دوازدهمین کنگره علوم خاک ایران
تبریز، 12 الی 14 شهریور 1390
(فیزیک خاک و رابطه آب خاک و گیاه)

تعیین بهترین مدل توزیع اندازه خاکدانه با استفاده از شاخص‌های آماری

محمدتقی تیرگرسلطانی[1]، منوچهر گرجی[2] و محمدحسین محمدی[3]

1- کارشناس ارشد گروه مهندسی علوم خاک، دانشگاه تهران stsoltani@ut.ac.ir 2- استادیار گروه مهندسی علوم خاک، دانشگاه تهران
3- استادیار گروه خاکشناسی دانشگاه زنجان

چکیده

یکی از بهترین روش‌های کمی کردن ساختمان خاک، بیان توزیع اندازه خاکدانه در قالب یک مدل ریاضی است. در تحقیق حاضر ضمن ارزیابی و مقایسه مطلوبیت مدل‌های جکی، توزیع نرمال، توزیع لاگ نرمال، رازین راملر، فردلاند، و فرکتالی در توصیف کمی منحنی توزیع تجمعی جرم- اندازه خاکدانه‌ها، بهترین مدل بر اساس کمترین تعداد پارامتر و مقدار خطا تعیین گردید. مدل‌های مذکور بر داده‌های توزیع تجمعی خاکدانه‌ها در کاربری‌های مختلف اراضی برازش داده شدند و علاوه بر مقایسه ضرایب تبیین رگرسیون، از آماره‌های مالو و اکاییک در ارزیابی مطلوبیت مدل‌ها و گزینش بهترین مدل استفاده شد. نتایج مقایسه ضرایب تبیین نشان داد که توزیع اندازه خاکدانه‌ها در اراضی کشاورزی و جنگل تخریب شده از مدل لاگ‌نرمال، و در خاک‌های جنگلی و مرتعی از مدل نرمال پیروی می‌کند. این نتایج نشان می‌دهند که مدل فرکتالی(یک پارامتری) و مدل فردلاند(سه پارامتری) در تمام خاک‌ها از دقت بیشتری نسبت به سایر مدل‌ها برخوردار می‌باشند. نتایج بررسی آماره‌ها در انتخاب بهترین مدل نیز نشان داد که مطلوبیت مدل فرکتالی در تمام خاک‌ها بیشتر از سایر مدل‌ها است. دقت برازش بالاتر، سادگی (یک پارامتری بودن)، و فیزیکی بودن مدل فرکتالی از دلایل این انتخاب بود.

کلمات کلیدی : آماره‌های ارزیابی مطلوبیت مدل، مدل‌های توزیع اندازه خاکدانه، مدل فرکتالی

مقدمه

با توجه به محدوده وسیع اندازه‌ی واحدهای ساختمانی در خاک‌های مختلف، به کارگیری روش‌های مختلف ارزیابی ساختمان خاک، نیازمند مطالعه روابط بین کلاس‌های اندازه خاکدانه، و فرایندهای مختلف خاکدانه‌سازی است (امزکتا1999). توزیع اندازه واحدهای ثانویه خاک(خاکدانه‌ها) با بسیاری از فرایندهای کشاورزی و زیست محیطی در ارتباط است، از این رو توصیف و کمی نمودن خصوصیات ساختمانی خاک از اهمیت ویژه‌ای برخوردار می‌باشد. یکی از روش‌های دقیق کمی کردن توزیع اندازه خاکدانه‌ها، بیان این توزیع به شکل یک مدل ریاضی است. بر این اساس مدل‌های مختلفی بر پایه نوع توزیع آماری خاکدانه‌ها(نرمال، لاگ‌نرمال، ویبول و...)، تجربی(مانند معادله بالداک و کی1987) و یا هندسه فرکتالی(مانند معادله تیلر و ویت‌کرافت1992) پیشنهاد شده‌اند. در بسیاری از این مدل‌ها، صرف نظر از مبانی و فرضیات پایه‌ای توسعه مدل، دقت برازش(ضریب تبیین)[1] بالاتر می‌تواند ناشی از بیشتر بودن تعداد پارامترهای آن مدل نیز باشد. این امر اغلب موجب پیچیدگی مدل و ایجاد ابهام در تفسیر نتایج حاصل از آن شده و از مطلوبیت مدل مربوطه می‌کاهد. بر همین اساس بوچان و همکاران (1993) اعلام کردند که R^2 معیار کافی برای انتخاب بهترین مدل نیست، و لازم است علاوه

