

## مقایسه روش‌های شبکه عصبی مصنوعی و رگرسیونی برای پیش‌بینی هدایت هیدرولیکی اشباع خاک‌های استان خوزستان

روح اله رضایی ارشد<sup>۱</sup>، غلامعباس صیاد<sup>۱\*</sup>، مسعود مظلوم<sup>۲</sup>، مهدی شرفا<sup>۳</sup> و علیرضا جعفرنژادی<sup>۴</sup>

(تاریخ دریافت: ۱۳۸۹/۴/۱۴؛ تاریخ پذیرش: ۱۳۹۱/۳/۹)

### چکیده

اندازه‌گیری مستقیم ویژگی‌های هیدرولیکی خاک وقت‌گیر و پرهزینه بوده و تا حدی به علت غیرهمگن بودن خاک و خطاهای آزمایشگاهی غیرقابل اعتماد است. در عوض ویژگی‌های هیدرولیکی خاک می‌تواند از جایگزینی داده‌های زودیافتی مانند بافت خاک و چگالی ظاهری با استفاده از توابع انتقالی به دست آید. شبکه‌های عصبی و رگرسیون آماری از جمله روش‌هایی هستند که برای تخمین توابع انتقالی خاک (PTFs) استفاده می‌شوند. در این پژوهش از شبکه عصبی نوع پرسپترون چندلایه (MLP) و مدل‌های رگرسیونی حذف تدریجی متغیرها و گام به گام به ورود متغیرها برای بسط این توابع برای تخمین هدایت هیدرولیکی اشباع خاک با استفاده از چگالی ظاهری، تخلخل کل و درصد توزیع اندازه ذرات خاک استفاده شد. داده‌ها از ۱۲۵ پروفیل خاک مربوط به مطالعات خاک‌شناسی و اصلاح اراضی موجود در سازمان آب و برق خوزستان تهیه شد. نتایج نشان داد که شبکه MLP با الگوریتم آموزشی بیزین با ضریب تعیین ( $R^2=0/65$ ) و خطای ( $RMSE=0/04$ ) نسبت به مدل‌های رگرسیونی کارایی بهتری در تخمین هدایت هیدرولیکی اشباع خاک داشت.

واژه‌های کلیدی: ویژگی‌های هیدرولیکی خاک، توابع انتقالی خاک، شبکه عصبی، مدل‌های رگرسیونی

۱. به ترتیب دانشجوی سابق کارشناسی‌ارشد و استادیار علوم خاک، دانشگاه شهید چمران اهواز

۲. مربی مهندسی کامپیوتر، دانشگاه شهید چمران اهواز

۳. استادیار مهندسی علوم خاک، دانشگاه تهران

۴. عضو هیئت علمی بخش خاک و آب مرکز تحقیقات کشاورزی و منابع طبیعی خوزستان، اهواز

\*: مسئول مکاتبات، پست الکترونیکی: gsayyad@scu.ac.ir

## مقدمه

اندازه‌گیری ویژگی‌های هیدرولیکی خاک مانند هدایت هیدرولیکی اشباع خاک که از مهم‌ترین ویژگی‌های فیزیکی خاک می‌باشد، در طراحی و امکان‌پذیر بودن فنی و اقتصادی پروژه‌های زه‌کشی نقش به‌سزائی دارد. اندازه‌گیری این ویژگی‌ها به دلیل نیاز به صرف هزینه و زمان زیاد و تغییرپذیری مکانی و زمانی فراوان آنها، اغلب با دشواری‌هایی همراه است (۳۰). بنابراین استفاده از روش‌هایی که بتوانند این ویژگی‌ها را با استفاده از پارامترهای زودیافت خاک مانند بافت، کربن آلی و چگالی ظاهری با دقت مناسب تخمین بزنند، لازم به نظر می‌رسد. استفاده از توابع انتقالی یکی از روش‌های غیرمستقیم می‌باشد که می‌توان با استفاده از آنها از ویژگی‌های زودیافت خاک، ویژگی‌های دیریافت خاک را تخمین زد (۱۳). بنابراین میزان علاقه به استفاده از توابع انتقالی برای تخمین ویژگی‌های هیدرولیکی خاک روزبه‌روز در حال افزایش است (۱۰ و ۱۵). از جمله مدل‌های برآوردکننده توابع انتقالی شامل مدل‌های رگرسیونی (۲۲) و شبکه‌های عصبی مصنوعی (۱۶ و ۱۷) می‌باشند. مدل‌های آماری یا معادلات رگرسیونی تابعی از فرضیات و ملاحظات ذاتی آزمایش‌ها در تحلیل است. در روش‌های آماری واقعیت پدیده به کمک یک معادله جبری بیان می‌شود که فرضیات اصلی مسأله در آن لحاظ شده است. از آنجا که برای پیش‌بینی پارامترهای هیدرولیکی در خاک، چندین متغیر ورودی و خروجی مؤثرند، تحلیل آماری این نوع مسائل مستلزم ارائه تعدادی روابط ریاضی است که اغلب برای استفاده و تفسیر دشوار است. اما امروزه با توسعه سریع فناوری‌های پردازش رایانه‌ای و ایجاد نرم‌افزارهای مربوطه، از مزایای فناوری هوش مصنوعی مانند شبکه‌های عصبی برای مسائل مربوط به مدل‌سازی سیستم‌ها و فرآیندها استفاده می‌شود (۱).

