

بعض طرق تقدير البيانات المكانية ذات النمط النقطي مع التطبيق العملي دراسة مقارنة

جعفر موسى محمد البياتي

قسم الإحصاء، كلية الإدارة والاقتصاد، جامعة كركوك، كركوك، العراق

Email:jaufarmousa@yahoo.com

الملخص

يتناول هذه البحث طرق تقدير البيانات المكانية ذات نمط النقطي المتمثل بطريقة الموزون المسافة العكسية وتحليل سطح الاتجاه من الرتبة الأولى وكريكنك الاعتيادي مع حساب القيمة المطلقة لمتوسط خطأ التقدير MAE و القيمة المطلقة لجذر التربيعي لمتوسط خطأ التقدير RMSE لهذه الطرق الثلاث كعيارين لتحديد الطريقة المثلى لكي يمكن اعتمادها في عملية التكهون (التنبؤ)، وقد تم استخدام بيانات حقيقية في هذا البحث في مجال نسبة الملوحة الذائبة في مياه الآبار في المحافظة كركوك والتي تم الحصول عليها من هيئة المياه الجوفية في المحافظة التابعة لوزارة الموارد المائية لسنة 2014، كما وان جميع العمليات الحسابية تم برمجتها باستخدام برنامج Matlab وقد حصلنا على نتائج جيدة بطريقة كريكنك الاعتيادي نسبة إلى الطريقتين الآخرين .

الكلمات الدالة: تحليل البيانات المكانية، كريكنك الاعتيادي، الموزون المسافة العكسية، المتغيرات الموقعية

المقدمة:

هذه الطرق من ناحية القيمة المطلقة لمتوسط خطأ التقدير MAE القيمة المطلقة للجذر التربيعي لمتوسط خطأ التقدير RMSE.

تعريف المتغيرات المكانية:

إن المتغيرات التي يتعامل بها علم الإحصاء المكاني تختلف عن المتغيرات الاعتيادية حيث أن لكل قيمة من قيم المتغير الموقعي إحداثيات تمثل موقع تلك النقطة سواء كانت على سطح الأرض في المستوي أو في باطن الأرض أو خارج الأرض (مثلا التلوثات الجوية بالغازات). لنفرض أن $z(x)$ يمثل المتغير المكاني في الموقع x .

ضمن المنطقة D في الفضاء الاقليدي $x \in D \subseteq R^p$ ، حيث أن $P=2$ في المستوي أو $P=3$ في الفضاء. إذ يمكن قياس هذا المتغير على عينة حجمها n من المواقع وهذه القياسات يرمز لها بالمتغير $z(x)$ وقيمته $z(x_1), z(x_2), \dots, z(x_n)$ في المواقع x_1, x_2, \dots, x_n التي يفصل بينها إزاحة h أي أن:

$$|h| = \sqrt{u^2(x) - v^2(x)} \dots \dots \dots (3)$$

حيث $u(x)$ إحداثيات باتجاه شرق-غرب و $v(x)$ إحداثيات باتجاه شمال - جنوب. [5]

البيانات المكانية Spatial Data

البيانات تكون مكانية عندما يكون الإطار الإحصائي مكانيا، [4]، أي تكون المفردات التي يتكون منها ذلك الإطار منتظمة في حيز مكاني لتمثل مواقع أو قطاعات أو مساحات محددة. يكتسب هذا النوع من العينات أهمية خاصة في علم الجغرافية، وهناك ثلاث أنواع من البيانات في الإحصاء المكاني وهي كالاتي:

أ- **البيانات النقطية:** عندما يكون الإطار الإحصائي عبارة عن توزيع مكاني، على هيئة نقاط، موزعة على خريطة جغرافية، مثل مواقع المصانع، أو المساكن الريفية المنعزلة أو المحلات التجارية المتفرقة أو غير ذلك .

يعد علم الإحصاء المكاني من العلوم المهمة المستمرة بالتطور وذلك لأهميته في مجال التتقيب عن المعادن وفي مجال البيئة في تقدير تلوث الهواء وكذلك في مجال الزراعة في تقدير المحاصيل الزراعية وكثافة الغابات، لذا فان معظم الظواهر الجغرافية ظواهر مكانية ويميز الجغرافيون بين بيانات العينة المكانية النقطية والخطية وأخرى المساحية [1]، كما وان هناك طرق مختلفة لتقدير النقاط لمثل تلك البيانات والتي تتميز عن بعضها من ناحية الدقة وسهولة الاستعمال، في هذا البحث تم دراسة ثلاثة طرق في التقدير للبيانات النقطية منها الموزون المسافة العكسية وتحليل سطح الاتجاه من الرتبة الأولى وأسلوب كريكنك الاعتيادية (التي تستعمل غالبا في الإحصاء الجيولوجي)، ومن ثم تطبيق هذه الطرق على بيانات حقيقية في تقدير نسبة ملوحة مياه الآبار في محافظة كركوك ومن خلال حساب Mean Absolute Error القيمة المطلقة لمتوسط خطأ التقدير MAE

$$MAE = \left| \frac{1}{N} \sum_{i=1}^N \hat{z}(x_i) - z(x_i) \right| \dots \dots \dots (1)$$

حيث $z(x)$ قيمة الحقيقية و $\hat{z}(x)$ قيمة المقدرة و N عدد نقاط النقاط المجاورة للنقطة المقدرة.

