

ارزیابی کارایی تابع مفصل در تخمین برخی از ویژگی‌های خاک

احسان غجه پور^۱، اعظم جعفری^{۲*}، وحیدرضا جلالی^۳، مجید محمودآبادی^۴

(تاریخ دریافت: ۱۳۹۹/۰۹/۰۳ تاریخ پذیرش: ۱۴۰۰/۰۱/۱۴)

چکیده

مطالعه پراکنش مکانی ویژگی‌های خاک با هدف تهیه نقشه پهنه‌بندی آن در جهت مدیریت خاک و برنامه‌ریزی صحیح به منظور بهره‌برداری مناسب از منابع تجدید ناپذیر خاک، از اهمیت ویژه‌ای برخوردار است. تابع مفصل، از تکنیک‌های درون‌یابی جدیدی است که امروزه در علوم مختلفی کاربرد وسیعی پیدا کرده است. در پژوهش حاضر سعی شده است تغییرات مکانی برخی از ویژگی‌های خاک با استفاده از تابع مفصل مورد ارزیابی قرار گیرد و نتایج حاصل از آن با تکنیک‌های زمین‌آماري مختلف مقایسه گردد. بدین منظور ۱۲۱ نمونه سطحی خاک به روش شبکه‌بندی منظم از یک منطقه ۴۸۴ هکتاری از غرب شهرستان بافت استان کرمان جمع‌آوری شد و برخی از ویژگی‌های خاک شامل ماده آلی و بافت خاک اندازه‌گیری شد. برای درون‌یابی از چهار تابع مفصل ارض‌میدسی شامل؛ توابع کلایتون، فرانک، گامبل و جو و تکنیک‌های زمین‌آماري شامل کریجینگ ساده، کریجینگ معمولی، کریجینگ شاخص و کریجینگ منفصل یا گسسته و روش وزن‌دهی عکس فاصله (IDW) استفاده شد. تحلیل نتایج با استفاده از شاخص‌های آماری $RMSE$ ، R^2 ، MAE و MBE صورت گرفت. به منظور برازش تابع مفصل بر داده‌ها، ابتدا تابع توزیع متغیرهای مورد مطالعه تعیین گردید. نتایج نشان داد توزیع هر یک از متغیرهای مورد مطالعه با هم تفاوت دارد و توسط توابع توزیع متفاوتی تشریح می‌شود. همچنین با افزایش فاصله مقدار همبستگی برای همه متغیرهای مورد مطالعه کاهش پیدا می‌کند به طوری‌که بعد از فاصله ۲۰۰۰ متری، هیچ همبستگی مکانی نشان نمی‌دهند. مقایسه روش تابع مفصل و تکنیک‌های زمین‌آماري براساس معیارهای ارزیابی نشان داد تابع مفصل در تخمین متغیرهای مورد مطالعه بهتر عمل کرده و مقادیر خطای حاصل از تخمین برای تابع مفصل کمتر محاسبه شد. به طور کلی نتایج این پژوهش نشان داد با توجه به ماهیت چولگی داده‌های خاک تکنیک‌های تخمینگر توابع مفصل توانایی بیانی کامل از وابستگی احتمالاتی را دارند و می‌توانند در مطالعات مکانی مورد توجه قرار گیرند.

واژه‌های کلیدی: تابع مفصل، تکنیک‌های درون‌یابی، همبستگی مکانی، پیش‌بینی

غجه پور، ا.، جعفری، ا.، جلالی، و.، محمودآبادی، م. ۱۴۰۱. ارزیابی کارایی تابع مفصل در تخمین برخی از ویژگی‌های خاک. مجله تحقیقات کاربردی خاک، جلد ۱۰، شماره ۱، صفحه: ۸۲-۹۷.

- ۱- دانش آموخته کارشناسی ارشد، گروه علوم خاک، دانشکده کشاورزی، دانشگاه شهیدباهنر کرمان، کرمان، ایران
 - ۲- استادیار، گروه علوم خاک، دانشکده کشاورزی، دانشگاه شهیدباهنر کرمان، کرمان، ایران
 - ۳- دانشیار، مجتمع آموزش عالی شیروان، خراسان شمالی، ایران
 - ۴- استاد، گروه علوم خاک، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان، کرمان، ایران
- *: پست الکترونیک: a.jafari@uk.ac.ir

مقدمه

ویژگی‌های خاک دارای تغییرات مکانی و زمانی در مقیاس‌های کوچک و بزرگ می‌باشند. پراکنش مکانی ویژگی‌های خاک برای مدیریت و برنامه‌ریزی مناسب آن به منظور بهره‌برداری بهتر، از اهمیت زیادی برخوردار است. مطالعه این تغییرات، در سطح وسیع بسیار هزینه‌بر است، به این منظور لازم است از روش‌های درون‌یابی به منظور پیش‌بینی مقادیر ویژگی‌های مدنظر استفاده گردد، ولی در بسیاری موارد این روش‌ها دقت‌های متفاوتی دارند. از این رو، روش‌های زمین‌آماری به دلیل در نظر گرفتن همبستگی و ساختار مکانی چنین داده‌هایی مهم هستند، زیرا یکی از خصوصیات مشترک ویژگی‌های محیطی از جمله خاک، پیوستگی مکانی آن‌ها است. مفهوم تغییرپذیری مکانی خاک‌ها این است که مقادیر برخی ویژگی‌های خاک در نقاط مجاور در مقایسه با نقاطی که فاصله آن‌ها بیشتر است، شباهت بیشتری به یکدیگر دارند (Khosravi & Abbasi, 2016; Yaghmaeian et al., 2019; Karimi et al., 2019). بررسی تغییرپذیری ویژگی‌های خاک از روش‌های زمین‌آماری که قادر به تحلیل جزء وابسته به مکان متغیرهای مکانی است، استفاده می‌شود. اما مهم‌ترین ضعف روش زمین‌آماری، شرط نرمال بودن داده‌هاست که در شرایط طبیعی کمتر مشاهده می‌شود. داده‌های خاک اغلب دارای چولگی و کشیدگی هستند که در آن‌ها فرض نرمال بودن توزیع داده‌ها رعایت نمی‌شود. حساسیت روش‌های زمین‌آماری به داده‌های پرت نیز از دیگر معایب این روش است. رویکردهای مرسوم در پهنه‌بندی با استفاده از روش‌های زمین‌آماری که نیاز به داده‌ی نرمال دارد، باعث کاهش دقت روش مذکور می‌گردد. لذا تابع مفصل به عنوان روشی نوین که نیاز به داده‌های نرمال ندارد و به داده پرت نیز حساس نیست می‌تواند هم‌خوانی بیشتری با داده‌های خاک داشته باشد، و باعث افزایش دقت پهنه‌بندی گردد. رویکرد تابع مفصل می‌تواند جایگزین تکنیک‌های مرسوم درونیایی شود. در سال‌های اخیر تابع مفصل در زمینه هیدرولوژی آماری کاربرد گسترده‌ای پیدا کرده است، اما هنوز در علوم خاک از تابع مفصل استفاده جامع نشده است. احمدی و همکاران (Ahmadi et al., 2017) در پژوهشی اقدام به کاربرد توابع مفصل ارشمیدسی و حدی در تحلیل چند متغیره

