکی از کلیدی‌ترین و مهم‌ترین مراحل در این مدل به شمار می‌آید و میزان کارایی این روش به طور قابل ملاحظه‌ای به کیفیت انتخاب نزدیک‌ترین (مشابه‌ترین) نمونه‌ها از داده‌های بانک مرجع بستگی دارد. اگر k مقداری کوچک فرض شود، نتایج نسبت به تک نقطه‌های نا متعارف مدل حساس می‌باشد و در صورتی که k مقداری بزرگ انتخاب گردد امکان قرار گرفتن نقاطی از دیگر کلاس‌ها در محدوده مورد نظر وجود دارد. معمولا بهترین مقدار k با استفاده از عملیات Cross-Validation قابل محاسبه می‌باشد (ژیندانگ و کومار، ۲۰۰۹).

تکنیک n-fold Cross-Validation یک روش آماری می‌باشد که در واقع کیفیت تخمین‌های یک مدل را براساس تعداد و نوع داده‌های ورودی مشخص می‌کند. در این روش داده‌های مدل به n قسمت مساوی تقسیم می‌شود. سپس برای n ایمن قسمت، مدل با توجه به داده‌های بقیه قسمت‌های مدل (n-1) فیت می‌شود و میزان خطای مدل برای پیش‌بینی قسمت nام ارزیابی می‌شود. این روند آنقدر ادامه می‌یابد که تمامی قسمت‌ها را شامل شود (هاستی و همکاران، ۲۰۰۸).

مواد و روش‌ها

در این مطالعه از آزمایشات و تجربیات صورت گرفته توسط ریندلز و گالیور (۱۹۸۳) استفاده شده است و با استفاده از الگوریتم k نزدیک‌ترین همسایگی نتایج حاصله بسط داده شده و محدوده‌ای به منظور تخمین خطر آسیب ناشی از وقوع جریان‌های گردابی مخرب در آبگیرهای قائم با توجه به مقادیر مختلف نسبت عرض به طول کانال ورودی آبگیر جهت استفاده در ضوابط طراحی آبگیرهای قائم تعیین شده است. داده‌های مورد مطالعه شامل محدوده مقادیر مختلف عدد فرود ($0.09 \leq Fr \leq 2.47$) و نسبت عمق استغراق به قطر آبگیر ($0.67 \leq S/D \leq 6.56$) و به ازای زوایای مختلف جریان α ($0°,7.5°,15°,30°$) برای نسبت‌های مختلف طول به عرض کانال (3.56, 2.25, 1.13) می‌باشد. در تحقیقات ریندلز و گالیور (۱۹۸۳) سه نوع گرداب مورد بررسی قرار گرفت. گرداب نوع سه (گرداب مستغرق که به سطح آب نمی‌رسد) ، گرداب نوع پنج (گرداب مستغرق که به سطح آب می‌رسد ولی به طور کامل تا دهانه آبگیر توسعه نمی‌یابد) و گرداب نوع شش (گرداب کاملا توسعه یافته تا دهانه آبگیر). همانگونه که در قسمت‌های قبل اشاره گردید، تجربیات و آزمایشات مختلف نشان داده است که مشکلات ناشی از جریان‌های گردابی در گرداب‌های نوع پنج و شش رخ می دهد (ورمین، ۱۹۹۹).

در این تحقیق، ابتدا اطلاعات مورد نظر ومقادیر مختلف Fr و S/D از منبع ذکر شده استخراج گردید. قبل از ورود به مرحله کار با مدل نزدیک‌ترین همسایگی، بهتر است داده‌ها نرمال شوند. برای این کار می‌توان با استفاده از رابطه(۳) ، داده‌هایی نرمال شده با میانگین صفر و انحراف معیار یک تبدیل کرد:

$$X' = \frac{x-\bar{x}}{\sigma(x)} \qquad (۳)$$

Archive of SID

www.SID.ir

که در آن X' داده‌های نرمال شده و $\bar{x}$ میانگین داده‌ها و $\sigma(x)$ بیانگر انحراف معیار داده‌ها می‌باشد.

