

## ارزیابی کارایی الگوریتم مدیریت داده‌ها به روش گروهی (GMDH) برای پیش‌بینی شاخص‌های نگهداری آب در خاک‌های شالیزاری

ناصر دوات‌گر<sup>1</sup>، علیرضا سپاسخواه، محمدرضا نیشابوری، لیلا رضائی،

حسین بیات و نادر نریمان‌زاده

استادیار پژوهش مؤسسه تحقیقات برنج کشور؛ n\_davatgar@yahoo.com

استاد دانشگاه شیراز؛ sepas@shirazu.ac.ir

استاد دانشگاه تبریز؛ neyshmr@hotmail.com

کارشناس ارشد آزمایشگاه فیزیک خاک، مؤسسه تحقیقات برنج کشور؛ lrezaee77@gmail.com

استادیار دانشگاه بوعلی سینا همدان؛ h.bayat@basu.ac.ir

استاد دانشگاه گیلان؛ nmzadeh@guilan.ac.ir

دریافت: 93/1/17 و پذیرش: 94/3/26

### چکیده

صحت توابع انتقالی در پیش‌بینی خواص هیدرولیکی خاک را می‌توان با استفاده از توابع پرنعطاف افزایش داد. این تحقیق به منظور ارزیابی کارایی توابع با قابلیت انعطاف متفاوت (رگرسیون‌های خطی و غیر خطی چند متغیره (MLR)، فیزیکی - تجربی آریا و پاریس (AP)، شبکه عصبی مصنوعی (ANN)، مدیریت داده‌ها به روش گروهی (GMDH) در پیش‌بینی مقدار آب خاک در حد ظرفیت مزرعه‌ای و نقطه پژمردگی دائم خاک‌های شالیزاری اجرا گردید. توابع انتقالی با استفاده از داده‌های توزیع اندازه ذرات، کربن آلی، چگالی ظاهری و مقدار رطوبت حجمی خاک در مکش‌های 33 و 1500 کیلوپاسکال در 136 نمونه خاک اراضی شالیزاری ساخته شد. الگوریتم GMDH در مقایسه با دیگر توابع انتقالی با کمترین ریشه میانگین مربعات خطا (RMSE)، کمترین محک آکایک (AIC) و بیشترین نمایه توافق (D) از صحت و اعتبار بیشتری در پیش‌بینی مقدار رطوبت حجمی خاک در مکش‌های 33 و 1500 کیلوپاسکال برخوردار بود. به نظر می‌رسد الگوریتم GMDH در تعیین روابط غیرخطی و پیچیده بین متغیرهای پایه مانند توزیع اندازه ذرات، چگالی ظاهری خاک و کربن آلی با مقدار رطوبت حجمی خاک در مکش‌های 33 و 1500 کیلوپاسکال موفق تر از دیگر توابع انتقالی عمل نموده است.

واژه‌های کلیدی: توابع انتقالی، حد ظرفیت مزرعه‌ای، حد پژمردگی دائم، رگرسیون خطی چند متغیره، شبکه عصبی مصنوعی

<sup>1</sup> نویسنده مسئول، آدرس: رشت، مؤسسه تحقیقات برنج کشور

## مقدمه

ارزیابی صحیح رژیم آب خاک گام مهم در تصمیم‌گیری‌های مدیریت آب است. با افزایش فراهمی نرم‌افزارهای محاسباتی، استفاده از مدل‌های هیدرولوژی و شبیه‌سازی رشد و نمو گیاه برای ارزیابی رژیم آب خاک نیز در حال افزایش است. قابلیت استفاده و کیفیت داده‌های ورودی از عوامل اساسی در مطالعات مدل‌سازی است. مدل‌های هیدرولوژی و شبیه‌سازی رشد گیاه از خواص هیدرولیک خاک و مؤلفه‌های مرتبط با آن به عنوان بخشی از پارامترهای ورودی استفاده می‌کنند. اندازه‌گیری مستقیم این خواص در مزرعه یا آزمایشگاه به علت وقت‌گیر بودن، لزوم استفاده از نیروی انسانی ماهر و هزینه زیاد و تغییرات مکانی زیاد در مقیاس ناحیه‌ای مشکل است. برای گسترش کاربرد مدل‌ها، نیاز است این پارامترها به خواص پایه خاک که در ارزیابی خاک‌ها در دسترس یا به آسانی قابل اندازه‌گیری هستند، ربط داده شود. در این شرایط لازم است خواص هیدرولیکی خاک با استفاده از توابع انتقالی برآورد شوند. از تابع انتقالی برای انتقال داده‌ها از آنچه داریم به آنچه نیاز داریم استفاده می‌شود (نوما، 1989). در این توابع می‌توان به صورت غیرمستقیم خواص هیدرولیکی خاک را از خواص پایه خاک که به آسانی اندازه‌گیری و در سطح وسیعی در دسترس هستند (مانند بافت، چگالی ظاهری و مواد آلی خاک) برآورد نمود (واستون و همکاران، 1995؛ میناسنی و مک براتنی، 2002a,b؛ میناسنی و همکاران، 2004).

پیش‌بینی خواص هیدرولیکی خاک از توزیع اندازه ذرات، به ویژه برای مطالعه وضعیت آب خاک در مقیاس حوضه آبریز و یا ناحیه‌ای سودمند است؛ زیرا اندازه‌گیری خواص هیدرولیکی به آسانی ممکن نیست. اما، توزیع اندازه ذرات ممکن است از راه بانک اطلاعات در دسترس باشد (اسکاگز و همکاران، 2001). این پیش‌بینی برپایه شباهت بین شکل توزیع اندازه ذرات و منحنی مشخصه آب خاک بر پایه روابط فیزیکی استوار است، به طوریکه بتوان منحنی مشخصه آب خاک را از داده‌های توزیع اندازه ذرات برآورد نمود. از معروف‌ترین این تقریب‌ها مدل آریا و پاریس (1981) است. این محققان تقریبی فیزیکی - تجربی که ترکیبی از فرضیات فیزیکی با روابط تجربی بود ارائه کردند و با استفاده از یک پارامتر ثابت مقیاس برای تمامی کلاس‌های بافت خاک، منحنی‌های توزیع اندازه ذرات و مشخصه آب خاک را به یکدیگر مرتبط نمودند. سپس آریا و همکاران (1999) سه روش برای پیش‌بینی پارامتر مقیاس پیشنهاد و به این ترتیب توانایی مدل را در برآورد منحنی مشخصه آب خاک بهبود

بخشیدند. هرچند اثر بافت، چگالی ظاهری و مواد آلی بر منحنی مشخصه آب خاک و هدایت هیدرولیکی از مدت‌ها پیش مشخص شده است، اما فرموله کردن تلویحی روابط بین خواص به آسانی اندازه‌گیری شونده و خواص هیدرولیکی خاک به علت هندسه پیچیده منافذ خاک با چالش روبرو است. به همین جهت پیش‌بینی خواص هیدرولیکی از بافت و دیگر خواص خاک به طور معمول با عدم قطعیت به نسبت زیاد همراه است (تایتچ و هنینگ، 1996؛ واگنر و همکاران، 2001). لوس ولت و همکاران (2011) نشان دادند که توابع انتقالی نقش کلیدی در ارزیابی عدم قطعیت شبیه‌سازی رطوبت خاک دارند و منشأ جغرافیایی بر شکل و دامنه توزیع احتمال پارامترهای هیدرولیک خاک مؤثرند. مردون (2010) نشان داد که تلفیق آمار چند متغیره و روابط ساختاری بین خواص هیدرولیک و خواص پایه خاک می‌تواند در پیش‌بینی صحیح خواص هیدرولیک کارآمدتر باشد.

روش‌هایی که برای ساخت توابع انتقالی استفاده می‌شود، عبارتند از: روش رگرسیون خطی و غیرخطی چندگانه (گوپتا و لارسون، 1979؛ باتجز، 1996؛ واستن و همکاران، 1995؛ بل و ون‌کولن، 1995؛ راولز و براکنسیک، 1985؛ مایر و جارویس، 1999؛ ساکستون و راولز، 2006)، شبکه عصبی مصنوعی (ANN)<sup>1</sup> (تامری و همکاران، 1996؛ شاپ و لیچ، 1998؛ میناسنی و مک براتنی، 2002؛ میناسنی و همکاران، 2004؛ جانا و همکاران، 2007)، مدیریت داده‌ها به روش گروهی (GMDH)<sup>2</sup> (پاچپسکی و همکاران، 1998؛ پاچپسکی و راولز، 1999) و طبقه‌بندی و رگرسیون درختی<sup>3</sup>، CART، (مکنزی و جکویر، 1997؛ راولز و براکنسیک، 1985).

