تحليل سطح الاتجاه للبيانات المكانية مع التطبيق

د. محمد نذير إسماعيل قاسم **أ. دينا أدونية عوديشو**

قسم الرياضيات / كلية التربية

جامعة الموصل

الاستلام **القبول**

٣١ / **05** / ٢٠٠٧ **05** / **11** / ٢٠٠٧

ABSTRACT

This paper investigates the problem concerning the analysis of trend surface of spatially correlated data, we have got the model of trend surface based on an assumed covariance function. This function contains parameters called variance components, these parameters are estimated by maximum likelihood estimation that depends on Newten-Raphson algorithms to obtain the convergent solutions.

The applied representation includes the analysis of real spatial data which represent the level of ground water in Al-Qaaim west of Iraq, the results is very encouraging.

All algorithms are programmed by MATLAB.

الملخص

يتناول البحث تحليل سطح الاتجاه للبيانات المكانية حيث حصلنا على نموذج سطح الاتجاه بعد افتراض دالة تغاير مكاني معينة والأخيرة دالة بدلالة معلمات تسمى مكونات التباين Variance components إذ تم تقدير هذه المكونات بواسطة طريقة الترجيح الأعظم التي اعتمدت خوارزمية نيوتن – رافسون التكرارية للوصول إلى الحل المتقارب.

ومن ناحية الجانب التطبيقي فقد تم تحليل بيانات مكانية حقيقية والتي تمثل ارتفاع مناسيب المياه الجوفية في منطقة القائم غرب العراق وكانت نتائج التقدير والتنبؤ مشجعة جدا.

جميع الخوارزميات تمت برمجتها بواسطة نظام الـ MATLAB.

1- المقدمة:

ان تحليل سطح الاتجاه يستعمل للحصول على وصف سطح لتمثيل نمط معين حقيقي للبيانات أو يسلط الضوء لتمثيل التغيرات الحقيقية التي تحدث في العمليات العشوائية المكانية.

ويتمثل النموذج الأساس عادة في تحليل سطح الاتجاه بتوفيق سطح الاتجاه بطريقة المربعات الصغرى الاعتيادية Ordinary least Square (OLS) والذي يفترض ان البواقي تكون مستقلة، ولكن إذا كانت البواقي مرتبطة فيما بينها مكانيا فان طريقة تقدير المربعات الصغرى تصبح غير كفوءة، وان الأخطاء المعيارية واختبار المعنوية تكون متحيزة، ان الارتباط الذاتي بين البواقي يمكن ان يظهر في البيانات المكانية مثلا في عينة من مواقع مياه جوفية في منطقة معينة.

ان النمو الكبير والمتفجر في البيانات المكانية واستعمالاتها الكبيرة تطلب ت الحاجة ا لكبيرة للمعلومات المكانية التي تتمثل في تقدير سطوح الاتجاه فمثلا تعدين البيانات المكانية Shekhar and Chawla (2003) ،Roddick and Spilioulou (1999) التي هي عبارة عن عمليات لاكتشاف معلومات مفيدة من معلومات موجودة غير معروفة والوصول إلى أنماط Patterns علمية مءعينة من قاعدة هذه البيانات.

ان النماذج الإحصائية غالباً ما تستعمل لتمثيل المشاهدات بدلالة المتغيرات العشوائية وهذه النماذج يمكن استعمالها للتقدير والوصف والتنبؤ المعتمد على النظرية الاحتمالية.

ان العلاقات الرياضية ضمن الأشياء المكانية غالبا ما تكون ضمنية مثلا التداخل والتقاطع وما وراءهما. ون إحدى الطرائق الممكنة للتعامل مع العلاقات المكانية الضمنية هي تحويل مادية العلاقات إلى أعمدة بيانات تقليدية ومن ثم تطبيق أساليب التحليل الإحصائي على البيانات التقليدية Quinlan (1993) ،Barnett and Lewis (1994) ، Agrawal and Srikant (1994) .

