

## پیش‌بینی مقاومت سایشی سنگ‌های ساختمانی با استفاده از مدل‌های رگرسیونی و الگوریتم جنگل تصادفی

علی بامری<sup>۱\*</sup>، ابراهیم قاسمی<sup>۲</sup>

۱- دانشجوی دکتری مهندسی معدن، دانشکده مهندسی معدن، دانشگاه صنعتی اصفهان، اصفهان، ایران.

۲- دانشیار دانشکده مهندسی معدن دانشگاه صنعتی اصفهان، اصفهان، ایران.

دریافت: ۱۴۰۳/۰۱/۲۵ پذیرش: ۱۴۰۳/۰۳/۳۰

(\*نویسنده مسئول: a.bamery@mi.iut.ac.ir)

### چکیده

این مطالعه با هدف پیش‌بینی شاخص باهمه ( $BAV$ )، به عنوان معیاری از مقاومت سایشی سنگ‌های ساختمانی، با استفاده از مدل‌های رگرسیونی و الگوریتم‌های یادگیری ماشین انجام شده است. برای این منظور، ۱۶۰ مجموعه داده شامل پارامترهای فیزیکی سنگ‌ها از جمله درصد تخلخل ( $N$ )، جذب آب ( $Wa$ ) و سختی شور ( $SHV$ ) جمع‌آوری و تحلیل شدند. نتایج نشان داد که هر سه پارامتر تأثیر قابل توجهی بر مقاومت سایشی دارند، به طوری که با افزایش تخلخل و جذب آب، مقاومت سایشی کاهش و با افزایش سختی شور، مقاومت سایشی افزایش می‌یابد. مدل‌های رگرسیونی چندمتغیره و الگوریتم جنگل تصادفی با دقت بالایی مقدار  $BAV$  را پیش‌بینی کردند، به ویژه مدل جنگل تصادفی که توانست با شاخص‌های عملکردی بالاتر، از جمله ضریب تعیین  $R^2$  (۰/۸۲۱)، خطای ریشه میانگین مربعات نرمال شده  $NRMSE$  (۰/۱۱۴)، واریانس خطا  $VAE$  (۰/۸۲۱) و میانگین قدر مطلق خطا  $MAE$  (۴/۶۷۵) دقت بهتری ارائه دهد. این مدل‌ها نه تنها نیاز به آزمایش‌های فیزیکی پرهزینه و زمان‌بر را کاهش می‌دهند، بلکه به مهندسان کمک می‌کنند تا در انتخاب مصالح ساختمانی مناسب برای شرایط پرتردد و محیط‌های سخت، تصمیم‌گیری بهتری داشته باشند.

### واژگان کلیدی

مقاومت سایشی سنگ، یادگیری ماشین، رگرسیون، الگوریتم جنگل تصادفی، تخلخل، درصد رطوبت، سختی شور

### ۱- مقدمه

پل‌ها به کار می‌روند. با این حال، مقاومت سنگ در برابر فرسایش و سایش یکی از مسایل اساسی در ارزیابی کیفیت و کاربرد سنگ‌های ساختمانی است، زیرا قرارگیری در معرض سایش می‌تواند بر عمر و عملکرد سازه‌های ساخته شده تأثیر منفی بگذارد [۱-۳]. برای ارزیابی مقاومت سنگ‌ها در برابر عوامل مختلف، روش‌های آزمایشگاهی گوناگونی توسعه یافته‌اند. هر یک

سنگ‌های ساختمانی، به عنوان یکی از مهم‌ترین مصالح ساختمانی با تاریخچه طولانی، ویژگی‌های منحصر به فردی دارند که آنها را به گزینه‌ای ایده‌آل برای کاربردهای سازه‌ای و معماری تبدیل می‌کند. این مصالح به دلیل دوام، مقاومت و زیبایی طبیعی خود، در ساختارهای متنوعی مانند کفپوش‌ها، دیوارپوش‌ها و

می‌گیرد. در این روش، ذرات ساینده بین نمونه و دیسک دوار فولادی قرار می‌گیرند. دستگاه از یک دیسک فولادی با قطر ۷۵۰ میلی‌متر و ضخامت ۴۰ میلی‌متر ساخته شده و نمونه سنگ را با یک گیره مخصوص نگه می‌دارد. شکل ۱ دستگاه مورد استفاده در این آزمایش را نشان می‌دهد. نمونه سنگ به صورت مکعبی با ابعاد ۷۱ میلی‌متر آماده می‌شود و پیش از آزمایش، تا دمای ۱۰۵ درجه سانتی‌گراد گرم شده و سپس در دمای محیط خنک می‌شود. در این آزمایش، حجم از دست رفته نمونه تحت تاثیر سایش به عنوان معیاری برای ارزیابی مقاومت سایشی سنگ‌های ساختمانی استفاده می‌شود. روش انجام آزمایش به اینصورت است که ابتدا ۲۰ گرم پودر کربندوم در مسیر چرخش نمونه روی دیسک پاشیده می‌شود. پس از قرار دادن نمونه در قسمت نگهدارنده دستگاه، دیسک فولادی با سرعت زاویه‌ای ۳۰ دور بر دقیقه شروع به چرخش می‌کند. یک نیروی ثابت به اندازه ۲۹۴ نیوتن به صورت عمودی به نمونه اعمال و دیسک پس از ۲۲ دور و طی مسافت لغزش ۳۱ متر متوقف می‌شود. در ادامه، مسیر تمیز شده و نمونه ۹۰ درجه چرخانده می‌شود؛ این عمل ۲۰ مرتبه برای هر نمونه تکرار می‌گردد. آزمایش برای هر نمونه در ۱۶ بازه زمانی و برای هر بازه شامل ۲۲ دور انجام می‌شود. در پایان، میانگین حجم از دست رفته نمونه بر حسب میلی‌متر مکعب محاسبه و با توجه به مقادیر به دست آمده، نرخ سایش طبق رابطه‌های ۱ و ۲ تعیین می‌گردد [۱۵-۱۰].

$$\Delta V = \left( \frac{M_0 - M_1}{\rho A} \right) 50 \quad (1)$$

$$BAV = \frac{\Delta V}{S} \quad (2)$$

از این روش‌ها برای سنجش خصوصیات خاصی از سنگ‌ها طراحی شده‌اند و اطلاعات دقیقی در مورد رفتار سنگ‌ها در شرایط مختلف ارائه می‌دهند. این داده‌ها به مهندسان و متخصصان کمک می‌کنند تا سنگ‌هایی را که برای کاربردهای خاصی مناسب‌تر هستند، انتخاب کنند و اطمینان حاصل کنند که مواد استفاده شده در پروژه‌های صنعتی و عمرانی، مقاومت کافی دارند. در همین راستا، دو آزمایش مهم از میان آزمون‌های ارزیابی مقاومت سایشی سنگ‌ها، آزمون سایش آملسر<sup>۱</sup> و آزمون سایش چرخ پهن<sup>۲</sup> هستند. این آزمایش‌ها به طور خاص به ارزیابی مقاومت سنگ در برابر سایش و خراش می‌پردازند و هر یک روش‌های منحصر به فردی برای اعمال نیرو و سایش روی نمونه‌های سنگی دارند. آزمون سایش آملسر بر پایه چرخ‌های دوار طراحی شده و نمونه را در تماس با سطح ساینده قرار می‌دهد تا بتوان مقاومت آن در برابر سایش را به صورت دقیق اندازه‌گیری کرد. آزمون سایش چرخ پهن نیز از یک چرخ پهن استفاده می‌کند که با چرخیدن روی سطح سنگ، مقاومت آن را در برابر سایش و خراش مورد سنجش قرار می‌دهد [۹-۴].

