



ارزیابی عملکرد روشهای داده کاوی در پیش بینی بالاروی امواج دریا در منطقه آبخوبی سواحل

فرزاد زند^{۱*}، محمد ذونعمت کرمانی^۲

۱- دانشجوی کارشناسی ارشد مهندسی سواحل، بنادر و سازه های دریایی، گروه عمران، دانشگاه آزاد اسلامی واحد سیرجان، سیرجان، ایران
farzad.zand.1991@gmail.com

۲- دانشیار، بخش مهندسی آب، دانشگاه شهید باهنر کرمان، کرمان، ایران zounemat@uk.ac.ir

چکیده

به علت پیچیدگی فرآیندهای هیدرودینامیک در نواحی ساحلی، ابزار داده کاوی برای حل این موضوع مورد توجه قرار گرفته اند. در این مقاله، قابلیت روش GEP (به عنوان یک روش داده محور) در پیش بینی بالاروی موج بررسی شده است. به این منظور از داده های تحقیق "وان در میر و استم (Van der Meer and Stam)" استفاده شده است. داده های به کار رفته شامل پارامترهای تشابه موج، شیب بستر، تناوب موج، ضریب نفوذپذیری، اندازه نسبی ذرات بستر و بالاروی بی بعد موج بوده اند تا تخمین بالاروی موج با مدل برنامه ریزی بیان ژن (GEP) ممکن گردد. میزان کارایی مدل سازی با شاخص های COE ، $RMSE$ و R ارزیابی گردیده است و نتایج حاکی از توانایی بالای الگوریتم GEP در پیش بینی بالاروی موج بوده اند.

واژه های کلیدی: داده کاوی، بالاروی امواج دریا، الگوریتم بیان ژن، مدیریت ساحل

۱- مقدمه

منطقه آبخوبی، ناحیه ای از ساحل است که انرژی امواجی که به سمت ساحل می آیند را از طریق بالاروی و پایین روی امواج مستهلک و یا منعکس می کند. میزان فاصله قائم میان بالاترین تراز سطح آب و تراز ایستابی به عنوان تراز بالاروی موج در منطقه آبخوبی شناخته می شود [۱]. برای محاسبه و پیش بینی مقدار بالاروی امواج در ساحل روش های مختلفی وجود دارد که عمدتاً مبنای آزمایشگاهی و تجربی دارند [۲]. فرمول های تجربی بر اساس داده های آزمایشگاهی و تحلیل ابعادی به دست آمده اند، پراکندگی هایی میان مقادیر پیش بینی شده و اندازه گیری شده از خود نشان می دهند که عدم قطعیت در فرمول های ارائه شده را سبب می گردد. عدم قطعیت موجود در روابط تجربی منجر به افزایش هزینه ساخت این سازه ها می شود. شیوه های دیگری مانند روش های عددی نیز وجود دارد. روش های عددی نیاز به صرف هزینه و خطای محاسباتی بالایی دارند [۳]. زیرا در حل عددی، منابع مختلف خطا مانند منفصل سازی معادلات حاکم، حل دستگاه معادلات حاصله و انتخاب تابع تقریب وجود دارند [۴]. محدودیت ها و مشکلات شیوه های یاد شده استفاده از روش های داده کاوی نظیر مدل برنامه ریزی بیان ژن را



الزامی می‌نماید چرا که مدل‌های مبتنی بر داده‌کاوی، نیازی به شناخت دقیق پدیده با جزییات کامل ندارند بلکه به وسیله داده‌های در دسترس، قابلیت یادگیری دارند.