در روش رگرسیون پارامترهای خواص هیدرولیکی با استفاده از خواص پایه خاک تخمین زده می‌شوند. از مزایای روش رگرسیون این است که ضروری‌ترین پارامترهای ورودی به صورت خودکار در رگرسیون گام به گام<sup>4</sup> وارد می‌شود. از معایب روش رگرسیون این است که هر معادله فقط شکل خاصی از وابستگی را می‌تواند نشان دهد و دیگر آنکه وقتی تعداد پارامترهای در دسترس خاک به عنوان ورودی زیاد می‌شوند، یافتن شکل خاص از وابستگی در بین همه آنها (یا بعضی از آنها) تقریباً غیرممکن می‌گردد (واستن و همکاران، 2001). به نظر

<sup>1</sup> Artificial neural network

<sup>2</sup> Group method data handling

<sup>3</sup> Classification and regression trees

<sup>4</sup> Stepwise regression

عصبی سخت تر است، چون تفسیر وزن‌های آن مشکل و برخلاف GMDH همه متغیرهای ورودی بدون در نظر گرفتن تأثیر آنها در ساختار شبکه پیش‌بینی‌ها باقی می‌مانند.

انگارو و همکاران (2005) نشان دادند که روش GMDH نسبت به ANN در هر دو زمینه توابع انتقالی (PTF) نقطه‌ای و پارامتریک (تخمین پارامترهای معادله بروکز و کوری) دارای صحت و مزیت نسبی بیشتری است و با توجه به نتایج ارزیابی اعتبارسنجی، استفاده از GMDH برای مدل‌سازی در خاک‌های مورد مطالعه آن محققین و محیط‌های مشابه توصیه‌پذیر است.

توماسلاو و همکاران (2003) با استفاده از الگوریتم GMDH کارایی توابع PTF نقطه‌ای و پارامتریک (برای تخمین پارامترهای معادله وان‌گونخن) را مقایسه کرده و نشان دادند که مقدار متوسط خطا در تخمین توابع PTF نقطه‌ای کمتر بوده است. بر پایه بررسی انجام شده تاکنون کارایی روش GMDH برای تخمین پارامترهای هیدرولیکی در اراضی شالیزاری ارزیابی نشده است.

این مطالعه با هدف ارزیابی کارایی توابع انتقالی با قابلیت انعطاف متفاوت (مانند رگرسیون‌های خطی چند متغیره، شبکه عصبی مصنوعی و مدیریت داده‌ها به روش گروهی) برای پیش‌بینی مقدار آب خاک در حد ظرفیت مزرعه‌ای و نقطه پژمردگی دائم در خاک‌های شالیزاری استان گیلان اجرا گردید.

#### مواد و روش‌ها

ابتدا در 136 نقطه از خاک‌های سطحی مزارع شالیزاری با توزیع جغرافیایی یکنواخت در نیمه شمالی دشت فومنات به دو شکل دست خورده و دست نخورده نمونه برداری شد. چگالی ظاهری خاک با استفاده از نمونه دست نخورده و به روش سیلندر (گراسمان و رینزچ، 2002) و کربن آلی به روش والکلی بلاک (احیایی و بهبهانی زاده، 1376) اندازه‌گیری شد. اندازه ذرات شامل ذرات شن با استفاده از غربال، سیلت و رس با استفاده از هیدرومتر (گی و اور، 2002) اندازه‌گیری شدند. مقدار آب حجمی خاک نیز در مکش‌های 0، 1، 2، 5 و 10 کیلوپاسکال به روش بار آویزان و در مکش‌های 33، 100، 200 و 1500 کیلوپاسکال با استفاده از دستگاه صفحه فشاری (دان و هاپمنز، 2002) اندازه‌گیری شدند.

#### ساخت توابع انتقالی

برای ساخت توابع انتقالی برای پیش‌بینی مقدار آب خاک در حد ظرفیت مزرعه‌ای (مکش 33 کیلوپاسکال) و نقطه پژمردگی دائم (مکش 1500 کیلوپاسکال) از انواع رگرسیون چند متغیره (MRL)، شبکه عصبی مصنوعی

می‌رسد تبدیل غیرخطی خواص پایه خاک (به عنوان مثال لگاریتمی یا نمایی) و استفاده از معادلات رگرسیون غیرخطی با افزایش قابلیت انعطاف در برآورد بهتر منحنی مشخصه آب خاک مفید باشند (ویلیامز و همکاران، 1992).

شبکه عصبی مصنوعی، ANN، در بازنمایی وابستگی پیچیده بین « ورودی - خروجی » انعطافی عالی نشان می‌دهد و توانایی تقلید رفتار سیستم‌های پیچیده از طریق تغییر در قدرت اجزای شبکه بر روی یکدیگر را دارد (واستن و همکاران، 2001). در این شبکه همه متغیرهای وابسته به صورت خود به خود از طریق متغیرهای مستقل تخمین زده می‌شوند. از مزایای این روش این است که می‌تواند مقادیر زیادی از داده‌های ورودی و خروجی را در یک مرحله به یکدیگر مرتبط کند و به این ترتیب سبب صرفه‌جویی در وقت و انرژی شود (مردون و همکاران، 2006). از دیگر مزایای شبکه عصبی این است که در آنها نیازی به مدل رگرسیون اولیه که داده‌های ورودی و خروجی سیستم را به یکدیگر مرتبط کند، نیست (مردون و همکاران، 2006؛ شاپ و لیچ، 1998). از معایب شبکه عصبی مصنوعی این است که قادر به تشخیص ورودی‌ها بر اساس اهمیت آنها نیست. در بعضی از موارد مشخص شده است که ANN در برآورد رطوبت نگهداری شده صحیح‌تر از روش رگرسیون عمل می‌نماید (پاچپسکی و همکاران، 1996؛ شاپ و بوتن، 1996). کوئک کوئک و بولتینگ (1999) نیز دریافته‌اند که ANN از روش رگرسیون بهتر عمل می‌کند اما، این تفاوت معنی‌دار نیست. مردون و همکاران (2006) نشان دادند که به استثنای رطوبت اشباع، روش رگرسیون در تخمین پارامترهای هیدرولیکی بهتر از ANN عمل نموده است اما این تفاوت معنی‌دار نیست.

اخیراً از تقریب مدیریت داده‌ها به روش گروهی (GMDH) به عنوان ابزاری قدرتمند برای بیان روابط پیچیده در « ورودی - خروجی » استفاده گردید (پاچپسکی و همکاران، 1998؛ فارو، 1984). در این روش شبکه‌های پلی‌نومیالی ساخته می‌شود که تعداد گره‌ها در آن کمتر از ANN اما انعطاف آنها بیشتر است (توماسلاو و همکاران، 2003؛ پاچپسکی و همکاران، 1998؛ واستن و همکاران، 2001). برطبق نظریه مولر و همکاران (1998) شبکه‌های عصبی روش‌های استقرایی هستند که قادر به تشخیص اهداف پیچیده نیستند، زیرا نیاز به مقدار زیادی اطلاعات اولیه دارد. اما، GMDH یک الگوریتم استنتاجی است و می‌توان آن را روشی برپایه رگرسیون دانست که بهترین‌ها را از هر دو روش رگرسیون و شبکه‌های عصبی ترکیب می‌کند. در مقایسه با GMDH ارزیابی شبکه‌های

با تعداد متغیرهای ورودی و خروجی در مدل است. تعداد گره‌های لایه پنهان می‌تواند به صورت آزاد تغییر کند. ماتریس وزن‌ها از راه فرایند واسنجی و آموزش بدست می‌آید که می‌توان از آنها برای برآورد داده مستقل استفاده کرد.

در این تحقیق از تابع فعال‌سازی تانژانت هیپربولیک برای لایه پنهان استفاده شد (میناسنی و همکاران، 2004). لایه ورودی داده‌ها شامل ترکیب‌های مختلف از شن، رس، سیلت، مواد آلی، چگالی ظاهری خاک و رطوبت حجمی اشباع خاک در هر مرحله از آموزش بود. انتخاب بهترین ترکیب لایه ورودی، اندازه لایه پنهان و تعداد گره‌ها به صورت آزمون و خطا برپایه کمترین ریشه میانگین مربعات خطا، RMSE، در داده‌های آموزش انجام گرفت. برای اجرای شبکه عصبی مصنوعی از نرم افزار Neuro Solution استفاده گردید.

#### الگوریتم مدیریت داده به روش گروهی

تقریب GMDH الگوریتمی است که در آن متغیرهای اصلی ورودی در یک شبکه قابل انعطاف از معادلات رگرسیونی نگهداری می‌شوند که می‌توان از آنها برای ارتباط بین متغیرهای ورودی (مستقل) و متغیرهای خروجی (وابسته) استفاده کرد. در این الگوریتم برآورد اولیه از متغیرهای خروجی با استفاده از معادلات رگرسیونی که شامل زیرمجموعه‌های کوچک از متغیرهای ورودی است (در هر زیرمجموعه فقط دو تا سه متغیر مستقل قرار دارد)، انجام می‌شود (فارو، 1984). برای فهم الگوریتم GMDH، با فرض وجود داده‌های اولیه که شامل ستونی از مقادیر مشاهده شده  $y$  و  $N$  ستون از مقادیر مشاهده شده متغیرهای مستقل  $x_1, x_2, \dots, x_n$  است، معادلات اولیه به صورت چند جمله‌ای درجه دوم در نظر گرفته می‌شوند:

$$Z = A + Bu + Cv + Du^2 + Ev^2 + Fu.v \quad (4)$$

در این رابطه  $A, B, C, D, E, F$  ضرایب و  $u$  و  $v$  زوج مقادیری از  $x$  هستند و  $Z$  بهترین برازش از متغیر  $y$  است. بهترین این برآوردها در مجموعه متغیرهای ورودی وارد و سپس زیرمجموعه‌های کوچکی از این متغیرها برای برآوردهای جدید در هر مرحله از چرخه سیر پذیر<sup>5</sup> استفاده می‌گردند [برای جزئیات بیشتر پیرامون روابط ریاضی و چگونگی اجرای این الگوریتم به فارو (1984) و پاچپسکی و همکاران (1998) مراجعه شود.

