



پیش‌بینی پارامترهای امواج ناشی از باد در بندر امیرآباد به کمک درخت‌های تصمیم رگرسیونی

جواد محجوبی^۱، امیر اعتماد شهیدی^۲

۱-۲- دانشگاه علم و صنعت ایران، دانشکده عمران، تلفن: ۰۲۱-۷۳۹۱۳۱۷۰

پست الکترونیکی: j_mahjoobi@civileng.iust.ac.ir

خلاصه

تخمین امواج ناشی از باد به دلیل انرژی بالای آن‌ها از اهمیت خاصی در مهندسی سواحل و بنادر برخوردار می‌باشد. در این مقاله پیش‌بینی پارامترهای امواج ناشی از باد در آب‌های دور از ساحل بندر امیرآباد با استفاده از درختان تصمیم رگرسیونی انجام شده است. بدین منظور از اطلاعات میدانی باد و موج ثبت شده در فاصله زمانی ۲۰۰۲/۳/۲۰ تا ۲۰۰۳/۳/۱۹ استفاده شده است. خصیصه‌هایی که برای ساخت مدل درختی استفاده شده است، شامل سرعت و جهت باد به عنوان ویژگی‌های پیش‌بینی کننده و ارتفاع و پرپود موج به عنوان خصیصه‌های هدف می‌باشند. بررسی‌های انجام شده نشان می‌دهند که درختان تصمیم رگرسیونی دارای دقت قابل قبول در پیش‌بینی مشخصات امواج ناشی از باد در بندر امیرآباد می‌باشند.

کلمات کلیدی: پیش‌بینی امواج، بندر امیرآباد، داده‌کاوی، درخت تصمیم

مقدمه

امواج از جمله پارامترهای مهم در طراحی سازه‌های ساحلی و دریایی محسوب می‌شوند. در حال حاضر، مشخصات امواج ناشی از باد در دریاهای کشور با استفاده از روش‌های تجربی، اطلاعات ماهواره‌ای و اجرای مدل‌های عددی تعیین می‌شود. روش‌های تجربی، ساده‌تر و سریع‌تر بوده و به همین دلیل در کشور ما نیز تاکنون بیشتر استفاده شده‌اند. مبنای اصلی روش‌های تجربی یک سری روابط بین پارامترهای بدون بعد موج است. در این روش‌ها بر اساس قوانین مکانیک سیالات، اعداد بدون بعدی ساخته شده و ضرایب تجربی روابط با استفاده از اندازه‌گیری‌های میدانی تعیین می‌شوند. روش‌های تجربی معمولاً در بررسی‌های اولیه در طراحی پروژه استفاده می‌شوند. در دهه‌های اخیر با توسعه رایانه و بالا رفتن توان پردازنده‌ها، مدل‌های ریاضی (عددی) پیچیده‌ای برای پیش‌بینی مشخصات امواج ناشی از باد تهیه شده‌اند. مدل‌های نسل به نسل طیفی بر پایه ایده‌های رشد طیفی موج بنا نهاده شده و فرآیندهای فیزیکی گسترده‌ای را در بر می‌گیرند. به دلیل ساده بودن و هزینه کم روش‌های تجربی، مهندسی مشاور و نهادهای مرتبط برای پیش‌بینی مشخصات امواج عموماً از این روش‌ها استفاده می‌کنند. در صورتیکه این روش‌ها چون بر اساس اطلاعات میدانی مکان‌های خاصی توسعه یافته‌اند، از دقت خوبی برخوردار نمی‌باشند. از طرف دیگر استفاده از مدل‌های عددی برای پیش‌بینی موج، نسبت به روش‌های تجربی زمان و هزینه بیشتری لازم داشته و انتظار می‌رود که دقت این مدل‌ها نیز با توجه به هزینه صرف شده برای اجرای مدل، از روش‌های تجربی بیشتر باشد. در دهه‌های اخیر ابزار محاسبات نرم و سیستم‌های هوشمند به عنوان ابزار جدید مدل‌سازی سیستم‌های پیچیده مهندسی جای خود را یافته‌اند و مهندسی عمران نیز از این قاعده مستثنی نمی‌باشد. بنابراین ابزار محاسبات نرم زیادی نظیر سیستم استنباط فازی و شبکه‌های عصبی مصنوعی در بخش‌های مختلف مهندسی عمران بکار گرفته شده است. یکی از بخش‌هایی که استفاده از این ابزار در آن‌ها اخیراً مورد توجه قرار گرفته‌است در پیش‌بینی مشخصات امواج می‌باشد که در آن شبکه عصبی، سیستم استنباط فازی و ترکیب شبکه عصبی و سیستم استنباط فازی بکار گرفته شده است. در زمینه بکارگیری مدل‌های درختی در پیش‌بینی پارامترهای امواج دریا تحقیقات زیادی صورت نگرفته است. در این مقاله با استفاده از مدل‌های درختی ارتفاع و پرپود امواج ناشی از باد در بندر امیرآباد پیش‌بینی شده است.

