

«یادداشت تحقیقاتی»

ارزیابی درختان رگسیونی و رگسیون تطبیقی چندمتغیره اسپلین در تخمین حداکثر عمق آبشستگی در پایین دست سرریز جامی شکل

مهرشاد صمدی^۱، ابراهیم جباری^{۲*}

۱- کارشناس ارشد سازه‌های هیدرولیکی، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران

۲- دانشیار، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران

* تهران، صندوق پستی ۱۶۷۶۵-۱۶۳

Jabbari@iust.ac.ir

چکیده- سرریزها در ساختمان یک سد به منظور تخلیه آب مازاد بر ظرفیت مخزن سد ساخته می‌شوند. در سرریزهای جامی شکل، جت آب به صورت مایل به بستر فرسایش‌پذیر پایین دست برخورد می‌کند که سبب ایجاد گودال آبشستگی می‌شود. توسعه این گودال آبشستگی ممکن است پایداری سد را به خطر بیندازد، از این رو تخمین دقیق و صحیح عمق آبشستگی یکی از مسائل مهم علم هیدرولیک است. در سال‌های اخیر ابزارهای محاسباتی نرم به منظور مدل‌سازی پدیده‌های پیچیده و غیرخطی کاربرد فراوانی داشته‌اند. لذا در این مطالعه با استفاده از الگوریتم‌های داده‌کاوی همچون درختان طبقه‌بندی و رگسیونی و رگسیون تطبیقی چندمتغیره اسپلین، برای تخمین حداکثر عمق آبشستگی در پایین دست سرریز جامی شکل استفاده شده است. برای این منظور با استفاده از ۹۵ داده آزمایشگاهی و پارامترهای بی‌بعد، این مدل‌ها توسعه یافته‌اند. نتایج به دست آمده نشان داد که نسبت بی‌بعد $q/\sqrt{gH^3}$ مهمترین پارامتر در تخمین عمق آبشستگی می‌باشد. علاوه بر این شاخص‌های آماری و نمودارهای پراکندگی نشان داد که مدل رگسیون تطبیقی چند متغیره اسپلین با داشتن بالاترین ضریب همبستگی $CC = 0.966$ و کمترین پارامترهای خطا $RMSE = 0.073$ و $MAE = 0.057$ دقت بالاتری نسبت به درختان رگسیونی در تخمین عمق آبشستگی در پایین دست سرریز جامی شکل دارد.

کلید واژگان: عمق آبشستگی، سرریز جامی شکل، درختان رگسیونی، رگسیون تطبیقی چندمتغیره اسپلین.

۱- مقدمه

درون سرریزها با اتلاف انرژی همراه است، ولی به دلیل سرعت و انرژی بالای جریان برخوردی به بستر فرسایش‌پذیر پایین دست سبب ایجاد گودال آبشستگی

سرریزها در ساختمان یک سد به منظور تخلیه آب مازاد بر ظرفیت مخزن سد ساخته می‌شوند. اگر چه عبور جریان از

پایین دست سرریزهای ریزشی آزاد به کار برده‌اند. نتایج به دست آمده از مطالعات آنها نشان داده است که این ابزارها با ارایه روابط ریاضی کاربردی نتایج رضایت بخشی برای تخمین عمق آب شستگی دارند.

هدف اصلی این مطالعه کاربرد روش درختان رگرسیونی و تکنیک جدید رگرسیون تطبیقی چند متغیره اسپلاین در تخمین عمق آب شستگی در پایین دست سرریزهای جامی شکل است.

۲- الگوریتم درختان طبقه بندی و رگرسیونی

الگوریتم درختان طبقه بندی و رگرسیونی^۱ برای پیش بینی مقادیر حقیقی بکار می‌رود. نتایج این الگوریتم به صورت یک درخت مرتبه‌ای دودویی^۲ است. در این مدل، درخت شاخه‌های خود را بصورت دوتایی و تنها بر اساس یک متغیر منفرد به دو گروه دیگر تفکیک می‌کند. بعضی متغیرها ممکن است به دفعات در ساختار درختی استفاده شوند و در مقابل بعضی دیگر ممکن است که اصلاً به کار نروند (Breiman et al., 1984).