أما هدف البحث فيتمثل في وصف مجموعة بيانات مكانية مترابطة، إذ تناولنا البيانات المكانية التي تمثل بيانات حقيقية لارتفاع مناسيب المياه لأبار جوفية وعددها ١٤ بئراً في قضاء القائم في العراق، وتقدير دالة التغاير للحصول على معلمات نموذج الانحدار لهذه البيانات للحصول على سطح الاتجاه للمتغير المكاني الذي يمثل ارتفاع مناسيب المياه في هذا القضاء، وسطح الاتجاه هذا يفيد في التنبؤ عن ارتفاع منسوب المياه في مواقع أخرى غير مقاسة في قضاء القائم.

2- المتغير الموقعي : Regionalized Variable

تدرس ظاهرة التعدين أو الظاهرة الفراغية بمتغير يختلف عن المتغيرات التي تستعمل في الإحصاء الكلاسيكي وقد تم تعريف هذا المتغير بالمتغير المكانيRegionalized Variable من قبل (1963) Matheron الذي يعد من الرواد الأوائل في هذا المجال.

ان المتغير المكاني يمثل كمية (درجة المعدن Ore Grade) الموجودة في موقع أو مواقع معينة في حقل معدن ما . ويرمز له بالرمز $Y(s)$ إذ $s \in R^r$ و R^r فضاء اقليدس ببعدين $r=2$ ، $S=(u(s),v(s))$ أو ثلاثة أبعاد $r=3$ ، $S=(u(s),v(s),w(s))$.

ان نظرية الإحصاء الفراغي تعتمد على دراسة الاختلافات الموجودة بين المتغيرات المكانية $Y(s)$ و $Y(s+h)$، إذ أن $Y(s)$ كمية الم عدن المتواجد في الموقع s وكمية المعدن $Y(s+h)$ في موقع يبعد عن الموقع s بإزاحة h و h تمثل المسافة الاقليدية بين الموقعين فقد تكون:

$$|h| = \sqrt{u^2(s)+v^2(s)}$$

أو

$$|h| = \sqrt{u^2(s)+v^2(s)+w^2(s)}$$

والمتغير المكاني $Y(s)$ يعد دالة متغير عشوائي لعدم معرفتنا بصورة مؤكدة عن قيمة هذا المتغير في الموقع المعين إلا بعد التنقيب عنه.

ان تحليل سطح الاتجاه يتضمن وصف البيانات المكانية بواسطة نموذج رياضي متعدد الحدود برتبة معينة ذات بعدين للعملية العشوائية $[Y(s), s \in D]$.

افرض ان مشاهدة في الموقع i يمكن كتابتها بالشكل:

$$Y_i = Y(s_i) = Y(u(s_i), v(s_i))$$

حيث ان:

$$S_i = (u(s_i), v(s_i))'$$

وان: u,v يمثلان نظام الإحداثيا ت الكارتيزية Cartesian Co-ordinates للمنطقة D قيد الدراسة، وبشكل عام فان تحليل سطح الاتجاه يفترض ان المشاهدة Y_i تتألف من مركبتين هما مركبة الاتجاه T_i المتمثلة بمعادلة متعدد الحدود ومركبة البواقي R_i أي ان:

$$Y_i = T_i + R_i$$

عندئذ تكون T قيمة محددة من:

$$T(u,v) = B_{00} + B_{10}u + B_{01}v + B_{20}u^2 + B_{11}uv + B_{02}v^2 + \ldots + B_{pq}u^p v^q$$

حيث ان $B_{00}, B_{10}, B_{01}, \ldots, B_{pq}$ تمثل معلمات النموذج والمطلوب تقديرها من المشاهدات.

وان $p+q$ تمثل رتبة سطح الاتجاه.

تمثل الرتبة الأولى (خطية) من سطح الاتجاه $p+q=1$

تمثل الرتبة الثانية (تربيعية) من سطح الاتجاه $p+q=2$

تمثل الرتبة الثالثة (تكعيبية) من سطح الاتجاه $p+q=3$

تمثل الرتبة الرابعة (ربلعية) من سطح الاتجاه $p+q=4$

وهكذا لسطوح ذات رتبة أعلى.