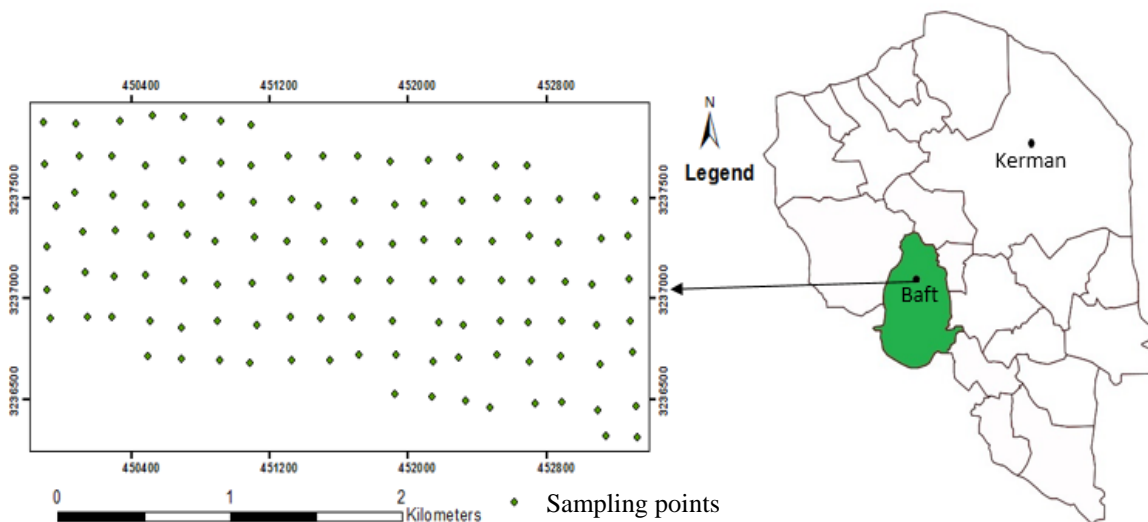
جریان‌های کمینه حوضه آبریز دز کردند. در این پژوهش جریان‌های کمینه حوضه آبریز دز در دوره آماری ۱۳۳۵-۱۳۹۱ با استفاده از توابع مفصل در محل اتصال سرشاخه‌ها مورد بررسی قرار گرفت. همچنین دانش‌زاده و همکاران (Daneshzadeh et al., 2016) اقدام به تحلیل دو متغیره خشکسالی هواشناسی سمنان با استفاده از تابع مفصل نمودند و نتیجه گرفتند که تابع مفصل گالامبوس مناسب‌ترین تابع مفصل برای تحلیل دو متغیره شدت و مدت خشکسالی در ایستگاه سینوپتیک سمنان است. در پژوهشی دیگر شفائی و همکاران (Shafaei et al., 2017) اقدام به مدلسازی مشخصات رویدادهای بارش با استفاده از مفصل دی-واین نمودند. در این پژوهش با استفاده از ساختارهای مختلف دی-واین ویژگی‌های مهم بارش شبیه‌سازی گردید. پژوهشی دیگر نیز توسط گنجعلیخانی و همکاران (Ganjalkhani et al., 2016) صورت گرفت. در این پژوهش اقدام به پهنه‌بندی کیفی آب زیرزمینی با استفاده از تابع مفصل نمودند. وی در این راستا داده‌های مربوط به غلظت بی‌کربنات در ۱۸ چاه مشاهده‌ای مربوط به دشت‌های کرمان و راور مورد بررسی قرار دادند. برای این منظور از چهار تابع مفصل ارشمیدسی شامل توابع کلایتون، فرانک، گامبل و جو استفاده کردند. نتایج به دست آمده از تابع مفصل با نتایج حاصل از روش‌های متداول پهنه‌بندی مورد مقایسه قرار گرفت. نتایج حاکی از دقت بالای این مدل نسبت به سایر مدل‌های مذکور بود. در پژوهشی دیگر، معظمی و همکاران (Moazami et al., 2014) میزان خطای حاصل از تخمین بارش با استفاده از تصاویر ماهواره‌ای را بررسی کردند و تجزیه و تحلیل عدم قطعیت با استفاده از تابع مفصل صورت گرفت. همچنین پژوهشی در مورد تخمین مقدار بارش در ایستگاه‌های مفقود توسط باردوسی و پگرام و همکاران (Bárdossy & Pegram, 2014) صورت گرفته است، که مقدار داده در ایستگاه مفقود بر اساس داده‌های ایستگاه-های مجاور با استفاده از تابع مفصل تخمین زده شد. نتایج این روش با روش‌هایی از جمله نزدیک‌ترین همسایگی، کریجینگ و فازی مقایسه شد که حاکی از دقت بالای این مدل نسبت به سایر مدل‌های مذکور بود. همچنین کنگ و همکاران (Kong et al., 2014) با به کارگیری توابع پیشینه‌ی آنتروپی مفصل، اقدام به پیش‌بینی جریان ماهانه‌ی رودخانه‌ای در چین نمودند، مقایسه نتایج بدست

مواد و روش‌ها

منطقه مورد مطالعه

منطقه مورد مطالعه در ۱۰ کیلومتری غرب شهرستان بافت در استان کرمان با طول جغرافیایی ۵۶ درجه و ۲۹ دقیقه تا ۵۶ درجه و ۳۱ دقیقه شرقی و عرض جغرافیایی ۲۹ درجه و ۱۵ دقیقه تا ۲۹ درجه و ۱۶ دقیقه شمالی واقع شده است (شکل ۱)، در منطقه مورد مطالعه سه کاربری کشاورزی، مرتع و صنعتی در مجاور هم واقع شده‌اند. کشت رایج منطقه گندم، جو، یونجه، حبوبات و باغات گردو، بادام و انگور می‌باشد. ارتفاع منطقه مورد مطالعه به طور متوسط ۲۲۷۰ متر از سطح دریا، میانگین دمای سالانه ۱۶ درجه سانتی‌گراد و میانگین بارش سالانه ۲۴۷ میلی‌متر می‌باشد. رژیم رطوبتی و حرارتی منطقه به ترتیب اریدیک ضعیف و ترمیک می‌باشد.

آمده از روش مورد استفاده با مقادیر مشاهداتی دال بر دقت بالای مدل مورد استفاده داشت. (Gräler & Pebesma, 2011) رویکرد جدیدی برای درونیابی کیفیت آب‌های زیرزمینی به کمک تابع مفصل بر اساس روش ارائه شده توسط عاس و همکاران (Aas *et al.*, 2009) برای تبدیل تابع مفصل Π متغیره به تابع مفصل دومتغیره ارائه دادند. در پژوهش حاضر سعی گردیده از خانواده‌های تابع مفصل نظیر خانواده ارشمیدسی که انعطاف پذیری بالایی دارند، در درونیابی برخی ویژگی‌های خاک استفاده شود. برای این منظور پهنه‌بندی برخی از ویژگی‌های خاک، در سه کاربری؛ کشاورزی، مرتع و صنعتی در منطقه بافت در استان کرمان مورد بررسی قرار گرفت، و در نهایت نتایج روش مذکور با نتایج حاصل از روش‌های انواع کریجینگ و وزن‌دهی عکس فاصله (IDW) مقایسه گردید.



شکل ۱- موقعیت منطقه مورد مطالعه و نقاط نمونه‌برداری

Figure 1. Location of the study area and sampling points

خاک شامل؛ ماده آلی (OC) به روش والکی و بلک (Black & Walkley, 1934) و بافت خاک به روش هیدرومتری (Bouyoucos, 1969)، اندازه‌گیری شد.

توصیف آماری داده‌ها

به منظور بررسی توزیع داده‌ها و دستیابی به خلاصه‌ای از اطلاعات آماری، توصیف آماری داده‌های مورد نظر صورت گرفت. توزیع فراوانی داده‌ها به لحاظ تاثیری که روی تخمین دارد از اهمیت زیادی برخوردار است. توزیع فراوانی با شاخص‌های میانگین، بیشینه، کمینه، ضریب

برای تعیین نقاط نمونه‌برداری از نرم‌افزارهای Google Earth و Arcgis10.3 استفاده شد. نمونه‌برداری در منطقه‌ای به وسعت ۴۸۴ هکتار صورت گرفت، بدین منظور با استفاده از روش شبکه‌بندی، منطقه به طور کاملاً منظم تقسیم‌بندی شد که مساحت هر شبکه ۴ هکتار بود. سپس از هر شبکه به طور تصادفی یک نمونه خاک و در مجموع ۱۲۱ نمونه از لایه سطحی خاک تهیه شد. نمونه‌های خاک پس از هوا خشک نمودن، از الک دو میلیمتری عبور داده شد، و برخی از ویژگی‌های شیمیایی

پکیج copula تحت محیط برنامه‌نویسی نرم‌افزار آماری R استفاده گردید. این پکیج توسط هوفرت و همکاران (Hofert et al., 2018) ارائه گردید.

تابع مفصل ایزاری برای مدل‌سازی همبستگی بین چند متغیر تصادفی می‌باشد (Schmidt, 2007). تابع مفصل به عنوان تابع توزیعی n -متغیره در بازه $[0, 1]$ تعریف می‌شود:

$$C: [0, 1]^n \rightarrow [0, 1] \quad (1)$$

اسکلار بیان کرد که بجای بیان تابع توزیع دو متغیره F به صورت تابعی از چندک‌ها، آن را به صورت تابعی از حاشیه‌ها یا احتمالات تجمعی u_i و u_j نوشته شود (Nelsen, 2007). همچنین نشان داد که اگر F کاملاً پیوسته باشد، تابع یگانه‌ای به صورت $C: [0, 1]^2 \rightarrow [0, 1]$ وجود دارد به نحوی که:

$$C(u_i, u_j) = F(F_i^{-1}(u_i), F_j^{-1}(u_j)) \quad (2)$$

که تابع C را تابع مفصل گویند. این نام به دلیل تاکید بر روشی که توابع توزیع مشترک را به حاشیه‌های تک تغییری آن متصل می‌کند، انتخاب گردیده است (رضاپور، ۱۳۹۰). بر اساس این تعریف، تابع مفصل دو متغیره بایستی دارای خصوصیات زیر باشد:

$$C(u_i, 1) = C(1, u_j) = 1 \quad (3)$$

$$C(u_i, 0) = C(0, u_j) = 0 \quad (4)$$

$$\text{رابطه (5)}$$

$$C(u_1, u_2) + C(v_1, v_2) - C(u_1, v_2) - C(v_1, u_2) \geq 0$$

$u_1 \geq v_1$ و $u_2 \geq v_2$

اگر H یک تابع توزیع n -متغیره و F_1 تا F_n توابع توزیع حاشیه‌ای یک متغیره‌ی آن باشند، تابع مفصل n -متغیره‌ی C به صورت زیر تعریف می‌شود.

