

مقایسه کارایی رگرسیون خطی چندگانه و رگرسیون درختی در برآورد هدایت آبی اشباع و پارامتر عکس طول درشت موئینگی خاک

سیروس دهقانی بانیانی^۱، شجاع قربانی دشتکی^{۲*}، جهانگرد محمدی^۳ و حبیب خداوردی لو^۴

چکیده

هدف از این پژوهش این بود که کارایی رگرسیون خطی چندگانه و رگرسیون درختی در ایجاد توابع انتقالی (PTFs) برای برآورد هدایت آبی اشباع (K_{fs}) و پارامتر عکس طول درشت موئینگی (α^*) مقایسه شود. بدین منظور K_{fs} و α^* در ۶۰ نقطه از دشت آزادگان واقع در منطقه شهرکرد با استفاده از تک استوانه و به روش بارتاب چندگانه اندازه گیری شد. با استفاده از برخی ویژگی های زود یافت اندازه گیری شده در دو لایه پدوژنیکی سطحی و زیر سطحی خاک به عنوان ورودی، PTF های با روش رگرسیون خطی چندگانه و رگرسیون درختی جهت برآورد دو پارامتر یاد شده پی ریزی شدند. اعتبار و صحت توابع پی ریزی شده با این دو روش با استفاده از آماره های جذر میانگین مربعات خطا (RMSE)، میانگین خطا (ME)، خطای نسبی (RE) و ضریب همبستگی پیرسن آزموده شد. نتایج نشان داد که رگرسیون درختی، در برآورد پارامترهای یاد شده معتبرتر است. مقدار RE و RMSE روش رگرسیون درختی در برآورد پارامتر α^* به ترتیب $-0/24$ و $0/19$ عکس سانتی متر بود. حال آنکه مقدار این دو شاخص در رگرسیون خطی چندگانه در بهترین حالت به ترتیب $-0/33$ و $0/05$ عکس سانتی متر بود. مقادیر ME نشان داد با رگرسیون درختی و رگرسیون خطی چندگانه، هدایت آبی اشباع به ترتیب بیشتر و کمتر از مقدار واقعی برآورد شد. مقایسه ی نتایج ارزیابی این دو روش نشان از برتری رگرسیون درختی در برآورد پارامترهای یاد شده دارد.

واژه های کلیدی: پارامتر عکس طول درشت موئینگی، توابع انتقالی خاک، نفوذ، هدایت آبی اشباع خاک.

ارجاع: دهقانی بانیانی س. قربانی دشتکی ش. محمدی ج. و خداوردی لو ح. ۱۳۹۰. مقایسه کارایی رگرسیون خطی چندگانه و رگرسیون درختی در برآورد هدایت آبی اشباع و پارامتر عکس طول درشت موئینگی خاک. مجله پژوهش آب ایران. ۵(۹): ۱۹۳-۲۰۴.

۱- دانشجوی کارشناسی ارشد خاکشناسی، دانشکده کشاورزی، دانشگاه شهرکرد

۲- استادیار گروه خاکشناسی، دانشکده کشاورزی، دانشگاه شهرکرد

۳- دانشیار گروه خاکشناسی، دانشکده کشاورزی، دانشگاه شهرکرد

۴- استادیار گروه خاکشناسی، دانشکده کشاورزی، دانشگاه ارومیه

* نویسنده مسئول: Shoja2002@yahoo.com

تاریخ دریافت: ۱۳۹۰/۰۲/۲۷ تاریخ پذیرش: ۱۳۹۰/۰۷/۳۰

مقدمه

نفوذ آب به خاک، فرآیندی بسیار کلیدی در چرخه‌ی آبی طبیعت است. شناخت کمی نفوذ آب به خاک لازمه مدیریت درست آب خاک و دستیابی به عملکرد بهینه زراعی و بهره‌برداری پایدار از خاک است. لذا، پژوهشگران همواره در پی یافتن روشی مناسب برای کمی‌سازی نفوذ آب به خاک بوده‌اند (قربانی دشتکی و همکاران، ۲۰۱۰). در این زمینه تلاش‌هایی فراوان صورت گرفته و روش‌هایی چند نیز توسط پژوهشگران ارائه شده است که هر کدام برتری‌ها و کاستی‌هایی دارند. از بین روش‌های ارائه شده برای اندازه‌گیری نفوذ آب به خاک، روش استوانه‌های دوگانه، بدلیل سادگی در کاربرد و هزینه کمتر، کاربردی گسترده دارد. در این روش، چنین می‌پندارند که استوانه بیرونی از ایجاد جریان‌های جانبی آب در خاک جلوگیری می‌کند و موجب جریان کاملاً عمودی آب به خاک می‌شود. با این حال، پژوهش‌ها نشان داده‌اند که برای این کار در عمل باید از استوانه‌هایی بسیار بزرگ استفاده کرد (رینولدز و همکاران، ۲۰۰۲). لذا، در سال‌های اخیر استفاده از روش تک استوانه‌ای برای اندازه‌گیری نفوذ آب به خاک و هدایت آبی‌اشباع پیشنهاد شده است. در روش تک استوانه‌ای، برای رسیدن به حالت یکنواخت سرعت نفوذ آب در استوانه نفوذسنج از معادله زیر استفاده می‌شود (رینولدز و همکاران، ۲۰۰۲).

$$\frac{q_s}{K_{fs}} = \frac{H}{C_1 d + C_2 a} + \frac{1}{\alpha^* (C_1 d + C_2 a)} + 1 \quad (1)$$

که در آن q_s نفوذ شبه پایدار آب به خاک، K_{fs} هدایت آبی اشباع خاک در شرایط مزرعه، H ارتفاع آب موجود بر سطح خاک، a شعاع استوانه، d عمق جایگذاری استوانه، C_1 و C_2 به ترتیب برابر $0/316$ و $0/184$ ضربدر عدد پی (π) و α^* پارامتر عکس طول درشت موینگی^۱ خاک است. پارامتر α^* بیانگر روابط بین پتانسیل ثقلی و موینگی خاک است به گونه‌ای که با افزایش مقدار α^* ، پتانسیل ثقلی بر پتانسیل موینگی چیره می‌شود. در حالی که کاهش مقدار این پارامتر بیانگر برتری موینگی بر ثقل است (رینولدز و همکاران، ۲۰۰۲). چون اندازه‌گیری مستقیم ویژگی‌های

هیدرولیکی خاک از جمله K_{fs} و α^* ، وقت‌گیر و پرهزینه است، در سال‌های اخیر استفاده از روش‌هایی همانند توابع انتقالی خاک (PTFs^۱) (بوما، ۱۹۸۹) برای اندازه‌گیری غیرمستقیم این ویژگی‌ها مرسوم شده است (قربانی دشتکی و همکاران، ۲۰۱۰؛ خداوردی‌لو و همکاران، ۲۰۱۱). PTF ها، که معمولاً معادلاتی رگرسیونی هستند، ویژگی‌های دیرپافت هیدرولیکی خاک را با استفاده از ویژگی‌های زودپافت خاک مانند توزیع اندازه‌ی ذرات، جرم ویژه ظاهری و ماده آلی خاک برآورد می‌کنند. نتایج حاصل از PTF ها به عنوان ورودی مدل‌های مختلف در مقیاس حوضه آبخیز و منطقه‌ای بسیار مفید می‌باشند. PTF ها را براساس نوع ورودی توابع به PTF های کلاسی^۲ و پیوسته^۳ تقسیم می‌کنند. ورودی توابع در PTF های کلاسی، کلاس بافت خاک و در PTF های پیوسته، داده‌های پایه‌ای مانند فراوانی نسبی ذرات خاک و جرم ویژه ظاهری هستند (قربانی دشتکی و همکاران، ۲۰۱۰).