ان سطح الاتجاه ذا الرتبة $p+q$ له m من المعلمات التي يمكن إيجادها من المعادلة:

$$m=\frac{1}{2}(p+q+1)(p+q+2)$$

وهذه المعلمات يمكن تقديرها فقط في حالة كون عدد المشاهدات n تحقق الشرط $n\geq m$. وإذا كانت $n=m$ فان هناك سطحاً من نوع خاص يوفق للبيانات والذي له الخاصية $T_i=Y_i$ وبدون بواقي، ولكن في معظم التطبيقات لتحليل سطح الاتجاه فان n اكبر بكثير من عدد المعلمات m. وان سطوح الاتجاه توفق للبيانات باستعمال طريقة المربعات الصغرى.

افرض ان Y تمثل قيمة ذات بعد n من المشاهدات وبشكل مصفوفة نكتب النموذج كالآتي:

$$Y=XB+e \qquad \ldots (1)$$

حيث ان : X مصفوفة ذات بعد $(n\times r)$ من الإحداثيات لمواقع المشاهدات و B متجه ذو بعد r من معلمات سطح الاتجاه أي ان:

$$X=\begin{bmatrix} 1 & u_1 & v_1 & u_1^2 & v_1^2 & u_1v_1 & \ldots & u_1^pv_1^q \\ 1 & u_2 & v_2 & u_2^2 & v_2^2 & u_2v_2 & \ldots & u_2^pv_2^q \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & u_n & v_n & u_n^2 & v_n^2 & u_n^2v_n^2 & \ldots & u_n^pv_n^q \end{bmatrix}$$

وان:

$$B=\left(B_{00}\ B_{10}\ B_{01}\ B_{20}\ B_{02}\ B_{11}\ \ldots\ldots B_{pq}\right)^t$$

المتجه u ذو بعد n من الأخطاء العشوائية بحيث:

$$E(e)=0$$
$$E(ee')=\Omega$$

حيث ان Ω: هي مصفوفة التغاير ذات بعد $(n\times n)$ متماثلة وأكيدة الايجابية، وفي اغلب التطبيقات العملية تكون Ω غير معلومة ويجب تقديرها من البيانات.

وفي الحالة القياسية $\Omega=\sigma^2 I$.

وان I: مصفوفة الوحدة وتكون البواقي مستقلة ولاحظ أيضاً

$$E(Y)=XB$$
$$Var(Y)=\Omega=\Omega(\theta)$$

حيث θ معلمة تمثل مكونات التباين غير معلومة والمطلوب تقديرها من البيانات.

3- نماذج مختلفة للتغاير المكاني: Different Models of Spatial Covariance

افرض ان $S_i = \left(u\left(s_i\right), v\left(s_i\right)\right)'$ متجه الإحداثيات الكارتيزية الذي يمثل مواقع المشاهدات $Y\left(s_i\right)$ وان لدينا n من المشاهدات:

$$\left(Y\left(s_1\right), Y\left(s_2\right), \ldots, Y\left(s_n\right)\right)$$

فان

$$h_{ij} = |h| = \sqrt{\left(u\left(s_i\right) - u\left(s_j\right)\right)^2 + \left(v\left(s_i\right) - v\left(s_j\right)\right)^2}$$

تمثل الإزاحة بين المشاهدة $Y\left(s_i\right)$ والمشاهدة $Y\left(s_j\right)$ وفيما يأتي بعض نماذج التغاير التي هي عبارة عن دالة للإزاحة h_{ij}، حيث ان كل نموذج خاص لدراسة معينة، مثلاً النموذج (2) يستعمل في الدراسات المائية مثل المياه الجوفية ، والنموذج (3) يستعمل في التعدين عن الذهب، والنموذج (5) يستعمل في الدراسات الزراعية والنباتات ، أما النموذج (6) يستعمل في التعدين عن الفضة والنحاس. انظر Worsley et al (1991)، وان دوال التغاير التي تعتمد على الإزاحة فقط تسمى دوال تغاير موحدة الخواص Isotropic.

