روش انتروپی حداکثر اریب (BME) در برآورد متغیرهای پیوسته خاک

نورایر تومانیان[۱]، نیکو حمزه پور[۲] و مصطفی کریمیان اقبال[۳]

[۱]عضو هیئت علمی مرکز تحقیقات کشاورزی اصفهان، [۲] دانشجوی دکتری گروه خاکشناسی، دانشکده کشاورزی، دانشگاه تربیت مدرس، [۳] دانشیار گروه خاکشناسی، دانشکده کشاورزی، دانشگاه تربیت مدرس

مقدمه

عقلانی‌ترین و موثرترین راه دستیابی به اطلاعات دقیق، در هر نقطه از منطقه مورد مطالعه، تلفیق داده‌های مطمئن (**hard**) و نامطمئن (**soft**) با یکدیگر است. به طور کلاسیک، کریجینگ معمول‌ترین الگوریتم برای آنالیز مجموعه داده‌های مکانی را ارائه می کند. متاسفانه چهارچوب تئوریکی این روش امکان استفاده از داده‌های نامطمئن، بدون تبدیل آنها با کدهای شاخص [۲ و ۳] و یا تفکیک کردن (**Stratification**) [۳] را نمی‌دهد. راه دیگر، دخالت دادن این داده‌ها در فرآیند تخمین، استفاده از متوسط بازه است، اما چنین تخمین ناپخته ای متناسب با ساختار مکانی و تغییرات داخل کلاس خاک ها نمی-باشد. در جهت گسترش مفهوم تخمین مکانی، کریستاکوس در اوایل دهه نود روش جدیدی را پیشنهاد نموده است [۳]. وی از طریق تلفیق انتروپی حداکثر و روش‌های اریب (**BME**)، فرمولی بسیار عمومی برای تخمین مکانی ایجاد نموده است. این روش از مزایای بی‌شماری نسبت به روش‌های مرسوم زمین آمار کلاسیک برخوردار است. خصوصیات گفته شده، امکان استفاده همزمان از طیف وسیعی از اطلاعات نامطمئن به صورت دامنه‌ها، توابع احتمال تراکم‌ها، مدل‌ها یا نمودارها را فراهم می‌کند. برخلاف روش‌های زمین آمار کلاسیک، در این روش نیازی به تبدیل یا کدگذاری این نوع اطلاعات نمی‌باشد. از طریق توابع احتمال تراکم‌های نهائی، انواع مختلف نقشه ها از طریق محاسبه شاخص‌هائی که متناسب با هدف نهائی، فراهم می‌شود.

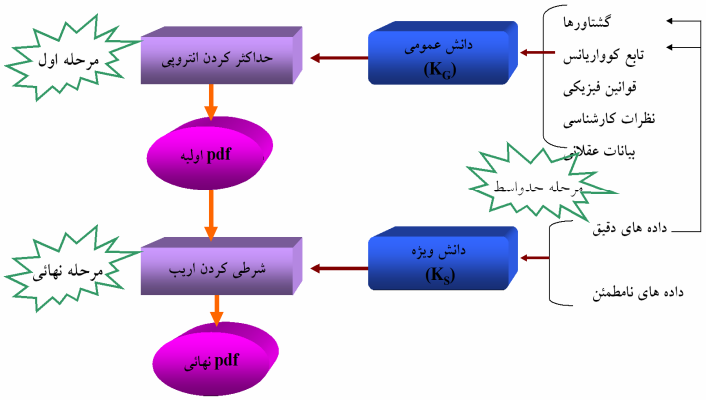
چهارچوب عمومی BME

BME شامل سه مرحله اساسی است: مرحله پیشین: این مرحله شامل یافتن عمومی‌ترین توزیع اولیه مقادیر با استفاده از تمام اطلاعات عمومی در دسترس و غیراختصاصی در رابطه با متغیر مورد مطالعه است. مرحله بینابینی در این مرحله اطلاعات ویژه‌ی در دسترس به وسیله مجموعه داده‌های مورد مطالعه، سازمان‌دهی و به روابط ریاضی قابل استفاده، ترجمه می‌شوند. مرحله نهائی: توزیع اولیه به دست آمده با توجه به اطلاعات جمع‌آوری شده از منطقه مورد مطالعه، به شکل قابل استفاده‌ای تبدیل می‌شوند. در خصوص **BME**، اولین قدم، از طریق حداکثر کردن انتروپی و مرحله سوم با استفاده از یک چهارچوب اریب (**Bayesian**) برای شرطی کردن توزیع اولیه نسبت به داده‌ها، به دست می‌آید (شکل). برای هر دو مرحله اول و سوم یاد شده، نمونه های عملی وجود دارد[۱]. اطلاعات مورد استفاده در این دو مرحله را می‌توان به دو گروه طبقه-بندی کرد: دانش عمومی و دانش ویژه. دانش عمومی ($\mathbf{K_G}$): دربرگیرنده تمام دانش زمینه، باورهای مورد قبول و تجارب پیشین می‌باشد. همه اینها می‌توانند به شکل اظهارات عقلانی و منطقی، قوانین فیزیکی، خصوصیات (برای مثال میانگین، تابع کوواریانس) یا حتی نظرات کارشناسی، ارائه شوند. $\mathbf{K_G}$ از این جهت به عنوان "عمومی" درنظر گرفته می‌شود که دربرگیرنده کلاس بزرگی از زمینه‌ها و موقعیت‌ها می‌باشد و به اندازه‌گیری خاص و ویژه ای در منطقه مورد مطالعه محدود نمی‌شود. دانش اختصاصی ($\mathbf{K_S}$)**:** از اطلاعات مربوط به یک منطقه خاص تشکیل شده است که اغلب مجموعه داده‌هائی‌اند که از منطقه مورد مطالعه جمع‌آوری شده‌اند و شامل داده‌های مطمئن و نامطمئن، چه به شکل پیوسته و یا طبقه‌بندی