برای کاربرد داده‌ها در روش GMDH، داده‌های ورودی و خروجی به گونه‌ای نرمال شدند که میانگین

(ANN)، الگوریتم مدیریت داده‌ها به روش گروهی (GMDH)، و مدل فیزیکی - تجربی آریا و پاریس (1981) استفاده گردید. برای ارزیابی نتایج حاصل از تجزیه و تحلیل آماری توابع، ابتدا داده‌های خاک به صورت تصادفی به دو زیرمجموعه تقسیم شدند. از زیرمجموعه اول برای ساخت و واسنجی توابع و از زیرمجموعه دوم، به عنوان مجموعه داده‌های معیار، برای ارزیابی اعتبار توابع استفاده شد.

#### رگرسیون خطی چندمتغیره و غیر خطی

رگرسیون خطی چندمتغیره<sup>1</sup> (MLR) ابزار آماری رایج در توسعه توابع انتقالی برای پیش‌بینی متغیر وابسته ( $Y$ ) از تعدادی متغیر مستقل ( $X_i$ ) است (وریکن و هربست، 2005):

$$Y = a + \sum_{i=1}^n b_i x_i + \varepsilon \quad i = 1 \text{ و } \dots \text{ و } n$$

که در آن  $a$  عرض از مبدأ،  $b_i$  ضرایب رگرسیون و  $\varepsilon$  خطا است. علاوه بر آن از رگرسیون غیرخطی با یک متغیر مستقل (چند جمله‌ای درجه دو):

$$Y = a + bx + cx^2 + \varepsilon \quad (2)$$

و توابع نمایی با یک متغیر مستقل استفاده شد:

$$Y = a + be^x \quad (3)$$

که در آنها  $b$  و  $c$  ضرایب رگرسیون می‌باشند. از شن (%). سیلت (%). رس (%). کربن آلی (%). چگالی ظاهری خاک ( $g \text{ cm}^{-3}$ ) و رطوبت حجمی اشباع ( $\text{cm}^3 \text{ cm}^{-3}$ ) به عنوان متغیرهای مستقل استفاده گردید. متغیرهای وابسته، رطوبت حجمی خاک ( $\text{cm}^3 \text{ cm}^{-3}$ ) در نقاط پتانسیلی 33 و 1500 کیلوپاسکال بودند.

#### بسط توابع انتقالی با استفاده از شبکه عصبی مصنوعی

در این تحقیق از شبکه عصبی مصنوعی نوع پرسپترون چند لایه<sup>2</sup> استفاده شد. هر مدل ANN از تعدادی عناصر محاسباتی ساده، نرون‌ها یا گره‌ها<sup>3</sup> تشکیل شده است که در تعدادی زیرگروه، لایه‌ها، قرار گرفته و به یکدیگر در شکل یک شبکه به وسیله وزن‌هایی متصل شده‌اند. یک مدل در حالت ایده‌آل از یک لایه ورودی، یک لایه خروجی و یک (یا چند) لایه پنهان<sup>4</sup> تشکیل شده است که دو لایه خروجی و ورودی را به هم وصل می‌کند. تعداد گره‌ها در لایه‌های ورودی و خروجی منطبق

<sup>1</sup> Multiple Linear Regression

<sup>2</sup> Multilayer perceptron

<sup>3</sup> Neurons or nodes

<sup>4</sup> Hidden layers

<sup>5</sup> Iteration cycle

(5)

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N}}$$

میانگین خطا (ME) و ضریب تعیین ( $R^2$ ):

(6)

$$ME = \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)}{N}$$

(7)

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

که در آن  $Y_i$  مقدار مشاهده شده،  $\hat{Y}_i$  مقدار پیش‌بینی شده،  $\bar{Y}_i$  میانگین مقدار مشاهده شده و  $N$  تعداد کل مشاهدات است.

هر اندازه RMSE کوچک باشند، صحت برآورد بیشتر است. ME مقدار اُریب را در برآوردها ارزیابی می‌کند. مقادیر مثبت و منفی آن، به ترتیب نشان‌دهنده برآورد بیشتر و کمتر از مقدار مشاهده شده است (کورنیلز و همکاران، 2001).

در توابع انتقالی توسعه یافته، تعداد پارامترها متفاوت بود. در این شرایط انگارو و همکاران (2005) برای مقایسه صحت و اعتبار توابع انتقالی از معیار اطلاعات آکایک<sup>3</sup> (AIC) استفاده نمودند:

(8)

$$AIC = N \ln \left[ \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N} \right] + 2n_p$$

که در آن  $N$  برابر با تعداد نمونه‌های خاک و  $n_p$  تعداد پارامترها در توابع انتقالی است. AIC معیاری از کیفیت برازش تابع است. تابعی که کوچک‌ترین AIC را داشته باشد از کیفیت بهتری برخوردار است.

از پارامترهای دیگر ارزیابی خطا می‌توان به نمایه توافق<sup>4</sup>،  $D$ ، اشاره نمود که توانایی مدل را برای پیش‌بینی عامل مورد نظر نشان می‌دهد. این پارامتر دارای ارزشی بین صفر و یک است. ارزش یک نشان‌دهنده توانایی کامل مدل در پیش‌بینی مشاهدات است (سپاسخواه و همکاران، 2004):

صفر و واریانس یک داشته باشند و از این متغیرهای نرمال شده در ساخت شبکه استفاده شد. در مرحله آخر، نتایج پس از خروج از حالت نرمال به مقیاس اولیه برگردانده شدند. برای اجرای الگوریتم GMDH از نرم افزار GEvoM (منبع 14) استفاده گردید.

مدل آریا و پاریس (1981)

آریا و پاریس (1981) مدلی فیزیکی- تجربی برای پیش‌بینی آب نگهداری شده توسط خاک با استفاده از توزیع اندازه ذرات، چگالی ظاهری و چگالی حقیقی خاک ارائه نمودند. این مدل بر پایه تشابه شکل توزیع اندازه ذرات خاک و منحنی مشخصه آب خاک توسعه یافته است که در آن داده‌های توزیع اندازه ذرات به توزیع اندازه منافذ تبدیل می‌شود و آب خاک از روی حجم منافذ و پتانسیل ماتریک نظیر آن از معادله موئینگی محاسبه می‌شوند. برای اندازه‌گیری توزیع اندازه ذرات، اجزاء سیلت و رس با استفاده از هیدرومتر و قرائت آن در زمان‌های 30 ثانیه، 1، 3، 10، 30، 60، 90، 120، 240 و 1440 دقیقه و اجزاء شن با استفاده از الک‌های 1000، 500، 250، 106 و 53 میکرومتر اندازه‌گیری شدند. آریا و همکاران (1999) از راه منحنی رشد لجستیک و رابطه خطی مشابه دو تقریب برای پارامتر  $\alpha$  پیشنهاد نمودند. در این تحقیق برای برآورد بهینه پارامتر مقیاس از  $\alpha$  ثابت آریا و پاریس،  $\alpha$  تقریب شده از راه منحنی رشد لجستیک و رابطه خطی مشابه استفاده شد [برای جزئیات بیشتر پیرامون روابط ریاضی این مدل و پارامترهای مقیاس به آریا و همکاران (1999) مراجعه شود].

معیارهای آماری ارزیابی توابع انتقالی

بعد از ساخت توابع جدید انتقالی صحت<sup>1</sup> و اعتبار<sup>2</sup> آنها ارزیابی گردیدند. صحت توابع انتقالی، برپایه انطباق بین داده‌های اندازه‌گیری شده و برآورد شده برای مجموعه داده‌هایی است که این توابع بر آن اساس ساخته شده‌اند. اعتبار توابع انتقالی براساس انطباق بین داده‌های اندازه‌گیری شده و برآورد شده برای داده‌هایی است که در ساخت توابع انتقالی استفاده نشده‌اند. از آماره‌های ارزیابی زیر برای ارزیابی و مقایسه صحت و اعتبار توابع انتقالی استفاده شد. در مدل آریا و پاریس به علت آنکه برازش پیش‌بینی مدل به داده‌های مشاهده شده در کل منحنی مشخصه آب خاک انجام می‌گیرد، از ضریب تعیین ( $R^2$ ) نیز استفاده گردید.