افرض ان:

$$\sigma_{ij} = c\left(h_{ij}\right)$$

1. نموذج Ripley (1981)

$$1.\ \sigma_{ij} = \sigma_1^2 - \sigma_2^2 h_{ij} \qquad i \neq j$$
$$= \sigma_1^2 + \sigma_3^2 \qquad i = j \qquad \ldots(2)$$

Ripley (1981)

2. نموذج Mardia and Marshall (1984)

$$2.\ \left.\begin{aligned} \sigma_{ij} &= \sigma^2\left(1 - (3/2)h_{ij}/a + (1/2)h_{ij}^3/a^3\right) && 0 \le h_{ij} < a \\ &= Zero && h_{ij} \ge a \end{aligned}\right\} \quad \ldots(3)$$

Mardia and Marshall (1984).

3. نموذج Brewer and Mead (1986)

$$3.\ \sigma_{ij} = \sigma^2 \exp - \left(h_{ij}/a\right)^t \qquad 0 \le t \le 2 \qquad \ldots(4)$$

Brewer and Mead (1986).

4. نموذج Mardia and Watkins (1989)

$$4.\ \left.\begin{aligned} \sigma_{ij} &= \sigma^2\left(1 - h_{ij}/a\right)^t && 0 \le h_{ij} < a \\ &= Zero && h_{ij} \ge 0 \end{aligned}\right\} \quad \ldots(5)$$

Mardia and Watkins (1989).

5. نموذج Haslett and Raftery (1989)

$$5.\ \sigma_{ij} = \left.\begin{array}{ll}\sigma^2 b\, exp\left(h_{ij}/a\right) & i \neq j \\ = \sigma^2 & i = j\end{array}\right\} \quad \ldots(6)$$

Haslett and Raftery (1989)

4- طريقة الإمكان الأعظم : Maximum Likelihood Method

عدّ Mardia and Marshall (1984) طريقة تقدير الإمكان الأعظم لمعلمات نموذج الانحدار للبيانات المكانية الذي تكون فيه الأخطاء العشوائية مترابطة ذاتياً.

التغاير الذاتي بين أية مشاهدتين من Y يمكن ان يعرف بواسطة نموذج معلمي Parametric model كدالة بدلالة الموقع بين هاتين المشاهدتين والذي يعتمد على المسافة بين موقعي المشاهدتين ومجموعة من المعلمات المجهولة θ.

كما نعلم ان مصفوفة التغاير الذاتي Ω يجب ان تكون مصفوفة أكيدة الايجابية.

نفترض أن العملية العشوائية $Y(s)$ هي عملية كاوسيان ولا يفترض ان تكون مستقرة أي ان:

$$E\left(Y(s)\right) = \mu(s) = f(s)'B$$

ودالة التغاير

$$cov\left(Y(s), Y(s^*)\right) = c\left(h, \theta\right)$$

حيث ان: B متجه معلمات الاتجاه بسعة $(k \times 1)$.

θ متجه معلمات دالة التغاير بسعة $r \times 1$.

$f(s)$ متجه معلوم من الدوال التي تمثل إحداثيات مواقع المشاهدات.

h المسافة الاقليدية بين المشاهدة $Y(s)$ و $Y(s^*)$.

المطلوب تقدير المعلمتين B و θ بطريقة الإمكان الأعظم، ولتقدير B و θ بطريقة الإمكان الأعظم نأخذ عينة عشوائية بحجم n من المشاهدات من العملية العشوائية $\{Y(s), s \in D\}$.

حيث ان D: هي مجال منطقة الدراسة أي ان العينة العشوائية ذات حجم n ستمثل بالمشاهدات $Y(s_1), Y(s_2), \ldots, Y(s_n)$ حيث ان $s_1, s_2, \ldots, s_n$ تمثل مواقع المشاهدات.

$$s_i \in D \quad , \quad \forall i \quad i = 1,2,\ldots,n$$

يمكن تمثيل المشاهدة الواحدة بالنموذج الخطي

$$Y(s) = f(s)'B + e(s), \quad s \in D \qquad \ldots(7)$$

حيث ان $e(s)$ الخطأ العشوائي الذي يتوزع توزيع كاوسيان بحيث

$$e(s) \sim N\left(0, c(h, \theta)\right)$$

افرض

$$Y = \left(Y(s_1), Y(s_2), \ldots, Y(s_n)\right)'$$

يمكن كتابة النموذج (7) بشكل مصفوفات وكما يأتي:

$$Y = XB + e$$
$$E(Y) = XB$$
$$Var(Y) = Var(XB + e)$$
$$= Var(e) = \Omega(\theta)$$

حيث ان:

B : متجه المعلمات وكما هو معرف سابقاً.