وكذلك من خلال حساب Root Mean Squared Error القيمة المطلقة للجذر التربيعي لمتوسط خطأ التقدير RMSE

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [\hat{z}(x_i) - z(x_i)]^2} \dots \dots \dots (2)$$

حيث أن الصيغة التي تكون اقل MAE و RMSE تمثل الأسلوب الأمثل الذي يمكن اعتماده في التنبؤ لنقطة غير مقاسة، [7].

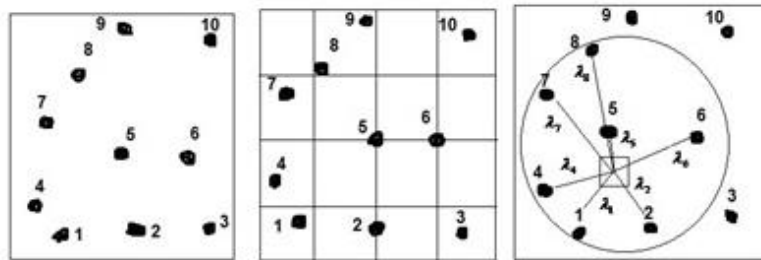
هدف البحث :

الهدف الأساسي من هذا البحث هو دراسة التقدير البيانات المكانية ذات النمط النقطي بأسلوب الموزون المسافة العكسية وتحليل سطح الاتجاه من الرتبة الأولى وأسلوب كريكنك الاعتيادية مع مقارنة بين

أ - الموزون المسافة العكسية

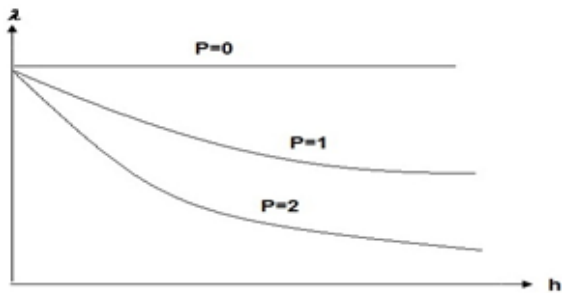
Inverse distance weighting

الموزون المسافة العكسية IDW يعتبر إحدى طرق التقدير الذي يجمع بين فكرة القرب التي تستند إليها طريقة مضلعات Thiessen Polygon مع التغيير التدريجي لسطح الاتجاه. هذا يمثل تطبيق الفرضية : أن مجموعة المواقع المقاسة القريبة من بعضها البعض لها تأثير أكبر من تلك المواقع المقاسة البعيدة في عملية التنبؤ (التنبؤ) للموقع غير المقاسة، لذا فإن IDW يستعمل القيم المقاسة القريبة أثناء عملية التنبؤ (التنبؤ) للنقطة غير المقاسة. إذ أن القيم المقاسة القريبة من موقع التنبؤ (التنبؤ) سيكون لها وزن أكبر من القيم البعيدة عنها وكذلك فإن هذا الوزن سيقبل مع بعد المسافة، ومن هنا جاء اسم معكوس المسافة الموزونة IDW كما في الشكل (1) أدناه.



شكل (1) يوضح طريقة معكوس المسافة الموزونة في التقدير

يكون متوسط القيم المقاسة وكلما ازداد p فإن الأوزان تتناقص بالنسبة للنقاط البعيدة ،وكما موضح في الشكل (2). فعندما تكون $p = 1$ فهذا يعني أن التقدير هو خطي بسيط بين النقاط . وقد وجد العديد من العاملين أنه إذا كان $p = 2$ يعطي نتائج جيدة للنقاط القريبة ذات الأوزان الكبيرة أكثر. [13]



الشكل (2) دوال الموزون المسافة العكسية

ب - تحليل سطح الاتجاه من الرتبة الأولى Teard surface analysis of first order

إن تحليل سطح الاتجاه يقوم بتقريب نقاط ذو قيم معلومة بمعادلة متعددة الحدود . وإن هذا التقدير يمكن بعد ذلك استخدامه في تقدير قيم بقية النقاط الأخرى . لذا فإن معادلة سطح الاتجاه من الرتبة الأولى تكون بالشكل الآتي وكما موضح في الشكل (3)

من الرتبة الأولى

$$z(x, y) = b_0 + b_1x + b_2y \dots \dots \dots (6)$$

ب- البيانات الخطية : من البيانات الجغرافية ما يختص بالتوزيعات المكانية الخطية (مثل خطوط أو شبكات المواصلات بأنواعها المختلفة والأنهار وروافدها ، أو أي ظاهرة جغرافية تتخذ شكل الامتداد الخطي)