ووستن و همکاران و سویراج و همکاران با استفاده از پارامترهایی هم‌چون تخلخل مؤثر، فراوانی نسبی ذرات خاک، ماده آلی و چگالی ظاهری، توابعی برای برآورد هدایت آبی اشباع خاک ارائه کردند (۲۶ و ۳۰). نتایج پژوهش‌های نوین،

پاچپسکی و همکاران، اسخاپ و بوتن، اسخاپ، تاماری و همکاران نشان دادند که شبکه‌های عصبی مصنوعی در اغلب موارد عملکردی بهتر از روش‌های رگرسیونی دارند (۶، ۱۹، ۲۳ و ۲۴). یکی از مزیت‌های شبکه‌های عصبی در مقایسه با توابع انتقالی رگرسیونی این است که این شبکه‌ها نیاز به مفاهیم قبلی در مورد ارتباط بین داده‌های ورودی و خروجی ندارند (۲). با شناخت توانایی سلول‌های عصبی در مغز انسان و مشابه‌سازی این روند در سیستم‌های کامپیوتری، شبکه‌های عصبی هوشمند مصنوعی برای نخستین بار در سال ۱۹۴۳ توسط مک کولا (عصب‌شناس) و پیت (ریاضیدان) به وجود آمد و از آنجایی که مفاهیم پایه‌ای مطرح شده در مباحث مربوط به آن از اساس ریاضی مستحکم و قابل اعتمادی برخوردار بود، بعدها به طور گسترده‌ای مورد اقبال عمومی پژوهش‌گران در مدل‌سازی تجربی فرآیندهای شدیداً غیرخطی قرار گرفت (۱۲). امروزه شبکه‌های عصبی مصنوعی به عنوان یک روش درون‌یاب و شبیه‌ساز پیشرفته که به خوبی توانایی شبیه‌سازی فرآیندهای پیچیده را دارد، به کار می‌رود. در واقع به کارگیری این روش افق جدیدی را در خانواده مهندسی و پس از آن در شاخه‌های دیگر علوم گشود.

شبکه عصبی شبیه‌سازی از دستگاه عصبی انسان است و در واقع تقلیدی از مغز و شبکه اعصاب انسان می‌باشد. در این شبکه سعی بر این است که ساختاری تهیه شود که همانند مغز، قدرت یادگیری، تعمیم‌دهی و تصمیم‌گیری داشته باشد (۵). در این گونه ساختارها هدف این است که با معرفی عملکرد یک سیستم دینامیکی، مدل را آموزش داده، چگونگی عملکرد سیستم را در حافظه مدل ذخیره و از آن برای مواردی که قبلاً با آن مواجه نشده است، استفاده شود. به دلیل توانایی این شبکه‌ها در مدل‌سازی، فرآیندهای بسیار پیچیده، که تعداد عوامل تأثیرگذار در آنها زیاد است، استفاده از آنها در علوم کشاورزی فراهم می‌باشد. مهم‌ترین بخش یک شبکه عصبی نرون (Neuron) است. نرون‌ها سلول‌های تشکیل‌دهنده دستگاه عصبی انسان هستند. هر شبکه عصبی

نسبی ذرات خاک به روش هیدرومتری، ماده آلی به روش والکلی- بلک، درصد سنگریزه به روش حجمی و تخلخل کل با استفاده از چگالی ظاهری و حقیقی از رابطه  $(n=1-\rho_b/\rho_s)$  محاسبه شد (۲). میانگین هندسی قطر ذرات و انحراف معیار هندسی آنها از روابط شیزاوا و کمپل به دست آمدند (۲۵). ابتدا داده‌های مورد نیاز به نرم‌افزار SPSS انتقال داده شد. این اطلاعات دسته‌بندی و با رسم باکس پلات مربوط به داده‌های هر پارامتر، داده‌های پرت شناسایی و حذف شد. سپس استخراج توابع انتقالی برای هدایت هیدرولیکی اشباع خاک با استفاده از ویژگی‌های توزیع اندازه ذرات، تخلخل کل، چگالی ظاهری، میانگین هندسی اندازه ذرات، انحراف استاندارد اندازه ذرات، درصد ماده آلی و درصد سنگریزه انجام شد. برای این کار از رگرسیون آماری و شبکه عصبی مصنوعی استفاده شد. برای ایجاد توابع انتقالی شبکه عصبی مصنوعی از نرم‌افزار MATLAB و برای رگرسیون آماری چندمتغیره خطی از نرم‌افزار SPSS استفاده گردید. شبکه‌های مورد استفاده در این پژوهش از شبکه عصبی پرسپترون چندلایه (MLP) با الگوریتم آموزشی بیزین (Bayesian) استفاده شد. روش ورود داده‌ها بدین‌گونه بود که به صورت تصادفی ۸۰ درصد داده‌ها (۱۰۰ داده) برای آموزش و ۲۰ درصد داده‌ها برای آزمون شبکه انتخاب شد. (۲۵)

#### نرمال‌سازی داده‌ها

پیش از آموزش شبکه عصبی، داده‌های ورودی به آن نرمال شدند. هدف از نرمال‌سازی این است که داده‌ها به اعدادی مابین صفر تا یک تبدیل شوند، زیرا در این پژوهش برای عناصر پردازش‌گر (نرون‌ها) در لایه مخفی، تابع آستانه سیگموئیدی انتخاب گردید که خروجی این تابع اعدادی بین صفر تا یک می‌باشد. بنابراین باید داده‌های ورودی به این تابع نیز اعدادی بین صفر و یک باشند (۵). برای ورودی‌های نزدیک به صفر تا یک، تغییرات وزن نرون‌ها حداقل خواهد بود، زیرا در این اعداد، عناصر پردازش‌گر به دلیل شکل تابع سیگموئید، کند عمل می‌کنند. ولی برای مقادیر ورودی‌های نزدیک به نیم، پاسخ نرون‌ها به سیگنال ورودی