پس از محاسبه فاصله بین داده‌های آزمایش و تعلیم و مرتب کردن نتایج به صورت صعودی، تعداد k از نزدیک‌ترین همسایه‌ها جهت براورد وضعیت هر گروه از داده‌ها انتخاب می‌گردد. نحوه تعیین k به این صورت می‌باشد که مدل را با توجه به مقادیر مختلف k مورد ارزیابی قرار می‌گیرد و با توجه به تکنیک Cross-Validation و تعیین مجموع مربعات خطا، مقداری از k که کمترین خطا را داشته باشد، به عنوان k انتخاب شده برای مدل تعیین می‌گردد و ومدل براساس آن مقدار k عمل می‌کند. در نهایت با استفاده از یک سری شاخص های آماری، عملکرد مدل مورد ارزیابی و اعتبارسنجی قرار می‌گیرد.

یکی از شاخص‌های آماری که برای ارزیابی مدل‌ها از آن استفاده می‌شود، ضریب همبستگی پیرسون می‌باشد که به صورت زیر تعریف می‌شود.

$$r=\frac{n[\sum_{i=1}^{n}y_i x_i]-[\sum_{i=1}^{n}y_i][\sum_{i=1}^{n}x_i]}{\sqrt{n[\sum_{i=1}^{n}y_i^2-(\sum_{i=1}^{n}y_i)^2][n\sum_{i=1}^{n}x_i^2-(\sum_{i=1}^{n}y_i)^2]}} \quad (٤)$$

که در آن y_i مقادیر پیش‌بینی شده برای داده iام و x_i مقادیر مشاهده شده برای داده iام و n تعداد داده ها می باشد.

مقادیر ضریب همبستگی همواره بین [۱- و ۱] می‌باشد و قضاوت براساس این ضریب آسان می‌باشد. با این حال باید توجه داشت که ضریب همبستگی نمی‌تواند به تنهایی شاخص مناسبی برای ارزیابی مدل باشد. چرا که ممکن است در یک مدل فرضی مقادیر پیش‌بینی شده و مشاهده شده دارای اختلاف قابل ملاحظه‌ای باشند ولی این اختلافات به گونه‌ای باشد که از یک روند یکنواختی پیروی نمایند.

لذا می‌توان از شاخص‌های آماری کمی دیگری نیز در کنار ضریب همبستگی برای براورد دقت مدل استفاده کرد. در این مطالعه از دو ضریب جذر میانگین مربعات خطا (RSME) و کارایی مدل (EF) نیز برای ارزیابی مدل استفاده شد. نحوه محاسبه این دو ضریب به ترتیب در روابط (۵) و (۶) آورده شده است.

$$RMSE=[\frac{\sum_{i=1}^{n}(x_i-y_i)^2}{n}]^{0.5} \quad (۵)$$

$$EF=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2-\sum_{i=1}^{n}(y_i-x_i)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad (۶)$$