۲- پیشینه پژوهش

یگانه‌بختیاری و ساوالان‌پور (۱۳۹۱) با استفاده از شبکه‌های عصبی مصنوعی بالاروی امواج ناشی از باد در منطقه سواحل ساحلی را مدل‌سازی کردند. این موضوع با نگاهی جامع‌تر نسبت به پدیده بالاروی امواج در منطقه آبشویی ساحلی با توجه به رخداد فرآیندهای مشابه در هر دو حالت ساحل (نفوذپذیر و نفوذناپذیر) و با انتقال دانش بالاروی امواج (با استفاده از اثرات نفوذپذیری و طیف چگالی انرژی موج و موارد دیگر) مطالعه شده است. در ابتدای شبیه‌سازی هر دسته داده با استفاده از پارامترهای ورودی مختلف، میزان تأثیر آنها بر مقادیر پیش‌بینی شده به دست آمده و بدین ترتیب پارامترهای حاکم و مؤثر بر پدیده بالاروی امواج شناسایی شده‌اند. با استفاده از کل داده‌ها نیز شامل بستر نفوذپذیر و نفوذناپذیر مدل‌سازی با شبکه‌های عصبی صورت گرفته است. نتایج نشان داد که پارامترهای ورودی شاخص نوع شکست، نفوذپذیری، ارتفاع موج نسبی و شیب بستر دارای بیشترین اثر در مقادیر پیش‌بینی شده بالاروی امواج هستند. در انتها مشخص شد که مدل‌های مبتنی بر شبکه عصبی دقت بیشتری از روابط تجربی داشته‌اند [۱]. نجفی جیلانی و همکاران (۱۳۹۲) بالاروی موج بر روی موج‌شکن توده‌سنگی را با به کارگیری شبکه عصبی پرسپترون چند لایه (MLP) ارزیابی کردند. نتایج بیانگر نزدیک بودن داده‌های آزمایشگاهی به خروجی‌های حاصل از آموزش شبکه عصبی بوده است و تمایل میانگین مربعات خطا به عدد صفر، دقت بالای مدل شبکه عصبی را نمایان ساخته است [۲]. (۲۰۱۱) Malekmohamadi et al. به بررسی کارایی ماشین‌های بردار پشتیبان (SVMs)، شبکه‌های بیزین (BNNs)، شبکه‌های عصبی مصنوعی (ANNs) و شبکه استنتاج عصبی- فازی تطبیقی (ANFIS) در پیش‌بینی ارتفاع امواج تولید شده توسط باد پرداختند. نتایج نشان داد که مدل‌های ANN، ANFIS و SVM ارتفاع امواج را با دقت قابل قبولی پیش‌بینی کرده‌اند در حالی که روش BNNs قابل اعتماد نبوده است [۵]. (۲۰۱۶) Poate et al. با آماربرداری ویدئویی از شش نقطه ساحلی به پیش‌بینی بالاروی موج در سواحل شنی مبادرت ورزیدند. داده‌ها در بیش از ۲ سال و با صرف انرژی و توان زیاد گردآوری شده‌اند. علاوه بر این داده‌های میدانی، داده‌های مصنوعی با مدل عددی XBeach-G، مخصوص سواحل شنی ساخته شده است و هر دو دسته داده (میدانی و مصنوعی) در جهت بررسی پارامترهای هیدرودینامیکی و مورفولوژی مؤثر بر بالاروی موج به کار رفته‌اند. یک معادله بالاروی با استفاده از داده‌های مصنوعی ایجاد و با داده‌های میدانی صحت‌سنجی شده است. پارامترهای معادله بالاروی به دست آمده، به ترتیب اهمیت شامل ارتفاع مشخصه موج (H_s)، پرپود میانگین ($T_m - 1.0$)، شیب ساحل ($\tan\beta$) و اندازه ذرات (D_{50}) هستند. معادله جدید بالاروی در ساحل شنی، برازش فوق‌العاده‌ای با داده‌های مصنوعی و میدانی داشته است به طوری که بر روابط بالاروی موجود، برتری قابل توجهی داشته است [۶]. (۲۰۱۷) Atkinson et al. پارامترهای اندازه‌گیری امواج متشکل از تجاوز بالاروی ۲ درصد ($R_{2\%}$) و حداکثر بالاروی موج (R_{max}) را از ۱۱ ساحل متفاوت در جنوب شرقی استرالیا جمع‌آوری کردند و با مدل‌های تجربی متداول برای پیش‌بینی بالاروی امواج به کار بستند. مطابق با داده‌های استفاده شده، هیچ مدلی به تنهایی قادر به بهترین تخمین بالاروی در تمام سواحل نبوده است. بی‌ثباتی در پیش‌بینی با کلیه مدل‌ها مسلم می‌باشد اما آشکار است که پیش‌بینی با مدل‌های واسنجی نشده برای یک ساحل مشخص در معرض خطای قابل توجهی قرار می‌گیرد و این موضوع در طی استفاده از چنین مدل‌هایی در مدیریت سواحل بایستی در نظر گرفته شود [۷].