روش‌های گوناگونی برای پی ریزی توابع انتقالی خاک وجود داشته (پاچپسکی و اسخاپ، ۲۰۰۴) که آنها را می‌توان در دو گروه کلی قرار داد. نخست، روش‌هایی که در آنها تنها روابط خطی بین داده‌های زودپافت و دیرپافت بررسی می‌شود. از جمله این روش‌ها می‌توان به میانگین‌گیری، رگرسیون ساده و رگرسیون خطی اشاره کرد. در روش دوم، افزون بر رابطه‌ی خطی بین داده‌ها، روابط غیرخطی بین آنها نیز مدنظر قرار می‌گیرد. بدین‌منظور از روش‌هایی چون رگرسیون غیرخطی، رگرسیون درختی و شبکه‌های عصبی مصنوعی برای تعیین رابطه‌ی بین ویژگی‌های زودپافت و دیرپافت استفاده می‌شود.

رگرسیون خطی چندگانه یکی از روش‌های بسیار متداول در پی‌ریزی توابع انتقالی خاک بوده که در منابع مختلف جزییات آن آمده است (یارویس و همکاران، ۲۰۰۲؛ پاچپسکی و اسخاپ، ۲۰۰۴؛ قربانی دشتکی و همکاران، ۲۰۱۰؛ خداوردی‌لو و همکاران، ۲۰۱۱).

1 - Pedotransfer functions(PTF)

2 -Class PTFs

3 -Continuous PTFs

خطی چندگانه و مقایسه کارایی دو روش اخیر در برآورد این پارامترها بود.

مواد و روش‌ها

این پژوهش در خاک‌های دشت شهرکرد، واقع در استان چهارمحال و بختیاری در سال ۱۳۸۸ انجام شد. به همین منظور تعداد ۶۰ نقطه در دانشگاه شهرکرد و دشت آزادگان واقع در منطقه بن انتخاب شد و در هر نقطه برخی ویژگی‌های زودریافت از جمله فراوانی نسبی ذرات به روش هیدرومتری (جی و بادر، ۱۹۸۶)، جرم ویژه‌ی ظاهری به روش سیلندر(بلیک و هارتج، ۱۹۸۶)، ماده‌آلی به روش والکلی بلک (والکلی و بلک، ۱۹۳۴)، میانگین وزنی (MWD) و میانگین هندسی (GMD) قطر خاکدانه‌ها به روش الک تر و به ترتیب با روابط ۲ و ۳، میانگین هندسی و انحراف معیار قطر ذرات به روش شیرازی و بورسما (۱۹۸۴)، کربنات کلسیم معادل به روش تیتراسیون (نلسن، ۱۹۸۲)، رطوبت ولیه و نهایی به روش وزنی و مقدار سنگریزه سطحی و عمقی اندازه‌گیری شد.

$$MWD = \sum_{i=1}^n X_i w_i \quad (2)$$

$$GMD = \exp \left[\sum_{i=1}^n w_i \log X_i \right] \quad (3)$$

در معادلات بالا، X_i متوسط قطر یا اندازه خاکدانه‌ها در هر گروه قطری و w_i نسبت وزنی خاکدانه‌ها در آن گروه قطری است.

برای اندازه‌گیری پارامتر عکس طول درشت‌موییگی (α^*) و هدایت‌آبی‌اشباع در مزرعه (K_{fs}) از روش بار ثابت چندگانه^۲ استفاده شد. بدین منظور، چندین ارتفاع آب درون تک‌استوانه اعمال شد و اجازه داده شد تا نفوذ آب در هر ارتفاع به نفوذ شبه‌پایدار برسد. سپس، K_{fs} و α^* با دو روش مختلف و با استفاده از معادلات زیر محاسبه شدند (کارتر و گرگوریچ، ۲۰۰۶).

$$\text{در روش اول به کمک معادلات زیر } K_{fs} \text{ و } \alpha^* \text{ تعیین شد.} \quad (4)$$

$$K_{fs} = [T(q_2 - q_1)] / (H_2 - H_1)$$

رگرسیون درختی یکی دیگر از روش‌های ایجاد PTF‌های برای برآورد ویژگی‌های دیریافت خاک است. در این روش، ویژگی‌های دیریافت خاک بر اساس متغیرهای زودریافت به شاخه‌های مختلفی تقسیم می‌شوند. به اولین گره، گره مادر گفته می‌شود که به گره‌های فرزند تقسیم می‌شود. تمام داده‌های اندازه‌گیری شده‌ی دیریافت خاک در گره اول یا گره مادر قرار می‌گیرند و گره مادر خود براساس ارتباط بین ویژگی‌های زودریافت و دیریافت خاک به گره‌های فرزند که بیشترین یکنواختی^۱ را دارند تقسیم می‌شود. بنابراین، در نخستین گره، ویژگی دیریافت مدنظر به کمک متغیری که بیشترین سهم را در توجیه تغییرات آن دارد به دو گره فرزند که در واقع میانگین کلی از متغیر دیریافت است تقسیم می‌شود. به همین ترتیب و براساس میزان یکنواختی تقسیم‌بندی دیگر گره‌ها صورت می‌گیرد. بهترین طبقه‌بندی زمانی انجام می‌شود که بهترین متغیر برای ایجاد بیشترین یکنواختی در گره‌های فرزند انتخاب شود. هرچه به سمت گره‌های پایین‌تر حرکت می‌کنیم، مقدار انحراف معیار و تعداد داده‌ها کاهش پیدا می‌کند. اگر طبقه‌بندی تا آخرین مرحله ادامه پیدا کند انحراف معیار به صفر نزدیک می‌شود. به عبارت دیگر، هرچه به سمت گره‌های پایین‌تر حرکت کنیم، همگنی گره افزایش یافته و آن گره برآوردی دقیق‌تر از ویژگی دیریافت مورد نظر به ما می‌دهد. از لحاظ فیزیکی، گره‌های ایجاد شده با ویژگی-های زودریافت خاک هم‌خوانی دارند. برای نمونه، گره‌هایی که مقدار رس بیشتری دارند، هدایت‌آبی‌اشباع کمتری نیز دارند (پاچپسکی و همکاران، ۲۰۰۶؛ تیموفیو، ۲۰۰۴). در پژوهش‌هایی چند برخی از پژوهشگران مانند کلارک و پری‌گیبن (۱۹۹۲) و پاچپسکی و همکاران (۲۰۰۶) نشان دادند که رگرسیون درختی نسبت به رگرسیون خطی چندگانه کارآمدتر است. بررسی منابع نشان داد تاکنون این روش در برآورد پارامترهای عکس طول درشت‌موییگی و هدایت‌آبی اشباع خاک مورد آزمون قرار نگرفته است. هدف از این پژوهش برآورد پارامتر عکس طول درشت‌موییگی و هدایت‌آبی اشباع به کمک رگرسیون درختی و رگرسیون