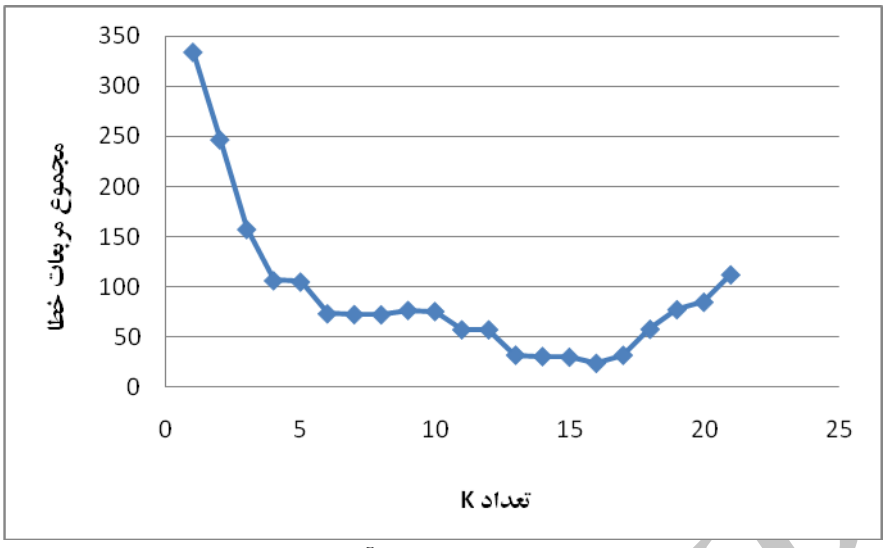
که در آن y_i مقادیر پیش‌بینی شده برای داده iام و x_i مقادیر مشاهده شده برای داده iام و n تعداد داده‌ها و $\bar{x}$ مقدار میانگین داده‌ها می‌باشد.

نتایج و بحث

با توجه به داده‌های موجود، برای مدل کردن آن‌ها بر مبنای مدل نزدیک‌ترین همسایگی از نرم‌افزار WEKA (ویتن و فرنک، ۲۰۰۵) استفاده شد. نرم‌افزار WEKA یک نرم‌افزار Open Source در زمینه داده کاوی می‌باشد که تحت زبان برنامه‌نویسی جاوا عمل می‌کند. اطلاعات مربوط به Fr، S/D، α، L/W، نوع گرداب و وضعیت خطر به صورت شاخص‌های مورد بررسی (Attributes) برای نرم‌افزار تعریف شد و پس از کدنویسی و تهیه فایل با فرمت arff (فرمت فایل ورودی نرم‌افزار WEKA)، از مدل k- نزدیک‌ترین همسایگی برای طبقه‌بندی داده‌ها استفاده شد. در ابتدا مقدار بهینه k با استفاده از تکنیک Cross-Validation تعیین شد. در شکل(۳) میزان دقت در تکنیک Cross-Validation جهت تعیین مقدار k بهینه براساس آماره مجموع مربعات خطا (SSE) نمایش داده شده است. همان‌طور که در شکل(۳) مشاهده می‌شود، مقدار K=16 دارای کمترین میزان خطا می‌باشد لذا این مقدار k برای مدل نزدیک‌ترین همسایگی استفاده می‌شود.