ث-البيانات المساحية (مضلعات) :

عندما تكون مفردات الإطار الإحصائي ممثلة في الخريطة على هيئة مناطق ذات مساحات محدودة . يكثر استخدام هذه العينات في الجغرافية الطبيعية لتحديد أنواع النباتات الطبيعية أو أنواع الترب أو درجة التضرس أو التصحر في منطقة جغرافية معينة . [1] إلا أن بحثنا يقتصر على تقدير البيانات ذات نمط النقطي وبالأساليب الآتية:

إن مبدأ IDW هو تحديد الأوزان الكبيرة للنقاط القريبة من تلك النقاط البعيدة والتي يعبر عنها بالشكل التالي:

$$\hat{z}(s_0) = \sum_{i=1}^n \lambda_i z(s_i) \dots \dots \dots (4)$$

حيث أن $\hat{z}(s_0)$ قيمة المتنبأ عنها في الموقع s_0 و n عدد النقاط العينة المقاسة القريبة من موقع التنبؤ (التنبؤ) . و λ_i الأوزان للمواقع المقاسة وهذه الأوزان سوف يقل للنقاط البعيدة و $z(s_i)$ يمثل قيم المقاسة في المواقع s_i ، $i = 1, 2, \dots, n$ الصيغة العامة لحساب الأوزان يكون كالآتي:

$$\lambda_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^n d_{i0}^{-p}} \dots \dots \dots (5) \quad \sum_{i=1}^n \lambda_i = 1$$

وعندما تصبح المسافة أكبر، فإن الوزن λ يقل عن طريق عامل p . المقدار d_{i0} هي المسافة بين موقع التنبؤ (التنبؤ) s_0 ، وبقية المواقع المقاسة القريبة s_i ، وأن أوزان المواقع المقاسة المستعملة في عملية التنبؤ (التنبؤ) مجموعها يساوي 1 .

إن المعلمة p لها تأثير على الأوزان المواقع المقاسة في موقع التنبؤ (التنبؤ) ، ولذلك كلما ازدادت المسافة فإن الأوزان تنخفض أضعافاً مضاعفة. وإن سرعة انخفاض أوزان يعتمد على قيمة p . إذا كان $p = 0$ فإن هذا يعني ليس هناك انخفاض في الوزن مع المسافة أي أن كل وزن λ_i يكون نفسه ، والتنبؤ (التنبؤ) سوف

في هذه الحالة فإن $z(x_i)$ يستعمل لتقدير $z(x_0)$ الغير المقاسة بذلك فإن مقدر كريكنك الاعتيادي يكون بالصيغة التالية :

$$\hat{z}(x_0) = m(x_0) + \sum_{i=1}^{n(x)} \lambda_i^{ok} [z(x_i) - m(x_0)]$$

$$= \sum_{i=1}^{n(x)} \lambda_i^{ok} z(x_i) + [1 - \sum_{i=1}^{n(x)} \lambda_i^{ok}] m(x_0)$$

[12]

وحتى يكون هذا المقدر غير متحيز يجب أن يكون مجموع أوزان كريكنك الاعتيادي λ^{ok} مساوية للواحد ، لذا فإن مقدر كريكنك الاعتيادي الغير المتحيز يكون كالآتي :

with $\sum_{i=1}^{n(x)} \lambda_i^{ok} = 1$

$$\hat{z}_{ok}(x_0) = \sum_{i=1}^{n(x)} \lambda_i^{ok} z(x_i) \dots \dots \dots (8)$$

ولكي نقلل خطأ التباين

$$\sigma_E^2 = \text{var}[\hat{z}(x_0) - z(x_0)] = \text{var}[z(x_0)] + \text{var}[\hat{z}(x_0)] - 2 \text{cov}[z(x_0), \hat{z}(x_0)].$$

تحت القيد $\sum_{i=1}^{n(x)} \lambda_i^{ok} = 1$ نقوم بإضافة مضروب لاكرانج للمقدار أعلاه بشكل الآتي :

$$L = \sigma_E^2 + 2\mu_{ok} [1 - \sum_{i=1}^{n(x)} \lambda_i^{ok}]$$

وبأخذ المشتقة الجزئية بالنسبة لـ λ و μ ومساواتها للصفر نحصل على :

$$\frac{1}{2} \frac{\partial L}{\partial \lambda} = \sum_{i=1}^{n(x)} \lambda_i^{ok} C_R(x_j - x_i) + \mu_{ok} - C_R(x_j - x_0) = 0$$

$$\frac{1}{2} \frac{\partial L}{\partial \mu} = 1 - \sum_{i=1}^{n(x)} \lambda_i^{ok} = 0$$

بذلك فإن نظام المعادلات كريكنك الاعتيادي يكون بالشكل الآتي :

$$j = 1, 2, \dots, n(x) \dots \dots \dots (9)$$

$$\begin{cases} \sum_{i=1}^{n(x)} \lambda_i^{ok} C(x_i - x_j) + \mu_{ok} = C(x_j - x_0) \\ \sum_{i=1}^{n(x)} \lambda_i^{ok} = 1 \end{cases}$$