ریشه میانگین مربعات خطا (RMSE):

<sup>3</sup> Akai Information Criteria

<sup>4</sup> Agreement index

<sup>1</sup> Accuracy

<sup>2</sup> Reliability

(9)

خاک‌های مطالعه شده دارای دامنه وسیعی از خواص فیزیکی بودند. متغیر شن در مجموعه آموزش در دامنه 1 تا 84 درصد، رس در دامنه 2 تا 61 درصد و چگالی ظاهری خاک در دامنه 0/7 تا 1/9 گرم بر سانتی‌متر مکعب قرار داشتند. علت احتمالی این دامنه وسیع ناشی از وجود خاک‌های آبرفتی در خاک‌های ناحیه مطالعاتی است. بیشتر این خاک‌ها دارای بافت به نسبت سنگین بوده و میانگین و مُد در متغیر رس در مجموعه داده آموزش به ترتیب 48/31 و 50 درصد است. آزمون t نشان داد که برای هر یک از متغیرها در دو مجموعه آموزش و آزمون اختلاف معنی‌دار وجود ندارد (نتایج نشان داده نشدند). توماسلاو و همکاران (2003) و والزاک و همکاران (2006) نشان دادند که در متغیرهای مجموعه آموزش و آزمون از نظر آزمون t نباید اختلاف معنی‌داری وجود داشته باشد.

$$D = 1 - \frac{\sum_{i=1}^N (P_i - O_i)^2}{\sum_{i=1}^N (|P_i'| + |O_i'|)^2} \quad P_i' = |P_i - \bar{O}| \quad O_i' = |O_i - \bar{O}|$$

که در آن  $D$  نمایه توافق،  $O_i$  مقادیر اندازه‌گیری شده،  $\bar{O}$  میانگین مقادیر اندازه‌گیری شده،  $P_i$  مقادیر شبیه‌سازی شده و  $N$  تعداد مشاهدات است.

### نتایج و بحث

#### آمار توصیفی داده‌های آموزش و آزمون خاک

آمار توصیفی خواص فیزیکی و هیدرولیکی خاک‌های مطالعه شده به تفکیک مجموعه داده‌های آموزش و آزمون که علاوه بر توسعه توابع انتقالی (MLR)، در ANN و GMDH نیز استفاده شده است؛ در جدول 1 نشان داده شدند.

جدول 1- آمار توصیفی خواص فیزیکی و هیدرولیکی خاک در مجموعه داده‌های آموزش و آزمون برای استفاده در توابع انتقالی

متغیر*	آموزش					آزمون				
	حداقل	حداکثر	میانگین	مُد	پراش	حداقل	حداکثر	میانگین	مُد	پراش
OC	0/2	5/6	1/82	0/4	1/35	0/04	3/3	1/42	0/9	0/917
Sa	1	84	20/08	18	195/48	2	3/8	14/11	4	100
Si	14	70	48/31	50	118/52	36	68	48/37	50	43/65
C	2	61	31/61	27	126/5	13	58	37/54	42	110/25
Bd	0/7	1/9	1/2	1	0/06	0/8	1/7	1/30	1/5	0/055
$\theta_{sv}$	0/37	0/73	0/56	0/53	0/006	0/4	0/7	0/54	0/46	0/005
$\theta_{33}$	0/3	0/68	0/45	0/43	0/005	0/36	0/56	0/45	0/45	0/003
$\theta_{1500}$	0/1	0/32	0/22	0/21	0/003	0/06	0/29	0/20	0/19	0/004

(\*) OC: کربن آلی (%); Sa: شن (%); Si: سیلت (%); C: رس (%); Bd: چگالی ظاهری خاک ( $\text{g cm}^{-3}$ );  $\theta_{sv}$ : رطوبت حجمی اشیاع ( $\text{cm}^3 \text{cm}^{-3}$ );  $\theta_{33}$ : مقدار آب حجمی خاک در مکش 33 کیلوپاسکال ( $\text{cm}^3 \text{cm}^{-3}$ );  $\theta_{1500}$ : مقدار آب حجمی خاک در مکش 1500 کیلوپاسکال ( $\text{cm}^3 \text{cm}^{-3}$ )

#### مدل آریا و پاریس (AP)

در رابطه با عدم تطبیق کامل مدل AP بر داده‌های اندازه‌گیری شده انتظار می‌رود که علاوه بر خطاهای نظام‌دار و تصادفی در هنگام نمونه‌برداری و اندازه‌گیری آزمایشگاه، محدودیت‌های تئوریک مدل AP بر نتایج اثرگذار باشند. شوه و همکاران (1988) در نظر نگرفتن مواد آلی و ساختمان خاک را به عنوان یکی از محدودیت‌های مدل اشاره می‌کنند. خاک‌های مورد مطالعه از نوع شالیزاری گلخراب بودند که ساختمان آنها در فرایند عملیات آماده‌سازی اراضی خاک تخریب شده‌اند. اما، بیشترین انحراف مدل در استفاده از هر سه نوع پارامتر مقیاس در یک نمونه از خاک‌های مورد مطالعه با کربن آلی بیش از 5/6 درصد مشاهده شد.

ارزیابی برازش مدل آریا و پاریس (AP) با استفاده از پارامترهای مقیاس  $\alpha$  خطی،  $\alpha$  ثابت و  $\alpha$  لجستیک برای پیش‌بینی مقدار رطوبت حجمی خاک در مکش‌های مختلف در مقابل مقدارهای اندازه‌گیری شده در جدول 2 نشان داده شده‌اند. مقایسه میانگین آماره‌ها نشان داد که در پارامترهای مقیاس  $\alpha$  ثابت آریا و پاریس (1981) و  $\alpha$  خطی آریا و همکاران (1999) اختلاف معنی‌داری وجود ندارد.

تشریح می‌نمایند. اثر استفاده از نقطه‌های منحنی مشخصه آب خاک (مانند رطوبت حجمی اشباع خاک) به همراه خواص پایه در افزایش صحت PTF در مطالعات دیگران نیز نشان داده شده است (شاپ و لیچ، 1998 و بورگسن و شاپ، 2005). کربن آلی به عنوان متغیر مستقل در هر دو تابع انتقالی  $\theta_{33}$  و  $\theta_{1500}$  حضور دارد. گزارش‌ها پیرامون اثر مواد آلی خاک در نگهداری آب خاک متناقض است. راولز و براکنسیک (1982) و راولز و همکاران (1983)، (2003) نقش کربن آلی را به عنوان متغیر پیش‌بینی‌کننده در فهرست توابع انتقالی  $\theta_{33}$  و  $\theta_{1500}$  نشان دادند. اما، دانالوتوس و همکاران (1994) و پاکت و همکاران (1985) هیچ‌گونه رابطه‌ای بین کربن آلی خاک و  $\theta_{33}$  و  $\theta_{1500}$  نیافتند. در هر دو مطالعه آخر مقدار کربن آلی خاک کم بود. توماسلاو و همکاران (2003) معتقدند که علاوه بر کمیت ماده آلی، کیفیت آن نیز در نگهداری آب خاک و قرار گرفتن در فهرست احتمالی متغیرها در PTF مؤثر است.

با توجه به ضریب رگرسیون لگاریتم کربن آلی در تابع انتقالی  $\theta_{33}$  (0/131029) در مقایسه با کربن آلی در تابع انتقالی  $\theta_{1500}$  (0/032) به نظر می‌رسد که نقش کربن آلی در مقدار آب خاک در مکش 33 کیلوپاسکال قوی‌تر است. این نقش توسط راولز و همکاران (2003) نیز تأیید شده است.

#### مدیریت داده‌ها به روش گروهی (GMDH)

در آموزش (واسنجی) مدیریت داده‌ها به روش گروهی از گروه متغیرهای ورودی مشابه با شبکه عصبی مصنوعی استفاده شد. توابع چند جمله‌ای توسعه‌یافته با الگوریتم GMDH برای  $\theta_{33}$ ،  $\theta_{1500}$  در جدول‌های 4 و 5 نشان داده شده است. این الگوریتم برای پیش‌بینی مقدار آب حجمی خاک در حد ظرفیت مزرعه‌ای ( $\theta_{33}$ ) از شش معادله چند جمله‌ای ساخته شده است (جدول 4). متغیرهای رس و رطوبت حجمی اشباع (مکش صفر کیلوپاسکال) به عنوان متغیر ورودی در الگوریتم GMDH برای پیش‌بینی مقدار آب خاک در مکش 33 کیلوپاسکال انتخاب نشدند. در رگرسیون خطی چند متغیره برای پیش‌بینی حد ظرفیت مزرعه‌ای نیز این متغیرها انتخاب نشده بودند (جدول 3). توماسلاو و همکاران (2003) نشان دادند که اجزای درشت بافت خاک (مانند شن) بر مقدار رطوبت خاک در مکش‌های کم مؤثر هستند.