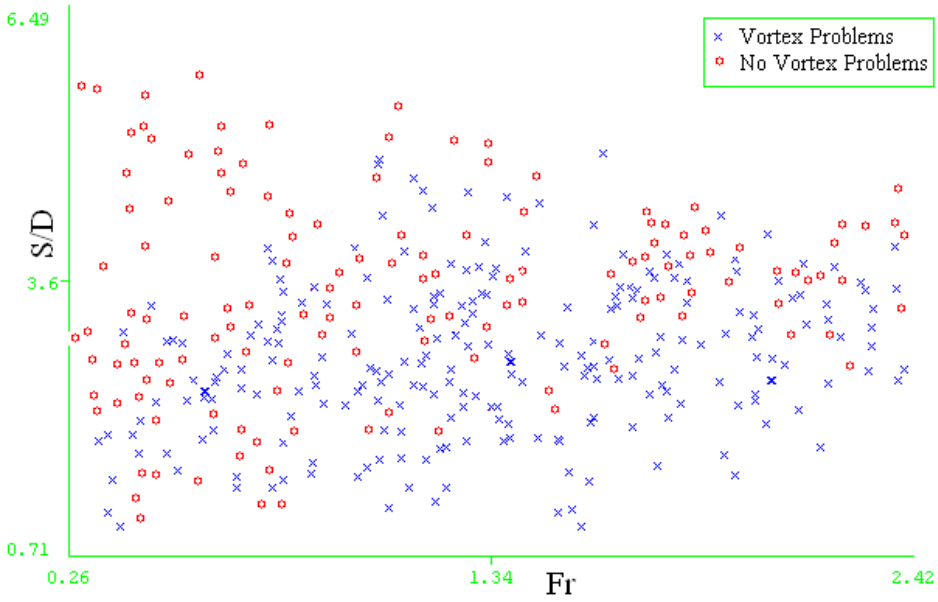
Archive of SID

www.SID.ir



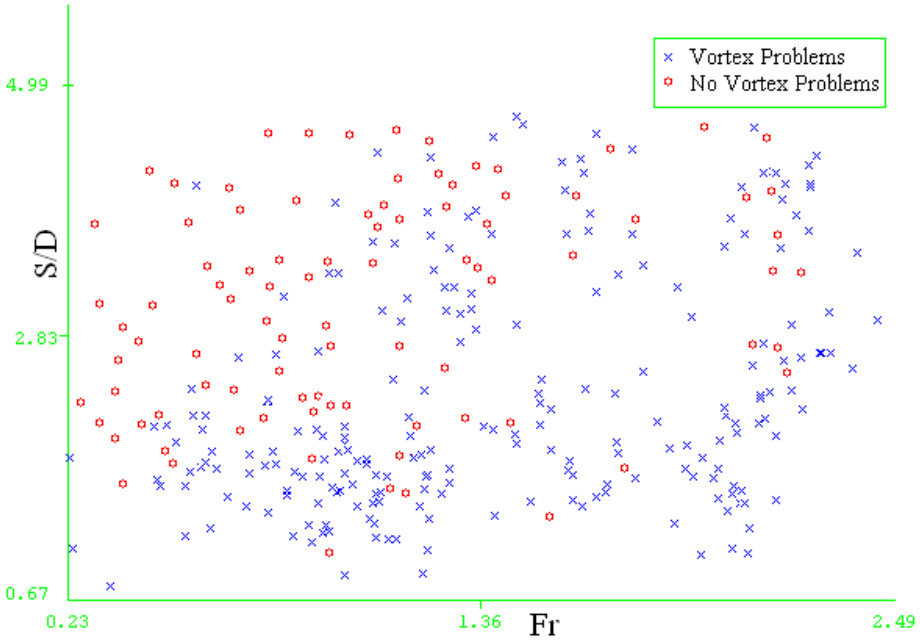
شکل(۳): تعیین میزان k بهینه براساس آماره مجموع مربعات خطا

پس از اتمام محاسبات، احتمال وقوع خسارات ناشی از جریان‌های گردابی برای آبگیر قائم تخمین زده شد. در شکل‌های ۴ تا ۶ وضعیت وقوع خطر ناشی از گرداب‌ها به ازای مقادیر مختلف Fr و S/D براساس نسبت‌های مختلف L/W نمایش داده شده است. همانگونه که از شکل‌های ارائه شده قابل مشاهده است، به ازای تمامی مقادیر L/W در نواحی بیشتری خطر آسیب ناشی از جریان‌های گردابی وجود دارد. با افزایش مقدار عدد فرود احتمال خطر وقوع گرداب‌های مخرب بیشتر می‌شود و در نواحی با مقادیر S/D و Fr بالا، خطر وقوع گرداب‌های مخرب بیشتر است.