والتي يمكن أن يكتب بالشكل الآتي

$$C\lambda = B$$

$$\lambda = C^{-1}B \dots \dots \dots (10)$$

عندما نحصل على أوزان كريكنك وعلى معلمة مضاعف لاكرانج فإن تباين مقدر كريكنك يكون بشكل الآتي :

$$\sigma_{ok}^2 = C(O) - \sum_{i=1}^{n(x)} \lambda_i C(x_i - x_0) - \mu_{ok} \dots \dots \dots (11)$$

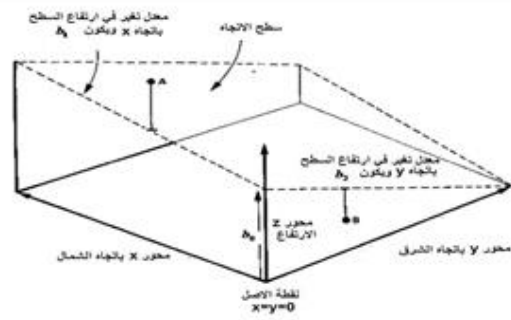
[8]

شبه دالة الفايروكرام Semivariogram Function

من الرتبة الثانية

$$z(x, y) = b_0 + b_1x + b_2y + b_3x^2 + b_4y^2 + b_5xy$$

حيث أن القيمة $z(x, y)$ هي دالة بإحداثيات ، b_0 يمثل ارتفاع النقطة عن السطح ، b_1 يمثل معدل تغير ارتفاع السطح باتجاه x ، b_2 يمثل معدل تغير ارتفاع السطح باتجاه y ، إلا أن دراستنا يقتصر على الرتبة الأولى .



الشكل (3) سطح الانحدار من الرتبة الاولى

[9]

تستعمل طريقة المربعات الصغرى العامة لحل المعاملات لـ b_0, b_1, b_2

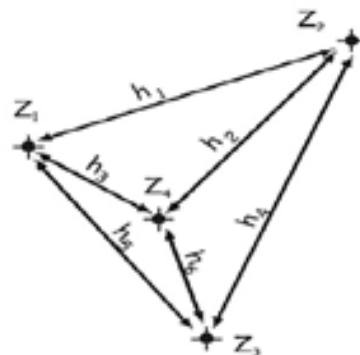
$$\begin{aligned} nb_0 + b_1 \sum x + b_2 \sum y &= \sum z \\ xb_0 + b_1 x^2 + b_2 xy &= \sum xz \\ yb_0 + b_1 xy + b_2 y^2 &= \sum yz \end{aligned}$$

$$\begin{bmatrix} n & \sum x & \sum y \\ \sum x & \sum x^2 & \sum xy \\ \sum y & \sum xy & \sum y^2 \end{bmatrix} * \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum z \\ \sum xz \\ \sum yz \end{bmatrix} \dots \dots \dots (7)$$

مع إمكانية إيجاد المعاملات بطريقة دالة الإمكان الأعظم. [2] ، [3]

ج - كريكنك الاعتيادي Ordinary kriging

في هذه الطريقة نفترض أن الوسط غير معلوم في منطقة الدراسة D ولكن ثابت في منطقة $A \subset D$ الجزئية من D ، أي في موقع النقاط المجاورة للنقطة المراد تقديرها وهذا يعني أن $\forall x \in A \subset D$ ، $m(x_i) = m(x_0)$ لقيم نقاط البيانات القريبة كما مبين في الشكل (4) أدناه :



الشكل (4) يوضح طريقة كريكنك في التقدير

وفي حالة الاستقرارية توجد علاقة تربط دالة الفايروكرام $\gamma(h)$ مع دالة التباين $C(h)$ والتباين $C(0)$ وهي :

$$\gamma(h) = C(0) - C(h) \dots \dots \dots (14)$$

وهناك صيغ رياضية متعددة لدالة الفايروكرام منها الفايروكرام الكروي spherical variogram

$$\gamma(h) = \begin{cases} \psi_0 & h = 0 \\ \psi_0 + \psi \left[\frac{3h}{2a} - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] & 0 < h \leq a \\ \psi_0 + \psi & h > a \end{cases} \dots \dots \dots (15)$$

وباستخدام العلاقة (14) نحصل على دالة التباين وهي كالآتي:

$$C(h) = \begin{cases} \psi_0 + \psi & h = 0 \\ \psi \left[1 - \frac{3h}{2a} + \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] & 0 < h \leq a \\ 0 & h > a \end{cases} \dots \dots \dots (16)$$

حيث أن ψ_0, ψ, a معاملات يتطلب تقديرها . [6] . [11]

الجانب التطبيقي :