۳- رگرسیون تطبیقی چندمتغیره اسپلاین

رگرسیون تطبیقی چندمتغیره اسپلاین^۳ روش داده‌کاوی نسبتاً جدیدی است که رگرسیون‌های خطی ساده را ترکیب می‌کند. این روش به وسیله تقسیم فضای جواب به بازه‌هایی از متغیرهای پیش‌بینی کننده (ورودی) و برازش یک اسپلاین (تابع پایه^۴) در هر بازه، مدل‌های رگرسیونی انعطاف‌پذیری را برای پیش‌بینی پارامتر هدف ایجاد می‌نماید. تابع پایه نشان دهنده اطلاعاتی در بردارنده یک متغیر مستقل است و ارتباط بین متغیرهای پیش‌بینی کننده و پارامتر هدف را بیان می‌کند و توسط یکی از روابط زیر بیان می‌شود (Friedman, 1991):

می‌شود که توسعه این گودال ممکن است پایداری سد را به خطر بیندازد؛ از این رو تخمین صحیح و دقیق عمق آب شستگی همواره یکی از موضوعات مهم و مورد توجه محققان علم هیدرولیک بوده است.

با توجه به اهمیت این موضوع و به دلیل مکانیزم پیچیده و گستردگی پارامترهای موثر بر آن، در طی چندین دهه مطالعات متعددی توسط محققین مختلف به منظور بررسی آب شستگی در پایین دست سازه‌های هیدرولیکی انجام شده است. از جمله مطالعات آزمایشگاهی انجام گرفته، می‌توان به بررسی آب شستگی در پایین دست کالورت لوله‌ای (نجفی، ۱۳۸۰)، بررسی آب شستگی در پایین دست جت‌های ریزشی (محبوبی، ۱۳۷۵؛ آذر، ۱۳۷۷ و رنجبر، ۱۳۸۶)، بررسی آب شستگی در پایین دست سرریز جامی شکل (اسدی سریزدی، ۱۳۷۷) و بررسی آب شستگی در پایین دست سرریز سیفونی (جعفری‌نیا، ۱۳۸۹) اشاره نمود.

امروزه، ابزارهای محاسباتی نرم به دلیل قابلیت انعطاف‌پذیری، تعمیم‌دهی و توانایی در تقریب و مدل‌سازی سیستم‌های پیچیده و غیرخطی، به طور گسترده‌ای در مسائل مختلف مهندسی استفاده می‌شوند. برای مثال از کاربردهای درختان رگرسیونی در هیدرولیک می‌توان به تخمین عمق آب شستگی در پایه‌های استوانه‌ای شکل (سبزیان‌پور و همکاران، ۱۳۸۷) و تخمین عمق آب شستگی در پایین دست سرریزها با استفاده از داده‌های میدانی (صمدی و جباری، ۱۳۹۰) اشاره کرد.

همچنین روش رگرسیون تطبیقی چند متغیره اسپلاین به عنوان یک ابزار جدید داده‌کاوی در مسائلی مانند پیش‌بینی رسوب معلق (ولی و همکاران، ۱۳۸۹)، پیش‌بینی رواناب در زیرحوضه‌ها (Sharda, et al., 2008) و تخمین میزان انتقال رسوب (معظمی گودرزی، ۱۳۹۰) به کار رفته است.

اخیراً (Samadi et al. (2012) دو الگوریتم CART و M5 درختان تصمیم را به منظور تخمین عمق آب شستگی در

1. Classification and Regression Trees (CART)
2. Binary Tree
3. Multivariate Adaptive Regression Splines (MARS)
4. Basis Function (BF)

$$d_s = f(q, H, R, \phi, d_{50}, d_w, g, \rho_s, \rho_w) \quad (۴)$$

که در این رابطه، d_s عمق آب‌شستگی نهایی اندازه‌گیری شده از سطح پایاب، q دبی در واحد عرض سرریز، H هدکل یا اختلاف سطح آب بالادست و پایین‌دست، R شعاع جام، ϕ زاویه خروجی جام نسبت به قائم، d_{50} قطر متوسط ذرات بستر، d_w عمق پایاب، g شتاب ثقل و ρ_w و ρ_s به ترتیب جرم مخصوص رسوب و آب می‌باشد (شکل ۱).