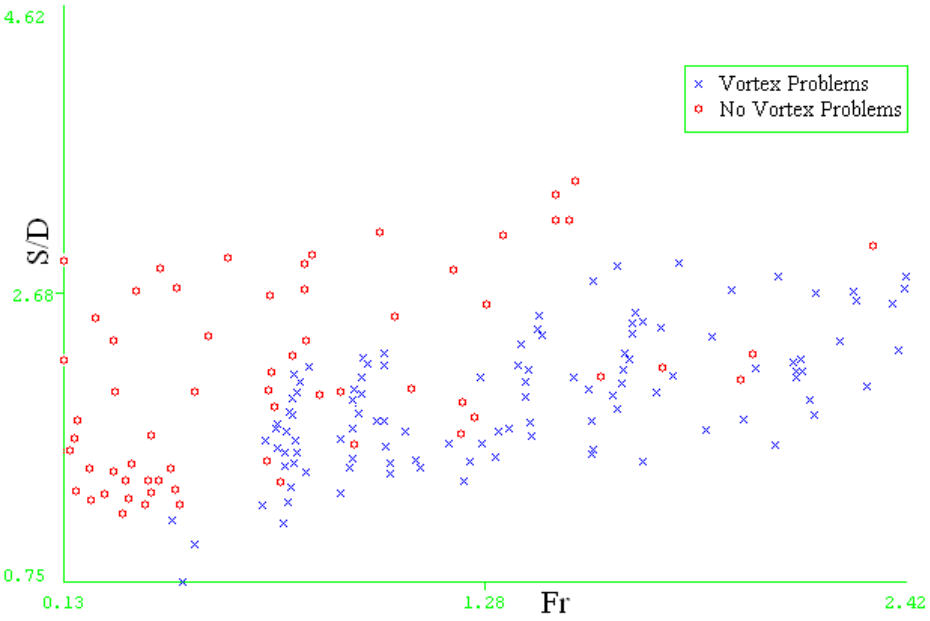


شکل(۴): وضعیت خطر وقوع جریان گردابی در آبگیر قائم به ازای L/w=1.13

www.SID.ir



شکل(۵): وضعیت خطر وقوع جریان گردابی در آبگیر قائم به ازای L/w=2.25



شکل(۶): وضعیت خطر وقوع جریان گردابی در آبگیر قائم به ازای L/w=3.56

به منظور ارزیابی و تعیین میزان کارایی و دقت مدل از سه پارامتر آماری ضریب همبستگی، جذر میانگین مربعات خطا و ضریب کارایی مدل استفاده گردید. مقادیر محاسبه شده این پارامترها در جدول(۱) آورده شده است.

جدول(۱): پارامترهای آماری محاسبه شده برای تعیین میزان کارایی و دقت مدل نزدیک ترین همسایگی

ضریب کارایی مدل (EF)	جذر میانگین مربعات خطا (RMSE)	ضریب همبستگی (r)
0.827	0.105	0.912

www.SID.ir

با توجه به ضریب همبستگی، هرچه مقدار این ضریب بیشتر از ۰/۵ باشد و به ۱ نزدیک‌تر باشد عملکرد بهتر مدل را نشان می‌دهد. با توجه به مقدار مناسب ضریب همبستگی و نیز مقدار بالای ضریب کارایی مدل، می‌توان همبستگی مناسب و نیز میزان کارایی قابل قبول مدل نزدیک‌ترین همسایگی را نشان داد و از طرف دیگر مقدار کم آماره جذر میانگین مربعات خطا، میزان خطای کم و دقت خوب مدل را نشان می‌دهد.

نتیجه‌گیری

در این مطالعه با استفاده از الگوریتم k نزدیک‌ترین همسایگی روشی برای پیش‌بینی احتمال خطر وقوع مشکلات ناشی از جریان‌های گردابی در دهانه آبگیرهای قائم مورد بررسی قرار گرفت. پارامترهای هیدرولیکی و هندسی موثر در وقوع جریان‌های گردابی مورد بررسی قرار گرفت و با الگوریتم نزدیک‌ترین همسایگی توسعه داده شد و وضعیت وقوع آسیب ناشی از گرداب‌های مخرب برای سه نسبت مختلف طول به عرض کانال ورودی آبگیر قائم (1.13, 2.25, 3.56) تخمین زده شد.