با توجه به فقدان اختلاف معنی‌دار در پارامتر مقیاس  $\alpha$  ثابت و خطی و اینکه استفاده از  $\alpha$  ثابت آریا و پاریس (1981) در محاسبات مدل ساده‌تر است، به کارگیری آنرا نسبت به پارامتر مقیاس  $\alpha$  خطی ترجیح داده و از این پارامتر برای پیش‌بینی مقدار رطوبت حجمی خاک در پتانسیل ماتریک 33- و 1500- کیلوپاسکال استفاده شد. نتایج این پیش‌بینی با توابع انتقالی توسعه‌یافته از راه رگرسیون چند متغیره خطی (MLR)، شبکه عصبی مصنوعی (ANN) و مدیریت داده‌ها به روش گروهی (GMDH) در ادامه مقایسه شدند.

#### رگرسیون خطی چند متغیره (MLR)

در توسعه توابع انتقالی (PTF) برای پیش‌بینی مقدار آب خاک در مکش‌های 33 و 1500 کیلوپاسکال (به ترتیب  $\theta_{33}$  و  $\theta_{1500}$ ) با خواص پایه خاک رابطه غیر خطی و نمایی قوی یا معنی‌دار بدست نیامد (نتایج نشان داده نشدند). اما، روابط رگرسیونی خطی چندمتغیره معنی‌دار بودند. توسعه توابع رگرسیون خطی چند متغیره با استفاده از خواص پایه خاک (و تبدیل لگاریتمی داده‌های آنها) به روش رگرسیون گام به گام پس‌رو انجام شد. در ایجاد این توابع فیزیکی - آماری<sup>1</sup> متغیرها به گونه‌ای انتخاب شدند که از یک سو همبستگی معنی‌دار با پارامترهای نقطه‌های منحنی مشخصه آب خاک ( $\theta_{33}$  و  $\theta_{1500}$ ) داشته باشند و از سوی دیگر همبستگی معنی‌دار بین متغیرهای پایه خاک وجود نداشته باشد. برای تشخیص همراستایی علاوه بر ضریب همبستگی بین متغیرهای مستقل، به حد واریانس تورمی در این متغیرها توجه شد و متغیرهایی که دارای واریانس تورمی بیش از پنج بودند از فهرست متغیرهای مستقل کنار گذاشته شدند. در نهایت توابعی که در سطح احتمال 0/01 معنی‌دار و از ضریب تعیین تعدیل شده بیشتر برخوردار بودند، انتخاب شدند (جدول 3).

مقدار آب خاک در مکش 33 کیلوپاسکال دارای رابطه مثبت با لگاریتم کربن آلی و رابطه منفی با شن است. این تابع رگرسیونی 56 درصد از واریانس متغیر  $\theta_{33}$  را توجیه می‌کند. متغیر شن به عنوان جزء درشت بافت خاک در پیش‌بینی مقدار آب در پتانسیل بالای آب خاک، مانند  $\theta_{33}$ ، مؤثر است؛ در حالیکه سیلت و رس از اجزای به ترتیب متوسط و ریز خاک در نگهداری آب در پتانسیل آب کم ( $\theta_{1500}$ ) نقش دارند (توماسلاو و همکاران، 2003). کربن آلی به همراه رس، سیلت، لگاریتم چگالی ظاهری خاک و رطوبت حجمی اشباع خاک 42 درصد از تغییرات مقدار آب خاک در مکش 1500 کیلوپاسکال را

<sup>1</sup> Statistical- Physical Functions

جدول 2- مقایسه میانگین آماره‌های ضریب تعیین، میانگین خطا و ریشه میانگین مربعات خطا در

پارامتر مقیاس ( $\alpha$ )	$R^2$	ME	RMSE
$\alpha$ ثابت آریا و پاریس (1981)	0/941 a	0/021 a	0/059 a
$\alpha$ خطی آریا و همکاران (1999)	0/939 a	0/010 a	0/058 a
$\alpha$ لجستیک آریا و همکاران (1999)	0/888 b	-0/045 b	0/095 b

$R^2$ : ضریب تعیین؛ ME: میانگین خطا و RMSE: ریشه میانگین مربعات خطا

جدول 3- توابع انتقالی توسعه‌یافته برای پیش‌بینی مقدار آب خاک در مکش‌های 33 و 1500 کیلوپاسکال

مشخصه‌های هیدرولیکی	تابع انتقالی	$R_{adj}^2$
$\theta_{33}$ (kPa)	$29 \text{ Ln OC} - 0/00506 \text{ Sand}$	0/56
$\theta_{1500}$ (kPa)	$\theta_{1500} = 0/475 + 0/002 \text{ Clay} + 0/032 \text{ OC} + 0/251 \text{ Ln Bd} + 0/233 \text{ Ln } \theta_s - 0/072 \text{ Silt}$	0/42

جدول 4- معادله‌های چندجمله‌ای و شمای الگوریتم GMDH برای پیش‌بینی مقدار آب خاک در مکش 33 کیلوپاسکال

$$Y_1 = 0/40411 + 0/073672 (\text{OC}) + 0/00252 (\text{Sand}) - 0/00681 (\text{OC}^4) - 1/24 \times 10^{-5} (\text{Sand}^4) - 0/00022 (\text{Sand} \times \text{OC})$$

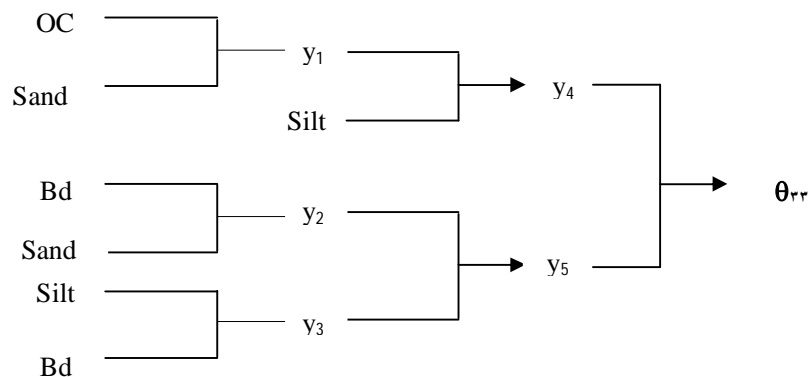
$$Y_2 = 0/649705 - 0/0030 (\text{Bd}) - 0/12326 (\text{Sand}) - 5/3 \times 10^{-6} (\text{Bd}^2) - 0/00764 (\text{Sand}^2) + 0/001186 (\text{Sand} \times \text{Bd})$$

$$Y_3 = 0/319403 + 0/009155 (\text{Silt}) + 0/05629 (\text{Bd}) - 6/5 \times 10^{-5} (\text{Silt}^2) - 0/00128 (\text{Bd}^2) - 0/00133 (\text{Bd} \times \text{Silt})$$

$$Y_4 = 0/37135 + 0/097233 (y_1) - 0/00746 (\text{Silt}) - 0/663 (y_1^2) - 6/4 \times 10^{-5} (\text{Silt}^2) + 0/031369 (y_1 \times \text{Silt})$$

$$Y_5 = 0/107013 + 1/521614 (y_2) + 1/02324 (y_3) - 6/55223 (y_2^2) - 4/40895 (y_3^2) + 0/11/55242 (y_2 \times y_3)$$

$$\theta_{33} = -0/39796 - 2/9342 (y_4) + 5/923924 (y_5) + 0/828853 (y_4^2) - 10/5072 (y_5^2) + 7/236054 (y_4 \times y_5)$$



متغیر ورودی در پیش‌بینی  $\theta_{1500}$  انتخاب شدند. توماسلا و همکاران (2003) در مقایسه توابع انتقالی برای پیش‌بینی مقدار آب حجمی خاک در حد پژمردگی دائم نشان دادند که چگالی ظاهری به همراه رطوبت اشباع خاک نیز در الگوریتم GMDH وارد شدند. آنان دو متغیر چگالی ظاهری و رطوبت اشباع را شاخص غیر مستقیم از اثر ساختمان خاک دانستند.