^۱ - دانشجوی کارشناسی ارشد سازه‌های هیدرولیکی

^۲ - دانشیار دانشکده عمران



مروری بر تحقیقات گذشته

در گذشته مطالعاتی در زمینه بکارگیری روش‌های مختلف در پیش‌بینی امواج ناشی از باد در دریای خزر انجام شده است، که به برخی از آنها اشاره می‌شود:

لاری در زمینه پیش‌بینی امواج ناشی از باد در دریای خزر، از یک شبکه عصبی برای پیش‌بینی امواج با استفاده از داده‌های سرعت باد، دوره تناوب، زمان تداوم، اختلاف دمای آب و هوا و سرعت جریان استفاده کرده است. نوع این شبکه پرسپترون چند لایه یا اصطلاحاً MLP^1 می‌باشد. این شبکه قادر به تعریف انواع مختلف توابع غیرخطی بوده و می‌تواند یک نگاشت غیرخطی بین ورودی‌ها (متغیر مستقل) و خروجی شبکه (ارتفاع موج) ایجاد کند. شبکه MLP مورد نظر دارای ۵ نرون در لایه ورودی، ۱ نرون در لایه خروجی و ۱۰ نرون در لایه مخفی دوم می‌باشد. داده‌های ورودی مورد استفاده، اطلاعات ثبت شده بویه موج‌نگار نکا به مدت ۱۰ ماه بوده که در مجموع بیش از ۲۲۰۰ گزارش بوده است. در این تحقیق، علاوه بر استفاده از شبکه عصبی، با تجزیه و تحلیل آماری و استفاده از نرم‌افزار SPSS، معادله رگرسیون ارتفاع امواج برحسب متغیرهای مستقل مشخص شده است. نتایج نشان داده است که ارتفاع امواج به پارامتر زمان تداوم وزش باد بستگی ندارند و این نشان‌دهنده حاکم بودن شرایط محدودیت طول موج‌گاه می‌باشد، در صورتی که این حالت در دریای خزر کمتر بوقوع می‌پیوندد [۱].

ساجدی اسکویی از اطلاعات ایستگاه‌های هواشناسی جهت مطالعه و پیش‌بینی امواج ناشی از باد برای منطقه دریایی بندر انزلی بهره گرفت. وی با استفاده از روش‌های آماری، و استفاده از آمار سه ساعته باد با محاسبه توزیع ارتفاع موج با دوره‌های بازگشت مختلف، و بهره‌گیری از نظریه گامبل چنین نتیجه گرفت که حداکثر ارتفاع امواج ۵٫۵ متر و پررود آن‌ها ۱۰٫۵ ثانیه است که مربوط به امواجی است که از جهت شمال ایجاد می‌شوند [۲].