نظرا لأهمية التكهن (التنبؤ) في الإحصاء المكاني في مجالات عملية كثير والذي يأخذ بنظر الاعتبار التركيب المكاني الأساسي الموجود في المعلومات الجيولوجية (المسافات بين المواقع)، وبواسطته يمكن إحصائيا الاستدلال عن التكهينات (التنبؤات) المثلى لمواقع معينة غير مقيسة لذا تم تطبيق هذا البحث على بيانات حقيقية في مجال درجة ملوحة مياه الآبار في محافظة كركوك لسنة 2014، وهذه البيانات قد تم الحصول عليها من هيئة المياه الجوفية في المحافظة والتابعة لوزارة الموارد المائية والمبين في الجدول (1) في الملحق حيث نلاحظ إن كل موقع في هذه البيانات يحتوي على إحداثيات $u(x)$ الذي يمثل شرق غرب و $v(x)$ الذي يمثل شمال جنوب وقيم $z(x_i)$ يمثل نسبة الملوحة الذاتية لكل جرام/لتر، إذ تم تطبيق هذه البيانات على طرق التقدير الثلاثة المذكورة في الجانب النظري لمعرفة أفضل هذه الطرق لكي يمكن اعتمادها في عمليه التكهن (التنبؤ) وكان العمل كالآتي:

1- طريقة الموزون المسافة العكسية: هذه الطريقة يعتمد على النقاط القريبة المجاورة، كما تم ذكره سابقا وبناء على ذلك تم اختيار خمسة مواقع متجاورة مختلفة حسب نقاط البيانات المراد تقديرها ومن ثم تطبيق المعادلة (5) لإيجاد وزن كل موقع مجاور للنقطة المراد تقديرها ومن ثم نستعمل المعادلة (4) للحصول على قيمة تقديرية لكل موقع، وكانت النتائج كما مبين في الجدول (2).

2- طريقة تحليل سطح الاتجاه من الرتبة الأولى: هذه الطريقة لا تشترط النقاط المجاورة للنقطة المراد تقديرها بل يستعمل جميع نقاط البيانات في عملية التقدير لذا فقد اعتمدنا على جميع نقاط البيانات في تقدير المعاملات بطريقة المربعات الصغرى كما في المعادلة (7) وقد كانت النتائج كالآتي:

b_0	b_1	b_2
20.5311	0.0150	-0.0490

اقترح Krige(1976) دالة شبه الفايروكرام بإزاحة h في الصيغة التالي:

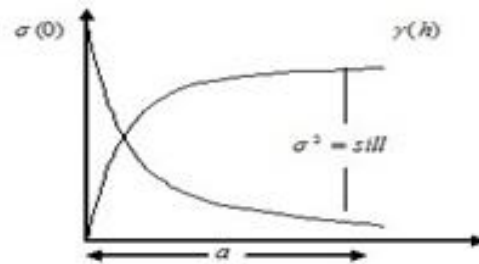
$$\gamma(h) = \frac{1}{2} \frac{1}{n(h)} \sum_{i=1}^{n(h)} (z(x_i) - z(x_i + h))^2 \dots \dots \dots (12)$$

وهي متوسط مربع الاختلافات الموجودة بين المشاهدات المكانية التي تبعد عن بعضها البعض إزاحة h ، إذ أن $n(h)$ تمثل عدد أزواج المشاهدات التي تفصل بين $z(x_i)$ ، $z(x_i + h)$ بإزاحة مقدارها h انظر [5]. أما إذا كتبت بالشكل:

$$2\gamma(h) = \frac{1}{n(h)} \sum_{i=1}^{n(h)} (z(x_i) - z(x_i + h))^2 \dots \dots \dots (13)$$

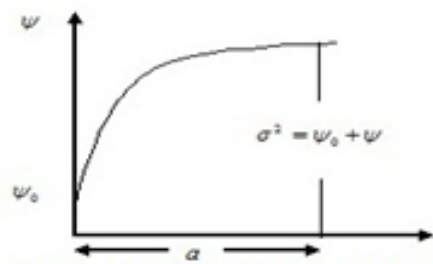
فإنها تسمى دالة الفايروكرام .

السبب في دراسة دالة الفايروكرام هو أن الصيغة الرياضية لها تمثل تباين الفروقات بين المشاهدات المكانية التي تبعد عن بعضها البعض إزاحة h . وكلما ازدادت الإزاحة h بين المشاهدات كلما أصبح التباين كبيرا حتى يستقر ارتفاعه عند مسافة معينة مثل $h = a$ وهذه المسافة a تسمى بالمدى وبعدها نلاحظ تلاشي التباين في دالة الفايروكرام حيث يستقر بقيمة ثابتة تساوي تباين σ^2 المشاهدات، وهذا التباين يدعى بـ (Sill) كما موضح بالشكل (5)



شكل (5) يوضح العلاقة بين دالتي الفايروكرام والتباين مع Sill

وعندما تقترب h من الصفر من الجهة اليمين فإن دالة شبه الفايروكرام لا تساوي صفر وإنما لها قيمة تساوي ψ_0 وهذه الظاهرة تمثل عدم الاستمرارية أو انقطاع دالة الفايروكرام عند $h = 0$ وتسمى في مجال الإحصاء المكاني Nugget Effect وهي تمثل الأخطاء العشوائية في وحدات القياس عندما تتغير فجأة الإزاحة h فجأة من وحدات المليمتر الى وحدات المتر أو الكيلومتر وكما موضح في الشكل (6)، [10] .