آزمون باهمه<sup>۳</sup> (BAV) که یکی دیگر از روش‌های پرکاربرد برای اندازه‌گیری مقاومت سایشی سنگ‌های ساختمانی است، به صورت استاندارد برای ارزیابی سنگ‌های مورد استفاده در کف‌پوش‌ها و سطوح پرتدد استفاده می‌شود. این آزمون با شبیه‌سازی شرایط سایشی و اندازه‌گیری میزان مواد ساییده شده از سطح سنگ، معیاری از کیفیت و دوام سنگ در برابر ساییدگی ارائه می‌دهد که می‌تواند در انتخاب سنگ‌ها برای کاربردهای با تردد بالا مانند فرودگاه‌ها و ایستگاه‌های مترو، بسیار موثر باشد. این آزمایش به عنوان یکی از قدیمی‌ترین روش‌ها برای سنجش مقاومت سنگ‌های ساختمانی و بتن به کار می‌رود و از مکانیزم سایش سه‌جسمی بهره

<sup>3</sup> Böhme Abrasion Value

<sup>1</sup> Amsler-Laffon

<sup>2</sup> Wide Wheel

استفاده از رگرسیون‌های چندگانه دقت پیش‌بینی‌ها را بهبود می‌بخشد. برای نمونه، کیلیج و تیمان در سال ۲۰۰۸ از یک مدل رگرسیونی چندگانه برای ترکیب سختی و سرعت موج فشاری سنگ‌ها استفاده کردند و به نتایج دقیقی در پیش‌بینی مقاومت سایشی دست یافتند [۱۸]. اهمیت این مدل‌ها به ویژه زمانی مشهود است که انجام آزمایش‌های عملی با محدودیت‌های زمانی یا هزینه‌ای مواجه باشد، زیرا می‌توان با داده‌های محدود و روابط ریاضی به نتایج معناداری رسید. این ویژگی‌ها باعث شده است که مدل‌های رگرسیونی به عنوان ابزاری کارآمد و سریع در مطالعات مربوط به ارزیابی و انتخاب سنگ‌های ساختمانی به کار روند و زمینه‌ساز توسعه روش‌های پیشرفته‌تر مانند یادگیری ماشین شوند [۱۹]. امروزه توسعه مدل‌های یادگیری ماشین، به ویژه جنگل تصادفی<sup>۱</sup>، به محققان این امکان را داده است تا با استفاده از الگوریتم‌های پیشرفته و داده‌های مختلف، مدل‌های پیش‌بینی با دقت بالاتر ایجاد کنند. به عنوان نمونه، بایرام در سال ۲۰۲۰ از تکنیک‌های داده‌کاوی مانند جنگل تصادفی و ماشین بردار پشتیبان (SVM) برای پیش‌بینی BAV استفاده کرد و نشان داد که این روش‌ها در مقایسه با مدل‌های سنتی، دقت بیشتری دارند [۱۵]. همچنین، محمد و همکارانش در سال ۲۰۲۱ با بررسی بیش از ۲۰ نوع سنگ ساختمانی، از جنگل تصادفی برای پیش‌بینی مقاومت سایشی سنگ‌ها استفاده کرده و به نتایج قابل‌اعتمادی دست یافتند که این مدل‌ها در کاهش زمان و هزینه آزمایش‌های عملی بسیار کارآمد بودند [۲۰]. این مدل‌های یادگیری ماشین، به ویژه الگوریتم جنگل تصادفی، در پیش‌بینی مقاومت سایشی سنگ‌ها کاربردهای گسترده‌ای دارند و می‌توانند دقت بالایی را نسبت به روش‌های سنتی فراهم کنند. استفاده از یادگیری ماشین این امکان را فراهم می‌کند که بدون نیاز به انجام آزمایش‌های پرهزینه و زمان‌بر، داده‌های موجود به صورت هوشمندانه تحلیل شده و پیش‌بینی‌های دقیقی ارائه شود. با استفاده از این مدل‌ها، مهندسان می‌توانند به سرعت و با اطمینان بیشتری سنگ‌های مناسب برای پروژه‌های مختلف ساختمانی را انتخاب کنند. به کارگیری الگوریتم‌هایی مانند جنگل تصادفی که به دلیل توانایی تجزیه و تحلیل داده‌های پیچیده شناخته شده‌اند، موجب کاهش هزینه‌ها، افزایش بهره‌وری و کاهش اتلاف منابع

که در رابطه‌های ۱ و ۲:

$\Delta V$ : حجم از دست رفته نمونه بر حسب سانتی‌متر مکعب

$M_0$ : جرم اولیه نمونه بر حسب گرم

$M_1$ : جرم نمونه بعد از سایش بر حسب گرم

$\rho$ : چگالی سنگ بر حسب گرم بر سانتی‌متر مکعب

$A$ : سطح سایش یافته بر حسب متر مربع

BAV: مقاومت به سایش سانتی‌متر مکعب بر متر

S: فاصله لغزش یافتگی بر حسب متر



شکل ۱- دستگاه آزمون باهمه (BAV)

استرالکوفسکی و کونک در سال ۲۰۲۲ با استفاده از داده‌های گردآوری شده از سنگ‌های مختلف ترکیه و با بهره‌گیری از شبکه‌های عصبی مصنوعی (ANN)، مدل‌هایی برای پیش‌بینی BAV طراحی کردند که نشان دادند مدل‌های آنها در کاهش نیاز به آزمایش‌های پیچیده و هزینه‌بر موثر است [۱۶].

مدل‌های رگرسیونی به عنوان یکی از روش‌های پرکاربرد در تحلیل و پیش‌بینی خواص مکانیکی سنگ‌ها، به ویژه مقاومت سایشی، توسعه یافته‌اند. این مدل‌ها با استفاده از روابط ریاضی میان متغیرهای فیزیکی سنگ مانند چگالی، جذب آب و سختی، سعی در پیش‌بینی پارامترهای آزمون BAV دارند که از انجام آزمایش‌های زمان‌بر جلوگیری می‌کنند. به عنوان مثال، یاشار و اردوغان در سال ۲۰۰۴ با استفاده از مدل‌های رگرسیونی خطی، رابطه‌ای میان شاخص SHV و BAV ارائه دادند که به پژوهشگران امکان تخمین BAV بر اساس سختی سنگ را می‌داد [۱۷]. همچنین تحقیقات دیگری نشان داده‌اند که ترکیب داده‌های مختلف با

<sup>1</sup> Random Forest

پیچیده کمک کرده و تصمیم‌گیری دقیق‌تری در انتخاب مواد و طراحی سازه‌ها ارائه دهند.

## ۲- مواد و روش

در این مطالعه، ۱۶۰ مجموعه داده شامل اطلاعات مربوط به مقاومت سایشی و دیگر ویژگی‌های فیزیکی سنگ‌های ساختمانی از مقالات معتبر پیشین جمع‌آوری گردید [۲۵-۳۴]. این داده‌ها شامل پارامترهایی مانند مقدار BAV، درصد تخلخل (N)، جذب آب (Wa) و سختی شور (SHV) سنگ‌ها هستند که همگی از عوامل مهم در تعیین مقاومت سایشی سنگ‌ها محسوب می‌شوند. این پارامترها از آن جهت انتخاب شده‌اند که به عنوان آزمایش‌های کم‌هزینه و غیرمخرب، به سادگی در محیط‌های آزمایشگاهی قابل انجام هستند و به کاهش هزینه‌ها و جلوگیری از تخریب نمونه‌ها کمک می‌کنند. هدف از جمع‌آوری این داده‌ها، استفاده از این پارامترها برای انجام مدل‌سازی و پیش‌بینی مقدار BAV است. به عبارت دیگر، این پارامترها به عنوان متغیرهای ورودی مدل‌های رگرسیونی و یادگیری ماشین به کار گرفته شده‌اند تا با استفاده از آنها بتوان مقدار BAV را به صورت دقیق‌تری پیش‌بینی کرد. در جدول ۱ مقادیر آماری توصیفی مربوط به هر پارامتر را ارائه می‌دهد. همچنین در ابتدا داده‌ها به دو قسمت تست (۲۰ درصد) و آموزشی (۸۰ درصد) تقسیم می‌شوند تا بتوانیم در انتها مدل‌ها را نسبت به هم بتوانیم مقایسه کنیم.

جدول ۱- مقادیر آماری پارامترهای مشخص شده

پارامترها	$BAV(cm^3/50cm^2)$	SHV	Wa [%]	N [%]
کمترین مقدار	19.2	5	0.01	0.01
میانگین	55.2	18.09	1.18	0.6
بیشترین مقدار	99	89.32	29.49	18.8
انحراف معیار	17.39	14.39	6.99	4.04

برای اطمینان از کیفیت داده‌ها و شناسایی الگوها، پیش‌پردازش با استفاده از نمودارهای جعبه‌ای-خطی<sup>۱</sup> و نمودار احتمال Q-Q انجام شده است. نمودارهای جعبه‌ای-خطی برای پارامترهای SHV، Wa و N، گسترده داده‌ها، میانه

می‌شود. در نتیجه، توسعه و استفاده از یادگیری ماشین در حوزه ساخت و ساز، نه تنها به تصمیم‌گیری بهتر کمک می‌کند، بلکه به ارتقای کیفیت و دوام پروژه‌های عمرانی نیز منجر خواهد شد [۲۱].

در مطالعه ای چن و همکاران در سال ۲۰۲۴ با استفاده از الگوریتم‌های درختی شاخص قابلیت آسیاب هاردگرو را در ۱۲۹ دسته داده زغال سنگ پیش‌بینی کردند. نتایج نشان داد که الگوریتم جنگل تصادفی از قابلیت مناسبی برای پیش‌بینی برخوردار است [۲۲].