الگوریتم GMDH برای پیش‌بینی مقدار آب حجمی خاک در حد پژمردگی دائم ( $\theta_{1500}$ ) از هفت معادله پلی‌نومیال ساخته شد (جدول 5). در این الگوریتم GMDH، تمام متغیرهای شن، سیلت، رس، کربن آلی، چگالی ظاهری و رطوبت حجمی اشباع به عنوان متغیر ورودی وارد شدند. در الگوریتم GMDH پیشنهادی پاجپسکی و راولز (1999) نیز شن، رس، کربن آلی و چگالی ظاهری به همراه نسبت گنجایش تبادل کاتیونی به رس به عنوان



جدول 5- معادله‌های چندجمله‌ای و شمای الگوریتم GMDH برای پیش‌بینی مقدار آب خاک در مکش 1500 کیلوپاسکال

$$Y_1 = 0/1884 - 0/0143 (\text{Silt}) + 0/0288 (\text{Clay}) + 0/0002 (\text{Silt}^2) - 0/0002 (\text{Clay}^2) - 0/0003 (\text{Clay} \times \text{Silt})$$

$$Y_2 = 0/8191 + 0/1855 (\text{OC}) + 1/3325 (\text{Bd}) - 0/0066 (\text{OC}^2) + 0/4241 (\text{Bd}^2) - 0/1187 (\text{OC} \times \text{BD})$$

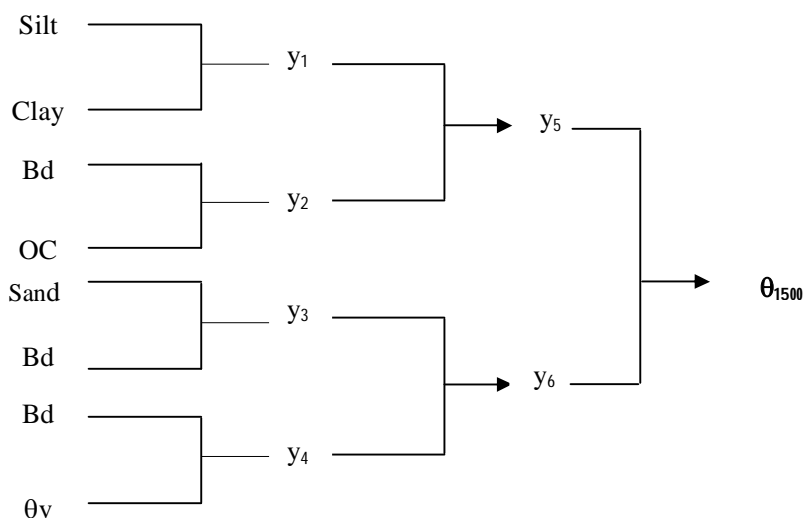
$$Y_3 = 0/2369 + 0/0055 (\text{Sand}) - 0/1266 (\text{Bd}) + 0/00001 (\text{Sand}^2) + 0/1004 (\text{Bd}^2) - 0/0053 (\text{Bd} \times \text{Sand})$$

$$Y_4 = 1/0411 - 0/7640 (\text{Bd}) - 2/3070 (\theta_v) + 0/1340 (\text{Bd}^2) + 1/4547 (\theta_v^2) + 1/1070 (\theta_v \times \text{Bd})$$

$$Y_5 = 0/5342 + 3/1803 (y_1) - 8/2522 (y_2) - 0/7604 (y_1^2) + 25/4096 (y_2^2) - 8/1391 (y_1 \times y_2)$$

$$Y_6 = -0/7823 - 2/1942 (y_3) - 42892 (y_4) - 9/5796 (y_3^2) - 4/4467 (y_4^2) + 31/9289 (y_3 \times y_4)$$

$$\theta_{1500} = 0/0233 + 1/4075 (y_5) - 0/8906 (y_6) + 0/6748 (y_5^2) + 4/6667 (y_6^2) - 3/6969 (y_5 \times y_6)$$



مجموعه داده‌های آموزش لزوماً دارای رابطه نزدیک با اعتبار آنها در مجموعه داده‌های آزمون نیست. این عدم قطعیت به ویژه در توابع انتقالی رگرسیون خطی چند متغیره زیاد دیده می‌شود.

#### مقدار آب خاک در مکش 33 کیلوپاسکال

کارایی توابع انتقالی ANN، MLR و GMDH برای پیش‌بینی مقدار آب خاک در مکش 33 کیلوپاسکال در مجموعه داده‌های آموزش در جدول 6 نشان داده شد. تحلیل صحت توابع انتقالی در مجموعه داده‌های آموزش نشان داد که پیش‌بینی شبکه عصبی مصنوعی برای  $\theta_{33}$  بدون اُریب بود. با در نظر گرفتن آماره RMSE صحت برآورد در الگوریتم GMDH بیشتر بود. ویلموت (1982) اعلام داشته است که RMSE در ارزیابی کارایی مدل‌ها بهترین هستند. اما، گیجسما و همکاران (2003) اعلام داشته است که RMSE به وجود مقادیر کرانه‌ای (داده‌های خیلی کوچک و یا خیلی بزرگ) حساس است و بنابراین خطاها را بیشتر از مقدار واقعی تخمین می‌زند. کمترین محک اطلاعات آکایک، AIC، در مجموعه داده‌های آموزش در GMDH مشاهده گردید. شاخص توافق (D-index) نیز که محکی استاندارد شده و نشان‌دهنده درجه

#### شبکه عصبی مصنوعی (ANN)

بهترین آرایه از تعداد لایه‌های پنهان و تعداد نرون‌ها در هر لایه پنهان در فرایند آموزش (واسنجی شبکه عصبی مصنوعی) برای مقدار رطوبت  $\theta_{33}$  با یک لایه پنهان و 4 نرون در لایه پنهان و در رطوبت  $\theta_{1500}$  با یک لایه پنهان و 7 نرون در لایه پنهان بدست آمد. متغیرهای ورودی برای برآورد هر دو متغیر مشابه و عبارت از کربن آلی، رس، شن، سیلت و چگالی ظاهری خاک بودند. این یافته‌ها برپایه آزمون و خطا (ارزیابی و مقایسه آماره‌های ریشه میانگین مربعات خطا، RMSE. هر آرایه در فرایند آموزش) انجام شد. نتیجه‌های این ارزیابی نشان داده نشدند و تنها آماره‌های بهترین آرایه در فرایند آموزش به همراه آزمون در جدول‌های 6 و 7 نشان داده شده‌اند.

#### آزمون صحت و اعتبار توابع انتقالی

آماره‌های ارزیابی صحت و اعتبار توابع انتقالی با استفاده از مجموعه داده‌های آموزش و آزمون در جدول‌های 6 و 7 نشان داده شده‌اند. پیش‌بینی‌ها در مجموعه داده آزمون در مقایسه با مجموعه داده‌های آموزش از عدم قطعیت بیشتری برخوردار بودند. پاچسکی و راولز (1999) نشان دادند که صحت PTF در

ارزیابی اعتبار مقدار آب خاک در مکش 1500 کیلوپاسکال با استفاده از مجموعه داده‌های آزمون نشان داد الگوریتم GMDH با دارا بودن کوچک‌ترین RMSE از اعتبار بیشتری برخوردار است. الگوریتم GMDH با شش و شبکه عصبی مصنوعی با پنج پارامتر ورودی از تعداد بیشتری پارامتر نسبت به مدل AP با یک و MLR با چهار پارامتر برخوردار بودند. با افزایش تعداد پارامترها RMSE کوچک‌تر می‌شوند. به این علت برای تمایز بیشتر بین مدل‌ها از محک AIC استفاده شد. الگوریتم GMDH از AIC کوچک‌تری نسبت به بقیه روش‌ها برخوردار بود. به همین ترتیب الگوریتم GMDH با توجه به شاخص D از بیشترین توانایی در پیش‌بینی 1500 $\theta$  در مجموعه داده‌های آزمون (مستقل) برخوردار است.

شبکه عصبی مصنوعی در فرایند آموزش و واسنجی در مجموعه داده‌های آموزش موفق‌تر از بقیه روش‌ها عمل نمود. اما، در مجموعه داده‌های آزمون در مقایسه با الگوریتم GMDH از اعتبار کمتری برخوردار بود. این نکته می‌تواند ناشی از پدیده بیش‌برازش<sup>1</sup> باشد که در هنگامی که فرایند واسنجی و آموزش شبکه با تکرار زیاد و برای مدت طولانی انجام می‌شود و یا مدل شبکه عصبی برای مسئله مطالعه شده پیچیده می‌گردد، بروز می‌کند (بورگسن و شاب، 2005).

از سوی دیگر آدلوی (2009) اعلام نموده است که عدم موفقیت ANN در برآورد (برون‌یابی) یک متغیر در مجموعه داده آزمون در مقایسه با مجموعه داده آموزش (که ANN بر پایه داده‌های آن و از طریق به حداقل رساندن تابع خطا تنظیم و مدل می‌گردد) این است که شبکه در هنگام مواجهه با داده‌های خارج از داده‌های آموزش توانایی ارتباط داده‌های جدید با سطح خطای مدل داده‌های آموزش را ندارد و بنابراین کارایی مطلوبی نشان نمی‌دهد.

در الگوریتم GMDH و ANN نیز همانند MLR از متغیرهای سیلت، رس، کربن‌آلی و چگالی ظاهری خاک به عنوان متغیر ورودی در الگوریتم استفاده شد. با این وجود GMDH و ANN از صحت و اعتبار بیشتری برخوردار بودند. بخشی از این برتری به علت توانایی الگوریتم GMDH در تبیین روابط غیرخطی و پیچیده است.

توانایی سیستم در پیش‌بینی بدون خطا است (ویلومت، 1981) و به نظر قدرت تمایز بیشتری در مدل‌ها دارد، در الگوریتم GMDH از ANN و MLR بیشتر است.