عظام با استفاده از سه دسته داده‌های آماری شامل موج‌نگار، داده‌های آماری و آمار بلند مدت ۳ ساعته ثبت شده باد در ایستگاه بندر انزلی به بررسی روش‌های پیش‌بینی امواج ناشی از باد پرداخت. وی سپس مدل خود را با استفاده از بررسی همزمان آمار امواج ثبت شده توسط بویه موج‌نگار و نتایج مدل، برای منطقه بندر انزلی واسنجی (کالیبره) کرد. سپس آمار دراز مدت امواج را با استفاده از مدل کالیبره شده استخراج کرد و نتایج مدل را با داده‌های ماهواره‌ای مقایسه و مورد تحلیل قرار داد. در نهایت از میان توزیع‌های آماری، توزیع مناسب برای محاسبه ارتفاع موج شاخص با دوره بازگشت‌های مختلف را ارائه کرد [۳].

همتی مشخصات امواج ناشی از باد را در بندر امیرآباد با استفاده از روش SPM و JONSWAP تعیین نمود. وی از آمار باد ایستگاه سینوپتیک بابلرس استفاده کرد. نتایج بدست آمده از پیش‌بینی‌های وی نشان می‌دهد که ۹۸٫۴۵ درصد امواج در شرایط محدودیت زمان وزش باد پیروی می‌کنند. او با استفاده از تحلیل مقادیر حدی، ارتفاع موج شاخص پیش‌بینی شده توسط روش SPM را به ازای دوره بازگشت‌های مختلف از توزیع ویبول محاسبه کرد و نتایج را با ارتفاع موج شاخص بویه و ماهواره به ازای دوره بازگشت‌های مختلف (بدون تفکیک جهت) مقایسه کرد. نتایج تحلیل حدی وی اختلاف کمی را بین مقادیر پیش‌بینی شده و اندازه‌گیری شده نشان می‌دهد [۴].

محجوبی و اعتماد شهیدی با استفاده از درختان تصمیم‌گیری و بکارگیری اطلاعات میدانی باد و موج ثبت شده منطقه نکا، ارتفاع امواج ناشی از باد را پیش‌بینی کردند. آن‌ها نتایج مدل درختی را با شبکه‌های عصبی مصنوعی مورد مقایسه قرار دادند. نتایج بدست آمده حاکی از دقت بیشتر مدل‌های درختی نسبت به شبکه‌های عصبی مصنوعی در منطقه مورد مطالعه می‌باشد [۵].

درخت‌های تصمیم‌گیری

درختان تصمیم‌گیری از نسل جدید تکنیک‌های داده‌کاوی بشمار می‌آیند که در دو دهه اخیر توسعه زیادی یافته‌اند. از این تکنیک‌ها هم می‌توان برای کشف و استخراج دانش از یک پایگاه داده‌ها و هم برای ایجاد مدل‌های پیش‌بینی استفاده نمود. درخت تصمیم‌گیری یکی از ابزارهای قوی و متداول برای دسته‌بندی و پیش‌بینی می‌باشد. درختان تصمیم‌گیر قادر به تولید توصیفات قابل درک برای انسان، از روابط موجود در یک مجموعه داده‌ای هستند و می‌توانند برای وظایف دسته‌بندی و پیش‌بینی بکار روند. این ساختار تصمیم‌گیری می‌تواند به شکل تکنیک‌های ریاضی و محاسباتی که به توصیف، دسته‌بندی و عام‌سازی یک مجموعه از داده‌ها کمک می‌کنند نیز معرفی شوند. [۶و۷]

از نقاط قوت درختان تصمیم‌گیری می‌توان به موارد زیر اشاره کرد:

- فهم مدل ایجاد شده توسط درخت تصمیم‌گیری آسان می‌باشد. عبارتی با اینکه ممکن است الگوریتم‌هایی که درخت را ایجاد می‌کنند چندان ساده نباشند ولی فهم نتایج آن آسان می‌باشد.

- درخت تصمیم‌گیری این توانایی را دارد که پیش‌بینی‌های خود را در قالب یکسری قوانین ارائه دهد.