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (6)$$

در حقیقت، تابع مفصل، یک تابع توزیع n -متغیره که توابع توزیع حاشیه‌ای آن یکنواخت است، تعریف می‌گردد (Schmidt, 2007).

تابع مفصل خانواده‌های متعددی دارد. در این میان تابع مفصل ارشمیدوسی به دلیل توانایی پوشش همبستگی مثبت و منفی داده‌ها، سهولت ایجاد آن‌ها و همچنین تعداد زیاد این خانواده کاربرد زیادی خصوصاً در علوم

تغییرات، چولگی، کشیدگی و انحراف معیار مورد بررسی قرار گرفت. با توجه به اینکه برای تجزیه و تحلیل زمین‌آماری نیاز به داده‌های با توزیع نرمال است بدین منظور با استفاده از آزمون کولموگراف-اسمیرنوف، نرمال بودن داده‌ها مورد بررسی قرار گرفت و از تبدیل Box-Cox برای نرمال سازی داده‌ها استفاده شد. بررسی وضعیت آماری ویژگی‌های خاک با کمک نرم افزار SPSS23.0 صورت گرفت.

تجزیه و تحلیل زمین‌آماری

از انواع کریجینگ شامل کریجینگ معمولی، کریجینگ ساده، کریجینگ شاخص و کریجینگ مفصل یا گسسته و روش IDW استفاده شد. بدین منظور پس از نرمال‌سازی داده‌ها، برای تعیین بهترین مدل با کمترین خطا، نیم‌تغییرنمای داده‌ها در محیط نرم‌افزار GS+ 5.1 ترسیم شد، سپس مقدار ناگت، آستانه و شعاع تاثیر به دست آمده به محیط نرم افزار Arcgis10.3 انتقال داده شد، و درونیابی با روش‌های ذکر شده صورت گرفت. توضیحات کامل در مورد روش‌های کریجینگ در منبع (Oliver & Webster, 2015) ارائه شده است.

تجزیه و تحلیل تابع مفصل

برازش تابع توزیع حاشیه‌ای

به منظور برازش تابع مفصل بر داده‌ها، ابتدا بایستی تابع توزیع حاشیه‌ای مناسب، بر داده‌ها برازش یابد. انواع مختلف توابع توزیع حاشیه‌ای وجود دارد که با توجه به تغییرات مکانی و ویژگی‌های آماری خصوصیت مورد بررسی، انواع تابع توزیع حاشیه‌ای بر آن برازش داده می‌شود و در نهایت براساس آزمون‌های نکوتی یا Goodness of fit test بهترین مدل انتخاب می‌شود. به این منظور، از سه آزمون کلموگروف-اسمیرنوف، آندرسون-دارلینگ و کای اسکوئر استفاده گردید و بر اساس این سه آزمون، تابع توزیع حاشیه‌ای مناسب برای هر پارامتر انتخاب گردید. در نهایت براساس تابع توزیع حاشیه‌ای خروجی، برای هر پارامتر کدنویسی مختص آن تابع در نرم‌افزار R انجام گرفت (کد نویسی تابع مفصل بر اساس توابع توزیع حاشیه‌ای تغییر می‌کند). آزمون‌های مذکور در محیط نرم‌افزار آماری EasyFit5.5 بر داده‌ها برازش شد.

تابع مفصل و توصیف ساختار متغیرهای مکانی

در پژوهش حاضر به منظور برازش تابع مفصل مناسب از

کشاورزی دارد (Grimaldi & Serilandi, 2006). خانواده توابع مفصل ارشمیدوسی شامل چهار تابع فرانک، جو، گامبل و کلایتون می‌باشد. در پژوهش حاضر از خانواده توابع مفصل ارشمیدوسی استفاده گردید (جدول ۱).

جدول ۱- برخی از توابع مفصل ارشمیدوسی (Nelsen, 2007)
Table 1. Some of the Archimedean copula functions (Nelsen, 2007)

Name	$C_\theta(u, v)$	Productive function (\emptyset)	Parameter range	Border states
Clayton	$\max\left(\left[u^{-\theta} + v^{-\theta} - 1\right]^{-1/\theta}, 0\right)$	$\frac{1}{\theta}(t^{-\theta} - 1)$	$[-1, \infty) \setminus \{0\}$	$C_{-1} = W$ $C_0 = \Pi$ $C_\infty = M$
Frank	$-\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right)$	$-\ln \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$	$(-\infty, \infty) \setminus \{0\}$	$C_{-\infty} = W$ $C_0 = \Pi$ $C_\infty = M$
Gumbel	$\exp\left(-\left[(\ln u)^\theta + (\ln v)^\theta\right]^{1/\theta}\right)$	$(-\ln t)^\theta$	$[1, \infty)$	$C_1 = \Pi$ $C_\infty = M$
Joe	$1 - \left[(1-u)^\theta + (1-v)^\theta - (1-u)^\theta (1-v)^\theta \right]^{1/\theta} - \ln \left[1 - (1-t)^\theta \right]$		$[1, \infty)$	$C_1 = \Pi$ $C_\infty = M$

در این معادله F_x تابع توزیع متغیر ناحیه‌ای Z می‌باشد که برای تمامی نقاط X یکسان می‌باشد. با توجه به اینکه با تغییر فاصله، تابع مفصل برازش یافته به داده‌ها تغییر می‌یابد، جهت برازش یک مدل مناسب به داده‌ها ابتدا بر روی فواصل، بهترین تابع مفصل برازش داده می‌شود. در ادامه با طبقه‌بندی فاصله، برای هر طبقه از تابع مفصل ترکیبی از دو تابع مفصل ابتدا و انتهای طبقه فاصله به دست می‌آید.

رویکرد ارائه شده در این پژوهش بر اساس روش ارائه شده توسط عاس و همکاران (Aas et al., 2009) می‌باشد. عاس و همکاران (Aas et al., 2009) روشی به منظور تبدیل تابع مفصل n متغیره به تابع مفصل دو متغیره ارائه کردند. این رویکرد امکان ترکیب توابع مفصل با خانواده‌های مختلف را فراهم می‌کند. در پژوهش فوق از دو نوع ساختار مختلف استفاده گردید که در این پژوهش از مدل موسوم به ساختار درختی متعارف استفاده گردید. ساختار کلی تابع چگالی مفصل ساختار درختی متعارف به صورت زیر می‌باشد (Aas et al., 2009).

رابطه (۸)

به منظور توصیف و تحلیل ساختارهای مکانی می‌توان از تابع مفصل به عنوان تابع توزیع دومتغیره برای نقاطی با فاصله مشخص استفاده نمود. بدین منظور لازم است که تابع توزیع حاشیه‌ای که بر متغیرهای ناحیه‌ای برازش یافته، برای تمامی نقاط در میدان تصادفی یکسان باشد (Li, 2010). بدین منظور ابتدا تابع توزیع حاشیه‌ای مناسب برای متغیرهای ناحیه‌ای Z به داده‌ها برازش داده می‌شود. سپس داده‌های مربوط به متغیرها که فاصله‌ی آن‌ها از یکدیگر $h + \Delta h$ است، جدا می‌گردد که در آن h فاصله مکانی دلخواه و Δh دامنه‌ی نوسان می‌باشد. در این پژوهش اعمال دامنه نوسان با توجه به این که فاصله نقاط تقریباً یکسان بوده الزامی نیست، اما برای اطمینان بیشتر و برای اینکه تمام نقاط در همه جهات تحت پوشش قرار بگیرد دامنه نوسان برای هر دسته از جفت نقاط لحاظ می‌شود.

در نتیجه تابع مفصل را برای هر جفت داده با فاصله $h \pm \Delta h$ می‌توان به صورت زیر در نظر گرفت.