[1] - Coefficient of Determination (R^2)

بر آن، تعداد پارامترهای مدل نیز مورد ارزیابی قرار گیرند. این محققین استفاده از آماره‌های F و مالو[1] (مالو1973) را در تعیین بهترین مدل توصیف کننده داده‌های توزیع اندازه ذرات[2] پیشنهاد نمودند. هوانگ و همکاران(2002) نیز آماره‌های مالو و اکائیک[3] (کاررا و نیومن1986) را در ارزیابی مطلوبیت مدل‌ها و تعیین بهترین مدل توصیف کمی منحنی PSD به کار بردند. چنین ارزیابی‌هایی بر روی داده‌های توزیع اندازه ذرات اولیه[4] انجام شده اما مطالعه‌ای که مطلوبیت مدل‌ها را برای داده‌های توزیع اندازه خاکدانه[5] مورد بررسی قرار داده باشد یافت نگردید. در تحقیق حاضر میزان مطلوبیت شش مدل مختلف در کمی نمودن ASD با استفاده از R^2 و آماره‌های C_p و AIC مورد ارزیابی و مقایسه قرار گرفته و بهترین مدل توصیف کمی منحنی ASD در کاربری‌های مختلف اراضی تعیین می‌گردد.

مواد و روش‌ها

نمونه برداری در غرب استان مازندران از تعدادی کاربری‌های زراعی، مرتعی، جنگل سوزنی برگ، جنگل پهن برگ و جنگل درختچه‌ای (مجموعا 14کاربری) انجام شد. این کاربری‌ها به گونه‌ای انتخاب شدند که تنوع وسیعی از وضعیت ساختمانی خاک در تشریح پروفیلی آنها مشاهده گردید. داده‌های منحنیPSD خاک‌ها به روش هیدرومتر(گی و بادر1986) و منحنی ASD نیز به روش الک تر(یودر1936) به دست آمدند. قطر خاکدانه‌ها(هر بخش اندازه‌ای) به عنوان متغیر مستقل (x_i)، و جرم تجمعی خاکدانه‌های کوچکتر از x_i به عنوان متغیر وابسته (y)، وارد نرم‌افزار MATLAB(7.10) شدند. در این تحقیق میزان مطلوبیت مدل‌های جکی(جکی1944)، نرمال(الن1997)، لاگ‌نرمال(گاردنر1956)، رازین‌راملر(ایرانی وکالیس1963)، فردلاند(فردلاند و همکاران2000)، و فرکتالی (تیلر و ویت‌کرافت1992) (جدول1) در توصیف منحنی ASD هر یک از کاربری‌ها مورد مقایسه قرار گرفت. به این منظور مقادیر R^2 و خطا[6] با استفاده از برازش غیر خطی مدل‌های مذکور بر داده‌های توزیع تجمعی جرم- اندازه خاکدانه‌ها در هر یک از کاربری‌ها، به دست آمدند (جدول3). سپس به منظور مقایسه دقت برازش مدل‌ها، R^2های به دست آمده در تمام نمونه‌های مورد مطالعه، دو به دو به کمک آزمونt جفتی مورد مقایسه قرار گرفتند.