شده، می‌باشند. K_G منجر به ایجاد توزیع اولیه از طریق حداکثر کردن انتروپی می‌شود در حالی که K_S امکان شرطی کردن این توزیع اولیه، برای تبدیل به توزیع نهائی را فراهم می‌آورد.

نظریه انتروپی و اطلاعات

اولین قدم در روش **BME**، شامل یافتن عمومی‌ترین توزیع اولیه‌ای است که دربرگیرنده تمام اطلاعات عمومی موجود (K_G) باشد. یکی از راه‌های ممکن، رسیدن به توزیعی است که انتروپی را تحت محدودیت‌های K_G، حداکثر می کند. در این بخش، ابتدا تعریفی از انتروپی ارائه می‌شود. سپس ارتباط بین انتروپی و اطلاعات، تشریح می‌شود و بخصوص اینکه چرا حداکثر کردن انتروپی توزیع اولیه، تحت محدودیت‌های K_G، منجر به حداکثر کردن اطلاعات مورد انتظار در توزیع اولیه می‌شود.



شکل ۱- طرح عمومی چهارچوب BME

تخمین BME

خلاصه‌ای از روش **BME** در شکل ۱ نمایش داده شده است. در طول مرحله اول، K_G با پایه دانش عمومی، از تمام اطلاعات عمومی قابل دسترس در منطقه مطالعه، تشکیل شده است. برای مثال، قوانین فیزیکی، نظرات کارشناسی، بیانات عقلانی و گشتاورهای جهانی. در این مرحله، K_G تنها از داده‌های مطمئن موجود در تمام منطقه مورد مطالعه و شامل میانگین m_z و تابع کوواریانس $C_z(h)$ می‌باشد. با حداکثر کردن انتروپی، K_G دارای توزیع (n_0+1) بُعدی می‌شود که در آن $n_0 = n_{0h}+n_{0s}$ است و نیز بُعد دیگری که توسط متغیر تصادفی در نقطه تخمین ایجاد شده است. این توابع احتمال تراکم اولیه، بدلیل استفاده از گشتاورهای تا رده دوم، دارای توزیع گوسین خواهند بود. در مرحله نهائی، قانون شرطی کردن اریب برای وارد کردن دانش ویژه K_S (داده‌های مطمئن و نامطمئن در همسایگی) استفاده می‌شود. این مرحله منجر به تشکیل توابع احتمال تراکم نهائی در نقطه X_0 می‌شود. چنین روشی هر بار برای هر منطقه تخمین تکرار می‌شود که در هر مرحله، منجر به یک توابع احتمال تراکم نهائی می‌شود. آخرین مرحله قبل از تولید نقشه نهائی، گزینه تخمین مرتبط با اهداف ویژه مطالعه است. با استفاده از توابع احتمال تراکم‌های نهائی مشابه، امکان تولید نقشه‌هائی که اطلاعات را به گونه‌ای خلاصه کنند که متناسب با اهداف نهائی مطالعه باشند، فراهم می‌شود.

منابع

[1] Christakos, G., Bogaert, P., and Serre, M.L. 2002. Temporal GIS. Advanced Functions for Field-Based Applications. Springer-Verlag, New York NY.
[2] Goovaerts, P. and Journel, A. G. 1995. Integrating soil map information in modeling the spatial variation of continuous soil properties. European Journal of Soil Science, 46(3):397-414.

[3] Stein, A., Hoogerwerf, M., and Bouma, J. 1988. Use of Soil-Map Delineations to Improve (Co)-Kriging of Point Data on Moisture Deficit. Geoderma, 43:163-177.