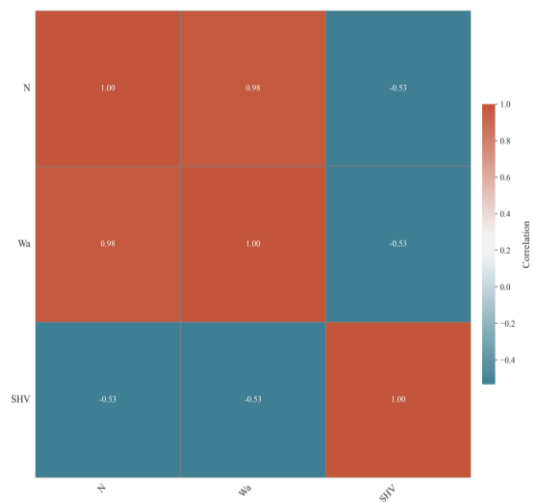
مائو و همکاران در سال ۲۰۲۴ برای تعیین مدول الاستیسیته سنگ، که برای تحلیل ژئومکانیکی و طراحی مهندسی حیاتی است، از الگوریتم جنگل تصادفی بهره گرفتند و با استفاده از ۳۵۰ نمونه سنگ رسوبی ایران (ماسه‌سنگ، سنگ آهک، مارن و گل‌سنگ) آن را با دقت بالا پیش‌بینی کردند. نتایج با ضریب تعیین ( $R^2$ ) بالا و دقت قابل توجه، قابلیت اطمینان جنگل تصادفی را تأیید کرد و نشان‌دهنده برتری آن نسبت به روش‌های سنتی بود. این مطالعه پتانسیل جنگل تصادفی را در مهندسی ژئوتکنیک برای طراحی ایمن و پایدار سازه‌ها برجسته ساخت [۲۳].

ژی و همکاران در سال ۲۰۲۵ با استفاده از الگوریتم جنگل تصادفی مقدار مقاومت فشاری تک محوره را با استفاده از پارامترها دستگاه حفاری اندازه‌گیری کردند که نتایج این تحقیق نشان دهنده دقت بالای استفاده الگوریتم جنگل تصادفی است [۲۴].

هدف اصلی این مقاله پیش‌بینی مقدار BAV با استفاده از مدل‌های مبتنی بر روش‌های رگرسیونی و الگوریتم جنگل تصادفی است. این پیش‌بینی در صرفه‌جویی در زمان و هزینه آزمایش‌های پیچیده و زمان‌بر موثر است. بر این اساس پس از جمع‌آوری ۱۶۰ سری داده از مقالات مرتبط در این حوزه مدل‌های رگرسیونی و الگوریتم جنگل تصادفی بر روی داده‌ها اعمال شد. این مدل به مهندسان کمک می‌کند تا بدون نیاز به آزمایش‌های عملی، مقاومت سایشی سنگ‌ها را پیش‌بینی کنند. استفاده از این مدل‌ها باعث تسریع در انتخاب سنگ‌های مناسب برای کاربردهای مختلف مانند کف‌پوش‌ها و سطوح پرتردد می‌شود. این روش‌ها می‌توانند به تحلیل داده‌های

<sup>1</sup> Box-and-Whisker Plot

همچنین برای درک ارتباط بهتر بین پارامترهای ورودی از ماتریس همبستگی داده‌ها استفاده شده است که در شکل ۴ مشاهده می‌شود.



شکل ۴- ماتریس همبستگی داده‌های ورودی

### ۳- توسعه مدل‌های رگرسیونی

در این مطالعه، برای پیش‌بینی مقدار BAV به عنوان شاخص مقاومت سایشی سنگ‌ها، از رگرسیون تک‌متغیره برای بررسی ارتباط بین BAV و پارامترهای مختلف مانند درصد N، درصد Wa و SHV استفاده شد. این پارامترها به دلیل ویژگی‌های فیزیکی مهم، به عنوان متغیرهای ورودی برای مدل‌سازی انتخاب شده‌اند. در جدول ۲ معادلات به دست آمده حاصل از رگرسیون تک‌متغیره آورده شده است.

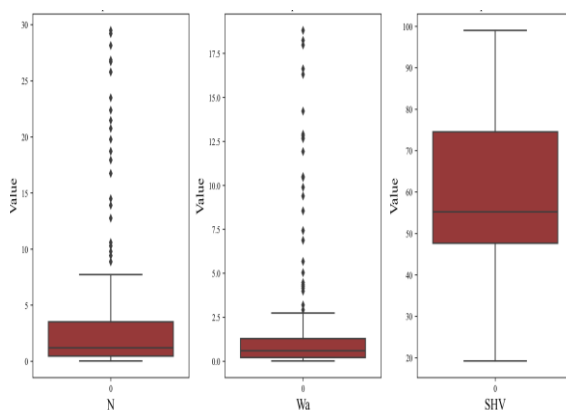
جدول ۲- معادلات رگرسیون تک‌متغیره برای پیش‌بینی

#### BAV

پارامترها	مدل رگرسیون	معادلات	$R^2$
تخلخل (N)	Linear	$BAV = 15.999N + 1.501$	0.683
جذب آب (Wa)	Linear	$BAV = 2.595Wa + 16.849$	0.679
سختی شور (SHV)	Power	$BAV = 8.554SHV^{-1.386}$	0.733

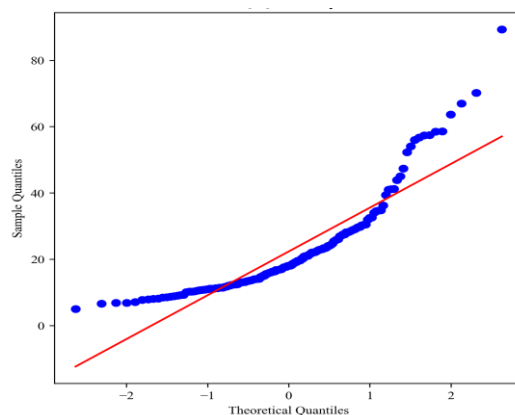
رگرسیون خطی بین BAV و درصد N نشان داد که این ارتباط مثبت و معنادار است. مقدار ضریب تعیین ( $R^2$ ) برابر با ۰.۶۸۳ است (شکل ۵). این رابطه بیانگر آن است که با افزایش N، مقدار BAV نیز افزایش و مقاومت سایشی سنگ‌ها کاهش می‌یابد.

و چارک‌های مقادیر این پارامترها را نمایش می‌دهند (شکل ۲) و به شناسایی موارد پرت که ممکن است مشکل‌ساز باشند، کمک می‌کنند. برای مثال، در نمودار مربوط به Wa و N، چندین مورد پرت مشاهده می‌شود که می‌توانند در پیش‌بینی نهایی تأثیر بگذارند. با وجود شناسایی این موارد پرت، داده‌ها برای مدل یادگیری حذف نشده‌اند و تمامی داده‌ها مورد استفاده قرار گرفته‌اند.



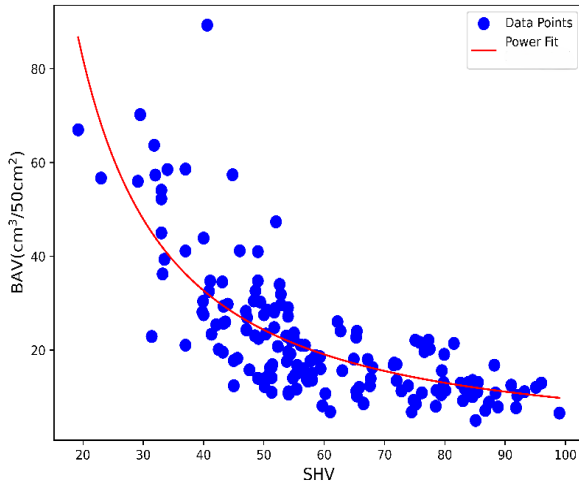
شکل ۲- نمودار جعبه‌ای-خطی SHV، Wa و N

همچنین، با توجه به شکل ۳، نمودارهای Q-Q به منظور بررسی نرمال بودن توزیع داده‌ها برای BAV استفاده شده است. نزدیک بودن نقاط به خط قطری نشان‌دهنده نزدیکی نسبی توزیع BAV به توزیع نرمال است؛ با این حال، انحراف‌هایی در نواحی انتهایی مشاهده می‌شود که بیانگر آن است که توزیع BAV به طور کامل نرمال نیستند. این انحراف ممکن است ناشی از وجود نقاط پرت یا توزیع نامتوازن داده‌ها باشد.



شکل ۳- نمودار Q-Q برای بررسی نرمال بودن توزیع باهمه

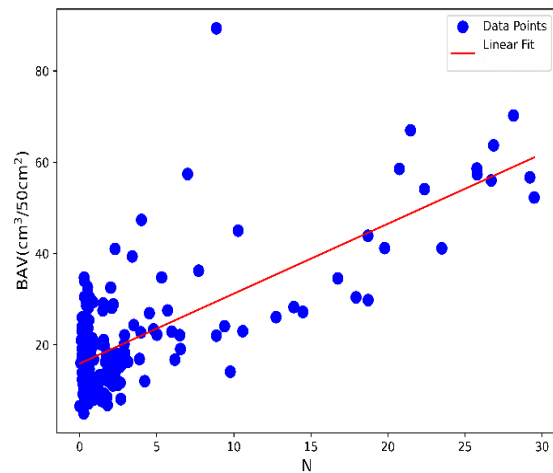
یافته‌ها بیانگر آن است که با کاهش  $SHV$ ، مقدار  $BAV$  افزایش و مقاومت سایشی سنگ‌ها کاهش می‌یابد.