مصنوعی از لایه‌های ورودی، پنهان و خروجی تشکیل شده است. در هر لایه، یک یا چند عنصر پردازش‌گر (نرون) وجود دارد که عملکرد آنها شبیه به عملکرد نرون‌های مغز انسان است. یک شبکه عصبی مجموعه‌ای از نرون‌هایی است که در آن هر نرون با همه نرون‌های لایه بعدی مرتبط است. هر یک از نرون‌های لایه ورودی در وزنی ضرب می‌شود. مقدار این وزن، همانند عملکرد تقاطعات سیناپتیک در نرون‌های زیستی، تعیین‌کننده تأثیر هر متغیر بر میزان عملکرد لایه ورودی است. هر عنصر پردازش‌گر یا نرون از دو بخش تشکیل شده است. در بخش نخست، مجموع وزنی مقادیر ورودی محاسبه می‌شود. این تابع ریاضی اصطلاحاً تابع آستانه یا تابع انتقال (Transfer function) نامیده می‌شود که عملکرد آن شبیه یک فیلتر غیرخطی است و موجب می‌شود تا خروجی نرون در یک محدوده عددی خاصی قرار بگیرد. مهم‌ترین موضوع در مدل‌های شبکه عصبی، انتخاب ورودی‌های مناسب به مدل برای رسیدن به خروجی‌های مد نظر می‌باشد (۱۵). علاوه بر آن، ساختار شبکه‌های عصبی مصنوعی و چگونگی انتخاب ارتباط بین نرون‌ها و وزنی که هر نرون به خود اختصاص خواهد داد، از اهمیت بالایی برخوردار است. ساختار یک شبکه عصبی متشکل از تعداد لایه‌ها، تعداد نرون‌ها در هر لایه، چگونگی ارتباط لایه‌ها با هم، روش آموزش شبکه و چگونگی توزیع پارامترها می‌باشد.

#### مواد و روش‌ها

در این پژوهش از پارامترهای هیدرولیکی و ویژگی‌های فیزیکی خاک نظیر چگالی ظاهری، تخلخل کل و درصد توزیع اندازه ذرات خاک موجود در مطالعات خاک‌شناسی و اصلاح اراضی استان خوزستان استفاده گردید. نمونه‌برداری از عمق ۳۰ سانتی‌متری خاک انجام شد. اندازه‌گیری هدایت آبی اشباع نمونه‌ها با استفاده از دستگاه نفوذسنج گلف مدل KI-2800 صورت گرفت. پارامترهای زودیافت چگالی ظاهری و حقیقی، به ترتیب از روش‌های کلوخه و پیکنومتر، فراوانی

### بهبود قدرت تعمیم‌دهی شبکه

برای تعمیم‌دهی مناسب شبکه، باید به گونه‌ای از آموزش بیش از اندازه (Over fitting) جلوگیری نمود. در مطالعه حاضر، برای جلوگیری از آموزش بیش از اندازه در شبکه‌های MLP از روش تنظیم خودکار (Automated Regularization) استفاده شد. در این روش تعیین پارامترهای کارایی به صورت اتوماتیک بسیار مطلوب می‌باشد. یک راه برای رسیدن به این فرآیند استفاده از الگوریتم آموزشی بیزین می‌باشد. در این الگوریتم، وزن‌ها و بایاس‌های شبکه، دارای مقادیر تصادفی با یک توزیع خاص فرض می‌شوند. پارامترهای تنظیم با یک واریانس نامعلوم از این توزیع‌ها مرتبط هستند. بنابراین می‌توان با استفاده از تکنیک‌های آماری آن را تخمین زد. یک ویژگی این الگوریتم این است که یک اندازه‌گیری از تعداد وزن‌ها و بایاس‌های مؤثر در شبکه ارائه می‌دهد. به این معنی که در پایان آموزش می‌توان تشخیص داد که چند پارامتر شبکه در آموزش مؤثر بوده‌اند (۳).

### طراحی مدل

در این پژوهش برای آموزش شبکه MLP از ۷ نرون (تعداد ورودی) در لایه ورودی، ۱ نرون در لایه خروجی (تعداد خروجی) و تعدادی نرون در لایه میانی (با سعی و خطا) استفاده شد. برای انتخاب تعداد مناسب نرون‌ها در لایه میانی و تعداد تکرار مطلوب و برای ارزیابی یادگیری شبکه‌های عصبی و حصول بهترین نتایج، از شاخص‌های  $R^2$  و RMSE استفاده شد. برای یافتن حالت بهینه شبکه‌ها از توابع آستانه مختلفی از جمله تابع سیگموئید لگاریتمی، تابع خطی و تانژانت هایپربولیک سیگموئید و هم‌چنین برای بهینه‌سازی وزن‌های شبکه‌های پرسپترون از الگوریتم آموزشی بیزین استفاده شد. برای هر شبکه (ANN)، ابتدا در یک ترکیب پیش‌فرض و با تکرارهای متفاوت، مقادیر ضریب تبیین ( $R^2$ ) و خطا (RMSE) مورد بررسی قرار گرفت. تعداد تکراری که به ازای آن، مقدار خطا (RMSE) داده‌های آزمون کمترین و ضریب تبیین ( $R^2$ ) بیشترین بود، به عنوان تعداد تکرار اولیه انتخاب شد.

سریع‌تر خواهد بود (۲۹). با در نظر گرفتن این واقعیت، نرمال‌سازی داده‌ها به گونه‌ای انجام گرفت که میانگین سری داده‌ها برابر ۵/۰ گردد. بدین منظور از رابطه زیر برای نرمال‌سازی استفاده گردید.

$$X_{\text{norm}} = 0/5 \left( \frac{X_0 - \bar{X}}{X_{\text{max}} - X_{\text{min}}} \right) + 0/5 \quad [1]$$

که در آن  $X_{\text{norm}}$  مقدار نرمال شده ورودی  $X$ ،  $\bar{X}$  میانگین داده‌ها،  $X_{\text{max}}$  و  $X_{\text{min}}$  به ترتیب حداکثر و حداقل داده‌ها می‌باشد.

### معیارهای ارزیابی

به منظور ارزیابی نتایج مدل شبکه عصبی و مقایسه آن با برآوردهای رگرسیونی از معیارهای مجذور میانگین مربعات خطا (RMSE)، و ضریب تبیین ( $R^2$ ) استفاده گردید. معادلات شاخص‌های مذکور به شرح زیر هستند (۲۲):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad [2]$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad [3]$$

در روابط بالا:

$y_i$ : مقادیر اندازه‌گیری شده متغیر،  $\hat{y}_i$ : مقدار پیش‌بینی شده متغیر،  $\bar{y}_i$ : میانگین مقدار پیش‌بینی شده متغیر و  $N$ : تعداد داده‌ها می‌باشند.