[1] - Mallows Statistic (C_p)
[2] - Particle Size Distribution (PSD)
[3] - Akaike's Information Criterion (AIC)
[4] - Particle Size Distribution (PSD)
[5] - Aggregate Size Distribution (ASD)
[6] - Sum of Squares Error (SSE)

جدول1- مدل‌های مورد استفاده و پارامترهای هر یک از آنها

نام مدل	معادله	پارامترها
جکی	$M(<x)=\exp\left\{-\frac{1}{P^2}\left[\ln(\frac{d}{d_0})\right]^2\right\}$	P: شاخص توزیع اندازه خاکدانه d_0 : قطر بزرگترین خاکدانه
توزیع نرمال	$M(<x)=\frac{1}{2}\left[1+erf\left(\frac{x-m}{s\sqrt{2}}\right)\right]$	μ: میانگین σ : انحراف معیار
توزیع لاگ نرمال	$M(<x)=\frac{1}{2}\left[1+erf\left(\frac{\ln x-m}{s\sqrt{2}}\right)\right]$	μ: میانگین σ : انحراف معیار
رازین راملر	$M(<x)=\exp\left[-(\frac{x}{a})^b\right]$	α : قطر در 36/78% جرم تجمعی β : مشابه انحراف معیار در توزیع نرمال
فردلاند	$M(<x)=\left\{\ln\left[\exp(1)+(\frac{a}{x})^n\right]\right\}^{-m}$	α : وابسته به نقطه شکست اول n : وابسته به شیب منحنی m : پارامتر شکل
تیلر و ویت کرفت	$M(<x)=(\frac{x}{x_0})^{3-D}$	x_0 : قطر بزرگترین خاکدانه D : بعد فرکتال

معیارهای به کار رفته جهت ارزیابی مطلوبیت مدل‌ها در جدول2 آمده است. روش مقایسه با استفاده از آماره‌های C_p و AIC این است که نخست یکی از مدل‌ها به عنوان مرجع انتخاب شده و مدل دیگر نسبت به آن سنجیده می‌شود. به این منظور لازم است ابتدا مقدار C_p یا AIC را برای مدل مرجع محاسبه کنیم (جدول2). اگر مقدار به دست آمده از سنجش مدل مرجع با مدل مقایسه شونده از 0/95 مقدار C_p یا AIC بدست آمده برای مرجع کمتر باشد، به این معنی است که مدل مقایسه شونده با اطمینان 95 درصد بهتر از مدل مرجع است و در غیر این صورت مدل مرجع ترجیح داده می‌شود(هوانگ و همکاران2002). هر چه مقادیر محاسبه شده این آماره‌ها برای مدل مقایسه‌ای کوچک‌تر باشد، مدل مربوطه بهتر است. شرایط اخیر با استفاده از آزمون t مشخص می‌گردد. در این مطالعه مدل توزیع نرمال به عنوان مرجع انتخاب شد و مطلوبیت سایر مدل‌ها نسبت به این مدل مورد ارزیابی و مقایسه قرار گرفته است.

جدول2- معیارهای تعیین بهترین مدل (هوانگ و همکاران 2002)

معیار سنجش	معادله	پارامترها
R^2	$R^2=\left\{\frac{N(\sum Y_O Y_P)-(\sum Y_O)(\sum Y_P)}{\sqrt{[N\sum Y_O{}^2-(\sum Y_O)^2]}\sqrt{[N\sum Y_P{}^2-(\sum Y_P)^2}}\right\}^2$	N: تعداد مشاهدات Y_O و Y_P: مقادیر مشاهده‌ای و پیش‌بینی
C_p	$C_p=\frac{SSE_c}{SSE_r/(N-P_r)}-(N-2P_c)$	N: تعداد مشاهدات P_c و P_r تعداد پارامترهای مدل مقایسه ای و مرجع
AIC	$AIC=N\left[\ln(2p)+\ln(\frac{SSE}{N-P})+1\right]+P$	N: تعداد مشاهدات P : تعداد پارامتر