- نیاز به محاسبات خیلی پیچیده‌ای برای دسته‌بندی داده‌ها ندارد.

- برای انواع مختلف داده‌ها از قبیل پیوسته و رده‌ای قابل استفاده می‌باشد.

- درخت تصمیم‌گیری نشان می‌دهد که کدام فیلد یا متغیرها تاثیرات مهمی در پیش‌بینی و دسته‌بندی دارند.



انواع درختان تصمیم

هنگامی که خروجی یک درخت، یک مجموعه گسسته از یک مجموعه مقادیر ممکن است؛ به آن درخت دسته‌بندی، گفته می‌شود (مثلاً مونث یا مذکر، برنده یا بازنده). این درخت‌ها تابع $X \rightarrow C$ را بازنمایی می‌کنند که در آن C مقادیر گسسته می‌پذیرد. اهداف اصلی درخت‌های تصمیم‌گیری دسته‌بندی کننده را می‌توان به صورت زیر نام برد:

- داده‌های ورودی را تا حد ممکن درست دسته‌بندی کنند.
- دانش یادگیری شده از داده‌های آموزشی را به گونه‌ای عام‌سازی کنند که داده‌های دیده نشده را با بالاترین دقت ممکن دسته‌بندی کنند.
- در صورت اضافه شدن داده‌های آموزشی جدید، بتوان به راحتی درخت تصمیم‌گیری را گسترش داد (دارای خاصیت افزایشی باشند).
- ساختار درخت حاصل به ساده‌ترین شکل ممکن باشد.

هنگامی که بتوان خروجی درخت را یک عدد حقیقی در نظر گرفت آن را، درخت همبستگی (رگرسیون)، می‌نامند، این درختان اعداد را در گره‌های برگ پیش‌بینی می‌کنند و می‌توانند از مدل رگرسیون خطی یا ثابت یا مدل‌های دیگر استفاده کنند. وظیفه‌ی یادگیری در درختان رگرسیون، شامل پیش‌بینی اعداد حقیقی بجای مقادیر دسته‌ای گسسته است. درخت $CART^1$ نامی است که به هر دو روال بالا اطلاق می‌شود. نام $CART$ مخفف کلمات درختان برگشت و دسته‌بندی است. [۸]

الگوریتم CART [۹]

در اینجا تعدادی رکورد وجود دارد که دسته آن‌ها از قبل معلوم می‌باشد (متغیر وابسته در آنها معلوم می‌باشد) هدف تهیه درختی است که بتوان بوسیله آن متغیر وابسته یا همان کلاس را برای یک رکورد جدید پیش‌بینی و تعیین نمود. روش $CART$ شاخه‌های خود را بصورت دو تایی و تنها بر اساس یک فیلد (متغیر مستقل) ایجاد می‌کند، یعنی هر گروه غیر برگ آن، به دو گروه دیگر تفکیک می‌شود. اولین قدم پاسخ به این سوال است که کدامیک از فیلدها بهترین شاخه را تولید می‌کند. بهترین ایجاد شاخه، هنگامی رخ می‌دهد که شاخه‌های حاصل طوری باشند که در هر شاخه یک کلاس بر سایر کلاس‌ها غلبه کند. معیاری که برای ارزیابی شاخه‌ها بکار می‌رود عبارت است از گوناگونی. برای محاسبه گوناگونی برای یک مجموعه از رکوردها روش‌های بسیاری وجود دارد. که در تمامی آنها گوناگونی زیاد عبارت است از مجموعه‌هایی که از کلاس‌های گوناگون در خود داشته باشند و گوناگونی کم عبارت است از مجموعه‌هایی که اعضای یک کلاس در آن بر سایر کلاس‌ها غلبه کند و بهترین نحوه ایجاد شاخه آن است که گوناگونی در مجموعه‌ها را تا حد امکان کم کند. در مرحله بعد دو شاخه وجود دارد که هر کدام دارای یکسری رکورد می‌باشند (هریک از رکوردهای گره بالاتر در یکی از شاخه‌ها قرار گرفته است). حال برای هر شاخه مثل قبل عمل می‌شود. یعنی برای هر یک از آنها دوباره یک فیلد طوری انتخاب می‌شود که بتوان بهترین شاخه‌های جدید را با حداقل گوناگونی ایجاد نمود. این مراحل آنقدر ادامه داده می‌شود تا در هر زیر شاخه گره‌ای تولید شود که ایجاد شاخه جدید در آن گره مقدار گوناگونی را کاهش قابل توجه‌ای ندهد. به این گره نهایی برگ گفته می‌شود.