### شبکه‌های پرسپترون چند لایه

شبکه‌های چند لایه پیش‌خور (Multi-layer perceptron)، یکی از مهم‌ترین ساختارهای شبکه‌های عصبی مصنوعی می‌باشد. این شبکه‌ها شامل مجموعه‌ای از واحدهای حسی (نرون‌های پایه) می‌باشند که متشکل از یک لایه ورودی، یک یا چند لایه پنهان و یک لایه خروجی می‌باشند. سیگنال ورودی در خلال شبکه و در مسیری رو به جلو به صورت لایه به لایه منتشر می‌شود. این نوع شبکه‌ها معمولاً با عنوان پرسپترون چند لایه (MLP) نامیده می‌شوند (۱۲). در پژوهش حاضر، برای مدل‌سازی با استفاده از شبکه‌های MLP از یک شبکه سه‌لایه‌ای پیش‌خور استفاده شد.

## نتایج و بحث

خلاصه آماری ویژگی‌های خاک‌های مورد استفاده برای آموزش و آزمون مدل‌ها در جدول ۱ نشان داده شده است. خاک‌های مورد بررسی دارای دامنه گسترده‌ای از هدایت هیدرولیکی اشباع و توزیع اندازه ذرات هستند و بیشترین انحراف استاندارد مربوط به این متغیرها می‌باشد.

### پیش‌بینی هدایت هیدرولیکی اشباع به وسیله رگرسیون خطی چندمتغیره

در این تحقیق برای به دست آوردن مدل رگرسیونی خطی چندگانه برای هدایت هیدرولیکی اشباع خاک با استفاده از متغیرهای مستقل، شامل: تخلخل کل، چگالی ظاهری، درصد ماده آلی، درصد سنگریزه، میانگین هندسی اندازه ذرات، انحراف استاندارد اندازه ذرات و درصد توزیع اندازه ذرات خاک از روش گام‌به‌گام ورود متغیرها (Stepwise) استفاده شد. آیمرن و همکاران بیان کردند که بهترین مدل رگرسیونی برای پیش‌بینی هدایت هیدرولیکی اشباع خاک‌های مزارع شالیزار یک مدل رگرسیونی Stepwise با متغیرهای ورودی درصد ماده آلی، چگالی ظاهری، درصد رس و میانگین هندسی اندازه ذرات می‌باشد (۸). در جدول ۲ مدل‌های مختلف ورود داده‌ها با استفاده از روش گام‌به‌گام ورود داده‌ها نشان داده شده است که مدل چهار با ویژگی‌های زود یافت درصد سنگریزه، تخلخل کل، درصد شن و درصد ماده آلی دارای بیشترین ضریب همبستگی و کمترین خطا برای برآورد هدایت هیدرولیکی اشباع می‌باشند. یکی از مفروضاتی که در رگرسیون مدنظر قرار می‌گیرد، استقلال خطاها (تفاوت بین مقدار واقعی و مقادیر پیش‌بینی شده توسط معادله رگرسیون) از یکدیگر است. در صورتی که فرضیه استقلال خطاها رد شود و خطاها با یکدیگر همبستگی داشته باشند امکان استفاده از رگرسیون وجود ندارد. به منظور بررسی استقلال خطاها از یکدیگر از آزمون دوربین-واتسون استفاده می‌شود. از آنجایی که مقدار

آماره دوربین-واتسون (۱/۷۳) در فاصله ۱/۵ تا ۲/۵ قرار دارد، فرض عدم وجود همبستگی بین خطاها رد نمی‌شود و می‌توان از رگرسیون استفاده کرد. فرض دیگر رگرسیون خطی بررسی همراستایی بین متغیرهای مستقل است. بدین منظور از فاکتور عامل تورم واریانس (Variance Inflation Factor (VIF) استفاده شد. متغیرهایی که VIF آنها کمتر از ۱۰ باشد مشکل همراستایی نخواهند داشت. در جدول مقادیر این فاکتور برای متغیرهای مستقل مدل چهار نشان‌دهنده این است که این متغیرها مشکل همراستایی ندارند. از فرضیات دیگر رگرسیون نرمال بودن متغیر وابسته می‌باشد که به همین منظور از آزمون کولموگروف-اسمیرنوف برای آزمون نرمال بودن استفاده شد. نتایج این تست نشان داد که هدایت هیدرولیکی اشباع دارای توزیع نرمال نمی‌باشد و از تبدیل لگاریتم برای نرمال کردن این متغیر استفاده شد.

معادله ۴ نشان می‌دهد که درصد شن، تخلخل کل، درصد ماده آلی و درصد سنگریزه تأثیر مثبتی روی هدایت هیدرولیکی اشباع داشته و از بین این خصوصیات تأثیر درصد تخلخل از بقیه متغیرها بیشتر و تأثیر درصد شن کمترین می‌باشد. بعد از تعیین معادله ۴ توسط داده‌های آموزشی کارایی این مدل در برآورد هدایت هیدرولیکی اشباع توسط داده‌های آزمون بررسی شد که نتایج آن در جدول ۳ و شکل ۱ ارائه شده است. مقادیر ضریب تبیین و مجذور میانگین مربعات خطا برای مرحله آزمون این مدل به ترتیب برابر ۰/۳۴ و ۰/۴۴ می‌باشند که نشان می‌دهد که مدل رگرسیونی خطی دقت خوبی در برآورد هدایت هیدرولیکی اشباع خاک ندارد.

$$+ (0.02 \times \text{درصد سنگریزه}) + 2/9 - = \text{لگاریتم هدایت}$$

$$(0.23 \times \text{درصد ماده آلی}) + (0.01 \times \text{درصد شن}) + (0.4 \times \text{تخلخل کل})$$

هیدرولیکی اشباع

[۴]