با استفاده از روش π باکینگهام و در نظر گرفتن پارامترهای H, ρ_s, g به عنوان متغیرهای تکرارشونده، معادله بدون بعد حاکم به صورت زیر قابل بیان است. همچنین با توجه به اینکه نسبت ρ_w / ρ_s ثابت است، می‌توان آن را از رابطه نهایی حذف کرد (رابطه ۵).

$$\frac{d_s}{H} = f\left(\frac{q}{\sqrt{gH^3}}, \frac{H}{d_w}, \frac{R}{H}, \frac{d_{50}}{H}, \phi\right) \quad (۵)$$

۶- توسعه مدل‌ها و ارزیابی آنها

برای توسعه مدل‌های داده‌کاو به دو مجموعه آموزش و آزمون نیاز است، بنابراین برای تعیین مجموعه آموزش دو سوم داده‌ها (۶۳ داده) به طور تصادفی انتخاب شده و یک سوم (۳۲ داده) باقی‌مانده نیز به مجموعه آزمون اختصاص داده شده است.

$$h_m = \max(0, c - X) \quad (۱)$$

$$h_m = \max(0, X - c) \quad (۲)$$

که در روابط بالا X متغیر پیش‌بینی کننده، c یک مقدار آستانه و h_m تابع اسپلاین است.

در نهایت فرم عمومی مدل MARS به شکل رابطه (۳) بیان می‌شود (Friedman, 1991).

$$Y = f(x) = \beta_0 + \sum_{m=1}^M \beta_m h_m(x) \quad (۳)$$

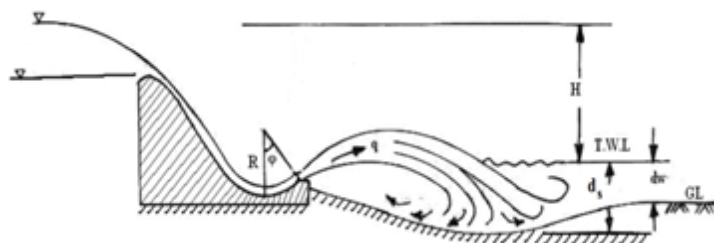
در رابطه بالا Y مقدار پیش‌بینی شده (پارامتر هدف) توسط تابع $f(x)$ می‌باشد که به صورت ترکیبی از یک مقدار ثابت اولیه β_0 و مجموع m عبارت که هر کدام از یک ضریب β_m و یک تابع پایه $h_m(x)$ تشکیل شده‌اند، تعریف می‌شود (Friedman, 1991).

۴- داده‌های مورد استفاده

در این مطالعه برای تخمین عمق آب‌شستگی در پایین‌دست سرریز جامی شکل از داده‌های آزمایشگاهی ارائه شده توسط Azmathullah, et al. (2005) (در مجموع ۹۵ داده) استفاده شده است.

۵- آنالیز ابعادی

با مروری بر مطالعات پیشین، می‌توان عمق آب‌شستگی نهایی در پایین‌دست سرریز جامی شکل را به صورت رابطه (۴) ارائه کرد (Azamathullah, et al. 2005).



شکل ۱ طرح‌واره آب‌شستگی در پایین‌دست سرریز جامی شکل (Agarwal et al., 2008)

آب‌شستگی می‌باشد. همچنین نمودار پراکنندگی برای داده‌های مشاهداتی و تخمین‌زده شده توسط این مدل در شکل ۳ رسم شده است.

۶-۲- توسعه مدل رگرسیون تطبیقی چندمتغیره اسپلاین

با استفاده از مدل MARS رابطه (۶) با ۶ تابع پایه به صورت زیر به دست آمد.