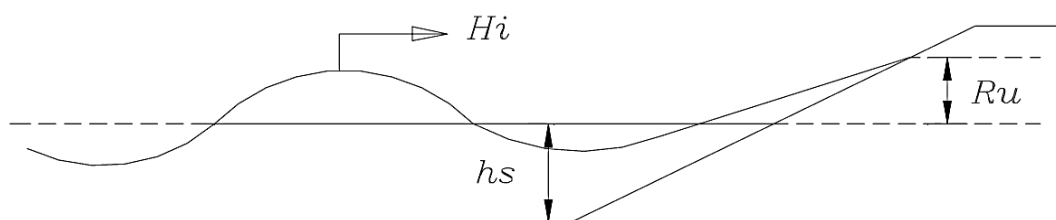
نتایج محققین در زمینه بالاروی امواج مشخص می‌سازد که روابط تجربی موجود (که عمدتاً بر اساس روش‌های برازشی بر روی داده‌های آزمایشگاهی ارائه شده‌اند) نتوانسته‌اند به طور کامل و میدانی بالاروی امواج را پیش‌بینی نمایند. در حالی که عدم نیاز به جزئیات کامل و دقیق مسأله در روش‌های داده‌کاوی (مانند مدل‌های شبکه عصبی مصنوعی، ماشین بردار پشتیبان و درخت تصمیم)، محققین را به استفاده از این روش‌ها در پیش‌بینی بالاروی موج معطوف کرده است.

۳- ساختار مدل سازی

شاخص نوع شکست موج (ξ_m)، شیب ساحل ($\cot \alpha$)، اندازه نسبی ذرات بستر (D_{85}/D_{15})، تناوب یا پرپود موج (T)، عامل نفوذپذیری بستر (SP) و بالاروی نسبی موج ($R_{u2\%}/H_s$) متغیرهای پژوهش را شامل شده‌اند که از داده‌های آزمایشگاهی مربوط به تحقیق "وان‌درمیر و استم" (Van der Meer and Stam, 1992) گردآوری شده‌اند [۸]. شکل ۱، بالاروی موج را نشان می‌دهد. مدل‌های مورد استفاده در این پژوهش، شامل روابط تجربی مستخرج شده توسط "وان‌درمیر و استم" و روش برنامه‌ریزی بیان ژن (GEP) بوده‌اند. روابط تجربی اشاره شده به تفکیک جنس بستر در معادله‌های ۱ (پوشش سنگی نفوذناپذیر) و ۲ (پوشش سنگی نفوذپذیر و همگن) آورده شده‌اند و برنامه‌ریزی بیان ژن یک روش داده‌محور فراکوشی می‌باشد که توان زیادی در شبیه‌سازی پدیده‌ها دارد.

$$\frac{R_{u2\%}}{H_s} = \begin{cases} 0.96 \xi_m & \text{for } 1.0 < \xi_m \leq 1.5 \\ 1.17 (\xi_m)^{0.46} & \text{for } \xi_m > 1.5 \end{cases} \quad (1)$$

$$\frac{R_{u2\%}}{H_s} = \begin{cases} 0.96 \xi_m & \text{for } 1.0 < \xi_m \leq 1.5 \\ 1.17 (\xi_m)^{0.46} & \text{for } 1.5 < \xi_m \leq 3.10 \\ 1.97 & \text{for } 3.10 < \xi_m < 7.5 \end{cases} \quad (2)$$



شکل ۱: طرح شماتیک از بالاروی موج [۹]