X : مصفوفة معلومة من إحداثيات المواقع.

e : متجه الأخطاء العشوائية بسعة $(n \times 1)$.

Ω : مصفوفة التغاير وهي دالة بمعلمات التغاير θ والمعرفة سابقاً.

على ضوء الافتراض بان توزيع $Y(s)$ كاوسيان فان:

$$Y \sim N\left(XB, \Omega(\theta)\right) \qquad \ldots(8)$$

ومن الجدير بالذكر بأننا سوف نستعمل $\Omega(\theta)$ للإشارة إلى Ω الذي هو دالة بدلالة المعلمات θ ، وان $Var(e) = Var(Y)$.

على ضوء التوزيع (8) يمكن كتابة دالة الإمكان للمعلمات θ, B وكما يأتي:

$$f(Y; B, \theta) = \frac{|\Omega|^{-\frac{1}{2}}}{(2\pi)^{n/2}} \; e^{-\frac{1}{2}(Y - XB)'\Omega^{-1}(Y - XB)} \qquad \ldots(9)$$

بأخذ $\log$ دالة الإمكان نحصل على:

$$L = log\, f(Y; B, \theta) = -\frac{n}{2} log(2\pi) - \frac{1}{2} log|\Omega| - \frac{1}{2}(Y - XB)'\Omega^{-1}(Y - XB) \quad \ldots(10)$$

نقوم باشتقاق L بالنسبة إلى B ومساواتها بالصفر فنحصل على:

$$\frac{\partial L}{\partial B} = (X'\Omega^{-1}Y) - (X'\Omega^{-1}X)B = 0$$

$$\therefore \hat{B} = (X'\,\Omega^{-1}X)^{-1}(X'\Omega^{-1}Y) \qquad \ldots(11)$$

بعد افتراض ان Ω معلومة حسب التقدير يمكن كتابة (11) كما يأتي:

$$\hat{B} = (X'\hat{\Omega}^{-1}X)^{-1}\, X'\hat{\Omega}^{-1}Y \qquad \ldots(12)$$

حيث ان المقدر $\hat{B}$ نحصل عليه من العلاقة (11) بعد تقدير مصفوفة التغاير Ω على اثر تقدير المعلمة θ.

نقوم باشتقاق L بالنسبة إلى θ ، $\theta = \left(\theta_1, \theta_2, \ldots, \theta_r\right)'$.

في البداية نذكر نتائج مهمة نستعملها في اشتقاق L بالنسبة إلى θ وكما يأتي:

1. $$\frac{\partial \log|\Omega|}{\partial \theta_i} = tr\,\Omega^{-1}\frac{\partial \Omega}{\partial \theta_i} \qquad \ldots(13)$$

2. $$\frac{\partial \Omega^{-1}}{\partial \theta_i} = -\Omega^{-1}\frac{\partial \Omega}{\partial \theta_i}\,\Omega^{-1} \qquad \ldots(14)$$

حيث tr : يمثل اثر المصفوفة وان لأيّ مصفوفة Z .

$$trZ = \sum_{i=1}^{n} Z_{ii}$$

لمزيد من التفاصيل حول (13) و (14) . انظر: Cressie (1993)

$$\frac{\partial L}{\partial \theta_i} = -\frac{1}{2} tr\ \Omega^{-1}\Omega_i + \frac{1}{2}W'\ \Omega^{-1}\Omega_i\Omega^{-1}W \qquad \ldots(15)$$

حيث ان

$$W = Y - XB \qquad , \qquad \Omega_i = \frac{\partial \Omega}{\partial \theta_i}$$

وعند إيجاد $\hat{B}$ يكون

$$W = Y - X\hat{B}$$

يمكن كتابة (15) بشكل أبسط وكما يأتي:

$$\frac{\partial L}{\partial \theta_i} = -\frac{1}{2} trE\ \Omega_i \qquad , \qquad i = 1,2,\ldots,r \qquad \ldots(16)$$

حيث ان

$$\left.\begin{aligned} E &= \Omega^{-1} - VV' \\ V &= \Omega^{-1}W \end{aligned}\right\} \qquad \ldots(17)$$

المعادلات (16) يمكن حلها تكرارياً وذلك باستعمال طريقة نيوتن – رافسون Newten-Raphson Method.