نتایج و بحث

بافت خاک‌های مورد مطالعه در محدوده کلاس‌های رسی، رسی سیلتی، و لوم رسی قرار گرفتند. بررسی‌های مقدماتی نشان داد که ضرایب کلیه رگرسیون‌ها از نظر آماری در سطح 1 درصد معنی‌دار می‌باشند. مقادیر R^2 برازش مدل‌ها بر داده‌های توزیع تجمعی خاکدانه‌ها در تمام خاک‌های مورد مطالعه در محدوده 0/8735 تا 0/9996 قرار داشتند، که کمترین آن مربوط به برازش مدل نرمال، و بیشترین آن به مدل فرکتالی (تیلر و ویت‌کرافت1992) تعلق داشت(جدول3). نتایج مقایسه R^2ها نشان داد که مدل‌های نرمال(دو پارامتری)، رازین راملر(دو پارامتری) و جکی (یک پارامتری) کمترین دقت برازش را نسبت به سایر مدل‌ها در خاک‌های مورد مطالعه داشتند(جدول3).

جدول3- مقادیر معیارهای سنجش دقت برازش مدل‌های مختلف بر توزیع اندازه خاکدانه ها در خاک‌های مورد مطالعه

مدل	جکی		نرمال		لاگ‌نرمال		رازین راملر		فردلاند		فرکتالی	
دقت برازش	R^2	SSE	R^2	SSE	R^2	SSE	R^2	SSE	R^2	SSE	R^2	SSE
میانگین	0/9524	0/051	0/9553	0/040	0/9745	0/0238	0/9569	0/040	0/9838	0/0128	0/9881	0/0137
انحراف	0/015	0/017	0/044	0/039	0/009	0/008	0/014	0/012	0/005	0/004	0/021	0/024
	0/9894	0/012	0/9904	0/009	0/9942	0/006	0/9902	0/010	0/9921	0/007	0/9996	0/0005
محدوده	0/9344	0/063	0/8735	0/108	0/9588	0/035	0/9313	0/058	0/9734	0/019	0/9189	0/091

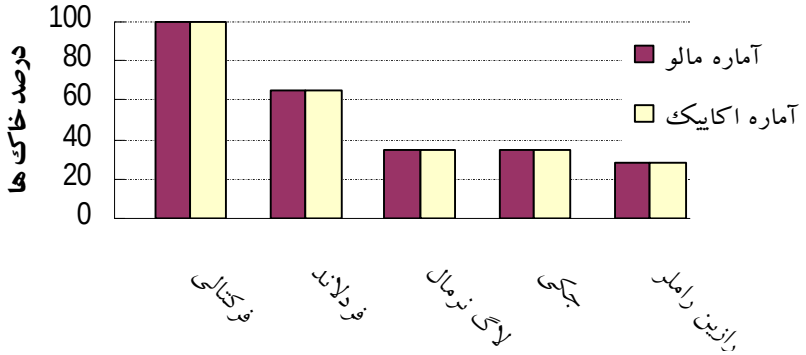
آزمون t جفتی R^2ها نیز تفاوت معنی‌داری($p=\%5$) را بین دقت برازش این مدل‌ها نسبت به هم نشان نداد. هوانگ و همکاران(2002) با بررسی داده‌های PSD 1400 خاک، مدل جکی(1949) را ضعیف‌ترین مدل از لحاظ دقت برازش اعلام نمودند. نتایج مطالعه حاضر نشان داد که دقت برازش مدل‌های فرکتالی و فردلاند در تمام کاربری‌ها بیش‌تر از سایر مدل‌ها است، اما تفاوت دقت برازش این دو مدل نسبت به هم از لحاظ آماری معنی‌داری نشد. همچنین این نتایج نشان دادند که در اراضی کشاورزی و جنگل تخریب شده، مدل لاگ‌نرمال از دقت برازش بیشتری نسبت به مدل نرمال برخوردار است، و در خاک‌های جنگلی و مرتعی مدل توزیع نرمال دارای دقت بالاتری است. چنین به نظر می‌رسد، در خاک‌های با ساختمان پایدارتر که در آزمون الک تر کم‌تر خرد می‌شوند، توزیع خاکدانه‌ها (پس از انجام الک تر) بیش‌تر از مدل نرمال پیروی می‌کند، و در خاک‌هایی با ساختمان ضعیف‌تر که در اثر تنش ناشی از الک تر، تخریب خاکدانه‌ها بیشتر و این واحدها به درجات پایین‌تر سلسله مراتب تشکیل ساختمان تنزل می‌یابند، انتهای منحنی توزیع تجمعی اندازه خاکدانه کمی خمیده‌تر شده و شکل این منحنی به منحنی PSD شباهت بیشتری پیدا می‌کند. گزارشات زیادی وجود دارند(شیرازی و بورسما1984) که نشان می‌دهند منحنی PSD اغلب از مدل لاگ نرمال تبعیت می‌کند و نزدیک بودن منحنی ASD به توزیع لاگ نرمال در کاربری زراعی و جنگل تخریب یافته مبین این واقیعت است که وضعیت کنونی ساختمان خاک که همواره حاصل برهم‌کنش عوامل تشکیل خاکدانه و تنش‌های محیطی است(پرفکت وکی1991)، در این کاربری‌ها در مراحل اولیه سلسله مراتب تشکیل قرار داشته به طوری که توزیع اندازه خاکدانه‌ها هنوز تا حد زیادی متاثر از توزیع اندازه ذرات اولیه می باشد. در حقیقت وجود چنین شرایطی می‌تواند اثرات شدید تخریبی زراعت و یا تخریب جنگل را بر کیفیت فیزیکی خاک منعکس نماید.