نرم افزار JMP برای ایجاد رگرسیون درختی استفاده شد. برای آزمودن صحت و اعتبار توابع α^* و K_{fs} ، و همچنین مقایسه کارآیی روش های رگرسیون درختی و رگرسیون خطی چندگانه در ایجاد توابع، از آماره های میانگین خطا (ME)، جذر میانگین مربعات خطا (RMSE)، خطای نسبی (RE) و ضریب همبستگی (R) بین داده های اندازه گیری شده و برآورد شده استفاده شد:

$$ME = \sum_{j=1}^n \frac{X(p)_j - X(m)_j}{n} \quad (8)$$

$$RMSE = \sqrt{\frac{\sum_{j=1}^n (X(p)_j - X(m)_j)^2}{n}} \quad (9)$$

$$RE = \sum_{j=1}^n \frac{|X(p)_j - X(m)_j|}{|X(m)_j|} \quad (10)$$

$$R = \frac{n \sum_{j=1}^n I(p)_j I(m)_j - (\sum_{j=1}^n I(p)_j)(\sum_{j=1}^n I(m)_j)}{\sqrt{[n \sum_{j=1}^n I^2(p)_j - (\sum_{j=1}^n I(p)_j)^2][n \sum_{j=1}^n I^2(m)_j - (\sum_{j=1}^n I(m)_j)^2]}} \quad (11)$$

که $X(p)_j$ مقادیر برآورد شده هر ویژگی از طریق توابع اشتقاق یافته و $X(m)_j$ مقادیر اندازه گیری شده برای خاک j می باشند. n نیز بیان کننده تعداد خاک های مورد استفاده در هر آزمون است (قربانی دشتکی و همکاران، ۲۰۰۹).

آماره R شاخصی برای اندازه گیری خطی بودن رابطه ی بین مقادیر اندازه گیری شده و برآورده شده است. هرچه مقدار آن به یک نزدیک تر باشد، رابطه ی بین دو متغیر به یک رابطه ی خطی نزدیک تر است. مقدار آماره ME بیان کننده آن است که تابع مورد نظر نفوذ آب به خاک را بیشتر (مقادیر مثبت) یا کمتر (مقادیر منفی) از واقعیت برآورد می کنند. مقدار RMSE همواره مثبت یا صفر است و با نزدیک شدن آن به صفر عملکرد توابع اشتقاق یافته افزایش می یابد. مقدار RE نیز همواره مثبت یا صفر بوده و بزرگی مقدار خطا را در مقایسه با مقادیر اندازه گیری شده نشان می دهد و با نزدیک شدن به صفر،

$$\alpha^* = (q_2 - q_1) / [q_1(H_2 + T) - q_2(H_1 + T)] \quad (5)$$

$$T = C_1 d + C_2 a \quad (6)$$

که در آنها، H_1 ، H_2 ارتفاع آب معادل q_1 ، q_2 و a شعاع داخلی استوانه، d عمق جایگذاری استوانه، C_1 و C_2 به ترتیب 0.316π و 0.184π است (کارتر و گرگوریچ، ۲۰۰۶).

در روش دوم با ثابت پنداشتن همه ی متغیرهای مستقل در رابطه ی (۱) به جز متغیر H ، شدت نفوذ شبه پایدار به عنوان تابعی از H و با استفاده از معادله ی رگرسیونی زیر به دست آمد:

$$q_s = (K_{fs} / T)H + K_{fs}[1/(\alpha^* T) + 1] \quad (7)$$

در این روش، شدت نفوذ شبه پایدار جریان در عمق های متناظر اعمال شده ی آب به عنوان تابعی از ارتفاع آب درون استوانه اندازه گیری و در یک نمودار رسم شد. سپس رابطه ی خطی بر داده ها برازش داده شد. همانگونه که رابطه ی γ نشان می دهد شیب خط برازش داده شده برابر با K_{fs}/T و عرض از مبدأ آن برابر با $K_{fs}[1/(\alpha^* T) + 1]$ است. بنابراین، براساس ضرایب خط یاد شده به طور هم زمان هر دو پارامتر K_{fs} و α^* تعیین شد. گفتنی است که در روش نخست سه مقدار برای K_{fs} و α^* بدست می آید که میانگین آنها تقریباً برابر با مقادیر حاصل از روش رگرسیونی است. اندازه گیری α^* با روش های مستقیم یاد شده بسیار دشوار و زمان بر بوده و اگر نفوذ اندازه گیری شده متناظر با ارتفاع آب های اعمال شده به نفوذ شبه پایدار نرسد (برای رسیدن به نفوذ شبه پایدار، شدت نفوذ باید در سه اندازه گیری متوالی برابر گردد)، مقادیری غیر واقعی (منفی) حاصل می شود (رینولدز و همکاران، ۲۰۰۲).

برای برآورد ویژگی های دیر یافت خاک از توابع انتقالی پیوسته و رگرسیون درختی استفاده شد. در این پژوهش، پارامترهای K_{fs} و α^* با استفاده از معادلات رگرسیونی به ویژگی های پایه ای خاک مانند توزیع اندازه ی ذرات، توزیع اندازه ی خاکدانه ها، ماده آلی و جرم ویژه ظاهری مرتبط گردید. از نرم افزارهای SPSS و Minitab برای ایجاد رگرسیون خطی چندگانه و از