$$h_1 = \max(0, q/\sqrt{gH^3} - 0.0635183)$$

$$h_2 = \max(0, 0.0635183 - q/\sqrt{gH^3})$$

$$h_3 = \max(0, d_{50}/H - 0.0209041)$$

$$h_4 = \max(0, 0.0209041 - d_{50}/H)$$

$$h_5 = \max(0, R/H - 0.358404)$$

$$h_6 = \max(0, q/\sqrt{gH^3} - 0.0335314)$$

$$d_s/H = 0.64299 - 1.37424 \times h_1 - 5.03133 \times h_2 -$$

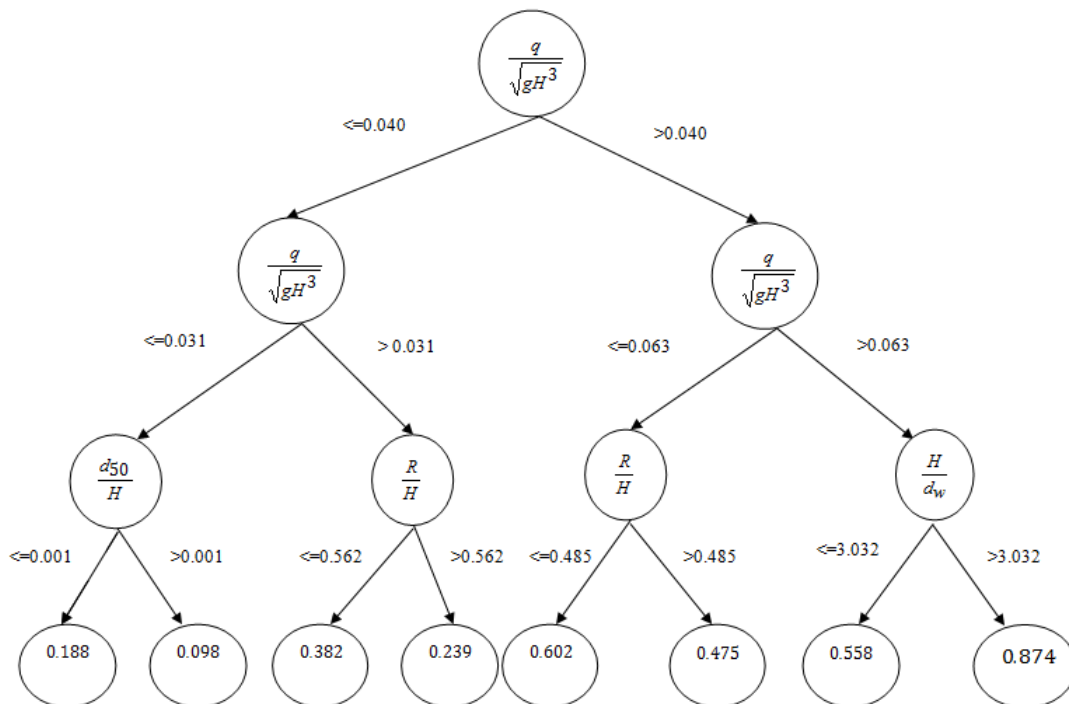
$$37.8933 \times h_3 - 12.4916 \times h_4 - 0.293061 \times h_5 +$$

$$7.29638 \times h_6 \quad (۶)$$

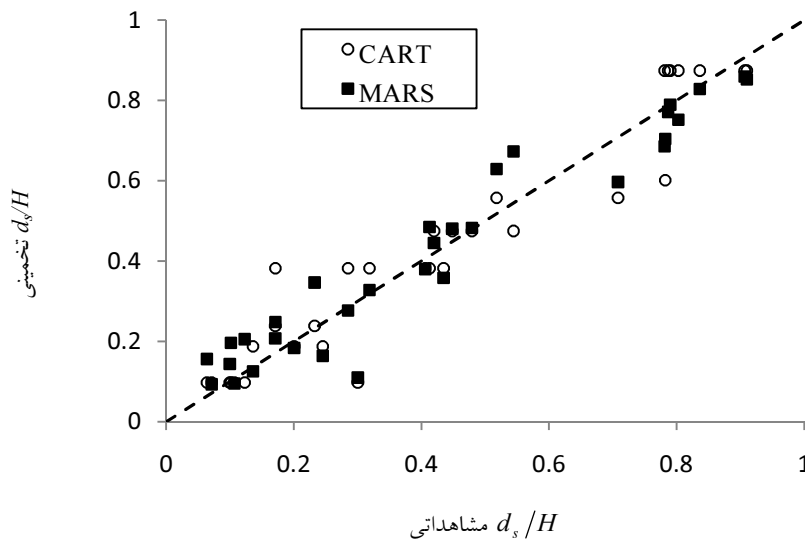
۶-۱- توسعه مدل درختان رگرسیونی

درختان رگرسیونی CART، برای تخمین پارامتر هدف در برگ‌های خود از مقادیر عددی استفاده می‌کنند و مجموعه‌ای از پارامترهای ورودی معرفی شده به مدل را در ساختار درختی به کار می‌برند. با استفاده از این الگوریتم مدلی با تعداد ۸ برگ نهایی مطابق شکل ۲ به دست می‌آید. در این شکل پارامترهایی که در داخل دایره‌ها (گره‌ها) هستند پارامترهای پایه‌ای هستند که به عنوان صفت (متغیر) آزمون انتخاب شده‌اند و اعدادی که خارج از گره‌ها روی شاخه‌ها نشان داده شده‌اند، معیارهای تقسیم هستند. علاوه بر این، اعدادی که در پایین‌ترین قسمت درخت، داخل دایره‌ها قرار دارند، بیانگر مقادیر تخمین‌زده شده d_s/H توسط مدل CART هستند.