جدول4، میانگینی از مقادیر عددی آمارهای C_p و AIC را در خاک‌های مورد مطالعه نشان می‌دهد. مقایسه مطلوبیت مدل‌ها نسبت به مدل نرمال(مرجع) با استفاده از آماره‌های فوق نشان داد که، مدل فرکتالی تیلر و ویت‌کرافت (1992) در عین سادگی و کمترین تعداد پارامتر، در تمام موارد بهتر از مدل نرمال می‌تواند تغییرات منحنی ASD را توجیه نماید. علاوه بر این، بعد فرکتالی خاکدانه‌ها (D) در این مدل، پارامتری با مفهوم فیزیکی است که می‌تواند جهت برآورد خصوصیاتی نظیر توزیع اندازه خلل و فرج، هدایت آبی، نفوذ آب در خاک، و منحنی رطوبتی خاک نیز مورد استفاده قرار گیرد (پرفکت و همکاران2002). پس از آن مدل سه پارامتری فردلاند (فردلاند و همکاران2000) علی‌رغم پیچیدگی فرمولی و تعداد پارامتر بیشتر، در 65% از خاک‌ها از مطلوبیت بالاتری نسبت به مدل توزیع نرمال برخوردار می‌باشد(شکل1). همچنین مدل‌های جکی(یک پارامتری) و لاگ نرمال(دو پارامتری)، هر یک در 35% و مدل رازین راملر(دو پارامتری) فقط در 28% از خاک‌های مورد مطالعه از مطلوبیت بیشتری نسبت به مدل نرمال برخوردارند(شکل1). چنان‌که شکل1 نشان می‌دهد در تمام خاکهای مورد مطالعه کاربرد هر یک از آماره های C_p و AIC جهت تعیین مطلوب‌ترین مدل به نتیجه گیری یکسانی منتهی می‌شود.