نتایج و بحث

دامنه تغییرات ویژگی‌های زودپافت خاک در دو لایه سطحی و زیرسطحی در جدول‌های ۱ و ۲ آمده است. دامنه تغییرات جرم ویژه ظاهری (ρ_b) از ۱/۱۲ تا ۱/۶۹ گرم بر سانتی‌متر مکعب بود. میانگین هندسی قطر خاکدانه‌ها (GMD) و میانگین وزنی قطر خاکدانه‌ها (MWD) به ترتیب از ۰/۷۹ تا ۱/۱۲ و ۰/۳ تا ۱/۹۰ میلی‌متر متغیر بود. با افزایش عمق میانگین وزنی و هندسی قطر خاکدانه‌ها کاهش یافته است که این موضوع احتمالاً به دلیل کاهش ماده آلی خاک و در نتیجه کاهش عامل سیمان‌کنندگی خاک با افزایش عمق خاک است. دامنه تغییرات درصد کربن آلی (OC) در لایه سطحی و عمقی خاک به ترتیب ۰/۷۶ تا ۳/۹ درصد و ۰/۱۷ تا ۱/۷ درصد که نشان‌دهنده کاهش ماده آلی با افزایش عمق خاک بوده است. مقدار کربنات کلسیم (CaCO_3) در لایه سطحی و زیرسطحی به ترتیب از ۵ تا ۴۷ درصد و ۰/۱۲ تا ۶۹ درصد متغیر بود. دامنه تغییرات درصد رس، شن، سیلت، میانگین هندسی قطر ذرات و انحراف معیار قطر ذرات در لایه سطحی به ترتیب ۱۴ تا ۶۴ درصد، ۱۰ تا ۵۸ درصد، ۲۶ تا ۶۰ درصد، ۰/۰۴ تا ۰/۱۲ میلی‌متر و ۴/۴۹ تا ۱۹/۸۶ میلی‌متر و در لایه زیرسطحی به ترتیب ۱۶ تا ۶۶ درصد، ۷ تا ۳۴ درصد، ۱۳ تا ۸۲ درصد، ۰/۰۳۵ تا ۰/۱۱ میلی‌متر و ۱۹/۳۵ تا ۴/۰۳ میلی‌متر بود. بیشینه، کمینه و میانگین نفوذ طی مدت ۱۱۰ دقیقه انجام آزمایش در خاک‌های مورد مطالعه به ترتیب ۱۵، ۴/۵ و ۶/۷ سانتی‌متر بود. همچنین، کمینه، بیشینه و میانگین شدت نفوذ شبه‌پایدار در خاک‌های مورد مطالعه به ترتیب ۰/۰۶۹، ۲/۱۷ و ۰/۳ سانتی‌متر بر دقیقه به دست آمد. در آزمون‌های آماری نخستین گام در تجزیه و تحلیل داده‌ها نرمال کردن داده‌ها است، زیرا آزمون‌های فرض خود براساس نرمال بودن داده‌ها پی‌ریزی شده است و نرمال نبودن داده‌ها این آزمون‌ها و تجزیه و تحلیل‌ها را نامعتبر می‌سازد (قربانی دشتکی و همکاران، ۲۰۱۰). بررسی توزیع داده‌های اندازه‌گیری شده نشان داد از بین داده‌های زودپافت و دیرپافت خاک تنها هدایت‌آبی اشباع خاک، توزیعی غیرنرمال دارد. یارویس و

عملکرد توابع افزایش می‌یابد (قربانی دشتکی و همکاران، ۲۰۱۰). همانگونه که در بخش مقدمه گفته شد در روش رگرسیون درختی، ویژگی‌های دیرپافت خاک بر اساس متغیرهای زودپافت به شاخه‌های مختلفی تقسیم می‌شود. به اولین گره، گره مادر گفته می‌شود که به گره‌های فرزند تقسیم می‌گردد. همگی خاک‌ها که هر کدام یک هدایت‌آب‌اشباع دارند در گره مادر قرار می‌گیرند. گره مادر بر اساس یکی از ویژگی‌های زودپافت خاک که بیشترین یکنواختی را در گره‌های فرزند ایجاد می‌کند تقسیم می‌شود. به عنوان مثال هدایت‌آبی اشباع خاک‌هایی که جرم ویژه ظاهری آن از یک مقدار کمتر است در یک گره و بقیه مقادیر در گره دیگر قرار می‌گیرند. آنگاه هر کدام از گره‌های فرزند به نوبه خود به دو گره دیگر تقسیم می‌شود. اساس تقسیم‌بندی همه گره‌ها، یکنواختی در گره است. بهترین طبقه‌بندی زمانی انجام می‌شود که بهترین متغیر برای ایجاد بیشترین یکنواختی در گره‌های فرزند انتخاب گردد. هرچه به سمت گره‌های پایین‌تر حرکت می‌کنیم، مقدار انحراف معیار و تعداد داده‌ها کاهش پیدا می‌کند. اگر طبقه‌بندی تا آخرین مرحله ادامه پیدا کند انحراف معیار به صفر نزدیک می‌شود. به عبارتی دیگر هرچه به سمت گره‌های پایین‌تر حرکت کنیم، همگنی گره افزایش یافته و آن گره برآوردی دقیق‌تر از ویژگی دیرپافت مورد نظر به ما می‌دهد. هر کدام از این گره‌های ایجاد شده یک مقدار میانگین از هدایت‌آب‌اشباع یا ویژگی دیرپافت مورد نظر به ما می‌دهد که این مقدار همان مقدار برآورد شده است. به عبارتی دیگر هر گره‌ها شامل خاک‌هایی است که نزدیکترین هدایت‌آبی اشباع را به هم دارند.

در روش رگرسیون درختی به منظور حصول شرایط بالا تابع ناخالصی^۱ یا $\hat{i}(t)$ تعریف شد. براساس این تابع، هر گره براساس یکنواختی به گره‌های فرزند تقسیم شده و بیشترین یکنواختی گره‌های فرزند برابر با بیشترین تغییرات تابع ناخالصی، $\Delta \hat{i}(t)$ است. معادله یاد شده به‌گونه‌ی زیر است:

$$\Delta \hat{i}(t) = \hat{i}(t_p) - E[\hat{i}(t_c)] \quad (12)$$

که در آن $\hat{i}(t_p)$ تابع ناخالصی گره مادر، $E[\hat{i}(t_c)]$ تابع ناخالصی گره‌های فرزند است (تیموفیو، ۲۰۰۴).

هدایت‌آبی‌اشباع را کمتر از مقدار واقعی برآورد نموده‌اند. با توجه به آماره‌های RMSE، RE و R می‌توان گفت که تابعی که در آن جرم ویژه ظاهری و میانگین هندسی قطر خاکدانه‌ها وارد شده است بهترین عملکرد را در تعیین هدایت‌آبی‌اشباع داشته است و رتبه ۱ به آن داده شده است. تابعی که در آن‌ها تنها از میانگین وزنی و میانگین هندسی قطر خاکدانه‌ها استفاده شده بود، بیشترین خطا را در تخمین هدایت‌آبی‌اشباع داشته‌اند. این نتایج نشان می‌دهد در خاک‌های مورد مطالعه در این پژوهش، جرم ویژه ظاهری و میانگین هندسی قطر خاکدانه‌ها بیشترین اثر را بر هدایت‌آبی‌اشباع داشتند.