شکل ۷- ارتباط بین مقاومت سایشی ( $BAV$ ) و سختی شور ( $SHV$ )

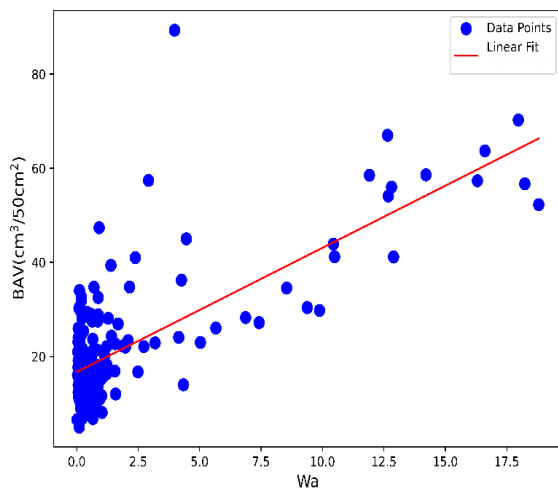
نتایج تحلیل رگرسیون حاکی از آن است که هر سه پارامتر  $N$ ،  $Wa$  و  $SHV$  تاثیر قابل توجهی بر  $BAV$  دارند و مدل‌های به دست آمده، بخش قابل توجهی از تغییرات مقاومت سایشی سنگ‌ها را پیش‌بینی می‌کنند. این تحلیل‌ها می‌توانند برای ارزیابی سریع و کارآمد سنگ‌ها در انتخاب مناسب‌ترین مصالح برای سطوح پرتردد استفاده شوند و نیاز به آزمایش‌های پرهزینه و زمان‌بر را کاهش دهند.

در ادامه این بخش، با استفاده از رگرسیون چندمتغیره، تاثیر پارامترهای فیزیکی سنگ‌های ساختمانی، شامل درصد  $N$ ،  $Wa$  و  $SHV$  بر روی مقاومت سایشی سنگ‌ها، با شاخص  $BAV$ ، بررسی شد. در جدول ۳ معادله تشکیل شده برای پیش‌بینی  $BAV$  آورده شده است. در این مدل،  $BAV$  به عنوان تابعی از سه پارامتر  $N$ ،  $Wa$  و  $SHV$  در نظر گرفته شده است. ضریب تعیین این مدل ۰٫۸۰۶ است که نشان می‌دهد این مدل توانایی بالایی در توضیح و پیش‌بینی  $BAV$  با استفاده از این سه پارامتر دارد. نتایج این مدل بیانگر آن است که با افزایش  $N$  و  $Wa$ ، مقدار  $BAV$  افزایش می‌یابد که به معنای کاهش مقاومت سایشی است، در حالی که افزایش  $SHV$  به کاهش  $BAV$  و در نتیجه افزایش مقاومت سایشی منجر می‌شود.



شکل ۵- ارتباط بین مقاومت سایشی ( $BAV$ ) و درصد تخلخل ( $N$ )

ارتباط خطی بین  $BAV$  و  $Wa$  نیز نشان‌دهنده ارتباطی مثبت و معنادار است (شکل ۶). ضریب تعیین ( $R^2$ ) این مدل برابر با ۰٫۶۷۹ است. نتایج حاکی از آن است که سنگ‌هایی با  $Wa$  بیشتر، مقاومت سایشی کمتری دارند که دلیل آن می‌تواند به افزایش نرمی و کاهش سختی در سنگ‌های با قابلیت بالاتر مرتبط باشد.

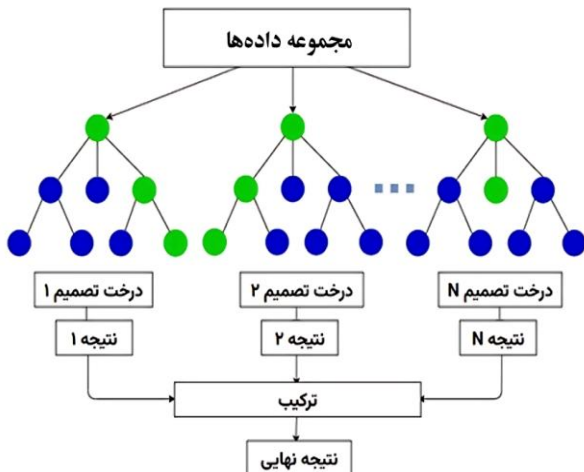


شکل ۶- ارتباط بین مقاومت سایشی ( $BAV$ ) و درصد جذب آب ( $Wa$ )

برای بررسی ارتباط بین  $BAV$  و  $SHV$  از مدل پاور<sup>۱</sup> استفاده شد و نتایج نشان‌دهنده ارتباط معنادار بین این دو پارامتر بود. ضریب تعیین این مدل نیز برابر با ۰٫۷۳۳ است (شکل ۷). این

<sup>۱</sup> Power

پیش‌بینی، هر یک از درختان به صورت مستقل و بر اساس داده‌های ورودی، نتیجه‌ای ارائه می‌دهند. در مسایل رگرسیونی، این نتایج با میانگین‌گیری از پیش‌بینی‌های همه درختان ترکیب می‌شوند و یک پیش‌بینی نهایی به دست می‌آید. در مسایل دسته‌بندی، از رای‌گیری اکثریت برای تعیین کلاس نهایی استفاده می‌شود. این ترکیب نهایی که نتیجه مشارکت همه درختان است، باعث افزایش پایداری و دقت مدل می‌شود و اطمینان می‌دهد که نتیجه پیش‌بینی از نظر آماری معتبرتر است و اثرات نویز و تغییرات ناخواسته در داده‌ها کاهش یافته‌اند (شکل ۸). در مجموع، جنگل تصادفی با فراهم آوردن امکان استفاده از ویژگی‌های متنوع و ساخت چندین درخت تصمیم مجزا، یک مدل قوی و پایدار ایجاد می‌کند که در بسیاری از کاربردهای پیچیده و داده‌های متغیر، عملکرد بسیار خوبی از خود نشان می‌دهد [۳۵].



شکل ۸- دیاگرام جنگل تصادفی

از سوی دیگر جنگل تصادفی ممکن است برای مسایل با داده‌های کوچک یا زمانی که نیاز به تفسیر مدل ساده و شفاف وجود دارد، مناسب نباشد. به عبارت دیگر مدل در جعبه سیاه قرار دارد. همچنین، به دلیل تعداد زیاد درختان تصمیم، مدل‌های جنگل تصادفی می‌توانند به راحتی قابل تفسیر نباشند که این می‌تواند در برخی کاربردها که نیاز به شفافیت و توضیح دقیق تصمیمات است، مشکل‌ساز باشد. در ادامه این مطالعه، از الگوریتم جنگل تصادفی برای پیش‌بینی مقدار  $BAV$  استفاده شد. جنگل تصادفی با ساخت مجموعه‌ای از درختان تصمیم که هر کدام با داده‌های مختلف و به صورت