جدول4- مقادیر معیارهای انتخاب بهترین مدل نسبت به مدل نرمال

مدل	جکی		لاگ‌نرمال		رازین راملر		فردلاند		فرکتالی	
دقت	C_p	AIC	C_p	AIC	C_p	AIC	C_p	AIC	C_p	AIC
میانگین	13/87	-19/65	5/52	-24/38	12/00	-19/55	1/48	-27/17	-5/59	-42/35
انحراف	17/14	4/08	8/97	4/29	13/49	4/04	2/99	2/52	1/52	14/20
محدوده	-6/17	-32/06	-4/61	-36/51	-4/34	-31/83	-2/54	-32/76	-6/69	-60/64
	46/78	-14/40	21/68	-20/22	37/17	-16/84	5/83	-22/68	-0/72	-13/80

* مدل نرمال به عنوان مدل مرجع می‌باشد.



شکل1- مطلوبیت مدل‌ها نسبت به مدل نرمال(مرجع) با استفاده از دو آماره مالو و آکاییک در خاک‌های مورد مطالعه

تاکنون مطالعه‌ای که مقایسات مطلوبیت مدل‌ها را برای داده‌های توزیع اندازه خاکدانه‌های کاربری‌های مختلف اراضی در کشور گزارش نموده باشد به دست نیامده است. لذا توصیه می‌شود مطالعات تکمیلی در این زمینه با تعداد بیش‌تری از خاک‌ها صورت گیرد و صحت نتایج به دست آمده کنترل گردد.

منابع :

Allen, T., 1967, Particle size measurement. Powder sampling and particle size measurement, Vol. 1, 5th edition. Chapman and Hall, London, 525 pp.

Baldock, J.A., and B.D. Kay, 1987. Influence of cropping history and chemical treatments on the water- stable aggregation of a silt loam soil. Can. J. Soil Sci. 67, 501-511.

Buchan,G.D., K.S.Grewal, and A.B. Robson. 1993. Improved models of particle-size distribution: An illustration of model comparison techniques. SSAJ. 57:901–908.

Carrera, J., and S.P. Neuman. 1986. Estimation of aquifer parameters under transient and steady state conditions: 1.Maximumlikelihood incorporating prior information. Water Resour. Res. 22:199–210.

Fredlund,M.D., D.G. Fredlund, and G.W.Wilson. 2000. An equation to represent grain size distribution. Can. Geotech. J. 37:817–827.

Gardner W.R. 1956. Representation of soil aggregates size distribution by a logarithmic-normal distribution. Soil Science Society of American Proceedings 20: 151-153.

Gee, G.W. & J.W. Bauder. 1986. Particle-size analysis. pp. 383-409. In Klute, A.(ed.). Methods of soil analysis. Part 1. Physical and mineralogical methods. 2nd ed. Agron. Monogr. 9. ASA AND SSSA, Madison, WI.

Hwang, Sang Il, Kwang. Pyo. Lee, Dong Soo Lee, and Susan E. Powers. 2002. Models for estimating soil particle-size distributions. SSAJ. 66:1143–1150.

Irani, R.R., C.F. Callis 1963. Particle size: Measurement, interpretation, and application.Wiley, New York,165 pp

Jaky, J. 1944. Soil mechanics. (In Hungarian.) Egyetemi Nyomda, Budapest.

Mallows, C.L. 1973. Some comments on Cp. Technimetrics 15:661–675.

Perfect, E. and B.D. Kay. 1991. Fractal theory applied to soil aggregate. SSAJ. 55:1552–1558

Perfect, E., M. Diaz-Zorita, and J.H. Grove, 2002. A prefractal model for predicting soil fragment mass-size distribution. Soil Till. Res. 64, 79-90.

Shirazi, M.A., and L. Boersma. 1984. A unifying quantitative analysis soil-plant systems. Elsevier, Amsterdam. of soil texture. SSAJ. 48:142–147.

Tyler S.W. and S.W. Wheatcraft., 1992. Fractal scaling of soil particle-size distributions: Analysis and limitations. SSAJ. 56, 362-369.

Yoder, R.E. 1936. A direct method of aggregate analysis of soils and a study of the physical nature of erosion losses. J. Am. Soc. Agron. 28:337-351.