رابطه (۷)

$$C_h(u_1, u_2) = P \left[F_z(Z(x)) < u_1, F_z(Z(x+h)) < u_2 \right] \\ = C \left(F_z(Z(x)), F_z(Z(x+h)) \right)$$

(RMSE)، ضریب تبیین (R^2)، میانگین خطای مطلق (MAE) و میانگین خطای انحراف (MBE) استفاده شد.

$$RMSE = \left[\frac{\sum_{i=1}^n (P_i - O_i)}{n} \right]^{0.5} \times \frac{100}{\bar{O}} \quad \text{رابطه (۱۲)}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (P_i - O_i)^2}{\sum_{i=1}^n (P_i - \bar{O})^2} \quad \text{رابطه (۱۳)}$$

$$MAE = \frac{\sum_{i=1}^n |P_i - O_i|}{n} \quad \text{رابطه (۱۴)}$$

$$MBE = \frac{\sum_{i=1}^n (P_i - O_i)}{n} \quad \text{رابطه (۱۵)}$$

که در آنها P_i مقادیر برآورد شده، \bar{O} مقدار متوسط متغیر مشاهده شده، O_i مقدار اندازه‌گیری شده و n تعداد نمونه‌ها است. در صورتی که مقدار اندازه‌گیری شده و برآورد شده با هم برابر باشند مقدار عددی RMSE، MAE و MBE برابر با صفر و مقدار عددی R^2 برابر با یک خواهد بود. همچنین در پایان نقشه پهنه‌بندی بر اساس مناسب‌ترین روش درون‌یابی رسم شد.

نتایج و بحث

توصیف آماری داده‌ها

جهت محاسبات آماری لازم است داده‌های مورد آزمایش که به صورت داده‌های خام هستند، منظم گردند. جدول (۲) آمار توصیفی ویژگی‌های اندازه‌گیری شده نمونه‌های خاک مورد مطالعه را نشان می‌دهد. پس از انجام تست نرمالیت کولموگراف-اسمیرنوف مشخص شد متغیرهای ماده آلی ($\text{Sig}=0/000$) و شن ($\text{Sig}=0/003$) دارای توزیع غیرنرمال و متغیرهای درصد رس ($\text{Sig}=0/200$) و سیلت ($\text{Sig}=0/078$) دارای توزیع نرمال می‌باشند. داده‌های غیرنرمال برای استفاده در تجزیه تحلیل زمین‌آماري با استفاده از روش Box-Cox نرمال شدند.

برای ارزیابی یک نقطه با استفاده از n نقطه‌ی اطراف آن بارسدوس (Bardossy, 2006) از تابع مفصل شرطی استفاده نمود و میانه‌ی تابع مفصل شرطی را به عنوان برآوردی از تابع توزیع در نقطه‌ی مجهول به صورت زیر در نظر گرفت.

رابطه (۹)

$$C^{-1}(U_1 | F(x_2), \dots, F(x_{k+1})) |_{U_1=0.5}$$

حال اگر F تابع توزیع متغیر ناحیه‌ای باشد، برآورد نقطه‌ی مجهول به صورت معکوس تابع F در نقطه‌ی مجهول می‌باشد،

رابطه (۱۰)

$$Z_{median}(x_1) = F^{-1}(C_{k+1}^{-1}(0.5 | F(x_2), \dots, F(x_{k+1})))$$

همچنین تخمین نقطه‌ای مقدار متوسط از مقدار نامشخص با استفاده از معادله‌ی زیر قابل محاسبه است.

رابطه (۱۱)

$$Z_{mean}(x_1) = \int_0^1 F^{-1}(u) c(u | F(x_2), \dots, F(x_{k+1})) du$$

پس از به دست آمدن بهترین تابع مفصل برای هر فاصله (براساس بیشترین شاخص MLE و پارامتر تابع مفصل (θ) ، با استفاده از روش پیشنهادی توسط عاس و همکاران (Aas et al., 2009)، توابع مفصل پنج متغیره با استفاده از معادله‌ی (۴) ساخته شدند و در نهایت با استفاده از معادلات (۵) و (۶) که به ترتیب برای تخمین نقاط مجهول بر اساس میانه و میانگین داده‌ها می‌باشد، پیش-بینی نقاط مجهول صورت می‌گیرد (Aas et al., 2009). در نهایت با استفاده از توابع مفصل ارشمیدوسی تابع مفصل میانه و تابع مفصل میانگین به دست آمد.

ارزیابی عملکرد مدل‌ها

به منظور ارزیابی عملکرد مدل‌ها، از روش ارزیابی متقابل حذف تک‌داده استفاده گردید. در این روش، در هر بار تکرار یکی از نقاط با مقدار معلوم، حذف شده و سپس مقدار آن با استفاده از روش مورد نظر تخمین زده می‌شود (Kitanidis, 1997). برای ارزیابی عملکرد مدل‌ها و انتخاب تکنیک درونیاب برتر از چهار شاخص آماری شامل؛ معیارهای میانگین ریشه دوم مربعات استاندارد

جدول ۲- توصیف آماری برخی از ویژگی‌های خاک‌های منطقه مورد مطالعه

Table 2. Statistical descriptive of some soil Chemical properties in the study area

Parameter	Unit	Minimum	Maximum	Average	Standard deviation	Coefficient of variation (%)	Skewness	Kurtosis
Organic matter	%	0.19	2.91	1.24	0.58	46.8	0.76	0.35
Clay	%	1.2	34.1	13.5	5.8	43.0	0.51	0.81
Sand	%	30.7	81.4	60.4	10.6	17.6	-0.32	3.3
Silt	%	1.4	56.0	26.1	8.3	31.9	0.12	1.9

در جدول (۳) مدل‌های برازش داده شده بر نیم تغییرنمای تجربی متغیرهای مورد مطالعه نشان داده شده است. نسبت اثر قطعه‌ای به آستانه را می‌توان برای ارزیابی ساختار مکانی داده‌ها مورد بررسی قرار داد. بر اساس تقسیم‌بندی جردانا و همکاران (Jordana et al., 2018) وقتی این نسبت کمتر از ۰/۲۵ باشد متغیر دارای ساختار مکانی قوی، بین ۰/۲۵-۰/۷۵ ساختار مکانی متوسط و بزرگتر از ۰/۷۵ ساختار مکانی ضعیف می‌باشد. بنابراین متغیرهای ماده آلی و سیلت دارای همبستگی مکانی قوی و متغیرهای درصد رس و شن دارای همبستگی مکانی متوسط می‌باشد

برای بیان تغییرپذیری خصوصیات خاک از شاخص ضریب تغییرات استفاده شد. بر اساس تقسیم‌بندی ارائه شده توسط ویل‌دینگ (Wilding, 1985) بر مبنای مقادیر ضریب تغییرات، ویژگی‌های خاک مورد مطالعه را می‌توان به سه گروه تقسیم کرد: ویژگی‌هایی با تغییرپذیری کم (ضریب تغییرات کمتر از ۱۵)، ویژگی‌هایی با تغییرپذیری متوسط (ضریب تغییرات ۱۵ تا ۳۵) و ویژگی‌هایی با تغییرپذیری زیاد (ضریب تغییرات بیشتر از ۳۵). بر این اساس، متغیرهای شن و سیلت در گروه تغییرپذیری متوسط و متغیرهای ماده آلی و رس در گروه تغییرپذیری زیاد قرار گرفتند.

تجزیه و تحلیل‌های زمین‌آماري

جدول ۳- مدل‌های برازش داده شده بر نیم تغییرنمای تجربی پارامترهای مورد مطالعه

Table 3. The models fitted on the variogram of the parameters

parameter	Unit	Model	C0	C0+C	A0	C0/(C0+C)	R ²	Spatial correlation
Organic Carbon	%	Exponential	0.06	0.42	600	0.14	0.31	Strong
Clay	%	Spherical	11	40	1000	0.27	0.31	Medium
Sand	%	Spherical	24	60	1000	0.40	0.39	Medium
Silt	%	Exponential	0.10	76	600	0.01	0.86	Strong