### جدول ۳- معادله رگرسیون چند متغیره برای پیش‌بینی

$BAV$	
$R^2$	معادله
0.806	$BAV = 0.493 N + 0.967 Wa - 0.351 SHV + 39.189$

### ۴- الگوریتم جنگل تصادفی

الگوریتم جنگل تصادفی یکی از روش‌های بسیار پرکاربرد و محبوب در یادگیری ماشین است که به دلیل دقت بالا و توانایی در جلوگیری از بیش‌برازش<sup>۱</sup> به انتخاب مناسبی برای تحلیل داده‌های پیچیده و متغیر تبدیل شده است. الگوریتم جنگل تصادفی به دلیل دقت بالا، توانایی جلوگیری از بیش‌برازش، تحلیل داده‌های پیچیده و متغیر، کاهش حساسیت به نویز و عملکرد عالی در مسایل رگرسیونی و دسته‌بندی، به عنوان یک مدل قوی و پایدار شناخته می‌شود. این الگوریتم با استفاده از درختان تصمیم مستقل و تصادفی، پایداری مدل را افزایش داده و پیش‌بینی‌ها را دقیق‌تر و معتبرتر می‌کند. جنگل تصادفی به ویژه در کاربردهای پیچیده و داده‌های متغیر عملکرد بسیار خوبی از خود نشان می‌دهد و می‌تواند در حل مشکلات مختلف در یادگیری ماشین و تحلیل داده‌ها موثر باشد. این الگوریتم از مجموعه‌ای از درختان تصمیم<sup>۲</sup> تشکیل شده که هر یک به صورت مجزا و مستقل آموزش می‌بینند. در آغاز فرآیند، داده‌ها به دو بخش اصلی داده‌های آموزشی و داده‌های تست تقسیم می‌شوند. داده‌های آموزشی برای ساخت و آموزش درختان تصمیم استفاده می‌شوند و داده‌های تست برای ارزیابی دقت و کارایی مدل به کار می‌روند. هر درخت تصمیم در جنگل تصادفی، به صورت تصادفی از داده‌های آموزشی ساخته می‌شود؛ به این معنا که هر درخت مجموعه متفاوتی از داده‌ها را برای آموزش دریافت می‌کند و این تنوع باعث افزایش پایداری مدل می‌شود. به علاوه، در ساخت هر درخت، تنها یک زیرمجموعه تصادفی از ویژگی‌ها انتخاب می‌شود و به این ترتیب، هر درخت فقط با بخشی از ویژگی‌های اصلی به آموزش می‌پردازد. این رویکرد، علاوه بر جلوگیری از بیش‌برازش، سبب می‌شود که مدل حساسیت کمتری به نویزهای موجود در داده‌ها داشته باشد و بتواند به صورت کلی‌تر و با دقت بیشتر بر روی داده‌های جدید عمل کند. پس از ساخت و آموزش تمامی درختان، برای انجام

<sup>2</sup> Decision Trees

<sup>1</sup> overfitting

*Learn* و *RandomForest* پیاده‌سازی شده است. در تنظیمات این مدل، هایپرپارامترهایی مانند تعداد درختان<sup>۳</sup>، عمق درخت<sup>۳</sup> و حداقل نمونه‌های لازم در برگ‌ها<sup>۴</sup> با استفاده از الگوریتم *Optuna* بهینه شده اند. *Optuna* به گونه‌ای طراحی شده است که تنظیم هایپرپارامترهای الگوریتم جنگل تصادفی به شکلی انجام شود تا مدل به حداکثر دقت و عملکرد برسد. همچنین، به منظور جلوگیری از بیش‌برازش، هایپرپارامترها بر اساس ویژگی‌های داده‌ها و معیارهای ارزیابی تنظیم شده‌اند. با استفاده از این تنظیمات، جنگل تصادفی به یک ابزار دقیق و کارآمد برای پیش‌بینی مقدار *BAV* تبدیل شده است و می‌تواند تاثیر پارامترهای *Wa*، *N* و *SHV* را به طور دقیق ارزیابی کند. در جدول ۴، مقادیر تنظیم شده برای هایپرپارامترهای مدل جنگل تصادفی ارائه شده که در دستیابی به پیش‌بینی‌های بهینه و افزایش کارایی مدل موثر بوده‌اند.

جدول ۴- هایپرپارامترهای بهینه *Optuna* برای مدل جنگل

تصادفی	
مقدار	هایپر پارامترها
۲۰۳	N-estimators
۵۷	Max-depth
۱۰	min_samples_leaf
۵	min_samples_split
<i>Squared error</i>	Objective

#### ۴-۲- دقت مدل جنگل تصادفی

در این بخش، دقت مدل‌های توسعه یافته با استفاده از جنگل تصادفی و ارزیابی آن‌ها با شاخص‌های مختلف بررسی می‌شود. معیارهای ارزیابی شامل خطای ریشه میانگین مربعات نرمال شده (*NRMSE*)، واریانس خطا (*VAF*) و میانگین قدر مطلق خطا (*MAE*) هستند که به عنوان ابزارهایی برای سنجش دقت و صحت پیش‌بینی‌های مدل مورد استفاده قرار می‌گیرند. این شاخص‌ها، به ویژه در مدل‌های یادگیری ماشین مانند جنگل تصادفی، عملکرد مدل را در پیش‌بینی متغیر وابسته یعنی مقدار *BAV* که نشان‌دهنده مقاومت سایشی سنگ‌ها است، ارزیابی می‌کنند.

تصادفی آموزش دیده‌اند، یک مدل قوی و دقیق ایجاد می‌کند. این الگوریتم به دلیل استفاده از روش بوت‌استرپ<sup>۱</sup> برای نمونه‌گیری و انتخاب تصادفی ویژگی‌ها در ساخت هر درخت، می‌تواند اثر نوسانات و نویز داده‌ها را کاهش دهد و از بیش‌برازش جلوگیری کند. با ترکیب پیش‌بینی‌های تمامی درختان، مدل نهایی به دست می‌آید که به ویژه در داده‌های پیچیده و متغیر، دقت بالایی را ارائه می‌دهد. همچنین، جنگل تصادفی به ما این امکان را می‌دهد که اهمیت هر ویژگی مانند *Wa*، *N* و *SHV* را در پیش‌بینی *BAV* بسنجیم. این قابلیت در بهبود دقت و کارایی انتخاب سنگ‌های مقاوم‌تر برای کاربردهای ساختمانی اهمیت دارد و به کاهش نیاز به آزمایش‌های متعدد و پرهزینه کمک می‌کند.

#### ۴-۱- توسعه مدل به روش جنگل تصادفی

در این بخش، برای بهبود دقت پیش‌بینی مقدار *BAV* و بررسی دقیق‌تر تاثیر عوامل مختلف، از الگوریتم جنگل تصادفی استفاده شده است. جنگل تصادفی به عنوان یک روش پیشرفته در یادگیری ماشین، با ایجاد مجموعه‌ای از درختان تصمیم‌گیری که به صورت مستقل آموزش داده می‌شوند، توانایی تحلیل داده‌های پیچیده و روابط غیرخطی را دارد. ترکیب پیش‌بینی‌های هر درخت و استفاده از نمونه‌های تصادفی از داده‌ها در هر مرحله، باعث می‌شود که جنگل تصادفی نسبت به روش‌های رگرسیونی در مواجهه با داده‌های ناهمگن عملکرد بهتری از خود نشان دهد و دقت پیش‌بینی را افزایش دهد.

در این مدل، داده‌های مربوط به پارامترهای *Wa*، *N* و *SHV* برای پیش‌بینی مقدار *BAV* مورد استفاده قرار گرفته‌اند. این پارامترها به دلیل ارتباط مستقیم با ویژگی‌های فیزیکی سنگ‌ها، نقش اساسی در پیش‌بینی مقاومت سایشی آن‌ها ایفا می‌کنند. ۲۰ درصد از داده‌ها برای آزمایش و ۸۰ درصد برای آموزش مدل اختصاص یافته‌اند. همچنین، برای اعتبارسنجی بهتر داده‌ها و اطمینان از پایداری عملکرد مدل، از روش اعتبارسنجی متقاطع *k-fold* استفاده شده است. این رویکرد به بهبود تعمیم‌پذیری مدل و کاهش احتمال بیش‌برازش کمک می‌کند. در مدل‌سازی مدل جنگل تصادفی با استفاده از زبان برنامه‌نویسی پایتون و بهره‌گیری از کتابخانه‌های *Scikit*

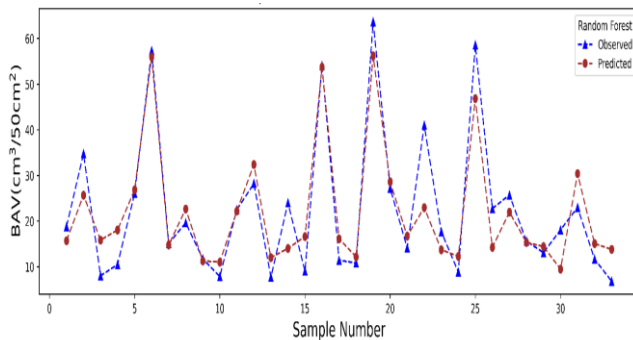
<sup>3</sup> max\_depth

<sup>4</sup> min\_samples\_leaf

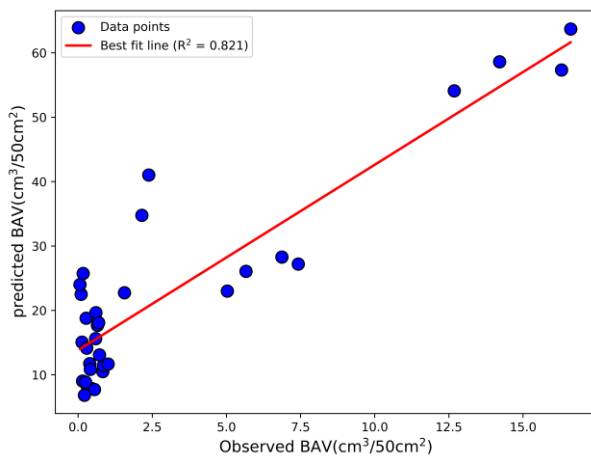
<sup>1</sup> Bootstrap

<sup>2</sup> n\_estimators

با توجه به نتایج جدول، مدل جنگل تصادفی بهترین عملکرد را در پیش‌بینی مقدار  $BAV$  نشان داده است. مقادیر کم  $MAE$  و  $NRMSE$  نشان‌دهنده خطای پایین مدل و مقادیر بالای  $VAF$  و  $R^2$  بیانگر توانایی مدل در توضیح و پیش‌بینی دقیق تغییرات داده‌های مربوط به مقاومت سایشی سنگ‌ها هستند. شکل ۹ و ۱۰ نشان‌دهنده مقادیر واقعی و پیش‌بینی شده است.