شكل (6) يوضح ظاهرة Nugget effect مع المدى

الاعتيايادي أولاً ثم طريقة ترجيح المسافة العكسية ثانياً وتحليل سطح الاتجاه ثالثاً ويرجع ذلك إلى أهمية استخدام جميع النقاط العينة مع أوزانها في التقدير التي تعتمدها طريقة كريكنك الاعتيادي مع إهمال النقاط البعيدة من خلال المدى $a = 18$ الذي يمثل نصف ابعاد مسافة بين النقطة المراد تقديرها وبقية النقاط العينة لان تأثير النقاط بعد قيمة المدى سوف يقل تدريجياً ويؤدي إلى أخطاء في التقديران لم يتم إهمالها، إما طريقة ترجيح المسافة العكسية رغم لها أفضلية أيضاً لان هناك تقارب كبير بينها وبين كريكنك الاعتيادي إلا أنها تكون في بعض الأحيان معرضة لتأثير نقاط شاذة عندما تستخدم فقط النقاط المجاورة وخاصة عندما تكون هذه النقاط قليلة وهذا ما جعلها في المرتبة الثانية، أما بالنسبة لطريقة تحليل سطح الاتجاه فأنها تعتمد على جميع نقاط العينة ولكن بدون استعمال الأوزان للمتغيرات وكذلك تأثير النقاط البعيدة وربما تكون شاذة وكذلك عدم دقة في قياس الإحداثيات للنقاط جعلها في المرتبة الثالثة.

الاستنتاجات:

1- يمكن استخدام المعيار مثل حساب قيمة المطلقة لمتوسط خطأ التقدير MAE و قيمة المطلقة لجذر التربيعي لمتوسط خطأ التقدير RMSE للمقارنة بين الطرق المستخدمة في هذا البحث ، حيث أن المقدار الأقل يمثل أسلوب الأمثل.

2- إن نقاط البيانات البعيدة عن النقطة المراد تقديرها لها تأثير سلبي على عملية التقدير كما لاحظنا في أسلوب تحليل سطح الاتجاه.

3- كريكنك الاعتيادي يمثل أفضل طريقة لعملية التنبؤ للبيانات المكانية ذات نمط النقطي حيث كانت متوسط خطأ التقدير MAE لها يساوي 27.77 وقيمة المطلقة لجذر التربيعي لمتوسط خطأ التقدير RMSE كانت 169.43 وهاتان القيمتان هما اقل قياسا بالطرق الأخرى المستخدمة .

لتوصيات :

1- تأكيد من طبيعة البيانات المكانية المستخدمة في الدراسة حيث أن هناك بيانات نقطية وخطية ومساحية ولكل منهم طريقة خاصة في التقدير .

2- يفضل أن يكون حجم العينة كبير لأنه كلما كان حجم العينة كبير كلما كانت النتائج أفضل هذا ما جعل طريقة كريكنك الاعتيادي أفضل على طرق الأخرى، وفي حالة كون العينات قليلة لظروف ما فيفضل استخدام أسلوب Co-kriging في التقدير ..

3- يفضل استعانة بمهندس المساح في قياس إحداثيات مواقع البيانات لان ذلك يؤثر على عملية التقدير .

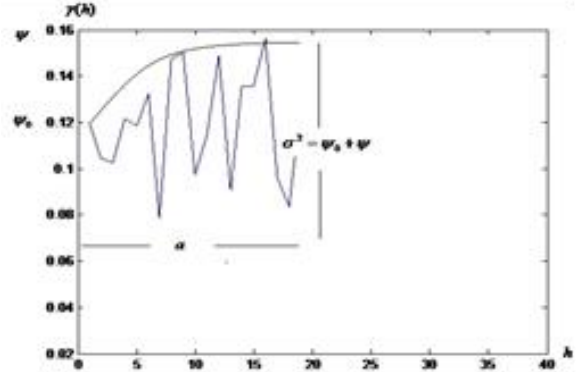
4- اختيار النموذج الملائم من خلال الرسم البياني بين الدالة الفايروكرام والإزاحة h وإلا سوف نحصل على نتائج غير دقيقة.

ومن هذه القيم حصلنا على معادلة الاتجاه الأتي :

$$z(x, y) = 20.5311 + 0.015x - 0.049y$$

بتعويض إحداثيات النقاط لكل قيمة معلومة نحصل على قيمة تقديرية لها وقد حصلنا على النتائج أيضاً مبينة في الجدول (2) .