همانطور که اشاره گردید وقوع جریان‌های گردابی یک پدیده شایع و محتمل در آبگیرها به شمار می‌آید و شناخت صحیح عوامل موثر در وقوع این پدیده و نیز پیش‌بینی خطر گرداب‌ها، می‌تواند در طراحی صحیح آبگیرها موثر واقع شود. مدل ارائه شده در این تحقیق با استفاده از پارامترهای آماری ضریب همبستگی، ضریب کارایی مدل و جذر میانگین مربعات خطا مورد بررسی قرار گرفت. مقادیر مناسب ضریب همبستگی ۰/۹۱۲، ضریب کارایی مدل ۰/۸۲۷ و جذر میانگین مربعات خطا ۰/۱۰۵، نشان دهنده عملکرد خوب و دقت مناسب مدل حاضر می‌باشد.

منابع

۱. جلای، و.ج. و م. همایی. ۱۳۹۰. ارئه مدلی غیرپارامتریک با استفاده از تکنیک k نزدیک‌ترین همسایگی در براورد جرم مخصوص ظاهری خاک. نشریه علوم آب و خاک، سال پانزدهم، شماره ۵۶، ص ۱۸۹-۱۸۱.

۲. زمردیان، م. و م. شجاعیان. ۱۳۸۳. تاثیرهندسه کانال تقرب بر قدرت چرخش گرداب و ضریب آبدهی آبگیرقائم. مجله علوم کشاورزی ایران، جلد ۳۵، شماره ۳ ، ص ۶۷۸-۶۶۹.

3. Constantinescu, G.S. and V.C. Patel. 2000. "Role of turbulence model in prediction of pump-bay vortices." J. Hydraul. Eng., 126(5), 387–391.

4. Hastie, T., R. Tibshirani and J. Friedman. 2008. The Elements of Statistical Lerning. Second edition, Springer series, California.

5. Hecker, G.E. 1981. "Model-Prototype Comparison of Free Surface Vortices." J. Hydraul. Eng., 107(10), 1243–1259.

6. Knauss, J. 1987. "Swirling Flow Problems at intakes", *IAHR Hydraulic structure Manual* No 1, A.A. *Balkema*, Rotterdam, The Netherlands.

7. Rindels, A.J. and J.S. Gulliver. 1983. "An experimental study of critical submergence to avoid free-surface vortices at vertical intakes." Project Rep. No. 224, Saint Anthony Falls Hydraulic Laboratory, Univ. of Minnesota, Minneapolis.

8. Sharif, M. and D. Burn. 2007. Improved *K*-Nearest Neighbor Weather Generating Model. Journal of Hydrologic Engineering, 137 (6), 42–51.

9. Travis, Q.B. and L.W. Mays. 2011. "Prediction of Intake Vortex Risk." J. Hydraul. Eng., 126(5), 701–705.

10. Vermeyen, T.B. 1999. "Glen Canyon Dam multi-level intake structure hydraulic model study." Rep. R-99-02 , U.S. Dept. of the Interior, Bureau of Reclamation, Water Resources Research Laboratory, Denver.

11. Witten, I.H. and E. Frank. 2005. "Data mining: Practical machine learn ing tools with java implementations, 2nd Ed., Morgan Kaufmann, San Francisco.

12. Xindung, W. and V. Kumar. 2009. Top Ten Algorithm in Data Mining. First edition, Taylor & Francis Group, USA.

Archive of SID

www.SID.ir

Prediction of Vortex on Vertical Intakes by K-nearest Neighbor Modeling

E. Fadaei, Gh. Barani, M.ghaeeni

Abstract
Using vertical intakes is a common method of water withdrawal from rivers or storages to supply potable or agriculture water. Formation of vortex at vertical intakes is one of serious problems encountered in vertical intakes. Vortex formation reduces the efficiency of structure and causes loss of life. In this study, according to hydraulic conditions and geometry of vertical intakes, vortex formation was predicted by the K-nearest neighbor modeling for different ratios of channel length to width (L/W). The efficiency and precision of the model was evaluated by some statistical coefficients. Appropriate values of the correlation coefficient 0.912, the efficiency of model 0.827 and the root mean square error 0.105 shows this model is suitable and efficient.

Keywords: K-nearest neighbor model, Vertical intake, Vortex, Prediction.

Archive of SID

www.SID.ir