لإيجاد جذور المعادلة (15) نحتاج إلى نتائج أخرى للاشتقاق

$$\frac{\partial^2 \Omega^{-1}}{\partial \theta_i \partial \theta_j} = S_{ij}\Omega^{-1} + S_{ji}\ \Omega^{-1} - R_{ij} \qquad \ldots(18)$$

حيث ان

$$\left.\begin{aligned} R_{ij} &= \Omega_{ij}\Omega^{-1} \\ S_{ij} &= \Omega^{-1}\Omega_i\Omega^{-1}\Omega_j \end{aligned}\right\} \qquad \ldots(19)$$

وان

$$\Omega_{ij} = \frac{\partial^2 \Omega}{\partial \theta_i \partial \theta_j}$$

المعادلة (18) نحصل عليها من اشتقاق (14) بالنسبة إلى θ_j. وعند اشتقاق (13) بالنسبة لـ θ_j ينتج:

$$\frac{\partial^2 log|\Omega|}{\partial\theta_i\partial\theta_j} = -tr\left(S_{ij} - R_{ij}\right) \qquad \ldots(20)$$

لأنه لأية مصفوفة من الثوابت C لدينا نتيجة

$$\frac{\partial}{\partial\theta_i} trC\Omega = trC\Omega_i$$

بتطبيق هذه النتيجة على (13) ينتج (20).

إذن باشتقاق (15) بالنسبة إلى θ_j واستعمال (14) ينتج:

$$\frac{2\partial^2 L}{\partial\theta_i\,\theta_j} = -tr\left(R_{ij} - S_{ij}\right) - W'\left(S_{ij} + S_{ji} - R_{ij}\right)\Omega^{-1}W \qquad \ldots(21)$$

وبما انه لأي متجه $Z \sim N(0\,,\,\Sigma)$ ولاية مصفوفة A

$$E(Z'AZ) = trAE(ZZ')$$
$$= trA\Sigma$$

إذن بعد اخذ التوقع للمعادلة (21) سنحصل على

$$E\left(\frac{-\partial^2 L}{\partial\theta_i\partial\theta_j}\right) = \frac{1}{2}trS_{ij}$$

حيث ان S_{ij} معرفة سابقاً في (19).

افرض ان:

$$A = \frac{1}{2}trS_{ij} \qquad , \qquad i, j = 1,2,\ldots,k \qquad \ldots(22)$$

يمكن التعويض عن A بمصفوفة المعلومات Information Matrix

انظر Mardia and Marshall (1984).

المعادلتان (22) و(16) يمكن حلهما تكرارياً إذ نفترض قيماً ابتدائية للمعلمة θ مثلا θ_0، ثم نجد Ω_0 ثم $\hat{B}$ وبعد ذلك نطبق طريقة نيوتن – رافسون الآتية:

$$\theta_{k+1} = \theta_k + A_k^{-1}\delta_k \qquad \ldots(23)$$

حيث ان A_k هي المصفوفة من المعادلة (22) محسوبة عند المعلمة θ_k.

و δ_k هي متجه المشتقات محسوبة عند θ_k و $\hat{B}_k$ من المعادلة (16)، أي ان:

$$\delta_k = -\frac{1}{2}trE\Omega_i$$

ونستمر في عملية التكرار إلى ان نحصل على التقارب في قيم التقدير للمعلمات θ_k و B_k.