شکل ۹- مقایسه مقادیر مشاهده و پیش‌بینی شده  $BAV$  با استفاده از مدل جنگل تصادفی



شکل ۱۰- مقادیر مشاهده و پیش‌بینی شده  $BAV$  در مدل جنگل تصادفی

در ادامه خروجی نمودار وابستگی جزئی<sup>۱</sup> ( $PDP$ ) که نشان‌دهنده تاثیر ویژگی‌های مختلف بر عملکرد الگوریتم است را بررسی می‌کنیم. این تحلیل به ما کمک می‌کند که ارتباط هر ویژگی با خروجی مدل را بهتر درک کنیم [۳۶]. شکل ۱۱ خروجی حاصل از تحلیل  $PDP$  است. نمودار اول که مربوط به پارامتر  $N$  است، نشان می‌دهد که با افزایش این ویژگی،

برای محاسبه این شاخص‌ها، مقادیر  $M_i$  و  $P_i$  به ترتیب به مقدار واقعی و مقدار پیش‌بینی شده اشاره دارند و  $A_{avg}$  میانگین مقادیر واقعی است. در این مطالعه، مقادیر  $NRMSE$ ،  $VAF$  و  $MAE$  برای ارزیابی میزان نزدیکی مقادیر پیش‌بینی شده به مقادیر واقعی محاسبه می‌شوند. فرمول‌های استفاده شده برای محاسبه این شاخص‌ها به صورت زیر هستند:

$$NRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (M_i - P_i)^2}}{A_{avg}} \quad (۳)$$

$$VAF = \left[ 1 - \frac{Var(M_i - P_i)}{Var(M_i)} \right] \times 100 \quad (۴)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |M_i - P_i| \quad (۵)$$

در این فرمول‌ها، مقدار بالاتر  $VAF$  نشان‌دهنده دقت بیشتر مدل در پیش‌بینی و مقادیر پایین‌تر  $NRMSE$  و  $MAE$  نشان‌دهنده کاهش خطای مدل است. در این مطالعه، مقادیر  $NRMSE$ ،  $VAF$  و  $MAE$  برای ارزیابی مدل جنگل تصادفی و مقایسه آن با مدل‌های رگرسیونی قبلی به کار می‌روند. این شاخص‌ها به ما امکان می‌دهند تا عملکرد مدل جنگل تصادفی را در پیش‌بینی مقدار  $BAV$  با دقت بیشتری بررسی کنیم و بسنجیم که آیا این روش در شناسایی پارامترهای کلیدی برای پیش‌بینی مقاومت سایشی، از مدل‌های رگرسیونی ساده و چندگانه بهتر عمل می‌کند یا خیر. در جدول ۵ و ۶ شاخص‌های آماری عملکرد برای دو حالت آموزش و تست قرار داده شده است.

جدول ۵- شاخص‌های آماری عملکرد مدل جنگل تصادفی و رگرسیون چندمتغیره برای داده‌های آموزش

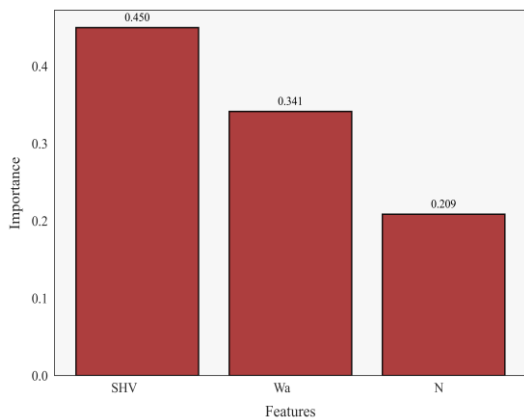
مدل	$R^2$	$NRMSE$	$VAF$	$MAE$
Random Forest	0.934	0.096	0.944	3.974
رگرسیون چندمتغیره	0.887	0.106	0.887	4.336

جدول ۶- شاخص‌های آماری عملکرد مدل جنگل تصادفی و رگرسیون چندمتغیره برای داده‌های تست

مدل	$R^2$	$NRMSE$	$VAF$	$MAE$
Random Forest	0.821	0.114	0.822	4.675
رگرسیون چندمتغیره	0.806	0.1187	0.806	4.8180

<sup>1</sup> Partial Dependence Plots

و ارزیابی دقیق‌تر اهمیت پارامترهای ورودی در مدل‌سازی فراهم می‌کند. این ویژگی با محاسبه میزان تأثیرگذاری هر پارامتر بر پیش‌بینی‌های مدل، امکان رتبه‌بندی آن‌ها را بر اساس اهمیتشان فراهم می‌سازد و به پژوهشگران کمک می‌کند تا پارامترهای ورودی را به دو دسته اصلی، یعنی مهم‌ترین و کم اهمیت‌ترین، تقسیم‌بندی کنند. در ادامه، با بهره‌گیری از قابلیت *Feature Importance* الگوریتم جنگل تصادفی، میزان اهمیت پارامترهای ورودی مورد بررسی قرار گرفته است. همان‌طور که در شکل ۱۲ مشاهده می‌شود، پارامتر *SHV* از اهمیت بیشتری نسبت به سایر پارامترهای ورودی برخوردار است. این نتیجه نشان‌دهنده آن است که *SHV* به‌عنوان یک عامل کلیدی در پیش‌بینی‌های مدل عمل می‌کند و تأثیر قابل توجهی بر خروجی‌ها دارد.

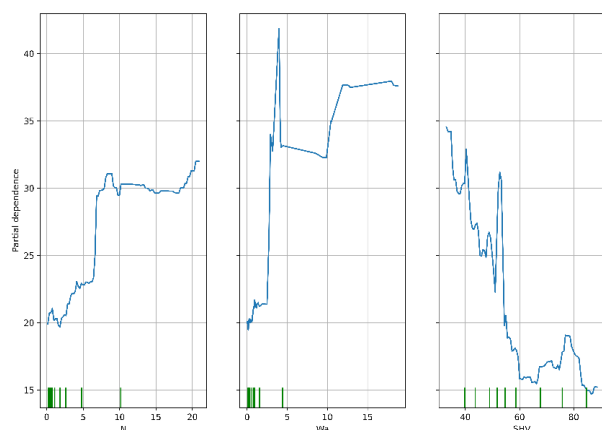


شکل ۱۲- *feature Importance* الگوریتم جنگل تصادفی

همچنین برای مشاهده مقایسه الگوریتم جنگل تصادفی و رگرسیون چندگانه از گراف تیپلور استفاده شده است. نمودار تیپلور ترکیبی از انحراف معیار، ضریب همبستگی و خطای میانگین مربعات ریشه مرکزی است که دیدگاه ارزیابی جامع‌تری را ارائه می‌دهد. هر چه یک مدل به نقطه رفرنس مشکلی نزدیکتر باشد، عملکرد آن بهتر است. با توجه به شکل ۱۳ مشاهده می‌شود که الگوریتم جنگل تصادفی در مقایسه با رگرسیون چندگانه بهتر عمل می‌کند.

خروجی مدل به طور پیوسته افزایش می‌یابد. در مقادیر پایین، تأثیر کمتری از  $N$  دیده می‌شود، اما با افزایش مقدار آن، تأثیر این ویژگی بر مدل بیشتر می‌شود. این موضوع حاکی از آن است که برای مدل، مقادیر بالاتر  $N$  اهمیت بیشتری دارند و می‌توانند بهبود خروجی مدل را به دنبال داشته باشند. پارامتر دوم،  $Wa$ ، در ابتدای نمودار تأثیر قابل توجهی بر مدل دارد. مشاهده می‌شود که با افزایش اولیه این ویژگی از مقدار صفر به حدود پنج، خروجی مدل به شدت افزایش می‌یابد؛ اما پس از رسیدن به این نقطه، نمودار به یک حالت ثابت‌تر می‌رسد. این بدان معناست که در مقادیر پایین،  $Wa$  تأثیر زیادی بر پیش‌بینی مدل دارد، اما پس از یک مقدار مشخص، افزایش بیشتر  $Wa$  تغییر چندانی در خروجی ایجاد نمی‌کند. به عبارتی، مدل نسبت به این ویژگی تا حدی به اشباع می‌رسد. پارامتر سوم،  $BAV$ ، تأثیر معکوسی بر مدل دارد. با افزایش مقدار این ویژگی، خروجی مدل به طور کلی کاهش پیدا می‌کند که نشان‌دهنده اثر منفی  $BAV$  بر پیش‌بینی مدل است. در برخی نقاط، تغییرات ناگهانی و شدید در نمودار دیده می‌شود که بیانگر حساسیت مدل به این ویژگی در بازه‌های خاصی است. روند کلی کاهش در این نمودار نشان می‌دهد که مقادیر بالاتر  $BAV$  می‌تواند با کاهش خروجی مدل همراه باشد و مدل ممکن است ترجیح دهد که این ویژگی در سطح پایین‌تری نگه داشته شود تا خروجی بهتری حاصل شود.