در مقایسه اعتبار، علاوه بر توابع توسعه‌یافته جدید MLR، ANN و GMDH از مدل فیزیکی - تجربی آریا و پاریس با استفاده از پارامتر مقیاس ثابت (آریا و پاریس، 1981) نیز استفاده شد. مقدار آریب (ME) در دو روش GMDH و ANN نسبت به MLR و AP کم بود و بر پایه آماره‌های RMSE، AIC و شاخص D در مجموعه داده‌های آزمون (مستقل) برآوردها در الگوریتم GMDH در مقایسه با دیگر روش‌ها از اعتبار بیشتری برخوردار است. به نظر می‌رسد الگوریتم GMDH در تعیین روابط غیرخطی و پیچیده بین متغیرهای پایه مانند توزیع اندازه ذرات، چگالی ظاهری خاک و کربن آلی با مقدار آب خاک در مکش 33 کیلوپاسکال موفق‌تر از دیگر توابع انتقالی عمل نموده است.

MLR: رگرسیون خطی چند متغیره؛ ANN: شبکه عصبی مصنوعی؛ GMDH: مدیریت گروهی داده‌ها؛ AP: مدل آریا و پاریس؛ ME: میانگین خطا؛ RMSE: ریشه میانگین مربعات خطا؛ AIC: معیار اطلاعات آکایک و D-index: شاخص توافق.

مدل فیزیکی - تجربی AP در پیش‌بینی مقدار آب در مکش 33 کیلوپاسکال از اعتبار کمتری نسبت به توابع توسعه یافته GMDH، ANN و MLR برخوردار بود. بخشی از عدم قطعیت در این مدل را می‌توان به در نظر نگرفتن کربن آلی نسبت داد. فرضیه این مدل بر پایه مشابهت توزیع اندازه ذرات به توزیع اندازه منافذ است. مواد آلی خاک در توزیع اندازه منافذ مؤثر است. در هنگام اندازه‌گیری توزیع اندازه ذرات برای استفاده در مدل AP، مواد آلی خاک سوزانده می‌شود. اما، راولز و همکاران (2003) نشان دادند که مقدار آب خاک در مکش 33 کیلوپاسکال به شدت تحت اثر کربن آلی قرار دارد.

#### مقدار آب خاک در مکش 1500 کیلوپاسکال

آماره‌های ارزیابی صحت پیش‌بینی مقدار آب خاک در مکش 1500 کیلوپاسکال با استفاده از توابع انتقالی در مجموعه داده‌های آموزش در جدول 7 نشان داده شده است. میانگین خطا در ANN ناچیز و برآوردهای آن بدون آریب هستند. علاوه بر آن ریشه میانگین مربعات خطا (RMSE) در ANN کمتر از الگوریتم GMDH و MLR است، بنابراین صحت آن بیشتر به نظر می‌رسد. شبکه عصبی مصنوعی دارای کوچک‌ترین AIC و با توجه به شاخص توافق (D-index) توانایی آن در برآورد 1500 $\theta$  در مجموعه داده‌های آموزش بیشترین است. اما، آماره‌های

<sup>1</sup> Over Fitting

جدول 6- آماره‌های نیکویی برازش توابع انتقالی در پیش‌بینی مقدار آب خاک در مکش 33 کیلوپاسکال

آموزش				
D- index	AIC	RMSE	ME	تابع انتقالی
0/85	-148	0/046	-0/002	MLR
0/88	-155	0/044	$-3/4 \times 10^{-5}$	ANN
0/94	-161	0/042	0/003	GMDH
آزمون				
D- index	AIC	RMSE	ME	تابع انتقالی
0/77	-93	0/042	-0/002	MLR
0/73	-82	0/046	$-3/3 \times 10^{-4}$	ANN
0/96	-87	0/041	$3/6 \times 10^{-4}$	GMDH
0/74	-80	0/066	-0/015	AP

جدول 7- آماره‌های نیکویی برازش توابع انتقالی در پیش‌بینی مقدار آب خاک در مکش 1500 کیلوپاسکال

آموزش				
D- index	AIC	RMSE	ME	تابع انتقالی
0/71	-122	0/048	0/001	MLR
0/88	-160	0/037	$8/3 \times 10^{-5}$	ANN
0/84	-134	0/043	-0/003	GMDH
آزمون				
D- index	AIC	RMSE	ME	تابع انتقالی
0/51	-48	0/059	0/026	MLR
0/62	-49	0/056	0/028	ANN
0/83	-134	0/037	0/019	GMDH
0/71	-96	0/06	-0/021	AP

ساختمان خاک در فرایند عملیات آماده‌سازی اراضی خاک تخریب می‌گردند اما، در پیش‌بینی‌ها در خاک-های با کربن آلی دارای انحراف است. شبکه عصبی مصنوعی در مقایسه با الگوریتم GMDH از اعتبار کمتری برخوردار بود. این نکته می‌تواند ناشی از پدیده بیش برازش و کارایی کم مدل در هنگام مواجه با داده‌های خارج از داده‌های آموزش است. الگوریتم GMDH با قابلیت انعطاف بیشتر و توانایی تبیین روابط غیر خطی و پیچیده بین مکش رطوبتی خاک و خواص پایه خاک از صحت و کارایی بیشتری نسبت به دیگر توابع انتقالی در خاک‌های مطالعه شده برخوردار است. علاوه بر آن نتایج نشان داد که بین صحت و کارایی توابع انتقالی همیشه رابطه نزدیکی وجود ندارد.

MLR: رگرسیون خطی چند متغیره؛ ANN: شبکه عصبی مصنوعی؛ GMDH: مدیریت گروهی داده‌ها؛ AP: مدل آریا و پاریس؛ ME: میانگین خطا؛ RMSE: ریشه میانگین مربعات خطا؛ AIC: معیار اطلاعات آکایک و D-index: شاخص توافق

### نتیجه‌گیری نهایی

در این مطالعه، توابع انتقالی مختلف (مدل فیزیکی تجربی آریا و پاریس، رگرسیون خطی چند متغیره، شبکه عصبی مصنوعی و الگوریتم مدیریت داده‌ها به روش گروهی) برای پیش‌بینی مقدار آب حجمی خاک در حد ظرفیت مزرعه‌ای (مکش 33 کیلوپاسکال) و نقطه پژمردگی دائم (مکش 1500 کیلوپاسکال) مقایسه شدند. مدل آریا و پاریس بر پایه توزیع اندازه ذرات است اما دارای محدودیت‌های نظری مانند در نظر نگرفتن مواد آلی و ساختمان خاک است. در خاک‌های شالیزاری

## فهرست منابع:

1. احمایی، م.، ع. و و ع. 1. بهبهانی‌زاده. 1376. شرح روش‌های تجزیه شیمیایی خاک. مؤسسه تحقیقات خاک و آب. نشریه شماره 823.
2. Adeloje. 2009. The relative utility of regression and artificial networks models for rapidly predicting the capacity of water supply reservoirs. *Environmental Modeling & Software*. 24: 1233- 1240.
3. Arya, L. M., and J. F. Paris. 1981. A physicoempirical model to predict soil moisture characteristic from particle –size distribution and bulk density data. *Soil Sci. Soc. Am. J.* 45: 1023-1030.
4. Arya, L. M., F. J. Leij, M. Th. Van Genuchten, and P. J. Shouse. 1999. Scaling parameter to predict the soil water characteristic from particle-size distribution data. *Soil Sci. Soc. Am. J.* 63: 510-519.
5. Batjes, N. H. 1996. Development of a world data set of soil water retention Properties using pedotransfer rules. *Geoderma*, 71: 31-52.
6. Bell, A. M., and H. Vankeulen. 1995. Soil pedotransfer functions for four Mexican soils. *Soil Sci. Am. J.* 59: 865-871.
7. Borgesens, C. D., and M. C. Schaap. 2005. Point and parameter pedotransfer functions for water retention predictions for Danish soils. *Geoderma*, 127: 154- 167.
8. Bouma, J. 1989. Using soil survey data for quantative land evaluation. *Adv. Soil Sci.* 9: 177-213.
9. Cornelis, W. M., J. Ronsyn, M. Van Meir Venne, and R. Hartmann. 2001. Evaluation of pedotransfer functions for predicting the soil moisture retention curve. *Soil Sci. Soc. Am. J.* 65: 638-648.
10. Danalatos, N. G., G. S.Kosmas, P. M.Driessen, and N. Yassoglou. 1994. Estimation of draining soil moisture characteristics from standard data as recorded in soil surveys. *Geoderma*, 64: 155- 165.
11. Dane. J. M., and J. W. Hopmans. 2002. Pressure plate extractor. p. 688- 690. *In* J. M. Dane and G. C. Topp. (eds.) *Methods of soil analysis. Part 4. Physical Method.* *Soil Sci. Soc. Am. Book series.* 5. Madison, Wisconsin, USA.
12. Farrow, S. J. 1984. The GMDH algorithm. *In* S. J.Farrow (ed.), *Self- organizing Methods in Modeling: GMDH type Algorithms.* Marcel Dekker. New York. pp:1-27.
13. Gee, G. W., and D. OR. 2002. Particle- Size analysis. *In*Dane, J. H., and G. C. Clake. (eds.), *Methods of soil analysis. Part 4. Physical Methods.* SSSA Book Series, no. 5. Madison, Wisconsin, USA. pp: 255- 294.
14. GevoMSoftware.2007. <http://research.guilan.ac.ir/gevom>
15. Gijmsan, A. J., S. S. Jagtap, and J. W. Jones. 2003. Wading through a SWAP of complete confusion, how to choose a method for estimating soil water retention parameters for crop models. *Eur. J. Agron.* 18: 77-106.
16. Grossman, R. B., and T. G. Reinsch. 2002. Bulk density and linear extensibility. *In* Dane, J. H., and G. C. Clake. (eds.). *Methods of soil analysis. Part 4. Physical Methods.* SSSA Book Series, no. 5. Madison, Wisconsin, USA. pp: 201- 228.
17. Gupta, S. C., and W. E. Larson. 1979. Estimating soil water characteristic from particle size distribution, organic matter percent and bulk density. *Water Resour. Res.* 15: 1633-1635.
18. Jana, R. B., B. P. Mohanty, and E. P. Springer. 2007. Multiscale pedotransfer functions for soil water retention. *Vadoes Zone J.* 6: 868-878.
19. Koek Koek, E. J. W., and H. Booltink. 1999. Neural network model to predict soil water retention. *Eur. J. Soil Sci.* 50: 489- 495.