برای ارزیابی درخت ایجاد شده توسط روش $CART$ یا هر روش دیگری معیارهایی وجود دارند. یکی از مهمترین و اصلی‌ترین این معیارها نرخ خطا در درخت می‌باشد. برای محاسبه نرخ خطا در درخت ابتدا باید نرخ خطا در هر شاخه بدست آید. نرخ خطا در هر برگ عبارت است از نسبت تعداد رکوردهایی که کلاس یا دسته آن‌ها درست پیش‌بینی نشده است. برای محاسبه نرخ خطای کل درخت مجموع وزنی نرخ خطاهای برگ‌ها بدست آورده می‌شود (وزن هر برگ در واقع نسبت جمعیت آن برگ به کل جمعیت رکوردها می‌باشد). کیفیت درخت حاصله نیز مهم می‌باشد. برای جلوگیری از تولید قانون‌های بی کیفیت در بعضی از شاخه‌ها قطع (هرس) صورت می‌گیرد این کار با آنکه نرخ خطا را افزایش می‌دهد ولی از ایجاد بعضی قانون‌های ناکارآمد جلوگیری می‌کند. همچنین باید به این نکته توجه داشت که باید قطع کردن به نحوی صورت گیرد که خطا از مقدار معینی بیشتر نشود.

موقعیت جغرافیایی بندر امیرآباد و داده‌های مورد استفاده

بندر امیرآباد در کناره جنوب شرقی دریای خزر، در مختصات جغرافیایی ۳۶ درجه و ۴۷ دقیقه شمالی و ۵۳ درجه و ۱۵ دقیقه شرقی واقع شده است. آمار باد استفاده شده در این تحقیق از ایستگاه هواشناسی بابلسر، که مشخصات باد شامل جهت و سرعت باد را در ارتفاع ۱۰ متری از سطح زمین ثبت کرده است، استخراج شده است. برای داده‌های موج نیز از آمار داده‌های بویه موج‌نگار بندر امیرآباد که در عمق ۱۷ متری و در موقعیت جغرافیایی ۳۶ درجه و ۵۵ دقیقه و ۲۷ ثانیه شمالی و ۵۳ درجه و ۲۴ دقیقه و ۳۷ ثانیه شرقی نصب شده است، استفاده گردیده است. مدت زمان ثبت امواج توسط این بویه از تاریخ ۲۰۰۲/۲/۱۹ تا ۲۰۰۳/۳/۱۹ بوده که این داده‌ها همانند داده‌های باد به صورت کاملاً ناپیوسته اندازه‌گیری شده‌اند. بازه‌ی زمانی اندازه‌گیری مشخصات امواج ۲ ساعت می‌باشد.