همکاران (۲۰۰۲) و وستن و همکاران (۲۰۰۱) نیز بیان داشتند هدایت‌آبی‌اشباع از توزیعی لگاریتمی دارد. به همین منظور برای نرمال‌سازی داده‌های هدایت‌آبی اشباع از تبدیل لگاریتمی استفاده شد. براساس روش‌های ذکر شده توابع انتقالی زیادی پی‌ریزی گردید که در اینجا روش‌هایی که بهترین عملکرد را در برآورد ویژگی‌های مورد نظر داشته‌اند ارائه شده‌اند. بهترین توابع اشتقاق یافته (۳ تابع اول) به همراه نتایج آزمون اعتبارسنجی و صحت‌سنجی آنها در جدول ۳ ارائه شده است. مقادیر میانگین خطا (ME) در این جدول نشان می‌دهد که همگی توابع اشتقاق یافته مقدار

جدول ۱- دامنه تغییرات ویژگی‌های زود یافت خاک در دو لایه سطحی و زیر سطحی

ویژگی‌های زود یافت خاک	ρ_b (g/cm^3)	OC (%)	MWD (mm)	GMD (mm)	CaCO ₃ (%)	سنگریزه (%)
لایه سطحی	۱/۱۲-۱/۶۹	۰/۰۸-۳/۹۰	۰/۳۰-۱/۹۰	۰/۷۹-۱/۱۲	۰ - ۴۷	۳-۳۵
لایه زیر سطحی	۱/۱۹-۱/۷۸	۰/۱۷-۱/۷۰	۰/۱۰-۱/۵۰	۰/۳۵-۰/۸۹	۰/۱۲-۶۹	۵-۴۰

جدول ۲- دامنه تغییرات توزیع اندازه‌های ذرات، میانگین و انحراف معیار هندسی قطر ذرات در دو لایه سطحی و زیر سطحی

توزیع اندازه‌های ذرات	(%) رسی	(%) ماسه	(سیلت) (%)	d_g (mm)	σ_g (mm)
لایه سطحی	۱۴-۶۴	۱۰-۵۸	۲۶-۶۰	۰/۰۰۴-۰/۱۲	۴/۴۹-۱۹/۸۶
لایه زیر سطحی	۱۶-۶۶	۷-۳۴	۱۳-۸۲	۰/۰۰۳۵-۰/۱۱	۴/۰۳-۱۹/۳۵

جدول ۳- آزمون صحت و اعتبار توابع پی‌ریزی شده برای برآورد هدایت‌آبی‌اشباع (K_{fs}) و پارامتر عکس طول درشت‌مویبگی (α^*)

تابع	آزمون صحت			آزمون اعتبار			رتبه		
	RE (-)	RMSE*	ME*	R	RE (-)	RMSE*		ME*	
$\log(K_{fs}) = 1.01 - 0.64\rho_b + 0.05MWD$	۰/۳۲	۰/۰۳۱	-۰/۰۰۵	۰/۸۶	۰/۲۵	۰/۰۹	-۰/۰۰۵	۸۳	۲
$\log(K_{fs}) = 1.24 - 2.30\rho_b + 0.671GMD$	۰/۲۵	۰/۰۲۹	-۰/۰۰۴	۰/۸۷	۰/۲۳	۰/۰۳۵	-۰/۰۰۴	۸۰	۱
$\log(K_{fs}) = 2.19 - 2.53\rho_b$	۰/۳۰	۰/۰۳۳	-۰/۰۰۶	۰/۸۵	۰/۲۷	۰/۰۴۲	-۰/۰۰۵	۷۹	۳
$\log(\alpha^*) = -0.194 - 6.374\log(\rho_b)$	۰/۲۸	۰/۰۴۸	-۰/۰۰۵	۰/۷۳	۰/۳۳	۰/۰۵۰	-۰/۰۰۶	۰/۶۸	۱
$\log(\alpha^*) = -0.890 + 3.95\log(GMD)$	۰/۴۹	۰/۰۵۸	-۰/۰۱۵	۰/۳۲	۰/۵۹	۰/۰۶۶	-۰/۰۱۶	۰/۲۸	۲

* واحد آماره‌های ارائه شده برای K_{fs} و α^* به ترتیب $(cm \cdot min^{-1})$ و (cm^{-1}) است.

بیشتر از مقدار واقعی نشان می‌دهد. این جدول همچنین نشان داد که رگرسیون درختی چه در آزمون صحت‌سنجی و چه در آزمون اعتباریابی نسبت به رگرسیون خطی چندگانه کاراتر بوده است. مقدار خطای نسبی (RE) و مجذور میانگین مربعات خطا (RMSE) به ترتیب ۰/۰۲۵ سانتی‌متر بر دقیقه و ۰/۲۱ است، که نسبت به رگرسیون خطی چندگانه، مقادیر کمتری هستند.

جدول ۴- آزمون اعتبار و صحت سنجی برای رگرسیون درختی، جهت برآورد هدایت‌آبی اشباع

نوع آزمون	R	ME (cm/min)	RMSE (cm/min)	RE (-)
اعتبار سنجی	۰/۸۹	۰/۱۱	۰/۰۲۵	۰/۲۱
صحت سنجی	۰/۹۰	۰/۰۹	۰/۰۲۳	۰/۱۹

توابع ارایه شده و نتایج آزمون‌های اعتبارسنجی و صحت‌سنجی آنها در برآورد پارامتر عکس طول‌درشت موینگی با استفاده از تمامی داده‌ها در جدول ۳ آمده است. این جدول نشان می‌دهد جرم ویژه ظاهری خاک به تنهایی بیش از ۵۰ درصد از تغییرات پارامتر عکس طول‌درشت موینگی را توجیه نموده است. تابع پی ریزی شده به کمک جرم ویژه ظاهری بهترین عملکرد را در برآورد این پارامتر داشته است. به گونه‌ای که این تابع هم در بخش اعتبارسنجی و هم در بخش صحت‌سنجی کمترین میزان خطای نسبی (به ترتیب ۰/۲۸ و ۰/۳۳) و جذر میانگین مربعات خطا را داشته است. مقادیر میانگین خطا (ME) نشان می‌دهد که تمامی توابع پی‌ریزی شده، پارامتر عکس طول‌درشت موینگی را کمتر از مقدار واقعی برآورد کرده‌اند.