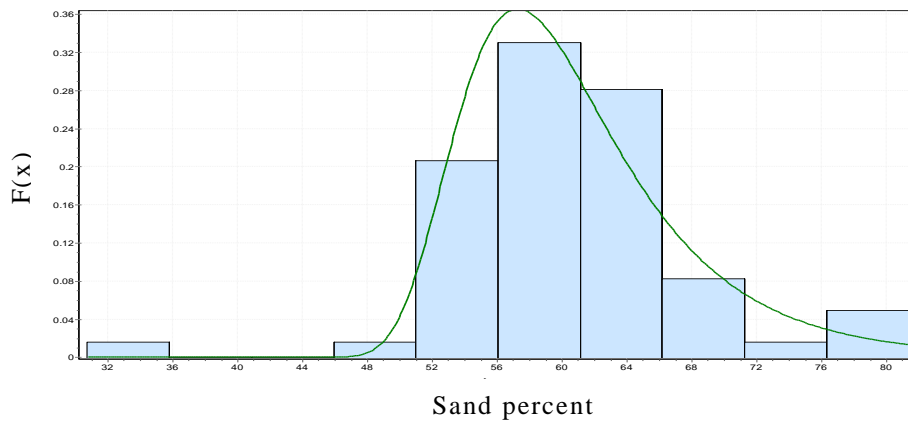
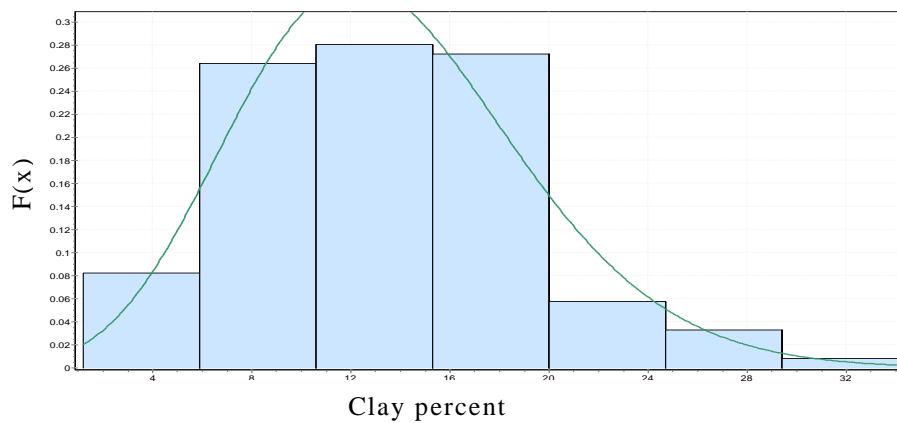
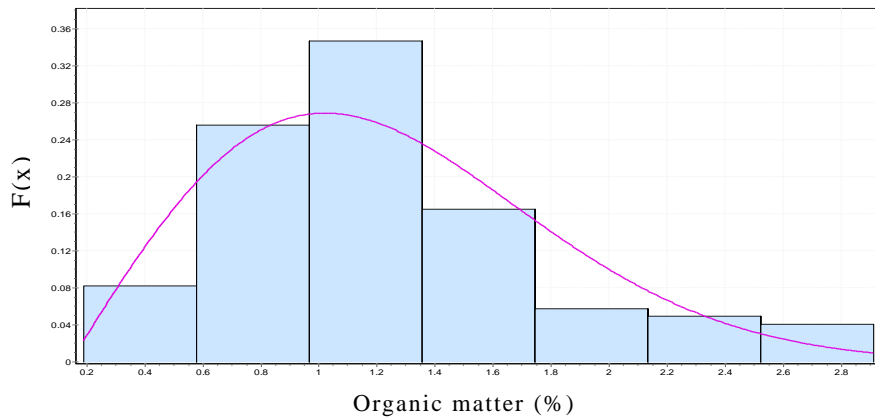
درصد ماده آلی، رس، شن و سیلت به ترتیب از تابع توزیع حاشیه‌ای Degam و Frechet، Beta، Reyleigh پیروی کردند. با توجه به توزیع و پراکندگی متغیرهای مورد بررسی، این توابع از بین توابع توزیع متعددی که برازش داده شدند، بهترین برازش را داشتند. شکل ۲ برازش توابع توزیع بر متغیرهای مورد مطالعه را نشان می‌دهد.

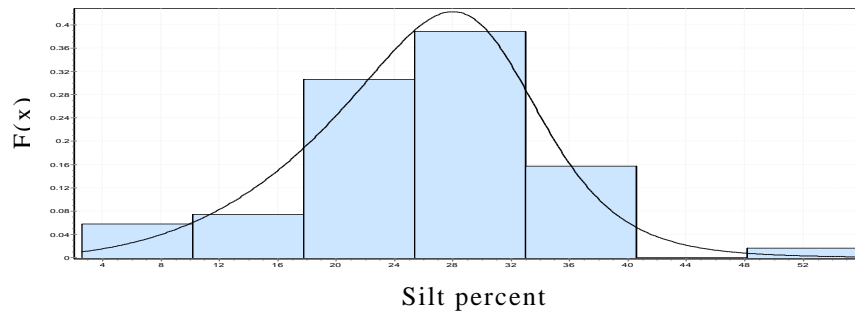
برازش تابع توزیع حاشیه‌ای

به منظور برازش تابع مفصل بر داده‌ها، ابتدا تابع توزیع حاشیه‌ای مناسب، بر داده‌ها برازش داده شد. جدول (۴) توابع توزیع حاشیه‌ای برازش داده شده بر داده‌ها را نشان می‌دهد. براساس مقدار آماره سه آزمون کلموگروف-اسمیرنوف، آندرسون-دارلینگ و کای اسکوئر، متغیر

جدول ۴- توابع توزیع حاشیه‌ای برآزش یافته بر داده‌ها
Table 4. The margin distribution functions fitted on data

Parameter	Unit	Margin distribution function	Amount of statistic in the Kolmogorov-smirnov test	Amount of statistic in the Anderson-darling test	Amount statistic in the chi-square test
Organic matter	%	Rayleigh	0.080	0.744	23.712
Clay	%	Beta	0.045	0.256	6.918
Sand	%	Frechet (3p)	0.102	1.842	13.17
Silt	%	Degam (4p)	0.017	0.991	10.6





شکل ۲- برازش تابع توزیع بر متغیرهای مورد مطالعه.

Figure 2. Distribution functions fitted on the studied variables

رسیده، از تابع مفصل مستقل که مربوط به داده‌های مستقل می‌باشد، برای برازش بر این نقاط استفاده شد. بهترین توابع مفصل و ضریب همبستگی بین جفت نقاط در هر طبقه برای هر متغیر در فواصل معین در جداول ۵ تا ۸ نشان داده شده است. طبق این جداول، مقدار شاخص MLE و پارامتر تابع مفصل (θ) با افزایش فاصله کاهش پیدا کرده است. کم بودن مقدار این شاخص‌ها در فواصل زیاد ناشی از کم بودن ضریب همبستگی است. آماره MLE معیاری برای برآورد پارامترهای یک مدل آماری و پارامتر تابع مفصل (θ) معیاری برای سنجش همبستگی با فاصله است. نتایج این پژوهش با نتایج گنجعلیخانی و همکاران (Ganjalkhani *et al.*, 2016)، باردوسی (Bárdossy, 2014)، لی (Li, 2010) و گالر و پبسا (Gräler & Pebesma, 2011) همخوانی دارد.

برازش تابع مفصل بر جفت نقاط

به منظور برازش تابع مفصل در جفت نقاط ابتدا بایستی جفت نقاط با فاصله مکانی $h+\Delta h$ جدا و دسته‌بندی گردند. همان طور که در بخش مواد و روش‌ها به آن اشاره گردید، با وجود اینکه فاصله تمام نقاط تقریباً یکسان بوده است، ولی دامنه نوسان برای هر دسته از جفت نقاط لحاظ می‌شود. لذا فاصله ۴۰۰ متر تا ۲۰۰۰ متر به عنوان فاصله نقاط در نظر گرفته شده است که بر اساس فاصله پایه و دامنه نوسان نقاط در این بازه تقسیم شدند. پس از تعیین فاصله نقاط از بین چهار تابع خانواده ارشمیدوسی بهترین تابع مفصل برای هر متغیر انتخاب گردید.