¹ - Classification and Regression Trees



ساخت مدل درختی پیش‌بینی مشخصات امواج و ارزیابی آن

با استفاده از الگوریتم CART درختان جداگانه‌ایی برای ارتفاع موج شاخص و میانگین دوره‌های تناوب عبور از صفر بالارونده ساخته شده است. از بین داده‌های موجود، دو سوم برای ساخت و یک‌سوم برای ارزیابی مدل‌ها انتخاب شده است. برای بررسی بیشتر داده‌های موجود، در جدول (۱) و (۲) مقادیر حداکثر، حداقل و میانگین پارامترهای مختلف ارائه شده است

جدول ۱- مقادیر حداکثر، حداقل و میانگین پارامترهای مختلف در داده‌های آموزشی

پارامترها	حداقل	میانگین	حداکثر
سرعت باد (متر بر ثانیه)	۰	۴,۵۱	۱۴,۹۱
جهت باد (درجه)	۰	۱۶۶,۶۶	۳۵۸,۵۵
ارتفاع موج شاخص (متر)	۰,۱۵۷	۰,۵۳۶	۲,۱۵۶
میانگین دوره‌های تناوب عبور از صفر بالارونده(ثانیه)	۱,۷	۳,۳۴۵	۷

جدول ۲- مقادیر حداکثر، حداقل و میانگین پارامترهای مختلف در داده‌های آزمون

پارامترها	حداقل	میانگین	حداکثر
سرعت باد (متر بر ثانیه)	۰	۴,۵۱	۱۴,۰۴
جهت باد (درجه)	۰,۱۱	۲۰۴,۲۲	۳۵۹,۸۵
ارتفاع موج شاخص (متر)	۰,۱۴۶	۰,۵۵۷	۱,۸۵۴
میانگین دوره‌های تناوب عبور از صفر بالارونده(ثانیه)	۱,۶۵۶	۳,۱	۵,۵۲۱

پس از ساخت درخت تصمیم رگرسیونی و ارزیابی درخت ساخته شده، بمنظور محاسبه میزان خطا در پیش‌بینی ارتفاع و پریود موج، از پارامتر انحراف (Bias) و شاخص پراکندگی (SI)، ضریب همبستگی (R) و جذر متوسط مربع خطاها (RMSE) استفاده شده است. این پارامترها مطابق روابط زیر محاسبه شده است:

$$\text{Bias} = \bar{y} - \bar{x} \quad (1)$$

$$\text{SI} = \frac{\sqrt{\frac{1}{n} \sum ((y_i - \bar{y}) - (x_i - \bar{x}))^2}}{\bar{x}} \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (y_i - x_i)^2} \quad (3)$$

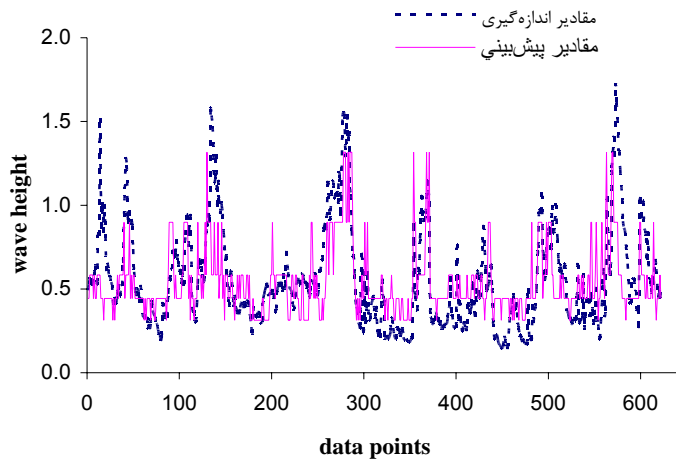
$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (4)$$

مقادیر محاسبه شده پارامترهای خطا در جدول (۳) و مقایسه مقادیر اندازه‌گیری شده و پیش‌بینی شده ارتفاع و پریود موج در اشکال (۱ و ۲) مشاهده می‌شود.