پارامترهای ξ_m ، $\cot \alpha$ ، D_{85}/D_{15} ، T_m و SP متغیرهای ورودی به مدل GEP را در بر گرفته‌اند تا محاسبه‌ی خروجی از مدل ($R_{u2\%}/H_s$) انجام گردد. T_m به معنای پرپود میانگین و ξ_m شاخص نوع شکست موج بر مبنای T_m هستند. داده‌های مورد استفاده در این پژوهش، ۲۵۶ ردیف بوده‌اند. جهت از بین بردن اثر توالی یکسان بودن برخی از مشخصات داده‌ها، در ابتدا مقادیر تصادفی (بین صفر و یک) به هر ردیف اختصاص داده شده است و بعد از آن، ردیف‌ها بر اساس مقادیر تصادفی از کوچک به بزرگ مرتب شده‌اند. در ادامه پس از به هم زدن k تایی ($k=4$)، ۳ دسته برای آموزش و دسته باقیمانده برای آزمون به کار رفته است و با ۴ بار تکرار هر ۴ دسته (کل داده‌ها) مورد آزمون قرار گرفته‌اند. با توجه به تعداد داده‌ها (۲۵۶ ردیف)، هر یک از ۴ دسته دارای ۶۴ ردیف بوده‌اند که هر ردیف، از پارامترهای ورودی به مدل و خروجی از آن تشکیل شده



است. بنابراین در هر تکرار، ۷۵ درصد داده‌ها (۳ دسته) برای آموزش و ۲۵ درصد باقیمانده (۱ دسته) جهت آزمون مدل به کار رفته است.

جهت ارزیابی روش‌های مورد بررسی در پژوهش حاضر از ضریب راندمان شبیه‌سازی (COE)، ریشه میانگین مربعات خطا (RMSE) و ضریب همبستگی (R) استفاده شده است (روابط ۳ تا ۵). مقدار اندازه‌گیری شده، مقدار پیش‌بینی شده با مدل، $(\bar{X}_m)_i$ میانگین مقادیر اندازه‌گیری شده، $(\bar{X}_p)_i$ میانگین مقادیر پیش‌بینی شده با مدل و n تعداد مقادیر اندازه‌گیری شده یا پیش‌بینی شده می‌باشند.

$$COE = 1 - \frac{\sum_{i=1}^n [(X_m)_i - (X_p)_i]^2}{\sum_{i=1}^n [(X_m)_i - (\bar{X}_m)_i]^2} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [(X_m)_i - (X_p)_i]^2} \quad (4)$$

$$R = \frac{\sum_{i=1}^n [(X_m)_i - (\bar{X}_m)_i] [(X_p)_i - (\bar{X}_p)_i]}{\left(\sum_{i=1}^n [(X_m)_i - (\bar{X}_m)_i]^2 \right)^{0.5} \left(\sum_{i=1}^n [(X_p)_i - (\bar{X}_p)_i]^2 \right)^{0.5}} \quad (5)$$

هر چه مقادیر خروجی مدل با مقادیر اندازه‌گیری شده انطباق بیشتری داشته باشند، شاخص‌های COE و R به یک نزدیکتر هستند. واضح است که قرابت بیشتر RMSE به صفر، خطای کمتر مدل‌سازی را نشان می‌دهد.

۴- تحلیل نتایج

در روابط تجربی مربوط به "وان‌درمیر و استم"، پارامتر ξ_m ، تنها ورودی بوده است در حالی که متغیرهای ξ_m ، $\cot \alpha$ ، T_m ، D_{85}/D_{15} و SP ساختار ورودی به مدل برنامه‌ریزی بیان ژن را ایجاد کرده‌اند. این تغییر (اضافه کردن پارامترهای $\cot \alpha$ ، T_m ، D_{85}/D_{15} و SP به ورودی مدل) به دلیل فراهم کردن فضای بهتر برای درک ارتباط مدل با خروجی از آن، افزایش کارایی مدل GEP در تخمین بالاروی امواج را موجب شده است. به طوری که مطابق با جدول ۱، میانگین شاخص‌های COE، RMSE و R کسب شده در مرحله آزمون با روش بیان ژن به ترتیب برابر با ۰/۶۲۵، ۰/۲۴۷ و ۰/۸۰۲ بوده است.