همان‌طور که در ساختار درختی ارائه شده توسط CART قابل مشاهده است، پارامتر $q/\sqrt{gH^3}$ در گره ریشه و مرتبه اول درخت به عنوان پارامتر تقسیم انتخاب شده است که نشان‌دهنده اهمیت این پارامتر در تخمین عمق



شکل ۲ مدل درخت رگرسیونی CART به منظور تخمین d_s/H در پایین دست سرریز جامی شکل



شکل ۳ مقایسه مقادیر تخمین زده شده و مشاهداتی d_s/H توسط مدل‌های CART و MARS

جدول ۱ شاخص‌های آماری مدل MARS و CART به منظور تخمین d_s/H در پایین دست سرریز جامی شکل برای داده‌های مجموعه آزمون

مدل	CC	RMSE	MAE
MARS	۰/۹۶۶	۰/۰۷۳	۰/۰۵۷
CART	۰/۹۵۵	۰/۰۸۴	۰/۰۶۳

۷- نتیجه گیری

در این تحقیق از ابزارهای محاسباتی نرم شامل دو مدل MARS و CART برای تخمین عمق آب‌شستگی در پایین دست سرریزهای جامی شکل استفاده شد. از ویژگی‌های بارز این مدل‌ها ارائه پیش‌بینی‌ها در قالب یک سری روابط ریاضی کاربردی است که می‌توان در محاسبات آب‌شستگی از آنها استفاده کرد. نتایج به دست آمده از این مطالعه نشان داد که این مدل‌ها به خوبی قابلیت تخمین سیستم‌های پیچیده و غیرخطی مانند آب‌شستگی را دارند. علاوه بر این شاخص‌های آماری نشان داد که مدل MARS علاوه بر ارائه روابط ریاضی انعطاف پذیر، دقت بالاتری در تخمین عمق آب‌شستگی در پایین دست سرریز جامی شکل دارد.

استفاده از رابطه فوق برای محاسبه آب‌شستگی به سادگی امکان‌پذیر است. برای مثال اگر نسبت‌های $q/\sqrt{gH^3}$ ، d_{50}/H و R/H به ترتیب کمتر از $۰/۰۳۳۵۳۱۴$ ، $۰/۰۲۰۹۰۴۱$ و $۰/۳۵۸۴۰۴$ باشند، این رابطه به صورت (۷) ساده می‌شود:

$$\frac{d_s}{H} = 0.64299 - 5.03133 \times (0.0635183 - q/\sqrt{gH^3}) - 12.4916 \times (0.0209041 - d_{50}/H) \quad (۷)$$

همان‌طور که در توابع پایه به دست آمده از مدل MARS نیز مشاهده می‌شود، پارامتر $q/\sqrt{gH^3}$ بیشترین تکرار را در توابع پایه دارد که بیانگر اهمیت این پارامتر در تخمین عمق آب‌شستگی می‌باشد.

در شکل ۳ مقادیر تخمین زده شده d_s/H توسط مدل MARS با مقادیر واقعی مقایسه شده است.

همان‌طور که در جدول ۱ مشاهده می‌شود، مدل MARS با داشتن بالاترین ضریب همبستگی و کمترین شاخص‌های خطا، بالاترین دقت را در تخمین d_s/H دارد. به طوری که مقادیر به دست آمده از مدل MARS در مقایسه با مدل CART سبب افزایش ۱/۲٪ در ضریب همبستگی و کاهش حدود ۱۳٪ و ۹/۵٪ در جذر میانگین مربعات خطا و میانگین مطلق خطا می‌شود.

۸- تشکر و قدردانی

نویسندگان مقاله از شرکت مدیریت منابع آب ایران برای پذیرش این پژوهش به عنوان یک طرح تحقیقاتی با کد پروژه RIV4-89107 دانشکده مهندسی عمران دانشگاه علم و صنعت نهایت تشکر و سپاسگزاری را دارند.