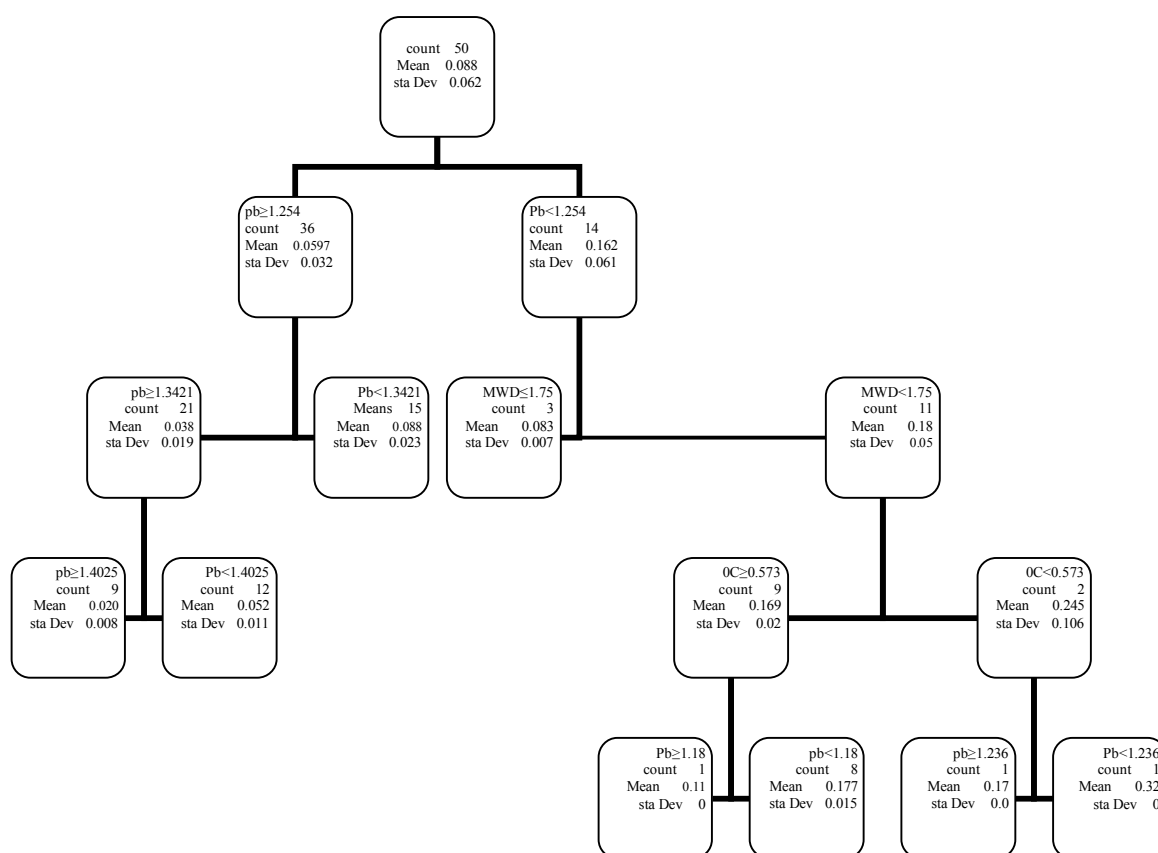
رگرسیون درختی پی‌ریزی شده برای برآورد پارامتر طول‌درشت موینگی در شکل ۲ ارایه شده است. پارامتر طول‌درشت موینگی در اولین تقسیم‌بندی بر اساس جرم ویژه ظاهری خاک‌ها به دو گروه تقسیم شده است. این موضوع نشان می‌دهد که جرم ویژه ظاهری مهمترین متغیر در برآورد پارامتر عکس طول‌درشت موینگی است. این واقعیت با آنچه که در رگرسیون

میانگین وزنی قطر خاکدانه‌ها نسبت به دو ویژگی دیگر (جرم ویژه ظاهری و میانگین هندسی قطر خاکدانه‌ها) کمترین مقدار از تغییرات هدایت‌آبی اشباع را در خاک‌های مورد مطالعه توجیه نموده است. این امر می‌تواند ناشی از این واقعیت باشد که توزیع خاکدانه‌ها به جای اینکه یک توزیع نرمال باشد، توزیعی لگ‌نرمال است. این نتایج نشان می‌دهد که بطور کلی با افزایش میانگین وزنی و هندسی قطر هدایت‌آبی اشباع افزایش و همچنین با کاهش جرم ویژه ظاهری هدایت‌آبی اشباع افزایش می‌یابد. این موضوع با مطالعات مطالعات وریکن و همکاران (۱۹۹۰)، وستن و همکاران (۲۰۰۱)، پاجپسکی و همکاران (۲۰۰۶)، چریکو و همکاران (۲۰۰۷)، سالازر و همکاران (۲۰۰۸) نیز هم‌خوانی دارد. رگرسیون درختی پی‌ریزی شده برای برآورد هدایت‌آبی اشباع در شکل ۱ آمده است. این شکل نشان می‌دهد که هدایت‌آبی اشباع در اولین تقسیم‌بندی بر اساس جرم ویژه ظاهری به دو گروه تقسیم شده است. در نخستین تقسیم‌بندی، نمونه‌هایی که جرم ویژه ظاهری آنها کمتر از ۱/۲۵ گرم بر سانتی‌متر مکعب بود در یک گره و آنهایی که جرم ویژه ظاهری بزرگتر مساوی ۱/۲۵ بود در گره دیگر قرار گرفتند. نفوذ آب به خاک به شدت تحت تاثیر بافت و ساختمان خاک است. در این پژوهش بیشتر خاک‌های مورد مطالعه در یک کلاس بافتی قرار داشتند، در نتیجه هدایت‌آبی اشباع بیشتر تحت تاثیر ویژگی‌های ساختمانی خاک قرار گرفت. شکل ۲ نیز موید این نکته می‌باشند. هر چه به سمت گره‌های پایین‌تر حرکت می‌کنیم مقدار انحراف معیار و RMSE کاهش پیدا می‌کند. بطوریکه مقدار RMSE در شاخه اول، دوم، سوم، چهارم و در آخرین شاخه به ترتیب ۰/۰۴۸، ۰/۰۳۵، ۰/۰۲۷، ۰/۰۲۲ و ۰/۰۱۵ بوده است. از شاخه پنجم به بعد تقسیم‌بندی بیشتر شاخه‌ها باعث کاهش چندانی در مقادیر RMSE نمی‌شود.

آماره‌های به‌دست آمده از ارزیابی رگرسیون درختی در برآورد هدایت‌آبی اشباع در جدول ۴ آمده است. مقدار میانگین خطا (ME) نشان داد که رگرسیون درختی هدایت‌آبی اشباع را برخلاف رگرسیون خطی چندگانه

می‌توان شاخه‌ها را دنبال کرد و برآوردی از ویژگی‌های دیرپافت مورد نظر به دست آورد. به عنوان مثال اگر در نمونه‌ی مدنظر جرم ویژه ظاهری از ۱/۲۵ گرم بر سانتی‌متر مکعب کمتر بود وارد شاخه سمت راست شده و میانگینی برای ویژگی مورد نظر در آنجا آمده است. با داشتن ویژگی بعدی یعنی MWD می‌توان به شاخه‌ی بعدی مراجعه و مقدار دقیق‌تری از ویژگی دیرپافت را بدست آورد. براین اساس، هرچه به سمت گره‌های پایین‌تر حرکت می‌کنیم، برآورد ما دقیق‌تر می‌شود.