لازم به ذکر است با توجه به اینکه میزان همبستگی بین نقاط در فواصل بالاتر از ۲۰۰۰ متر به مقدار تقریباً ثابتی

جدول ۵- بهترین تابع مفصل و ضریب همبستگی جفت نقاط در فواصل مشخص برای ماده آلی

Table 5. The best copula and correlation coefficient of pairs of points at the specified intervals for organic matter

Distance (m)	Number of pairs of points	Pearson correlation coefficient	copula	Mean Least Error (MLE)	copula parameter (θ)
400	215	0.75	Gumbel	39.11	6.171
800	188	0.65	Gumbel	30.45	5.989
1200	160	0.35	Clayton	19.25	3.561
1600	134	0.20	Clayton	9.14	1.961
2000	95	0.08	Independent	-	-

جدول ۶- بهترین تابع مفصل و ضریب همبستگی جفت نقاط در فواصل مشخص برای رس

Table 6. The best copula and correlation coefficient of pairs of points at the specified intervals for clay

copula parameter (θ)	Mean Least Error (MLE)	copula	Pearson correlation coefficient	Number of pairs of points	Distance (m)
5.355	24.28	Frank	0.61	215	400
3.357	17.01	Joe	0.44	188	800
1.800	14.10	Gumbel	0.23	160	1200
1.128	11.25	Frank	0.13	134	1600
-	-	Independent	0.05	95	2000

جدول ۷- بهترین تابع مفصل و ضریب همبستگی جفت نقاط در فواصل مشخص برای شن

Table 7. The best copula and correlation coefficient of pairs of points at the specified intervals for sand

copula parameter (θ)	Mean Least Error (MLE)	copula	Pearson correlation coefficient	Number of pairs of points	Distance (m)
5.124	24.34	Frank	0.60	215	400
4.351	18.97	Joe	0.45	188	800
1.508	15.75	Gumbel	0.27	160	1200
1.123	12.92	Frank	0.16	134	1600
-	-	Independent	0.08	95	2000

جدول ۸- بهترین تابع مفصل و ضریب همبستگی جفت نقاط در فواصل مشخص برای سیلت

Table 8. The best copula and correlation coefficient of pairs of points at the specified intervals for sand

copula parameter (θ)	Mean Least Error (MLE)	copula	Pearson correlation coefficient	Number of pairs of points	Distance (m)
4.329	24.28	Frank	0.60	215	400
2.741	17.75	Joe	0.46	188	800
1.495	13.61	Gumbel	0.27	160	1200
1.213	10.81	Frank	0.16	134	1600
-	-	Independent	0.07	95	2000

همانطور که در جداول ۵ تا ۸ مشخص است با افزایش فاصله مقدار ضریب همبستگی کاهش پیدا کرده است. با اجرای تابع مفصل، به طور خودکار میزان همبستگی به عنوان تابعی از فاصله تعیین و در هر فاصله میانگین خطای تخمین هم محاسبه می‌شود. کاربر می‌تواند با توجه به روند تغییرات پارامتر با فاصله، فواصل را تغییر دهد به شرط اینکه بیشترین همبستگی را در آن فاصله به دست آورد. در اکثر مواقع، با اجرای تابع مفصل و انتخاب خودکار فاصله، بیشترین میزان همبستگی حاصل می‌شود. هر چند که کاربر می‌تواند تمام مراحل اجرای تابع مفصل را بررسی کند.

در نهایت پس از بدست آمدن بهترین تابع مفصل برای هر فاصله، با استفاده از این توابع، تابع مفصل میانه و تابع مفصل میانگین محاسبه شد.

نتایج اعتبارسنجی مدل‌ها

بر اساس معیارهای اعتبارسنجی $RMSE$ ، R^2 ، MAE و MBE بهترین مدل تخمینگر برای هر متغیر انتخاب شد. نتایج اعتبارسنجی مدل‌ها در جداول ۹ تا ۱۲ نشان داده شده است.

برای هر فاصله چهار تابع کلایتون، فرانک، جو و گامبل برازش داده شد. هر کدام از این توابع که براساس شاخص MLE و پارامتر تابع مفصل θ (براساس بیشترین مقدار) برازش بهتری داشته باشد، برای آن فاصله انتخاب می‌شود. در مورد ماده آلی، به دلیل اینکه برای هر فاصله، توابع کلایتون و گامبل نسبت به سایر توابع برازش بهتری داشته‌اند، این توابع انتخاب شده است (جدول ۵). لازم به ذکر است میزان همبستگی بین نقاط در فواصل بالاتر از ۲۰۰۰ متر به مقدار تقریباً ثابتی رسیده، به عبارت دیگر، در این فاصله، جفت نقاط هیچ‌گونه همبستگی نشان ندادند، لذا امکان برازش تابع بر این نقاط وجود نداشت و از تابع مفصل که مربوط به داده‌های مستقل می‌باشد، برای برازش بر نقاط با فاصله‌ی ۲۰۰۰ متر استفاده گردید. همانطور که در این جداول نشان داده شده است میزان همبستگی داده‌ها در فاصله‌های مختلف مشخص شده است، در حالی که در محاسبه‌ی مقادیر واریوگرام، تنها یک عدد معرف میزان همبستگی و یا به عبارت دیگر اختلاف داده‌ها در آن فاصله است. با توجه به اینکه هر فاصله تابعی مختص به خود دارد، امکان مشاهده و بررسی میزان همبستگی داده‌ها در چندک‌های مختلف وجود دارد.

جدول ۹- نتایج اعتبارسنجی تکنیک‌های درونیابی مختلف برای پیش‌بینی متغیر ماده آلی

Table 9. The validation results of different interpolation techniques for prediction of the organic matter

Interpolation techniques	RMSE	R ²	MAE	MBE
IDW	33.58	0.47	0.30	0.006
Ordinary Kriging	29.18	0.66	0.26	-0.004
Simple Kriging	30.19	0.53	0.27	0.005
Universal Kriging	26.14	0.60	0.24	-0.003
Disjunctive Kriging	33.85	0.41	0.31	0.007
Median copula	26.10	0.71	0.14	-0.001
Average copula	26.12	0.70	0.15	-0.002

جدول ۱۰- نتایج اعتبارسنجی تکنیک‌های درونیابی مختلف برای پیش‌بینی متغیر رس

Table 10. The validation results of different interpolation techniques for prediction of clay

Interpolation techniques	RMSE	R ²	MAE	MBE
IDW	31.93	0.44	2.64	-0.028
Ordinary Kriging	31.04	0.47	2.49	-0.007
Simple Kriging	30.88	0.47	2.48	-0.006
Universal Kriging	31.39	0.46	2.52	-0.020
Disjunctive Kriging	34.72	0.35	2.72	-0.029
Median copula	20.60	0.53	2.17	-0.003
Average copula	20.65	0.50	2.20	-0.005

جدول ۱۱- نتایج اعتبارسنجی تکنیک‌های درونیابی مختلف برای پیش‌بینی متغیر شن

Table 11. The validation results of different interpolation techniques for prediction of the sand percent

Interpolation techniques	RMSE	R ²	MAE	MBE
IDW	9.53	0.43	3.15	-0.18
Ordinary Kriging	9.19	0.47	3.02	-0.027
Simple Kriging	8.62	0.54	2.72	0.018
Universal Kriging	8.67	0.50	2.78	-0.022
Disjunctive Kriging	8.63	0.53	2.73	0.019
Median copula	8.41	0.60	2.30	0.011
Average copula	8.46	0.58	2.33	0.014

جدول ۱۲- نتایج اعتبارسنجی تکنیک‌های درونیابی مختلف برای پیش‌بینی متغیر سیلت

Table 12. The validation results of different interpolation techniques for prediction of the silt percent

Interpolation techniques	RMSE	R ²	MAE	MBE
IDW	22.47	0.50	3.80	0.21
Ordinary Kriging	22.31	0.55	3.50	0.07
Simple Kriging	20.98	0.56	3.44	-0.05
Universal Kriging	20.70	0.58	3.42	0.05
Disjunctive Kriging	20.79	0.57	3.43	-0.06
Median copula	19.30	0.70	3.24	-0.02
Average copula	20.41	0.63	3.35	-0.04

زمین‌آماری و تابع مفصل در تخمین متغیرهای شن و سیلت عملکرد بهتری داشتند و در این مورد، تابع مفصل، تخمین دقیق‌تری را ارائه کرده است. احتمالاً مزایای تابع مفصل از جمله حساس نبودن به داده پرت، عدم نیاز به

با وجود اینکه تابع مفصل به ویژه با استفاده از میانه برآورد بهتری نسبت به تکنیک‌های مرسوم زمین‌آماری داشته اما اختلاف فاحشی از نظر خطای تخمین نسبت به تکنیک‌های زمین‌آماری ندارد. نتایج نشان می‌دهد تکنیک‌های

(Ganjalikhani *et al.*, 2016). امکان پوشش مکانی همه تغییرات متغیر مورد بررسی با برازش تابع توزیع حاشیه‌ای وجود دارد. به عبارت دیگر، انتخاب تابع توزیع حاشیه‌ای که تشریح کننده تغییرات متغیر مورد بررسی است، عملکرد تابع مفصل را تحت تاثیر قرار می‌دهد. بنابراین احتمال دارد یکی از عوامل عدم وجود اختلاف فاحش بین عملکرد تابع مفصل و تکنیک‌های درون‌یابی دیگر در تخمین خصوصیات خاک، سهولت برازش مدل‌های زمین‌آماری و پیچیدگی برازش تابع مفصل باشد. از دلایل دیگر کارایی نسبی تابع مفصل، می‌توان به بیان همبستگی بین داده‌ها در چندک‌های مختلف به صورت جداگانه در برازش تابع مفصل اشاره کرد (Bárdossy, 2014). منطقی به نظر می‌رسد که توانایی و کارایی تابع مفصل در تخمین خصوصیات خاک با تغییرپذیری و چولگی زیاد مورد آزمون قرار بگیرد تا بتوان در مورد عملکرد آن با عدم قطعیت کمتری صحبت کرد.