۹- فهرست علائم

β_0	مقدار ثابت اولیه
c	مقدار آستانه
d_s	عمق آب شستگی نهایی اندازه گیری شده از سطح پایاب
d_{50}	قطر متوسط ذرات بستر
d_w	عمق پایاب
$f(x)$	تابع
g	شتاب ثقل
H	هد کل
h_m	تابع اسپلاین
N_t	تعداد رکوردها در گره برگ t
ρ_s	جرم مخصوص رسوب
ρ_w	جرم مخصوص آب
q	دبی در واحد عرض سرریز
R	شعاع جام
X	متغیر ورودی
Y	پارامتر هدف
$y_i(t)$	مقدار خروجی در گره برگ t
$\bar{y}(t)$	میانگین مقادیر متغیر هدف در گره برگ t می باشد.
ϕ	زاویه خروجی جام سرریز نسبت به قائم

۱۰- منابع

آذر فرادنبه، ا. (۱۳۷۷). "اثر دانه بندی مصالح بر میزان آب شستگی پایین دست سرریزهای ریزشی آزاد"، پایان نامه کارشناسی ارشد، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس.

اسدی سریزدی، م. ح. (۱۳۷۷). "تعیین اثر شعاع جام پرتابی ساده در عمق آب شستگی پایین دست سرریزها"، پایان نامه کارشناسی ارشد، دانشکده فنی و مهندسی، دانشگاه تربیت

مدرس.

جعفری نیا، ر. (۱۳۸۹). "بررسی ابعاد حفره آب شستگی در پایین دست سرریز سیفونی"، رساله دکتری، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات.

رنجبر، ح. ر. (۱۳۸۶). "تجزیه و تحلیل و بررسی آب شستگی پایین دست جت های ریزشی با مصالح غیریکنواخت"، پایان نامه کارشناسی ارشد، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس.

سزبان پور، ع.، محجوبی، ج. و جباری، ا. (۱۳۸۷). "پیش بینی عمق آب شستگی در پایه های استوانه ای شکل با استفاده از درخت های رگرسیونی و مقایسه آن با روش های تجربی"، هفتمین کنفرانس هیدرولیک ایران، دانشگاه صنعت آب و برق شهید عباسپور، تهران.

صمدی، م.، جباری، ا. (۱۳۹۰). "تخمین عمق آب شستگی در پایین دست سرریزهای جامی شکل با استفاده از داده های میدانی"، دهمین کنفرانس هیدرولیک ایران، دانشگاه گیلان، رشت.

محجوبی، ا. (۱۳۷۵). "اثر قطر مصالح بر میزان آب شستگی ناشی از جت های ریزشی آزاد"، پایان نامه کارشناسی ارشد، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران.

معظمی گودرزی، پ. (۱۳۹۰). "تخمین میزان انتقال رسوب با استفاده از مدل های درختی و مقایسه آن با سایر روش های داده کاوی"، پایان نامه کارشناسی ارشد، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران.

نجفی، ج. (۱۳۸۰). "آب شستگی در پایین دست کالورت لوله ای"، پایان نامه کارشناسی ارشد، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس.

ولی، ع.، معیری، م.، رامشت، م. ح. و موحدی نیا، ن. (۱۳۸۹). "تحلیل مقایسه عملکرد شبکه های عصبی مصنوعی و مدل های رگرسیونی پیش بینی رسوب معلق مطالعه موردی: حوضه آبخیز زاینده رود"، پژوهش های جغرافیای طبیعی، شماره ۷۱، ص.ص. ۲۱-۳۰.

Agarwal M., Goyal M. and Deo M. C. (2008).

Samadi., M., Jabbari., E. and Azamathulla., H. (2012)."Assessment of M5' model tree and classification and regression trees for prediction of scour depth below free overfall spillways. " Neural Computing and Applications: (in press)

Sharda, V. N., Prasher, S. O., Patel, R. M., Ojasvi. P. R. and Chandra. P. (2008). "Performance of Multivariate Adaptive Regression Splines (MARS) in predicting runoff in mid-Himalayan micro-watersheds with limited data", Hydrological Sciences Journal, 53(6), pp. 1165-1175.

Wang., Y. and Witten, I. H. (1997). "Induction of model trees for predicting continuous classes", In: Proceedings of the Poster Papers of the European Conference on Machine Learning, University of Economics, Faculty of Informatics and Statistics, Prague.

"Locally weighted projection regression for predicting hydraulic parameters. "Civil Engineering and Environmental Systems 27(1), pp. 71-80.

Azmathullah, H. M. and Deo, M. C., and P.B. Deolalikar, P, B. (2005). "Neural networks for estimation of scour downstream of a ski-jump bucket", Journal of Hydraulic Engineering, ASCE, 131(10), pp . 898-908.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and regression trees*, Belmont, CA: Wadsworth Statistical Press.

Friedman, J. H. (1991). "Multivariate adaptive regression splines", Annals of Statistics 19(1), pp. 1-67.