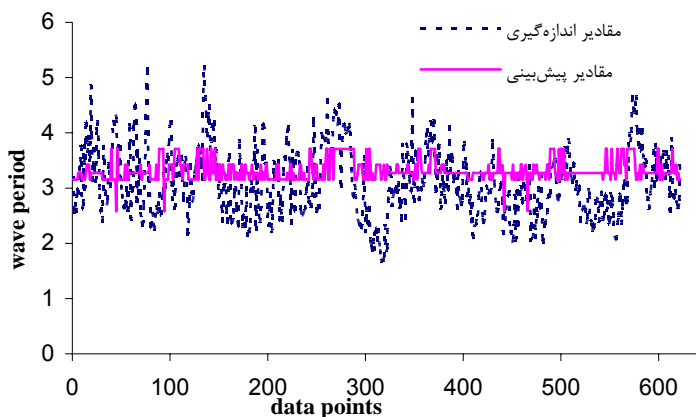


جدول ۳- مقادیر پارامترهای مختلف خطا در پیش‌بینی ارتفاع و پریرود موج

پارامترهای خطا	Bias	SI (%)	R	RMSE
مقدار خطای حاصل از مدل درخت تصمیم در پیش‌بینی ارتفاع موج	-۰,۰۲	۴۳	۰,۶۱	۰,۲۴
مقدار خطای حاصل از مدل درخت تصمیم در پیش‌بینی پریرود موج	۰,۲۱	۲۱	۰,۲	۰,۶۶



شکل ۱- سری زمانی ارتفاع موج حاصل از درخت تصمیم در مقایسه با داده‌های بویه امیرآباد



شکل ۲- سری زمانی پریرود موج حاصل از درخت تصمیم در مقایسه با داده‌های بویه امیرآباد

در مجموع روش درخت‌های تصمیم‌گیری در پیش‌بینی مشخصات امواج ناشی از باد در بندر امیر آباد ارتفاع موج را کمی دست پایین و پریرود موج را دست بالا پیش‌بینی می‌کنند. به نظر می‌رسد مهمترین منبع ایجاد خطا در پیش‌بینی مشخصات امواج استفاده از باد منطقه بابلسر برای منطقه دریایی امیرآباد که به فاصله ۶۶ کیلومتری از آن قرار دارد، می‌باشد.

در جدول (۴) خطای پیش‌بینی ارتفاع امواج برای امواج با ارتفاع کمتر از ۱ متر و امواج با ارتفاع بیشتر از ۱ متر به تفکیک آورده شده است. با مشاهده خطاهای ذکر شده در جدول (۴) و با مقایسه سری زمانی مقادیر محاسبه شده و داده‌های بویه‌ای مشاهده می‌شود که ارتفاع محاسبه شده از روش درخت تصمیم‌گیری در مقادیر پایین (ارتفاع کمتر از ۱ متر) تطابق بهتری با داده‌های بویه‌ای دارند، در حالی که در مقادیر بالاتر (ارتفاع بیشتر از ۱ متر) اختلاف ارتفاع موج محاسبه شده با مقادیر اندازه‌گیری شده بیشتر است.



جدول ۴- مقادیر پارامترهای مختلف خطا در پیش‌بینی ارتفاع امواج

پارامترهای خطا	Bias	SI (%)	R	RMSE
مقدار خطای پیش‌بینی ارتفاع امواج کمتر از ۱ متر	۰,۰۲	۰,۴۰	۰,۴۸	۰,۱۹
مقدار خطای پیش‌بینی ارتفاع امواج بیشتر از ۱ متر	۰,۳۸۷	۰,۴۰	۰,۱۸	۰,۴۸

برای مقایسه نتایج به دست آمده از این تحقیق از مرجع [۱۰] استفاده شده است که در آن دقت و صحت روش‌های نیمه‌تجربی SMB، CEM، SPM و روش طیفی JONSWAP در آب‌های دور از ساحل بندر امیرآباد مورد ارزیابی قرار گرفته است. داده‌های استفاده شده در مرجع [۱۰] با داده‌های به کار برده شده در این تحقیق تقریباً یکسان می‌باشند. نتایج حاکی از آن است که به طور میانگین خطای حاصل از پیش‌بینی مشخصات امواج توسط روش‌های تجربی حدود دو برابر خطای مدل درختی می‌باشد. به عنوان مثال جذر متوسط مربع خطاهای (RMSE) روش SPM که در مرجع مذکور به عنوان مناسبترین روش جهت پیش‌بینی مشخصات امواج انتخاب شده است، در پیش‌بینی ارتفاع موج ۰,۲۶ متر و دوره تناوب ۲ ثانیه می‌باشد.