20. Loosvelt, L., V. R. N. Pavwels, W. M. Connelis, G. J. M. De Iannoy and N. E. C. Verhoest. 2011. Impact of soil hydraulic parameter uncertainty on soil moisture modeling. *Water Resources Research*. Doi:10. 1029/ 2010WR009204.
21. Mayer, T., and N. J. Jarvis. 1999. Pedotransfer functions to estimate soil water retention parameters for a modified Brooks-Corey type model. *Geoderma*, 91:1-9.
22. McKenzie, N. J., and D. W. Jacquier. 1997. Improving the field estimation of saturated hydraulic conductivity in soil survey. *Aust. J. Soil Res.* 35: 803-825.
23. Merdun, H. 2010. Alternative methods in the development of pedotransfer functions for soil hydraulic characteristics. *Eur. J. Soil Sci.* 43: 62-71.
24. Merdun, H., O. Cinar, R. Meral, and M. Apan. 2006. Comparison of artificial neural network and regression pedotransfer functions for prediction of soil water retention and saturated hydraulic conductivity. *Soil and Till. Res.* 90: 108- 116.
25. Minasny, B., J. W. Hopmans, T. Harter, S. O. Eching, A. Tuli, and M. A. Denton. 2004. Neural networks prediction of soil hydraulic functions for alluvial soils using multistep out flow data. *Soil Sci. Soc. Am. J.* 68: 417-429.
26. Minasny, B., and A. B. Mc Bratney, 2002. FUZ ME version 3.0. Australian centre for precision agriculture, the University of Sydney, NSW. 2006.
27. Minasny, B., and A. B. Mc Bratney. 2002. The neuro-m method for fitting neural network parametric pedotransfer functions. *Soil Sci. Soc. Am. J.* 66: 352-361.
28. Muller, J. A., A. G. Ivachnenko, and T. O. Lemke. 1998. GMDH algorithms for complex systems modelling. *Mathematical and computer modelling of dynamical systems.* 4: 175-316.
29. Pachepsky, Ya, A., W. J. Rawls, D. Gimenez, and J. P. C. Watt. 1998. Use of soil penetration resistance and group method of data handling to improve soil water retention estimates. *Soil Till. Res.* 49: 117-128.
30. Pachepsky, Ya. A., and W. J., Rawls. 1999. Accuracy and reliability of pedotransfer functions as affected by grouping soils. *Soil Sci. Soc. Am. J.* 63: 1748-1757.
31. Pachepsky, Ya. A., D., Timlin, and Varallyay. 1996. Artificial neural networks to estimate soil water retention from easily measurable data. *Soil Sci. Soc. Am. J.* 60: 727-773.
32. Puckett, W. E., J. H. Dane, and B. F. Hajek. 1985. Physical and mineralogical data to determine soil hydraulic properties. *Soil Sci. Soc. Am. J.* 49: 831- 836.
33. Rawls, W. J., Y. A. Pachepsky, J. C. Ritchie, T. M. Sobeck, and H. Blood worth. 2003. Effect of organic carbon on soil water retention. *Geoderma*, 116: 61- 76.
34. Rawls, W. J., and D. L. Brakensiek. 1982. Estimating soil water retention from soil properties. *J. Irrig. Drain. Div., Proc. ASCE* 198 (IR2): 166- 171.
35. Rawls, W. J., and D. L. Brakensiek. 1985. Prediction of soil water properties for hydrologic modeling. In: Jones, E., Ward, T. J. (eds.). *Watershed Manag. Eighties. Proceedings of symposium ASCE, Denver, Co, 30April- 2May 1985, ASCE, New York.*
36. Rawls, W. J., D. L. Brakensiek, and B. Soni. 1983. Agricultural management effects on soil water processes: part I. Soil water retention and Green- Ampt parameters. *Trans. ASAE.* 26: 1747- 1752.
37. Saxton, K. E., and W. J. Rawls. 2006. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci. Soc. Am. J.* 70: 1569-1578.
38. Schaap, M. G., and F. J. Leij. 1998. Database-related accuracy and uncertainty of pedotransfer functions. *Soil Sci.* 163: 765-779.
39. Schaap, M., and W., Bouten. 1996. Modelling water retention curves of sandy soils using neural networks. *Water Resour. Res.* 32(10): 3033-3040.
40. Schuh, W. M., R. L. Cline, and M. D. Sweeney. 1988. Comparison of a Laboratory procedure and a textural model for predicting in situ soil water retention. *Soil Sci. Am. J.* 52: 1218-1227.

41. Sepaskhah, A. R., R. Amin, and R. Estakhri. 2004. Watershed sediment yield predication for soils containing rock fragments. *Iranian Journal of Science and Technology* 28: 145- 151.
42. Skaggs, T. H., M. Arya, P. J. Shuse, and B. P. Mohany. 2001. Estimating Particle- size distribution from limited soil texture data. *Soil. Sci. Soc. Am. J.* 65: 1038-1044.
43. Tamari, S., J. H. M. Wosten, and J. C. Ruiz-suarez. 1996. Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Sci. Soc. Am. J.* 60: 1732-1741.
44. Tietje, O., and M. Tapkenhinrichs. 1993. Evaluation of pedo-transfer functions. *Soil Sci. Soc. Am. J.* 57: 1088-1095.
45. Tietje, O., and V. Hennings. 1996. Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO Textural classes. *Geoderma*, 69: 71-84.
46. Tomasella, J., Ya. Pchepsky, S. Crestana, and W. J. Rawls. 2003. Comparison of two techniques to develop pedotransfer functions for water retention. *Soil Sci. Soc. Am. J.* 67: 1085-1092.
47. Ungaro, F., C. Calzolari, and E. Busoni. 2005. Development of pedotransfer functions using a group method of data handling for the soil of Pianura Padano-Veneta region of North Italy: Water retention properties. *Geoderma*, 124: 293-317.
48. Vereecken, H., and M. Herbst, 2005. Statistical regression .In: Pachepsky, Y., Rawls, W. J. (eds.). Development of pedotransfer functions in soil hydrology. *Developments in soil science.* 30: PP.3-19.
49. Wagner, B., V. R. Tamawski, V. Hennings, U. Muller, G. Wessolek, and R. Plagge. 2001. Evaluation of pedo-transfer function for unsaturated soil hydraulic conductivity using an independent data set. *Geoderma*, 102: 215-297.
50. Walczak, R. T., F. Moreno, C. Slawinski, E. Fernandez, and J. L. Arrue. 2006. Modeling of soil water retention curve using soil solid phase parameters. *Journal of Hydrology*, 329: 527- 533.
51. Williams, J., P. Ross, and K. Bristow. 1992. Prediction of the Campbell water retention functions from texture. Structure and organic matter. In: Van Genuchten, M, Th., Leji, F. J., Lund, L. J. (eds.). Indirect methods for estimating the hydraulic properties of unsaturated soils. Proceedings of the international workshop on indirect methods for estimating the hydraulic properties of unsaturated soils, Riverside, California, P: 11-13 October 1989.
52. Willmott. C. J. 1981. On the validation of models. *Phys. Geogr.* 2: 184- 194.
53. Willmott. C. J. 1982. Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.* 63: 1309- 1313.
54. Wosten, J. H. M., P. A. Finke, and M. J. W. Jansen. 1995. Comparison of class and continuous pedotrasfer functions to generate soil hydraulic characteristics *Geoderma*, 66: 227- 237.
55. Wosten, J. H. M., Ta. A. Pachepsky, and W. R. Rawls. 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of hydrology*, 251: 123- 150.