جدول ۱. خلاصه آماری داده‌های مورد استفاده در آموزش و آزمون شبکه عصبی و رگرسیون

| داده‌های آموزش تعداد ۱۰۰ |        |         |                  | داده‌های آزمون تعداد ۲۵ |        |         |                  |                 |
|--------------------------|--------|---------|------------------|-------------------------|--------|---------|------------------|-----------------|
| حداقل                    | حداکثر | میانگین | انحراف استاندارد | حداقل                   | حداکثر | میانگین | انحراف استاندارد |                 |
| ۰/۱۸                     | ۲/۸    | ۱       | ۰/۵              | ۰/۱                     | ۱/۷    | ۰/۹     | ۰/۴              | ماده آلی        |
| ۰/۳                      | ۰/۶    | ۰/۴     | ۰/۰۵             | ۰/۳                     | ۰/۵    | ۰/۴     | ۰/۰۵             | تخلخل کل        |
| ۳                        | ۸۴     | ۳۲      | ۱۸/۶             | ۱۸                      | ۷۰     | ۳۵      | ۱۱               | درصد شن         |
| ۰/۱                      | ۶۳     | ۱۱/۴    | ۱۲/۲             | ۰/۱                     | ۴۱     | ۹       | ۸                | سنگریزه         |
| ۰/۱                      | ۹۸     | ۹       | ۱۳/۸             | ۰/۴                     | ۲۸     | ۶/۷     | ۸                | هدایت هیدرولیکی |

جدول ۲. خلاصه مدل

| مدل | ضریب همبستگی | ضریب تبیین | ضریب تبیین تعدیل شده | خطای استاندارد تخمین | دوربین-واتسون |
|-----|--------------|------------|----------------------|----------------------|---------------|
| ۱   | ۰/۵          | ۰/۲۵       | ۰/۲۴                 | ۰/۵۵                 |               |
| ۲   | ۰/۵۹         | ۰/۳۵       | ۰/۳۴                 | ۰/۵۱                 |               |
| ۳   | ۰/۶۷         | ۰/۴۵       | ۰/۴۴                 | ۰/۴۷                 |               |
| ۴   | ۰/۶۹         | ۰/۴۸       | ۰/۴۶                 | ۰/۴۶                 | ۱/۷۳          |

مدل ۱: درصد سنگریزه

مدل ۲: درصد سنگریزه و تخلخل کل

مدل ۳: درصد سنگریزه، درصد تخلخل و درصد شن

مدل ۴: درصد سنگریزه، تخلخل کل، درصد شن و درصد ماده آلی

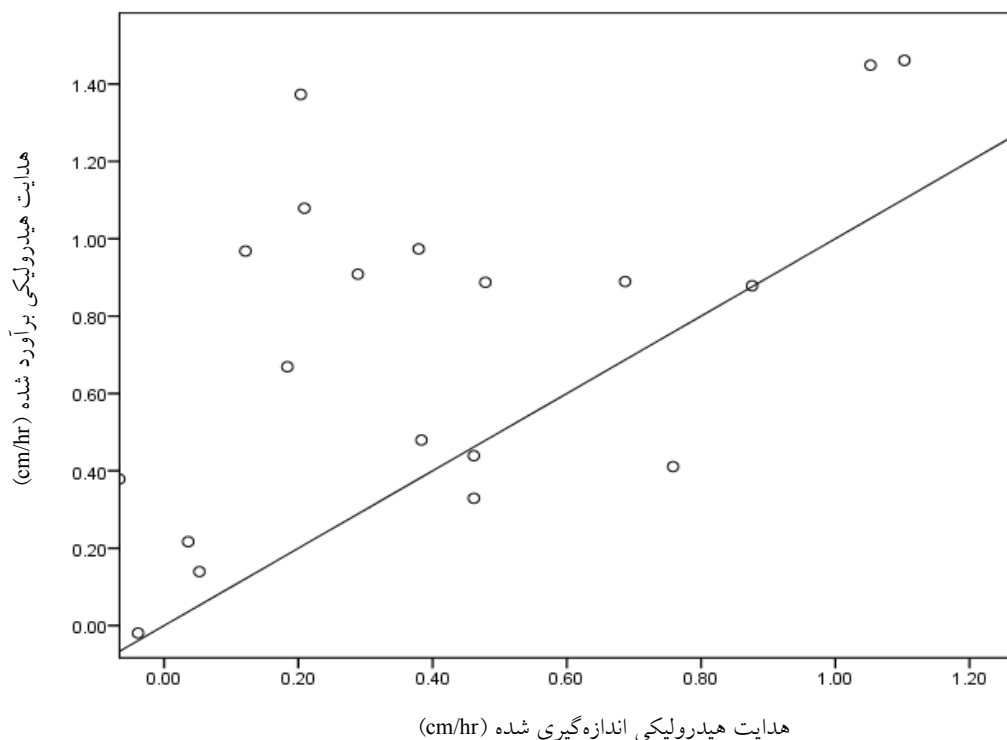
جدول ۳. نتایج آموزش و آزمون مدل رگرسیون خطی

| مرحله آموزش |                | مرحله آزمون |                |             |
|-------------|----------------|-------------|----------------|-------------|
| RMSE        | R <sup>۲</sup> | RMSE        | R <sup>۲</sup> |             |
| ۰/۴         | ۰/۳۴           | ۰/۴۸        | ۰/۴۸           | مدل رگرسیون |

### نتایج شبکه عصبی

برای تعیین مؤثرترین متغیرهای مستقل در تخمین هدایت هیدرولیکی اشیاع توسط شبکه‌های عصبی، آنالیز حساسیت توسط نرم‌افزار نروسولوشن (Neurosolution) نسخه پنج انجام شد (۲۱). نتایج آنالیز حساسیت در شکل ۲ نشان می‌دهد که همانند رگرسیون هدایت هیدرولیکی بیشترین حساسیت را به