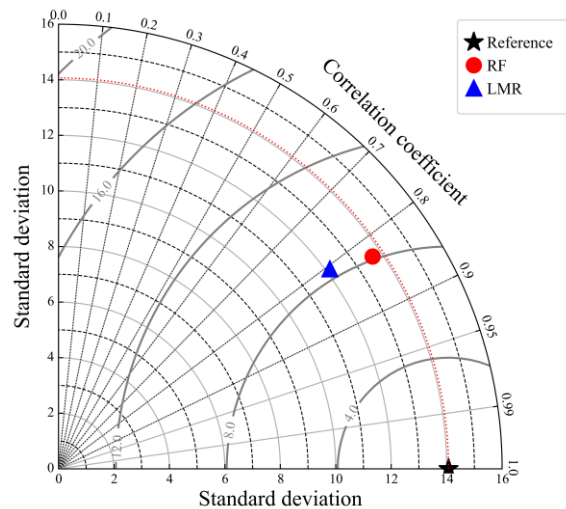
شکل ۱۱- تحلیل تأثیر ویژگی‌ها بر مدل رگرسیون جنگل تصادفی با استفاده از نمودارهای وابستگی جزئی



الگوریتم جنگل تصادفی از قابلیت کلیدی به نام *Feature Importance* برخوردار است که ابزاری قدرتمند برای تحلیل

مدل‌های رگرسیونی نشان داد. این الگوریتم توانست با کاهش حساسیت به اختلال و جلوگیری از بیش‌برازش، دقت پیش‌بینی را به طور قابل توجهی افزایش دهد. شاخص‌های ارزیابی این مدل نشان‌دهنده عملکرد برتر آن در پیش‌بینی *BAV* است. تحلیل وابستگی جزئی (*PDP*) نشان داد که پارامتر *N* تاثیر بیشتری بر *BAV* دارد و با افزایش آن، شاخص *BAV* به طور پیوسته افزایش می‌یابد. *Wa* نیز در مقادیر پایین تاثیر قابل توجهی دارد، اما در مقادیر بالاتر به نقطه اشباع می‌رسد. از سوی دیگر، *SHV* با کاهش *BAV* مرتبط است و افزایش آن به بهبود مقاومت سایشی کمک می‌کند.

مقایسه عملکرد مدل‌ها نشان داد که جنگل تصادفی به دلیل توانایی در تحلیل داده‌های پیچیده و روابط غیرخطی، بهترین ابزار برای پیش‌بینی *BAV* است. این روش با مقادیر *MAE* برابر با ۰/۶۷۵، *VAF* برابر با ۰/۸۲۲، *NRMSE* برابر با ۰/۱۱۴ و  $R^2$  برابر با ۰/۸۲۱، علاوه بر دقت بالا، به دلیل ترکیب ویژگی‌های تصادفی در ساخت درختان تصمیم و میانگین‌گیری از نتایج آن‌ها، پایداری بیشتری در مواجهه با داده‌های متغیر ارایه کرد.



شکل ۱۳- نمودار تیلور برای الگوریتم جنگل تصادفی و رگرسیون چندگانه

#### ۵- نتیجه‌گیری

این تحقیق با هدف پیش‌بینی شاخص مقاومت سایشی سنگ‌ها (*BAV*) به عنوان معیاری اساسی برای انتخاب مصالح ساختمانی، به توسعه مدل‌های رگرسیونی و الگوریتم‌های پیشرفته یادگیری ماشین پرداخته است. با جمع‌آوری داده‌های ۱۶۰ نمونه سنگ و استفاده از پارامترهای فیزیکی شامل درصد تخلخل (*N*)، جذب آب (*Wa*) و سختی شور (*SHV*)، مدلی ارایه شد که امکان پیش‌بینی سریع و دقیق این شاخص را فراهم می‌کند.

در مرحله نخست، مدل‌های رگرسیونی تک‌متغیره روابط معنادار بین هر یک از پارامترهای ورودی و *BAV* را نشان دادند. تحلیل‌ها حاکی از آن بود که افزایش *N* و *Wa* باعث افزایش *BAV* و در نتیجه کاهش مقاومت سایشی سنگ‌ها می‌شود، در حالی که *SHV* تاثیر معکوسی دارد؛ افزایش آن به کاهش *BAV* و بهبود مقاومت سایشی منجر می‌شود. این مدل‌ها با ارایه ضرایب تعیین قابل قبول، شفافیت خوبی در تحلیل تاثیرات اولیه داشتند. مدل رگرسیون چندمتغیره، با ترکیب این متغیرها، توانست دقت پیش‌بینی را افزایش دهد. با این حال، پیچیدگی داده‌ها و روابط غیرخطی میان متغیرها ضرورت استفاده از روش‌های پیشرفته‌تر را نشان داد. الگوریتم جنگل تصادفی، به عنوان یکی از روش‌های قدرتمند یادگیری ماشین، با ترکیب درختان تصمیم و استفاده از نمونه‌گیری تصادفی از داده‌ها و ویژگی‌ها، عملکرد بهتری نسبت به

## ۶- منابع

- [1] Ghannam, S., Najm, H., & Vasconez, R. (2016). Experimental study of concrete made with granite and iron powders as partial replacement of sand. *Sustainable Materials and Technologies*, 9, 1–9.
- [2] Amran, M., Murali, G., Khalid, N. H. A., Fediuk, R., Ozbakkaloglu, T., Lee, Y. H., Haruna, S., & Lee, Y. Y. (2021). Slag uses in making an ecofriendly and sustainable concrete: A review. *Construction and Building Materials*, 272, Article e121942.
- [3] Pappalardo, G. (2015). Correlation between P-Wave velocity and physical–mechanical properties of intensely jointed dolostones, Peloritani Mounts, NE Sicily. *Rock Mechanics and Rock Engineering*, 48, 1711–1721.
- [4] Germinario, L., Siegesmund, S., Maritan, L., & Mazzoli, C. (2017). Petrophysical and mechanical properties of Euganean trachyte and implications for dimension stone decay and durability performance. *Environmental Earth Sciences*, 76, Article 739.
- [5] Yılmaz, N. G., Göktan, R. M., & Kibici, Y. (2011). Relations between some quantitative petrographic characteristics and mechanical strength properties of granitic building stones. *International Journal of Rock Mechanics and Mining Sciences*, 48, 506–513.
- [6] Karaca, Z., Güneş Yılmaz, N., & Göktan, R. M. (2012). Considerations on the European standard EN 14157 test methods: Abrasion resistance of natural stones used for flooring in buildings. *Rock Mechanics and Rock Engineering*, 45, 103–111.
- [7] Özvan, A., & Direk, N. (2021). The relationships among different abrasion tests on deteriorated and undeteriorated rocks. *Bulletin of Engineering Geology and the Environment*, 80, 1745–1756.
- [8] Yaşar, E., & Erdoğan, Y. (2004). Estimation of rock physicochemical properties using hardness methods. *Engineering Geology*, 71, 281–288.
- [9] Engin, I. C. (2012). A correlation for predicting the abrasive water jet cutting depth for natural stones. *South African Journal of Science*, 108(9/10), 1–11.
- [10] Kılıç, A., & Teymen, A. (2008). Determination of mechanical properties of rocks using simple methods. *Bulletin of Engineering Geology and the Environment*, 67, 237–244.
- [11] Çobanoğlu, İ., & Çelik, S. B. (2017). Assessments on the usability of Wide Wheel (Capon) test as reference abrasion test method for building stones. *Construction and Building Materials*, 151, 319–330.
- [12] Bozdağ, A. (2020). Estimation of Böhme abrasion resistance through the physicochemical properties of natural building stones. *Journal of Engineering Sciences and Design*, 8, 1062–1071.
- [13] Çobanoğlu, İ., & Çelik, S. B. (2017). Assessments on the usability of Wide Wheel (Capon) test as reference abrasion test method for building stones. *Construction and Building Materials*, 151, 319–330.
- [14] Yavuz, A. B., Türk, N., & Koca, M. Y. (2002). The mineralogical, chemical, physical and mechanical properties of Muğla region marbles. *Jeoloji Mühendisliği*, 26, 281–288.