3- طريقة كريكنك الاعتيادية : فهي تقترض إن الوسط ثابت في منطقة الدراسة A وتستعمل جميع نقاط البيانات مع أوزانها في التقدير ، حيث إن تقدير الأوزان المتغيرات يتم الحصول عليها من دالتي الفايروكرام والتغاير للبيانات، إذ يتم حساب دالة شبة الفايروكرام التجريبي حسب المعادلة (12) ومن ثم رسمنا العلاقة بينها وبين المسافة h وقد حصلنا على الشكل الأتي :



الشكل (7) الرسم البياني لدالة الفايروكرام التجريبي

وقيم دالة الفايروكرام المحسوبة مبينة في الجدول (3) في الملحق . إن الشكل (7) هو اقرب إلى النموذج الرياضي الكروي الموضح في (15) ، [5] وقد حددنا المعلمات من الرسم إذ كانت $a = 18$ ، $\psi_0 = 0.12$ ، $\sigma^2 = C(0) = 0.16$ ، $\psi = 0.4$ إذ أن $\sigma^2 = \psi_0 + \psi$ وباستعمال المعادلة (14) نحصل على دالة التغاير الذي يكون بالشكل الأتي :

$$C(h) = \begin{cases} 0.16 & h = 0 \\ 0.4 \left[1 - \frac{3h}{36} + \frac{1}{2} \left(\frac{h}{18} \right)^3 \right] & 0 < h \leq a \\ 0 & h > a \end{cases} \dots \dots \dots (17)$$

من المعادلة (17) اعلاه نحصل على مصفوفة التغاير C والمتجه B الذي يمثل متجه التغاير بين الموقع المراد تقديره وبقية المواقع ، ومن الجدير بالذكر أن مصفوفة التغاير C يجب أن تكون موجبة التعريف Positive definition، وبعد إيجاد معكوسها نحصل على الأوزان λ_i وذلك بتطبيق الصيغتين (10) و(8) تعطينا قيمة تقديرية للنقطة. وبتكرار هذه العملية لجميع النقاط المعلومة نحصل على القيم التقديرية لكل نقاط البيانات ، والنتائج التي حصلنا عليها مبينة في الجدول (2).

وبعد تطبيق المعادلتين (1) و (2) على الطرق الثلاثة ظهرت النتائج المبينة في الجدول (4) ، هذه النتائج تؤكد أفضلية طريقة كريكنك

الملحق

جدول (1) بيانات أبار الجوفية تابعة لوزارة الموارد المائية هيئة المياه الجوفية في كركوك

الموقع	$u(x)$	$v(x)$	الأملاح الكلية الذائبة بالجرام / لتر	الموقع	$u(x)$	$v(x)$	الأملاح الكلية الذائبة بالجرام / لتر
1	35.46	44.38	2.89	20	35.58	44.17	3.33
2	35.45	44.38	2.98	21	35.46	44.13	3.32
3	35.46	44.4	2.59	22	35.43	44.13	3.35
4	35.44	44.37	3.17	23	35.49	44.63	2.36
5	35.44	44.4	2.51	24	35.59	44.13	3.28
6	35.44	44.41	2.91	25	35.73	44.47	2.47
7	35.41	44.39	3.31	26	35.67	44.01	2.92
8	35.43	44.33	2.91	27	35.38	43.94	3.4
9	35.41	44.41	3.22	28	35.3	44.8	3.24
10	35.45	44.32	2.85	29	35.71	43.89	3.15
11	35.4	44.4	3.19	30	35.44	43.78	2.89
12	35.51	44.34	2.55	31	34.96	43.9	3.21
13	35.38	44.35	2.68	32	35.64	45.13	2.91
14	35.38	44.35	2.95	33	35.02	45.12	2.54
15	35.39	44.29	2.79	34	34.58	44.38	2.98
16	35.6	44.34	2.51	35	35.41	45.38	2.63
17	35.3	44.37	3.39	36	35.09	45.37	2.43
18	35.63	44.4	2.36	37	33.13	44.45	2.67
19	35.3	44.51	2.51				

جدول (2) نتائج القيم التقديرية لطرق الثلاثة

الموقع	IDW	TSA	OK	الموقع	IDW	TSA	OK
1	2.83	2.91	2.83	20	2.92	3	2.92
2	2.81	2.91	2.81	21	3.23	3.02	3.23
3	2.89	2.9	2.89	22	3.11	3.01	3.12
4	2.86	2.92	2.86	23	2.74	2.81	2.74
5	2.91	2.9	2.91	24	3.09	3.02	3.09
6	2.92	2.9	2.92	25	2.66	2.88	2.66
7	3	2.91	3	26	3.29	3.07	3.29
8	2.91	2.93	2.91	27	3.2	3.09	3.19
9	3.02	2.9	3.02	28	2.74	2.74	2.74
10	2.9	2.94	2.9	29	3.15	3.12	3.15
11	2.91	2.9	2.91	30	3.23	3.16	3.23
12	2.96	2.93	2.96	31	3.28	3.1	3.27
13	3.03	2.92	3.03	32	2.65	2.61	2.65
14	2.97	2.92	2.97	33	2.91	2.6	2.9
15	3.01	2.95	3.01	34	2.99	2.9	2.99
16	2.67	2.93	2.67	35	2.72	2.5	2.72
17	3.07	2.91	3.07	26	2.93	2.5	2.92
18	2.83	2.91	2.83	37	2.73	2.85	2.74
19	3.12	2.86	3.12				