جدول ۱: نتایج روابط تجربی و روش برنامه‌ریزی بیان ژن در برآورد بالاروی موج

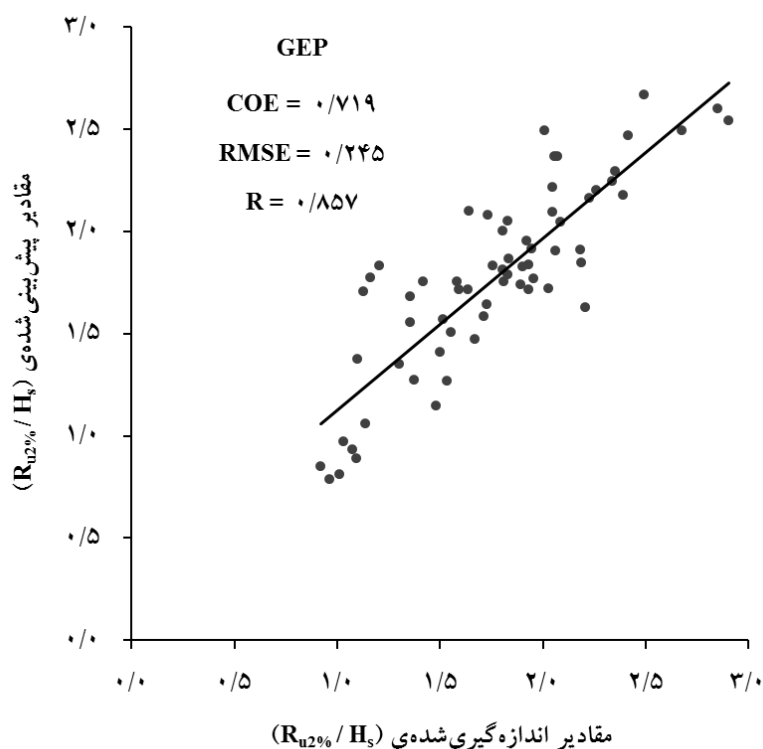
مرحله آزمون			مرحله آموزش			دسته	مدل
R	RMSE	COE	R	RMSE	COE		
۰/۸۳۵	۰/۲۲۲	۰/۵۸۹	۰/۷۹۲	۰/۲۸۲	۰/۵۶۷	۱	Van der Meer and Stam
۰/۷۵۱	۰/۲۶۲	۰/۵۵۶	۰/۷۷۴	۰/۲۶۲	۰/۵۹۹	۲	
۰/۷۱۵	۰/۳۲۴	۰/۴۲۳	۰/۸۲۹	۰/۲۴۷	۰/۶۲۶	۳	
۰/۸۳۸	۰/۲۷۶	۰/۶۴۲	۰/۷۸۱	۰/۲۶۶	۰/۵۳۸	۴	
۰/۷۸۴	۰/۲۷۱	۰/۵۵۲	۰/۷۹۴	۰/۲۶۴	۰/۵۸۲	میانگین دسته‌ها	
۰/۸۰۶	۰/۲۱۷	۰/۶۰۸	۰/۸۰۹	۰/۲۵۳	۰/۶۵۱	۱	GEP
۰/۷۹۴	۰/۲۴۲	۰/۶۲۱	۰/۸۰۳	۰/۲۵۰	۰/۶۳۷	۲	



۰/۷۵۴	۰/۲۸۵	۰/۵۵۴	۰/۸۴۱	۰/۲۱۹	۰/۷۰۶	۳
۰/۸۵۷	۰/۲۴۵	۰/۷۱۹	۰/۸۱۴	۰/۲۳۳	۰/۶۴۴	۴
۰/۸۰۲	۰/۲۴۷	۰/۶۲۵	۰/۸۱۶	۰/۲۳۸	۰/۶۵۹	میانگین دسته‌ها

شایان توجه است که بر اساس میانگین شاخص‌های ارزیابی (COE ، $RMSE$ و R)، مدل برنامه‌ریزی بیان ژن در هر دو مرحله‌ی آموزش و آزمون عملکرد مناسب‌تری نسبت به روابط تجربی داشته است. این موضوع به اهمیت آموزش و درک مدل GEP از مدل‌سازی اشاره دارد که سبب شده است دقت بیشتری در تخمین بالاروی موج در مرحله آزمون نیز داشته باشد.

هر چه تخمین بالاروی موج دقیق‌تر باشد مقادیر اندازه‌گیری شده و پیش‌بینی شده با مدل، فاصله کمتری از خط رگرسیون دارند. بنابراین در شکل ۲، مدل برنامه‌ریزی بیان ژن بهترین عملکرد در برآورد بالاروی نسبی موج را داشته است به طوری که دارای بیشترین مقدار ضریب راندمان ($COE=0.719$) در مرحله آزمون بوده است.