پارامترهای درصد شن، میانگین هندسی قطر ذرات و انحراف استاندارد، درصد سنگریزه، درصد ماده آلی و تخلخل خاک دارد. پس از تعیین تعداد لایه، تعداد نرون، توابع هدف و الگوریتم آموزش، اقدام به آموزش نهایی شبکه شده و وزن‌های نهایی نرون‌ها تعیین گردید. شکل ۳ نشان می‌دهد که تا تکرار ۵۰ خطا رو به کاهش می‌باشد ولی با افزایش تکرار تغییرات



شکل ۱. رابطه بین مقادیر اندازه‌گیری شده و پیش‌بینی شده هدایت هیدرولیکی اشباع به وسیله رگرسیون خطی

جدول ۴. نتایج حاصل از آموزش و آزمون شبکه عصبی MLP

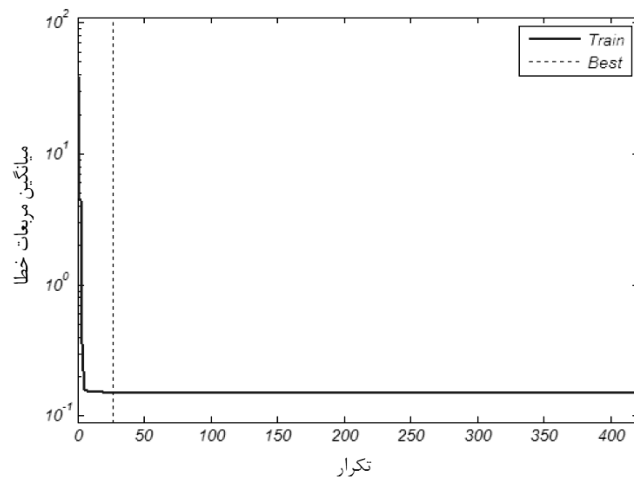
| مرحله آزمون |                | مرحله آموزش |                | توابع آستانه   | معماری شبکه | نوع شبکه |
|-------------|----------------|-------------|----------------|----------------|-------------|----------|
| RMSE        | R <sup>۲</sup> | RMSE        | R <sup>۲</sup> |                |             |          |
| ۰/۰۴        | ۰/۶۵           | ۰/۰۳        | ۰/۶۵           | tansig-purelin | ۶-۳-۱       | MLP      |

ترسیم گردید و با برازش بهترین خط عبوری از میان آنها (شکل‌های ۴ و ۵) مقادیر ضریب تبیین و خطا محاسبه شدند (جدول ۴). مقادیر زیاد ضریب تبیین و خطای کم در مراحل آموزش و آزمون شبکه، مؤید این است که مدل شبکه عصبی به خوبی توانسته است هدایت هیدرولیکی اشباع خاک را با استفاده از ویژگی‌ها زودیافت خاک تخمین بزند. قنبریان و همکاران با استفاده از ویژگی‌های درصد توزیع اندازه ذرات، چگالی ظاهری، تخلخل ویژه و منحنی رطوبتی خاک هدایت هیدرولیکی اشباع خاک ۳۱۱ نمونه خاک از پایگاه

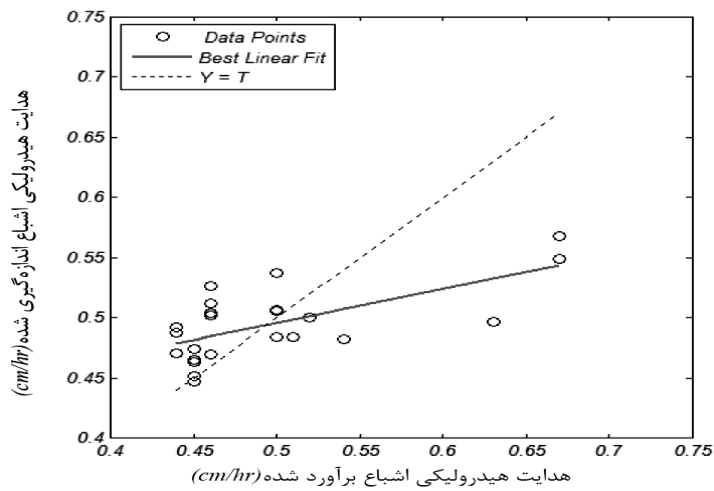
خطای شبکه در طی مرحله آموزش به یک روند ثابت می‌رسد. برای آموزش شبکه عصبی مصنوعی MLP از الگوریتم آموزشی بیزین استفاده شد که در نهایت از روش آزمون و خطا برای تعیین بهترین معماری برای این شبکه استفاده گردید. معماری بهینه انتخاب شده برای این شبکه دارای ۶ نرون در لایه ورودی، سه نرون در لایه پنهان با تابع هدف تانژانت سیگموئید و یک نرون در لایه خروجی با تابع هدف خطی با تعداد تکرار ۵۰۰ بود (جدول ۴). نتایج حاصل از این مدل در برابر داده‌های اندازه‌گیری شده



شکل ۲. آنالیز حساسیت هدایت هیدرولیکی اشباع خاک نسبت به متغیرهای مستقل

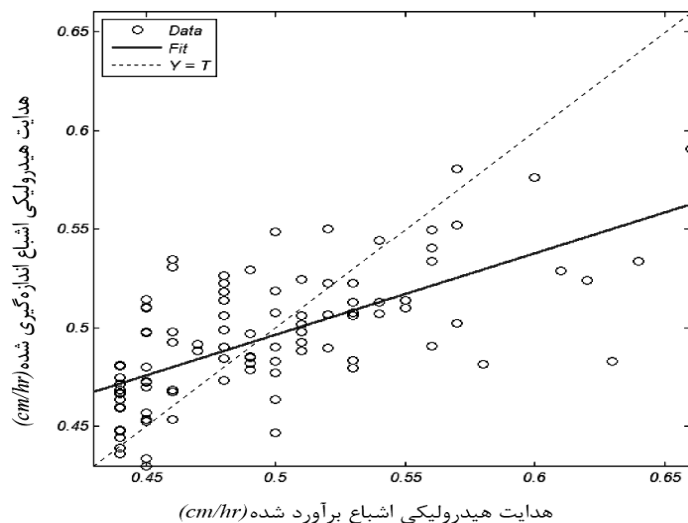


شکل ۳. تغییرات خطای شبکه بر حسب تعداد تکرار روند آموزش



شکل ۴. رابطه بین مقادیر هدایت هیدرولیکی اشباع برآورد شده به وسیله شبکه عصبی MLP با مقادیر واقعی اندازه گیری شده در مرحله آموزش