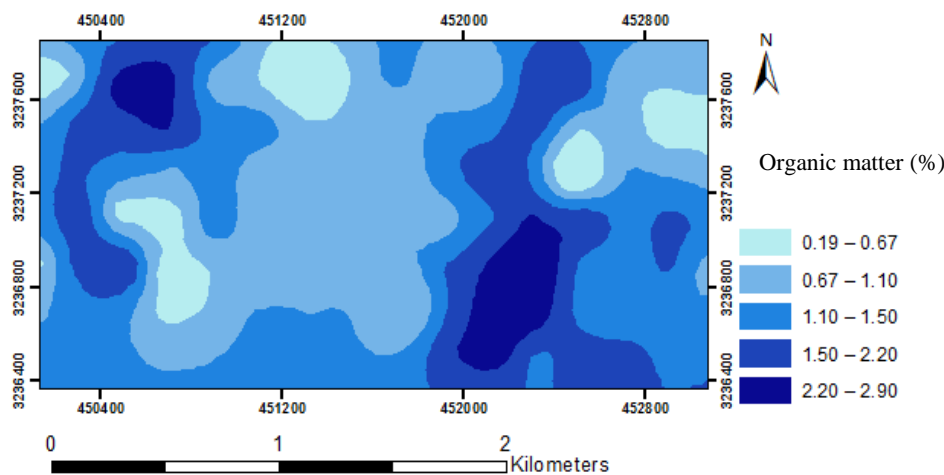
پهنه‌بندی با استفاده از تابع مفصل

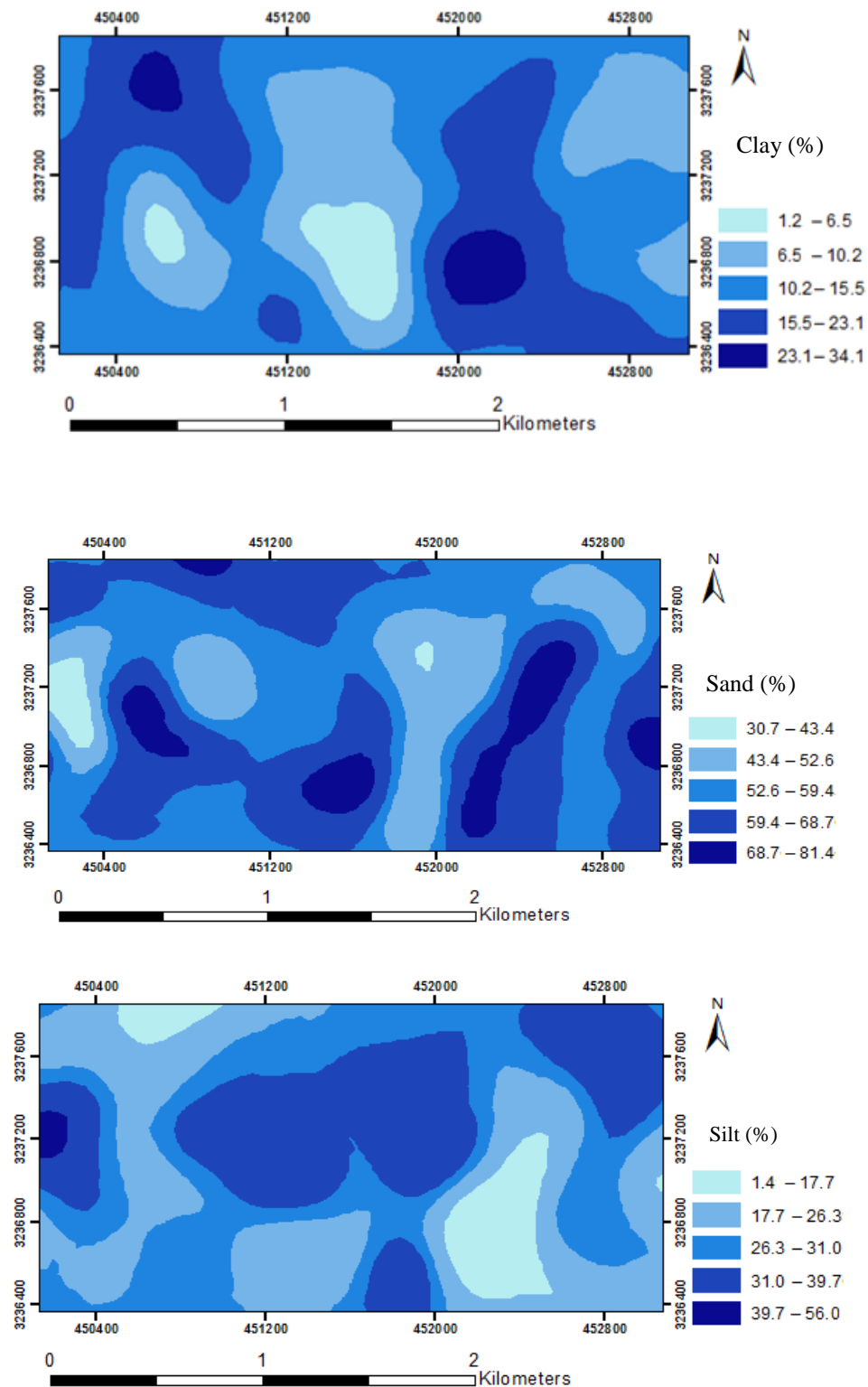
نقشه پهنه‌بندی متغیرهای مورد مطالعه بر اساس تابع مفصل میانه در شکل ۳ نشان داده شده است.

داده‌های با توضیح نرمال و همخوانی بیشتر با داده‌های خاک از دلایل برتری نسبی تابع مفصل در تخمین متغیرهای مورد مطالعه می‌باشد. نتایج حاصل از این پژوهش با نتایج حاصل از پژوهش‌های گنجعلیخانی و همکاران (Ganjalikhani *et al.*, 2016)، باردوسی (Bárdossy, 2014)، لی (Li, 2010) و گالر و پیسما (Gräler & Pebesma, 2011) که از تابع مفصل در پهنه‌بندی کیفیت آب‌های زیرزمینی استفاده کرده بودند، همخوانی دارد.

بالا بودن خطای برآورد توسط تکنیک‌های زمین‌آمار و تابع مفصل احتمالاً به دلیل آنالیز همزمان نمونه‌های خاک از کاربری‌های مختلف کشاورزی، صنعتی و مرتع باشد. مدیریت‌های مختلف در این کاربری‌ها به عنوان فاکتور خارجی تاثیرگذار بر شرایط خاک احتمالاً باعث تغییر خصوصیات خاک در فواصل کم شده و بر توانایی تخمین تکنیک‌های درون‌یابی تاثیر گذاشته است.

برازش تابع توزیع حاشیه‌ای را می‌توان یکی از دلایل عملکرد مناسب تابع مفصل در مقایسه با سایر تکنیک‌های مرسوم درون‌یابی را دانست چرا که در مدل‌های زمین‌آماری امکان برازش تابع توزیع حاشیه‌ای وجود ندارد

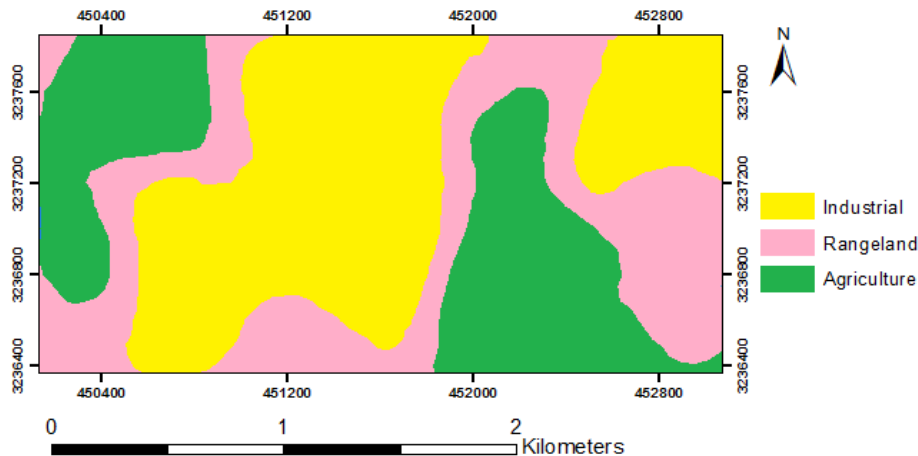




شکل ۳- نقشه پهنه‌بندی متغیرهای مورد مطالعه بر اساس تکنیک تابع مفصل میانه
Figure 3. Zoning map of the studied variables based on the Median copula

بیشترین مقادیر آنها در کاربری کشاورزی متمرکز شده است.