التطبيق

طريقة تحليل سطح الاتجاه:

بيانات ارتفاع مناسيب المياه الجوفية في قضاء القائم في العراق مبينة في الجدول (1)، انظر عبد الرحمن (2001). إذ تم اختيار نموذج تغاير ملائم لهذه البيانات، ومن خلال تحليل هذه البيانات تم التوصل إلى تقدير معلمات نموذج الانحدار لهذه البيانات المكانية للحصول على سطح الاتجاه للمتغير المكاني الذي يمثل اربفاع مناسيب المياه في هذا القضاء.

$u(s)$	$v(s)$	$Y(s)$
25	125	220.04
125	125	220.54
220	125	219.56
325	125	221.26
25	75	220.28
125	75	219.81
225	75	219.3
325	75	219.92
25	25	220.45
125	25	220.96
240	25	220.87
325	25	223.04
0	150	220
350	0	223.3

الجدول (1)

ارتفاع مناسيب المياه الجوفية $Y(s)$ بوحدة المتر والإحداثيات $u(s)$ ، $v(s)$
لمنطقة القائم في العراق.

ان نموذج التغاير الملائم لهذه البيانات هو نموذج (2).

$$\left.\begin{aligned} \sigma_{ij} &= \sigma_1^2 - \sigma_2^2 \;\; h_{ij} \quad \ldots \; \mathrm{i} \neq \mathrm{j} \\ &= \sigma_1^2 + \sigma_3^2 \qquad\quad \ldots \; \mathrm{i} = \mathrm{j} \end{aligned}\right\} \quad \ldots(24)$$

حيث σ_1^2 تمثل تغاير تأثير المشاهدة.

σ_2^2 تمثل تغاير التأثير المكاني للمشاهدة.

σ_3^2 هو تأثير النكت Nugget Effect

σ_1^2 ، σ_2^2 ، σ_3^2 تمثل معلمات دالة التغاير التي لا بد من تقديرها للحصول على نموذج التغاير التقريبي للبيانات من اجل استخدامها في عملية التنبؤ عن العملية العشوائية المكانية التي في دراستنا هذه تمثل المياه الجوفية وقد تكون معادن أو تلوث بيئة في دراسات أخرى. اما تأثير النكت فهو عبارة عن مصطلح يستخدم في الإحصاء المكاني والذي يمثل حد الخطأ العشوائي الذي يحدث بسبب كل من الأخطاء في القياس والأخطاء الدقيقة التي لا يمكن السيطرة عليها وتمثل ضعف في استمرارية الظاهرة المكانية قيد الدراسة مثلاً في التنقيب عن المعادن والمياه الجوفية الموجودة في باطن الأرض وغيرها من الدراسات المكانية.

نلاحظ من النموذج (24) عندما $i = j$ فان $\sigma_1^2 + \sigma_3^2$ تمثل التباين في القطر الرئيس من مصفوفة التغاير، وعندما $i \neq j$ فان $\sigma_1^2 - \sigma_2^2 h_{ij}$ تمثل التغاير الذي يعتمد على المسافة بين المشاهدة i والمشاهدة j .

يمكن كتابة النموذج (24) بشكل مصفوفة التغاير الآتية:

$$\Omega = \sigma_1^2 J - \sigma_2^2 H + \sigma_3^2 I \qquad \ldots(25)$$

نفرض ان $\sigma_1^2 = \theta_1$ ، $\sigma_2^2 = \theta_2$ ، $\sigma_3^2 = \theta_3$ فتصبح المعادلة (25) بالشكل

$$\Omega = \theta_1 J - \theta_2 H + \theta_3 I \qquad \ldots(26)$$

حيث أن:

$J(n \times n)$ مصفوفة من الوحدات.

$H(n \times n)$ مصفوفة المسافة h_{ij} والتي تحسب من القانون:

$$h_{ij} = \left((u_i - u_j)^2 + (v_i - v_j)^2\right)^{\frac{1}{2}} \qquad , \qquad i, j = 1, 2, \ldots, n$$

$I(n \times n)$ مصفوفة الوحدة.