شکل ۵. رابطه بین مقادیر هدایت هیدرولیکی اشباع برآورد شده به وسیله شبکه عصبی MLP با مقادیر واقعی اندازه‌گیری شده در مرحله آزمون

جدول ۵. نتایج حاصل از آموزش و آزمون شبکه‌های عصبی و رگرسیون خطی چندگانه

| مرحله آموزش |                | مرحله آزمون |                | نوع مدل             |
|-------------|----------------|-------------|----------------|---------------------|
| RMSE        | R <sup>2</sup> | RMSE        | R <sup>2</sup> |                     |
| ۰/۰۴        | ۰/۶۵           | ۰/۰۳        | ۰/۶۵           | شبکه عصبی MLP       |
| ۰/۴۴        | ۰/۳۴           | ۰/۴۶        | ۰/۴۸           | رگرسیون خطی چندگانه |

مقادیر ضریب تبیین و خطا برای شبکه MLP به ترتیب برابر با ۰/۶۵ و ۰/۰۴ و مقادیر این شاخص‌ها برای رگرسیون خطی چندگانه به ترتیب برابر با ۰/۳۴ و ۰/۴۴ شد که نشان‌دهنده برتری شبکه عصبی می‌باشد. کارایی بهتر شبکه‌های عصبی را می‌توان به این دلیل دانست که این مدل‌ها نیاز به یک پیش‌بینی هدایت هیدرولیکی ندارند (۷). آگیار و همکاران مقادیر ضریب همبستگی ۰/۶۹ و خطای ۰/۳۱ را برای برآورد هدایت هیدرولیکی توسط شبکه‌های عصبی به دست آوردند. هم‌چنین بیان کردند که شبکه‌های عصبی در صورت به کارگیری داده‌های کافی و متغیرهای حساس بهترین نتیجه را خواهد داد (۷).

داده‌ای UNSODA استفاده کردند. مقادیر ضریب همبستگی ۰/۹۸ و میانگین مربعات خطای ۰/۰۰۱ به دست آمد. بدیهی است که هر چه مقدار ضریب تبیین بیشتر (به یک نزدیک‌تر باشد) و مقدار خطا کمتر باشد (به صفر نزدیک‌تر باشد)، دقت پیش‌بینی مدل بیشتر خواهد بود (۱۱). ناوین و همکاران مقدار خطای ۰/۰۲ تا ۰/۰۶ را برای پیش‌بینی هدایت هیدرولیکی اشباع با استفاده از شبکه عصبی مصنوعی به دست آوردند (۱۸).

#### مقایسه شبکه عصبی با روش رگرسیون

شبکه عصبی کارایی بیشتری در تخمین هدایت هیدرولیکی اشباع خاک در مرحله آزمون مدل داشت (جدول ۵).

## نتیجه گیری

با دقت بیشتری نسبت به مدل‌های رگرسیونی خطی تخمین بزند. همچنین، هوشمند بودن روش تجزیه و تحلیل داده‌ها و عدم نیاز به آزمون‌های آماری را می‌توان از دیگر برتری‌های شبکه عصبی به شمار آورد.

شبکه عصبی MLP با الگوریتم آموزشی بیزین و تابع آستانه تانژانت سیگموئید با سه نرون در لایه پنهان و تابع آستانه خطی با یک نرون در لایه خروجی و شش نرون در لایه ورودی، توانست هدایت هیدرولیکی اشباع خاک در استان خوزستان را

## منابع مورد استفاده

۱. امیری چایجان، ر. ۱۳۸۸. برآورد متغیرها و بررسی اهمیت عوامل مؤثر در فرآیند خشک شدن بستر ثابت شلتوک به کمک شبکه‌های عصبی مصنوعی. مجله پژوهش صنایع غذایی ۱۹(۱): ۵۵-۷۳.
۲. جعفری حقیقی، م. ۱۳۸۲. روش‌های تجزیه خاک نمونه‌برداری و تجزیه‌های مهم فیزیکی و شیمیایی. انتشارات ندای ضحی، صفحه ۲۳۶.
۳. صیادی، ح. ۱۳۸۸. مقایسه عملکرد شبکه‌های عصبی RBF و MLP در برآورد تبخیر و تعرق گیاه مرجع. مجله دانش آب و خاک ۱۹(۱): ۱-۱۲.
۴. کیا، م. ۱۳۸۷. شبکه‌های عصبی در *MATLAB*. چاپ دوم، خدمات نشر کیان رایانه سبز.
۵. منهای، م. ب. ۱۳۷۷. مبانی شبکه‌های عصبی، چاپ اول، مرکز نشر پروفیسور حسابی.
۶. نوایان، م. ۱۳۸۶. مقایسه توابع انتقالی شبکه عصبی مصنوعی و رگرسیونی در تخمین هدایت آبی اشباع. مجموعه مقالات دهمین کنگره علوم خاک ایران. کرج. صفحات ۹۶۷-۹۶۹.
7. Agyare, W.A. and S.J. Park. 2007. Artificial neural network estimation of saturated hydraulic conductivity. *Vadose Zone J.* 6: 423-431.
8. Aimrun, W. and M.S.M. Amin. 2009. Pedo-transfer function for saturated hydraulic conductivity of lowland paddy soils. *Paddy Water Environ.* 7: 217-225.
9. Anonymous, 2007. *Neural Network Toolbox 5, User's guide, 9th printing version 5.* The Mathworks Inc. Massachusetts, USA.
10. Cosby, B.J., G.M. Hornberger, R.B. Clapp and T.R. Ginn. 1984. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.* 20: 682-690.
11. Ghanbarian-Alavijeh, B., A.M. Liaghat and S. Sohrabi. 2010. Estimating Saturated Hydraulic Conductivity from Soil Physical Properties using Neural Networks Model. *World Academy of Science, Engineering and Technology.* 62.
12. Haykin, S. 1999. *Neural Networks: A Comprehensive Foundation.* Prentice-Hall Inc., NJ.
13. Kao, C.S. and J.R. Hunt. 1996. Prediction of wetting front movement during one-dimensional infiltration into soils. *Water Resour. Res.* 9(2): 384-395.
14. Kumar, M., N.S. Raghuwanshi and R. Singh. 2002. Estimating evapotranspiration using artificial neural network. *Irrig Drain Eng.* 128(4): 224-233.
15. Leij, F., M.G. Schaap and L.M. Arya. 2002. Water retention and storage: Indirect methods. PP. 1009-1045. *In: J.H. Dane and G.C. Topp (Ed.), Methods of Soil Analysis. Part 4. SSSA Book Ser. No. 5. SSSA, Madison, W*
16. Minasny, B., J. Hopman, W.T. Harter, S.O. Eching, A. Toli and M.A. Denton. 2004. Neural networks prediction of soil hydraulic functions for alluvial soils using multistep outflow data. *Soil Sci. Soc. Amer. J.* 68: 417-429.
17. Minasny, B., A.B. McBratney and K.L. Bristow. 1999. Comparison of different approaches to the development of pedotransfer functions for water retention curves. *Geoderma* 93: 225-253.
18. Navi, K., C. Twarakavi, J. Šimůnek and M.G. Schaap. 2009. Development of Pedotransfer Functions for Estimation of Soil Hydraulic Parameters using Support Vector Machines. *Soil Sci. Soc. Am. J.* 73:1443-1452.
19. Pachepsky, Y.A., D. Timlin and G. Varallyay. 1996. Artificial neural networks to estimate soil water retention from easily measurable data. *Soil Sci. Soc. Amer. J.* 60: 727-733.