براساس تطابق دو نقشه پهنه‌بندی (شکل ۳) و نقشه کاربری اراضی (شکل ۴)، مقدار ماده آلی و درصد رس متناسب با تغییر کاربری‌ها تغییر کرده است به طوری که



شکل ۴- نقشه کاربری اراضی منطقه مورد مطالعه

Figure 4. Land use map of the study area

تخمین ویژگی‌های مورد مطالعه با خطای کمتری همراه بوده است. احتمالاً مزایای تابع مفصل از جمله حساس نبودن به داده پرت، عدم نیاز به داده‌های با توضیح نرمال و همخوانی بیشتر با داده‌های خاک از دلایل برتری نسبی تابع مفصل در تخمین پارامترهای مورد مطالعه می‌باشد. با توجه به نتایج حاصله از این پژوهش می‌توان به توانایی تابع مفصل در تخمین ویژگی‌های خاک پی برد.

نتیجه‌گیری کلی

در پژوهش حاضر تخمین برخی از ویژگی‌های خاک به کمک تابع مفصل صورت گرفت و نتایج حاصل از تابع مفصل با انواع تکنیک‌های زمین‌آماري مورد مقایسه قرار گرفت. براساس معیارهای اعتبارسنجی R^2 ، RMSE، MAE و MBE مشخص شد که تابع مفصل نسبت به تکنیک‌های مرسوم زمین‌آماري دقت بهتری در برآورد ویژگی‌های خاک مورد مطالعه داشته است، همچنین

References

- Aas K., Czado C., Frigessi A., and Bakken H. 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44(2): 182-198.
- Ahmadi F., Radmanesh F., Parham G.A., and Mirabbasi Najafabadi R. 2017. Application of archimedean and extreme values copula functions for multivariate analysis of low flows in dez basin. *Journal of Water and Soil*, 31(4): 1031- 1045. (In Persian)
- Bárdossy A. 2006. Copula-based geostatistical models for groundwater quality parameters. *Water Resources Research*, 42:1-12.
- Bárdossy A., and Pegram G. 2014. Infilling missing precipitation records – A comparison of a new copula-based method with other techniques. *Journal of Hydrology*, 519, Part A (0): 1162-1170.
- Daneshzadeh M., karami H., farzin S., sanikhani H., and Mousavi S.F. 2016. Bivariate analysis of meteorological drought in semnan using copula functions. *Applied Research of Water Sciences*. 2(3): 21-34. (In Persian)
- Ganjalikhani M., Zounemat Kermani M., Rezapur M., and Rahnama M. 2016. Evaluation of copula performance in groundwater quality zoning case study: kerman and ravar region. *Iranian Journal of Soil and Water Research*, 47(3): 551-560. (In Persian)
- Gräler B., and Pebesma E. 2011. The pair-copula construction for spatial data: a new approach to model spatial dependency. *Procedia Environmental Sciences*, 7(0): 206-211.
- Grimaldi S., and Serinaldi F. 2006. Asymmetric copula in multivariate flood frequency analysis. *Advances in Water Resources*, 29(8): 1155-1167.
- Hofert M., Kojadinovic I., Maechler M., Yan J., and Nešlehová G. 2018. Copula: Multivariate Dependence with Copulas. R package (Version 0.999-19).
- Jordana S.G., Jannourab R., Jordana G., Buerkerta A., and Joergensenb R.G. 2018. Spatial variability of soil properties in the floodplain of a river oasis in the mongolian Altay Mountains. *Geoderma*, 330: 99–106.

- Karimi A., Moezzi A.A., Chorom M., and Enayatizamir N. 2019. Influence of Sugarcane bagasse biochar on nutrition availability and biological properties of a calcareous soil. *Applied Soil Research*, 8(1): 1-17. (In Persian)
- Kitanidis, P. K. 1997. Introduction to Geostatistics: applications in hydrogeology, Cambridge University Press. UK, 247p.
- Khosravi Y., and Abbasi E. 2016. Spatial Analysis of Environmental Data Using Geostatistics, 280p. (In Persian)
- Kong X.M., Huang G.H., Fan Y.R., and Li Y.P. 2014. Maximum entropy- gumbel-hougaard copula method for simulation of monthly streamflow in Xiangxi River, China. *Stochastic Environmental Research and Risk Assessment*. 14(2): 1-14.
- Li J. 2010. Application of copulas as a new geostatistical tool. PhD thesis, University of Stuttgart, Faculty of Civil and Environmental Engineering.
- Moazami S., Golian S., Kavianpour M.R., and Hong Y. 2014. Uncertainty analysis of bias from satellite rainfall estimates using copula method. *Atmospheric Research*, 137(0): 145-166.
- Nelsen R.B. 2007. An Introduction to Copulas. Springer Series in Statistics. 2nd Edition. Springer. 272 p.
- Oliver M.A., and Webster R. 2015. Basic Steps in Geostatistical: The Variogram and Kriging. Springer. 100 p.
- Schmidt T. 2007. Coping with Copulas. In: Rank J. (Ed.), Copulas-From Theory to Application in Finance, Risk Books, London, pp. 3-34
- Shafaei M., Fakheri Fard A., Dinpajouh Y., and Mirabbasi R. 2017. Modeling rainfall event characteristics using D-vine copulas. *Journal Water and Soil Resources Conservation*. 6(2): 46-58. (In Persian)
- Walkley A.J., and Black I.A. 1934. An examination of Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Science*. 37: 29-38.
- Wilding L.P. 1985. Spatial Variability: Its Documentation Accommodation and Implication to Soil Surveys. In: Nielsen D.R. and Bouma J. (Eds.) *Soil Spatial Variability*. Pudoc. The Netherlands. pp. 166-194.
- Yaghmaeian Mahabadi N., Samiei K., Zavvareh M. and Rmezanpour H. 2019. Spatial variation of some soil properties and their relationships with tea yield in Fouman Region, Guilan. *Applied Soil Research*, 7(2): 82-96. (In Persian)

Evaluation of the Performance of Copula Function in Estimating Some Soil Properties

Ehsan Ghojehpour¹, Azam Jafari^{2*}, Vahid Reza Jalali³, Majid Mahmoodabadi⁴

(Received: November 2020 Accepted: April 2021)

Abstract

The study of spatial distribution of soil properties for optimal soil management and proper utilization of non-renewable soil resources is of particular importance. The copula function is one of the new interpolation techniques that are widely used in various sciences such as hydrology. Thus, the aim of this study was to evaluate the spatial variation of some soil properties using the copula function and to compare with geostatistics techniques. Sampling by regular networking was done in a 484 ha area in the west of Baft city, Kerman province, and 121 surface soil samples were collected. After air drying and passing through a 2 mm sieve, the percentage of organic matter and clay were determined in soil samples. To interpolate, four functions of the Archimedean copula including the Clayton, Frank, Gumbel and Joe functions, and geostatistics techniques including simple kriging, ordinary kriging, universal kriging and disjunctive Kriging and the Inverse Distance Weighting (IDW) method were used. The results were analyzed using Root Mean Square Error (RMSE), determination coefficient (R^2), mean absolute error (MAE), and Mean Bias error (MBE). In order to fit the copula function on the data, the distribution function of the studied variables was determined. The results showed that the distribution of each of the studied variables is different and is explained by different distribution functions. Also, with increasing distance, the value of correlation for all studied variables decreased so that after a distance of 2000 meters, they do not show any spatial correlation. Comparison of the Copula function and geostatistical techniques based on evaluation criteria showed that the Copula function had a better performance in estimating the studied variables and the estimation error for the Copula function were calculated less. In general, the results of this study showed that due to the skewed nature of soil data, Copula function have the ability to fully express the probabilistic dependence and can be considered in spatial studies.

Key words: Copula function, Interpolation techniques, Spatial correlation, Estimation

Ghojehpour E., Jafari A., Jalali V.R., and Mahmoodabadi M. 2022. Evaluation of the performance of copula function in estimating some soil properties. *Applied Soil Research*, 10(1):82-97.

1 - M.Sc. Graduated. Student of Soil Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran

2 - Assistant Professor, Department of Soil Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran

3 - Associate Professor, High Education Complex of Shirvan, North Khorasan, Iran

4 - Professor, Department of Soil Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran

*: Corresponding Author Email: a.jafari@uk.ac.ir