خطی چندگانه آمده است هم‌خوانی دارد. همان‌گونه که از شکل پیداست در نخستین تقسیم‌بندی نمونه‌هایی که جرم ویژه ظاهری آنها کمتر از ۱/۲۵ گرم بر سانتی‌متر مکعب بود در یک گره و آنهایی که جرم ویژه ظاهری بزرگتر مساوی ۱/۲۵ بود در گره دیگر قرار گرفتند. بعد از جرم ویژه ظاهری میانگین هندسی قطر، درصد سیلت، شن و درصد رس بیشترین تاثیر در ایجاد گره‌های جدید دارند. برای استفاده از این نوع رگرسیون با در دست داشتن ویژگی‌های زودپافت مورد نیاز



شکل ۱- استفاده از رگرسیون درختی برای بدست آوردن هدایت‌آبی اشباع در تمامی نمونه‌های خاک (count, mean, sta Dev) به ترتیب تعداد، میانگین و انحراف معیار هدایت‌آبی اشباع در هر گره است)

آنها بیشتر است، پارامتر عکس طول درشت موینگی نیز در آنها بیشتر است. با افزایش میزان شن احتمالاً درشت موینه‌های خاک و در نتیجه پارامتر عکس طول درشت موینگی افزایش می‌یابد. از سوی دیگر گره‌هایی که درصد رس کمتری دارند پارامتر عکس طول درشت موینگی بیشتری نیز دارند. آماره‌های

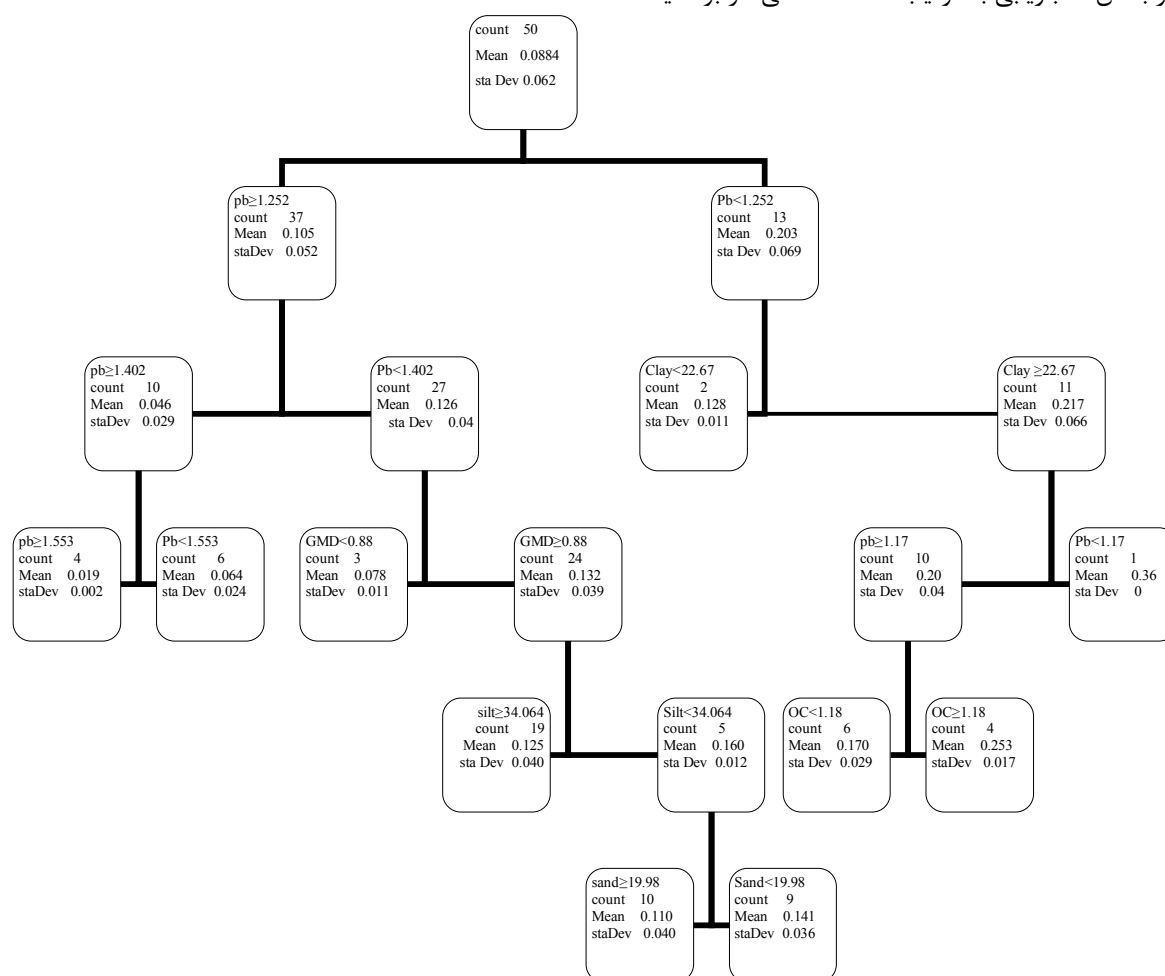
شکل ۲ همچنین نشان می‌دهد که در برآورد پارامتر عکس طول درشت موینگی به کمک رگرسیون درختی علاوه پارامترهای موثر در ساختمان، توزیع اندازه‌های ذرات، مانند درصد شن و سیلت هم موثر بوده است. این امر احتمالاً به دلیل حساسیت بیشتر این پارامتر به توزیع اندازه‌های ذرات است. گره‌هایی که درصد شن در

و ۰/۲۴ است، که نسبت به کارآترین مدل اشتقاق یافته از رگرسیون خطی چندگانه، مقادیر کمتری هستند.

جدول ۵- آزمون اعتبار و صحت سنجی برای رگرسیون درختی، جهت برآورد عکس طول درشت‌موییگی

RE (-)	RMSE (cm ⁻¹)	ME (cm ⁻¹)	R	نوع آزمون
۰/۲۴	۰/۰۱۹	-۰/۰۰۶۲	۰/۸۶	اعتبار سنجی
۰/۲۱	۰/۰۱۸۶	-۰/۰۰۰۶	۰/۸۸	صحت سنجی

به‌دست آمده از ارزیابی رگرسیون درختی در برآورد پارامتر عکس طول درشت‌موییگی در جدول ۵ آمده است. مقدار میانگین خطا (ME) نشان داد که رگرسیون درختی پارامتر عکس طول درشت‌موییگی نیز کمتر از مقدار واقعی برآورد کرده است. این جدول همچنین نشان داد که رگرسیون درختی چه در آزمون صحت‌سنجی و چه در آزمون اعتباریابی نسبت به رگرسیون خطی چندگانه کارآتر بوده است. مقدار خطای نسبی (RE) و مجذور میانگین مربعات خطا (RMSE) در بخش اعتباریابی به ترتیب ۰/۰۱۹ سانتی‌متر بر دقیقه



شکل ۲- استفاده از رگرسیون درختی برای بدست آوردن طول درشت‌موییگی (count, mean, sta Dev) به ترتیب تعداد، میانگین و انحراف معیار هدایت آبی اشباع در هر گره است

استفاده کردند دریافتند که رگرسیون درختی نسبت به رگرسیون خطی چندگانه از دقت بیشتری برخوردار است.