- [15] Bayram, F. (2020). Data mining techniques for the prediction of Böhme surface abrasion rates from rock properties. *Journal of Testing and Evaluation*, 48, 323–332.
- [16] Strzałkowski, P., & Köken, E. (2022). Assessment of Böhme abrasion value of natural stones through artificial neural networks (ANN). *Materials*, 15, Article 2533.
- [17] Yaşar, E., & Erdoğan, Y. (2004). Estimation of rock physico-mechanical properties using hardness methods. *Engineering Geology*, 71, 281–288.
- [18] Kılıç, A., & Teymen, A. (2008). Determination of mechanical properties of rocks using simple methods. *Bulletin of Engineering Geology and the Environment*, 67, 237–244.
- [19] Ibrahim, D. (2016). An overview of soft computing. *Procedia Computer Science*, 102, 34–38.
- [20] Mohammed, A. A. A., Fener, M., Comakli, R., İnce, İ., Balcı, M. C., & Kayabalı, K. (2021). Investigation of the relationships between basic physical and mechanical properties and abrasion wear resistance of several natural building stones used in Turkey. *Journal of Building Engineering*, 42, Article e103084.
- [21] Jang, H., & Topal, E. (2014). A review of soft computing technology applications in several mining problems. *Applied Soft Computing*, 22, 638–651.
- [22] Chen, Y., Khandelwal, M., Onifade, M., Zhou, J., Lawal, A. I., Bada, S. O., & Genc, B. (2025). Predicting the hardgrove grindability index using interpretable decision tree-based machine learning models. *Fuel*, 384, 133953.
- [23] Mao, Y., Li, Y., Teng, F., Nanekaran, Y. A., Azarafza, M., & Zhang, M. (2024). Predicting the elasticity modulus of sedimentary rocks using Deep Random Forest Optimization (DRFO) algorithm. *Environmental Earth Sciences*, 83(16), 487.
- [24] Xie, Y., Li, X., & Min, Z. (2025). Comparison of machine learning models for rock UCS prediction using measurement while drilling data. *Scientific Reports*, 15(1), 8434.
- [25] Korkanç, M. (2013). Deterioration of different stones used in historical buildings within Niğde province, Cappadocia. *Construction and Building Materials*, 48, 789–803.
- [26] Ozguven, A., & Ozelik, Y. (2014). Effects of high temperature on physico-mechanical properties of Turkish natural building stones. *Engineering Geology*, 183, 127–136.
- [27] Sert, M., & Özkahraman, H. T. (2016). The importance of welded tuff stones in construction industry according to their physico-mechanical properties. *Research Article, Afyon Kocatepe University*.
- [28] Özvan, A., & Direk, N. (2021). The relationships among different abrasion tests on deteriorated and undeteriorated rocks. *Bulletin of Engineering Geology and the Environment*, 80, 1745–1756.
- [29] Yaşar, E., & Erdoğan, Y. (2004). Estimation of rock physico-mechanical properties using hardness methods. *Engineering Geology*, 71, 281–288.
- [30] Kılıç, A., & Teymen, A. (2008). Determination of mechanical properties of rocks using simple methods. *Bulletin of Engineering Geology and the Environment*, 67, 237–244.

- [31] Engin, I. C. (2012). A correlation for predicting the abrasive water jet cutting depth for natural stones. *South African Journal of Science*, 108(9/10), Article 692.
- [32] Ozdemir, A. C., & Kahraman, E. (2023). Performance comparison of training algorithms for the estimation of Böhme abrasion resistance using neural networks. *Journal of Mountain Science*, 20(12), 3732–3742.
- [33] Sarı, S. A., & Yavuz, A. B. (2023). Predicting the abrasion resistance value before and after deterioration by freeze–thaw of limestones based on the initial material properties: A case study from Manisa area western Türkiye. *Environmental Earth Sciences*, 82, Article 353.
- [34] Güneş, N., & Gökten, R. M. (2023). Abrasion resistance prediction of true marbles using chemical composition data: A new approach. *Journal of Engineering Sciences and Design*, 11(4), 1245–1256.
- [35] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [36] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232.



## Predicting the Abrasive Resistance of Building Stones Using Regression Models and Random Forest Algorithm

Ali Bameri<sup>1\*</sup> and Ebrahim Ghasemi<sup>2</sup>

1- PhD Candidate of Isfahan University of Technology, Isfahan, Iran.

2- Associate Professor of Mining Engineering, Isfahan University of Technology, Isfahan, Iran.

Received: 13 April 2024      Accepted: 19 June 2024  
(\*Corresponding Author: a.bameri@mi.iut.ac.ir)

### **Keywords**

**Abrasive resistance of stone  
Böhme abrasion value (BAV)  
machine learning  
multivariate regression  
random forest  
physical and mechanical properties**

### **Final English Extended Abstract**

This study aims to predict the Böhme abrasion value (BAV) as a measure of the abrasive resistance of building stones using regression models and machine learning algorithms. A total of 160 datasets, including physical and mechanical parameters of stones such as porosity (N), water absorption (Wa), and Shore hardness value (SHV), were collected and analyzed. The results indicated that all three parameters significantly impact abrasive resistance: increased porosity and water absorption lead to lower abrasive resistance, while higher Schmidt hardness results in greater abrasive resistance. Multivariate regression models and the Random Forest algorithm were used to predict BAV with high accuracy, particularly the Random Forest model, which achieved superior performance with higher precision indicators, including a normalized root mean square error (NRMSE) of 0.110 and a variance account for (VAF) of 0.839. These models not only reduce the need for costly and time-consuming physical tests but also support designers and engineers in making more informed decisions when selecting suitable building materials for high-traffic and harsh environmental conditions.

### **Summary**

This study focuses on predicting the Böhme abrasion value (BAV), an indicator of the abrasive resistance of building stones, using regression models and the Random Forest algorithm. The research leverages 160 datasets comprising physical parameters such as porosity (N), water absorption (Wa), and Shore hardness (SHV) to develop predictive models. The methodology involves both multivariate regression and machine learning techniques, with the Random Forest algorithm demonstrating superior accuracy (NRMSE = 0.110, VAF = 0.839) compared to regression models. Key findings reveal that higher porosity and water absorption reduce abrasive resistance, while increased Shore hardness enhances it. The study concludes that these models effectively minimize the need for costly, time-intensive physical tests, aiding engineers in selecting durable building stones for high-traffic applications.

### **Introduction**

The abrasive resistance of building stones is a critical factor in determining their suitability for construction applications, particularly in high-traffic areas such as flooring and pavements. Traditional laboratory tests, while accurate, are often expensive and time-consuming, necessitating alternative predictive approaches. This research aims to predict the Böhme abrasion value (BAV) using accessible physical properties of stones, employing regression models and advanced machine learning techniques. The primary objective is to provide a fast, cost-effective tool for engineers to assess stone durability, addressing the challenge of balancing accuracy with practicality. By exploring the relationship between stone properties and abrasive resistance, this study offers insights into improving material selection processes and sets the stage for future advancements in predictive modeling for construction materials.



### Methodology and Approaches

The study utilized 160 datasets extracted from prior literature, encompassing BAV and physical parameters (porosity, water absorption, and Shore hardness). Data were split into 80% training and 20% testing sets. Two main approaches were employed: (1) multivariate regression, including single-variable and multi-variable models, to establish linear and power relationships between BAV and input parameters; and (2) the Random Forest algorithm, implemented in Python using Scikit-Learn and RandomForest libraries. Hyperparameters such as the number of trees (100), maximum depth (10), and minimum samples per leaf (1) were optimized to enhance model performance. Statistical metrics, including normalized root mean square error (NRMSE), variance account for (VAF), and mean absolute error (MAE), were used to evaluate and compare model accuracy.

### Results and Conclusions

The results demonstrated that both regression and Random Forest models effectively predict BAV, with the latter outperforming the former. Key observations include:

1. **Parameter Influence:** Increased porosity (N) and water absorption (Wa) correlate with higher BAV values (indicating lower abrasive resistance), while higher Shore hardness (SHV) reduces BAV, enhancing resistance.
2. **Model Performance:** The Random Forest model achieved an NRMSE of 0.110, VAF of 0.839, and  $R^2$  of 0.838, surpassing the multivariate regression model (NRMSE = 0.1187, VAF = 0.8104,  $R^2 = 0.806$ ).
3. **Practical Implications:** These models reduce reliance on labor-intensive physical tests, offering a rapid, reliable alternative for assessing stone durability. The study concludes that the Random Forest algorithm, with its ability to handle complex, non-linear relationships and noisy data, is the most effective tool for BAV prediction. It recommends integrating such models into engineering workflows to optimize material selection for construction projects, with potential for further refinement using larger datasets or additional stone properties in future research.