شکل ۲: نمودار پراکنش مقادیر اندازه‌گیری شده و پیش‌بینی شده با مدل برتر (مدل GEP، دسته ۴، مرحله آزمون)

۵- نتیجه‌گیری

برآورد بالاروی امواج دریا با روش‌های داده‌کاوی، موضوع مطالعه حاضر را در بر گرفته است. اولین گام در استفاده از روش‌های داده‌کاوی، تأمین داده‌های مربوط به پدیده مورد بررسی است. از آنجا که اندازه‌گیری داده‌های بالاروی موج به امکانات گسترده‌ی سخت‌افزاری و انسانی وابسته است، گردآوری داده‌های گزارش شده توسط دیگر محققین، پایه و گام نخست



تحقیق حاضر بوده است تا داده‌های مورد نیاز و مناسب انتخاب شوند. نتایج تخمین بالاروی موج، نشان از برتری قابل توجهی مدل برنامه‌ریزی بیان ژن بر روابط تجربی پیشنهاد شده توسط "وان‌درمیر و استم" داشته است به نحوی که با کسب بیشترین میانگین ضریب راندمان، بهترین عملکرد را از خود به جای گذاشته است.

مراجع

- [۱] یگانه‌بختیاری، ع.، ساوالان‌پور، ت. مدل‌سازی بالاروی امواج در منطقه آبشویی ساحلی با استفاده از شبکه‌های عصبی مصنوعی، نهمین همایش بین‌المللی سواحل، بنادر و سازه‌های دریایی، تهران، ۸-۱۰ آذر، ۱۳۹۱.
- [۲] نجفی‌جیلانی، ع.، ذاکری‌نیری، م.، ربیعی، ا. کاربرد شبکه عصبی مصنوعی در پیش‌بینی بالاروی موج در موج‌شکن‌های توده‌سنگی، دوازدهمین کنفرانس هیدرولیک ایران، دانشگاه تهران، ۷-۸ آبان، ۱۳۹۲.
- [۳] یگانه‌بختیاری، ع.، حمزه‌ضیابری، س.م. پیش‌بینی بالاروی امواج در منطقه سواش ساحلی با استفاده از الگوریتم M5 درخت تصمیم، هشتمین کنگره ملی مهندسی عمران، دانشکده صنعتی نوشیروانی بابل، ۱۷-۱۸ اردیبهشت، ۱۳۹۳.
- [۴] دولتشاهی‌پیروز، م.، ظهوری، ش. حل تحلیلی بالاروی امواج بلند ناشکنا روی سواحل چند شیبه، نشریه دانشکده فنی، جلد ۳۸، شماره ۲، صص. ۲۱۹-۲۲۹، تیر ۱۳۸۳.
- [5] Malekmohamadi I, Bazargan-Lari MR, Kerachian R, Nikoo MR, Fallahnia M. Evaluating the efficacy of SVMs, BNs, ANNs and ANFIS in wave height prediction. *Ocean Engineering*; 38(2-3): 487-۴۹۷, ۲۰۱۱.
- [6] Poate TG, McCall RT, Masselink G. A new parameterisation for runup on gravel beaches. *Coastal Engineering*; 117: 176-190, 2016.
- [7] Atkinson AL, Power HE, Moura T, Hammond T, Callaghan DP, Baldock TE. Assessment of runup predictions by empirical models on non-truncated beaches on the south-east Australian coast. *Coastal Engineering*; 119: 15-31, 2017.
- [8] van der Meer JW, Stam CJM. Wave runup on smooth and rock slopes of coastal structures. *Journal of Waterway, Port, Coastal, and Ocean Engineering*; 118(5): 534-550, 1992.
- [۹] شیریان، ن.، شفیعی‌فر، م.، آق‌تومان، پ.، چگینی، و. مطالعه مدل آزمایشگاهی و ارائه روابط جدید جهت تعیین میزان پایین‌آمدگی امواج نامنظم از موج‌شکن‌های شکل‌پذیر، مجله تحقیقات منابع آب ایران، سال ۳، شماره ۱، صص. ۴۵-۵۵، بهار ۱۳۸۶.