نتیجه‌گیری

در این مقاله با استفاده از درخت‌های تصمیم رگرسیونی ارتفاع و پرپود امواج ناشی از باد در آب‌های دور از منطقه ساحل بندر امیرآباد پیش‌بینی شده است. روش درخت‌های تصمیم‌گیری در پیش‌بینی مشخصات امواج ناشی از باد ارتفاع موج را کمی دست پایین و پرپود موج را دست بالا پیش‌بینی کرده‌اند. همچنین ارتفاع محاسبه شده از روش درخت تصمیم‌گیری در مقادیر پایین (ارتفاع کمتر از ۱ متر) تطابق بهتری با داده‌های بویه‌ای دارد، در حالی که در مقادیر بالاتر (ارتفاع بیشتر از ۱ متر) اختلاف ارتفاع موج محاسبه شده با مقادیر اندازه‌گیری شده بیشتر است. جذر متوسط مربع خطاهای (RMSE) روش درختی در پیش‌بینی ارتفاع موج ۰,۲۶ متر و دوره تناوب موج ۰,۶۶ ثانیه می‌باشد. مهمترین منبع ایجاد خطا در پیش‌بینی مشخصات امواج استفاده از باد منطقه بابلسر برای منطقه دریایی امیرآباد که به فاصله ۶۶ کیلومتری از آن قرار دارد، می‌باشد. همچنین بررسی‌های انجام شده، نشان داده است که دقت مدل‌های درختی در حدود ۲ برابر روش‌های تجربی و طیفی می‌باشد.

مراجع

۱. لاری، ک؛ (۱۳۷۶) "پیش‌بینی امواج ناشی از باد در دریای خزر". پایان‌نامه کارشناسی ارشد، دانشگاه تربیت مدرس.
۲. ساجدی، ا. (۱۳۸۲) "پیش‌بینی مشخصه‌های امواج ناشی از باد با استفاده از روش‌های آماری برای منطقه دریایی انزلی"، پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی، واحد تهران شمال، دانشکده علوم و فنون دریایی.
۳. عظام، م. (۱۳۸۳) "ارزیابی روش‌های پیش‌بینی امواج ناشی از باد در آب‌های دور از ساحل بندر انزلی با استفاده از داده‌های موجود و ارائه روش مناسب"، پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی، واحد تهران شمال، دانشکده علوم و فنون دریایی.
۴. همتی، س. (۱۳۸۴) "پیش‌بینی امواج ناشی از باد در منطقه دور از ساحل بندر امیرآباد با استفاده از اطلاعات بویه موج‌نگار"، پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی، واحد تهران شمال، دانشکده علوم و فنون دریایی.
۵. محجوبی، ج.، اعتماد شهیدی، ا. (۱۳۸۶) "تخمین ارتفاع امواج ناشی از باد در نکاء به کمک درختان تصمیم رگرسیونی"، اولین کنفرانس ملی داده‌کاوی، دانشگاه امیرکبیر.

6. Hand, D, Heikki, M, Padhraic, S. (2001). "Principles of Data Mining", The MIT press.
7. Kantardzic, M. (2003). "Data Mining: Concepts, Models, Methods, and Algorithms", John Wiley & Sons.
8. Mitchell, T.M. (1997) "Machine Learning" McGraw-Hill.
9. Breiman, L., Friedman, J., Olshen, R., & Stone, C. "Classification and regression trees", Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software, 1984.
۱۰. مرکز ملی اقیانوس‌شناسی، (۱۳۸۵) "تعیین اقلیم امواج ناشی از باد در آب‌های دور از ساحل بندر امیرآباد".