20. Parasuraman, K. 2006. Estimating saturated hydraulic conductivity in spatially variable fields using neural network ensembles. *Soil Sci. Soc. Amer. J.* 70: 1851–1859.
21. Principe, J.C., N.R. Euliano and W.C. Lefebvre. 2000. *Neural and adaptive systems: Fundamentals through simulations*. John Wiley & Sons, New York.
22. Rawls, W.J., T.J. Gish and D.L. Brakensiek. 1991. Estimating soil water retention from soil physical properties and characteristics. *Adv. Soil Sci.* 9: 213–234.
23. Schaap, M.G. 1998. Using neural networks to predict soil water retention and soil hydraulic conductivity. *Soil Till Res.* 47: 37–42.
24. Schaap, M.G. and W. Bouten. 1996. Modeling water retention curves of sandy soils using neural networks. *Water Resour. Res.* 32: 3033–3040.
25. Shiozawa, S. and G. S. Campbell. 1991. On the calculation of mean particle diameter and standard deviation from sand, silt and clay fractions. *Soil Sci.* 152(6): 427–431.
26. Sobieraj, J.A., H. Elsenbeer and R.A. Veressy. 2001. Pedotransfer functions for estimating saturated hydraulic conductivity implications for modeling stormflow generation. *J. Hydrol.* 251: 202–220.
27. Sudheer, K.P. and S.K. Jain. 2003. Radial basis function neural network for modeling rating curves. *J. Hydrol. Engin.* 8(3): 161–164.
28. Tamari, S., J.H.M. WoËsten. and J.C. Ruiz-SuaÁrez. 1996. Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Sci. Soc. Amer. J.* 60: 1732–1741.
29. Wosten, J.H.M., A. Lilly, A. Nemes and C. le Bas. 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma* 90: 169–185.
30. Wosten, J.H.M., Ya.A. Pachepsky and W.J. Rawls. 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *J. Hydrol.* 251: 123–150.

## Comparison of Artificial Neural Networks and Regression Pedotransfer Functions for Predicting Saturated Hydraulic Conductivity in Soils of Khuzestan Province

R. Rezae Arshad<sup>1</sup>, GH. Sayyad<sup>1\*</sup>, M. Mazloom<sup>2</sup>, M. Shorafa<sup>3</sup> and A. Jafarnejady<sup>4</sup>

(Received : Jul. 5-2010 ; Accepted : May. 29-2011)

### Abstract

Direct measurement of soil hydraulic characteristics is costly and time-consuming. Also, the method is partly unreliable due to soil heterogeneity and laboratory errors. Instead, soil hydraulic characteristics can be predicted using readily available data such as soil texture and bulk density using pedotransfer functions (PTFs). Artificial neural networks (ANNs) and statistical regression are two methods which are used to develop PTFs. In this study, the multi-layer perceptron (MLP) neural network and backward and stepwise regression models were used to estimate saturated hydraulic conductivity using some soil characteristics including the percentage of particle size distribution, porosity, and bulk density. Data of 125 soil profiles were collected from the reports of basic soil science and land reclamation studies conducted by Khuzestan Water and Power Organization. The results showed that MLP neural network having Bayesian training algorithm with the greater coefficient of determination ( $R^2=0.65$ ) and the lower error (RMSE =0.04) had better performance than multiple linear regression model in predicting saturated hydraulic conductivity.

**Keywords:** Soil hydraulic characteristics, Pedotransfer functions, Neural network, Regression models.

- 
1. Former MSc. Student and Assis. Prof. of Soil Sci., Respectively, College of Agric., Shahid Chamran Univ. of Ahvaz, Ahvaz, Iran.
  2. Instructor of Comp. Eng., Shahid Chamran University of Ahvaz, Ahvaz, Iran.
  3. Assis. Prof. Soil Sci. Eng., University of Tehran, Tehran, Iran.
  4. Scientific Member of Soil and Water Research Department, Agricultural and Natural Resources Research Center of Khuzestan.

\*: Corresponding Author, Email: gsayyad@Scu.ac.ir