کلارک و پری گین (۱۹۹۲)، سالچو و همکاران (۱۹۹۶) و پاچپسکی و همکاران (۲۰۰۶) که از رگرسیون درختی برای برآورد نقاطی از منحنی رطوبتی

- Science Society of America, Madison. WI, 363-375.
- 2-Bouma J. 1989. Using soil survey data for quantitative land evaluation. *Soil Science Society of America Journal* 9: 177-213.
- 3-Carter M. R. and Gregorich E.G. 2006. Soil sampling and method of analysis. Second edition. Part 77. Saturated hydrolic properties: Ring infiltrometer. pp. 1043-1056.
- 4-Chirico G.B. Medina H. and Romano N. 2007. Uncertainty in predicting soil hydraulic properties at the hillslope scale with indirect methods. *Journal of Hydrology* 334: 405-422.
- 5-Clark L.A. and Pregibon D. 1992. Tree-based models. In: Hastie, T. J.(Ed.), *Statistical Models*. S. Wadsworth, Pacific Grove, California, 377-419.
- 6-Carles M. R. 2008. Applicability of site-specific pedotransfer functions and Rosetta model for the estimation of dynamic soil hydraulic properties under different vegetation covers. *Soils Sediments* 8(2):137-145
- 7-Ghorbani Dashtaki Sh. Homae M. Mahdian M.H. and Kouchakzadeh M. 2009. Site-dependence performance of infiltration models. *Water Resour Manage. Springer Science + Business Media B.V.*
- 8-Ghorbani Dashtaki Sh. Homae M. and Khodaverdiloo H. 2010. Derivation and validation of pedotransfer functions for estimating soil water retention curve using a variety of soil data. *Soil Use and Mangement* 26: 68-74.
- 9-Jarvis N.J. Zavattaro L. Rajkai K. Reynolds W.D. Olsen P.A. McGechan M. Mecke M. Mohanty B. Leeds-Harison P.B. and Jacques D. 2002. Indirect estimation of near saturated hydraulic conductivity from readily available soil information. *Geoderma* 108:1-17.
- 10-Khodaverdiloo H. Homae M. van Genuchten M.Th. and Ghorbani Dashtaki Sh.. 2011. Deriving and Validating Pedotransfer Functions for some Calcareous Soils. *Journal of Hydrology* 399: 93-99.
- 11-Mckenzie N. Coughlan K. and Cresswell H. 2002. Soil physical measurement and interpretatation for land evaluation. Part 7. SBS, PVT. LTD. 108-119.
- 12-Nelson R.E. 1982. Carbonate and Gypsum. In: Miller R.H. and Keeney D.R. (Eds) *Methods of Soil Analysis*. Part 2. 2nd ed. Agron. Monogr. 9. ASA, Madison, WI. 181-197.
- 13-Pachepsky Ya. and Schaap M.G. 2004. Data mining and extrapolation techniques. In: Pachepsky Ya. and Rawls W.J. (Eds), *development of pedotransfer functions in soil*

نتیجه گیری

روش های مختلفی برای اندازه گیری همزمان هدایت آبی اشباع و پارامتر عکس طول درشت موینگی وجود دارد که اساس کار آنها بیشتر بر پایه استوانه های مجاور چندتایی است. یکی از این روش ها، روش بار ثابت چندگانه است. نتایج این پژوهش نشان می دهد که روش بار ثابت چندگانه یک روش مناسب در برآورد همزمان هدایت آبی اشباع و پارامتر عکس طول درشت موینگی است. از سوی دیگر پارامتر عکس طول درشت موینگی و هدایت آبی اشباع بیشتر تحت تاثیر ویژگی های ساختمانی خاک از جمله میانگین هندسی قطر خاکدانه ها (GMD) و میانگین وزنی قطر خاکدانه ها (MWD) و بویژه جرم ویژه ظاهری (ρ_b) خاک بودند. نتایج همچنین نشان داد که رگرسیون درختی برآورد بهتری از ویژگی های دیر یافت خاک نسبت به رگرسیون خطی چندگانه دارد. از سوی دیگر در رگرسیون درختی نیاز به نرمال کردن و تقسیم بندی داده ها نیست. همچنین در رگرسیون درختی نیازی به توجه کردن به فاکتورهایی مانند همراستایی چندگانه و ضریب تورم واریانس نیست. با توجه به برتری روش رگرسیون درختی بر رگرسیون خطی در این پژوهش، پیشنهاد می شود دیگر روش های غیرخطی داده کاوی از قبیل شبکه های عصبی مصنوعی و داده گردانی گروهی¹ به منظور پی ریزی توابع انتقالی خاک و برآورد پارامترهای یاد شده در این پژوهش بررسی و آزمایش شود.

سپاس گذاری

بدین وسیله از دانشگاه شهرکرد که اعتبار مالی لازم برای انجام این پژوهش را فراهم نمود تشکر و قدردانی می شود.

منابع

- 1-Blake G.R. and Hartge K.H. 1986. Bulk density. In: Klute A. (Ed) *Methods of Soil Analysis*. Part 1. 2nd ed. Agron. Monogr. No. 9.: American Society of Agronomy and Soil

- pedotransfer function model. AGWAT-2612:1-9
- 19-Timofeev R. 2004. Classification and Regression Trees (CART) Theory and Applications. CASE - Center of Applied Statistics and Economics Humboldt University, Berlin. 1:39
- 20-Vereecken H. J. Maes and Feyen J. 1990. Estimating unsaturated hydraulic conductivity from easily measured soil properties. Soil Science 149(1): 1-12.
- 21-Walkley A. and Black I.A. 1934. An examination of the Degtjareff method for determining soil organic matter and a proposed modification of the chronic acid titration method. Soil Science 37: 29-39.
- 22-Wosten J.H.M. Pachepsky Y.A. and Rawls W.J. 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. Journal of Hydrology 251: 123-150.
- hydrology. Elsevier, Amsterdam, The Netherlands. 21-32.
- 14-Pachepsky Ya., Rawls W.J. and Lin H.S. 2006. Hydropedology and pedotransfer functions. Geoderma 131:308-316.
- 15-Reynolds W.D. Elrick D.E. and Youngs E.G. 2002. Ring or Cylinder Infiltrometers (vadose zone). In: Dane J.H. and G.C. Topp (Eds.), Methods of soil analysis, Part 4. Physical methods, SSSA, Wisconsin, USA. 818-826.
- 16-Salchow E. Lal R. Fausey N. R. and Ward A. 1996. Pedotransfer functions for variable alluvial soils in southern Ohio. Geoderma 73: 165-181.
- 17-Shirazi M.A. and Boersma L. 1984. A unifying quantitative analysis of soil texture. Soil Science Society of America Journal 48: 142-147.
- 18-Slazar O. Wesstrom I. and Joel A. 2008. Evaluation of DRAINMOD using saturated hydraulic conductivity estimated by a