جدول (3) نتائج دالة الفايروكرام التجريبي

$\gamma(h)$	h	$\gamma(h)$	h	$\gamma(h)$	h
0.08	25	0.09	13	0.12	1
0.12	26	0.14	14	0.1	2
0.07	27	0.14	15	0.1	3
0.08	28	0.16	16	0.12	4
0.11	29	0.1	17	0.12	5
0.06	30	0.08	18	0.13	6
0.06	31	0.13	19	0.08	7
0.07	32	0.04	20	0.15	8
0.05	33	0.08	21	0.15	9
0.06	34	0.08	22	0.1	10
0.08	35	0.06	23	0.11	11
0.02	36	0.08	24	0.15	12

جدول (4) نتائج MAE و RMSE

	IDW	TSA	OK
MAE	27.79	27.84	27.77
RMSE	169.43	169.45	169.43

المصادر

- 8 - Cressie, N .:"Spatial prediction and ordinary kriging. Mathematical" Geology **20**, . (1988)، 405-421,
- 9- David J. Unwin "AN INTRODUCTION TO TREND SURFACE ANALYSIS (University of Leicester) , (1978) .
- 10- Krige, D .G. Some Basic Consideration in the Application of Geostatistics to the valuation of ore in south African Gold Mines , Journal of the South African Institute of Mining and Metallurgy, (1976)، 383-391.
- 11 - Malvic, T. & Đurekovic, M: Application of methods: inverse distance weighting, ordinary kriging and collocated cokriging in porosity evaluation, and comparison of results on the Benièanci and Stari Gradac Fields in Croatia, Nafta (Zagreb) ,, (2003), 54, 9, 331-340
- 12 -Steven, S.C. and Cressie, N A comparison of Geostatistics methodologies used to estimate snow water equivalent (Water resources bulletin). Journal of the American water resources association،(1996) ، Vol.32,No.2.,
- 13 - Sheimy, N.E): "Digital Terrain Modelling". The University of Calgary, (1999) .Available: <http://www.geomatics.ucalgary.ca/~nel/shei/lecture.htm>
- 1- خديجة عبد الزهرة حسين لأطر النظرية لنظم المعلومات الجغرافية" مجلة آداب البصرة ، جامعة البصرة ، (2007)، العدد (42).
- 2- محمد نذير إسماعيل القاسم وشيما رياض ذنون : " تحليل سطح الاتجاه المكاني لمنسوب المياه الجوفية في قضاء سنجار " المجلة العراقية للعلوم الإحصائية " (2014) ، العدد (26) ص [51-81]، جامعة الموصل.
- 3- محمد نذير إسماعيل القاسم و دينا أذونية عوديشو : " تحليل سطح الاتجاه للبيانات المكانية مع التطبيق" مجلة التربية والعلم ، (2009)، المجلد (22) العدد (2) ، جامعة الموصل .
- 4- نعمان عابد شحادة : "التحليل الإحصائي في الجغرافية والعلوم الاجتماعية" الطبعة الأولى (2011) ، دار صفاء للنشر والتوزيع ، الجامعة الأردنية- عمان.
- 5 - Allan A. Nielsen, "2D semivariograms," in Proceedings of the Fourth South African Workshop on Pattern Recognition, Paul Cilliers, Ed., Simon's Town, South Africa, (1993) , 25–26 November, pp. 25–35.
- 6- Allan A. Nielsen "Kriging," lecture note, Technical University of Denmark, (2004), Internet : <http://www.imm.dtu.dk/pubdb/p.php?3479>.
- 7 - Burrough, P.A. and Mcdonnell, R.A. Principles of geographical information systems. Clarendon press, Oxford، (1998) , 346 pp.

Some methods of estimating the spatial pattern point data with application (comparative study)

Jaufar Mousa Mohammed

Dep. of Statistics, College of Administration and Economics, Kirkuk University, Iraq

Abstract

The idea of this paper is to estimate spatial data points pattern by Inverse distance weighting, Teard surface analysis of first order and Ordinary kriging method, we calculate the mean absolute error MAE and root mean square error RMSE of these methods as standard to determine the best method to be adopted in the forecasting process, we have been using real data in this paper in the field of salinity dissolved in water wells in the province of kirkuk government , which has been obtained from the groundwater body in the Ministry of water Resources 2014 , and that all the calculations have been processed using matlab program with respect to another to methods

Key word: Spatial data analysis, Ordinary kriging , Inverse distance weighting , Regionalized variables