$\theta_1, \theta_2, \theta_3$ معلمات مجهولة يمكن تقديرها بإحدى طرائق التقدير المعروفة وهي طريقة الترجيح الأعظم أو طريقة بيز أو طريقة مقدر اصغر معيار تربيعي غير متحيز أو طريقة المربعات الصغرى، وسنستخدم طريقة الترجيح الأعظم التي اعتمدت خوارزمية نيوتن- رافسون في تقديرها باستخدام القيم الابتدائية الآتية:

$$\theta_1 = 0.0098 \ , \ \theta_2 = 0.042 \ , \ \theta_3 = 0 \qquad \ldots(27)$$

والتي أخذت من المصدر عبد الرحمن (2005).

من القيم الابتدائية أعلاه يمكن بسهولة إيجاد دالة التغاير Ω من المعادلة (26) أعلاه، حيث J, H, I معرّفة سابقاً.

بعد ذلك إيجاد معكوس المصفوفة Ω ثم إيجاد المصفوفة $\hat{B}$ من خلال

$$\hat{B} = \left(X'\hat{\Omega}^{-1}X\right)X'\hat{\Omega}^{-1}Y \qquad \ldots(28)$$

وتكوين المصفوفة S_{ij} عن طريق

$$S_{ij}=\Omega^{-1}\Omega_i\Omega^{-1}\Omega_j \qquad , \qquad \begin{matrix} i=1,2,3 \\ j=1,2,3 \end{matrix}$$

وان $\Omega_1=J$ ، $\Omega_2=-H$ ، $\Omega_3=I$

ومن خلال المصفوفة S_{ij} يمكننا تكوين المصفوفة

$$A=\frac{1}{2}trS_{ij} \qquad \ldots(29)$$

ثم إيجاد معكوس المصفوفة A ، كذلك حساب

$$\delta_k=\frac{-1}{2}trE\Omega_i \qquad \ldots(30)$$

علماً بأن: $E=\Omega^{-1}-\Omega^{-1}(Y-XB)(Y-XB)'\Omega^{-1}$

بعد إيجاد قيم A و S_k وقيم θ_i الابتدائية المعروفة يتم تطبيق قانون نيوتن – رافسون

$$\theta_{k+1}=\theta_k+A_k^{-1}\delta_k \qquad \ldots(31)$$

وتكرر العملية في المعادلة (31) و (28) إلى أن يتم الحصول على النتائج المقبولة لكل من B,θ علما بأننا لم نحصل على التقارب السريع أثناء التكرارات للمعلمات $\theta_1,\theta_2,\theta_3$ ، ولكن بعد تعويضنا عن المصفوفة A بالمصفوفة $A+\lambda diag\ A\,,\ \lambda>0$.

حيث λ تسمى معلمة Levenberg – Marquardt حصلنا على التقارب السريع عند التكرار العاشر وكانت قيمة $\lambda=0.3$.

انظر Mardia and Marshall (1984).

والنتائج التي حصلنا عليها موضحة في الجدول (2).

المعلمات \ التكرار	1	2	3	4	5	6	7	8	9	10
$\hat{B}_o$	221.4193	221.3660	221.2941	221.2440	221.1977	221.2222	221.2257	221.2264	221.2266	221.2266
$\hat{B}_1$	0.0057	0.0051	0.0046	0.0043	0.0041	0.0042	0.0042	0.0042	0.0042	0.0042
$\hat{B}_2$	-0.0088	-0.0100	-0.0105	-0.0107	-0.0108	-0.0107	-0.0107	-0.0107	-0.0107	-0.0107
θ_1	0.0098	0.0957	0.3590	0.4176	0.3428	0.4306	0.4470	0.4507	0.4516	0.4518
θ_2	0.042	0.0112	0.0047	0.0031	0.0023	0.0026	0.0027	0.0027	0.0027	0.0027
θ_3	0	0.4013	0.4613	0.4864	0.5303	0.4946	0.4899	0.4981	0.4889	0.4888

الجدول (2)

نتائج تقدير معلمات دالة التغاير $\theta_1,\theta_2,\theta_3$ ومعلمات نموذج الانحدار $\hat{B}_o,\hat{B}_1,\hat{B}_2$

وعند تعويض قيم $\hat{B}$ التي حصلنا في معادلة الانحدار الآتية:

$$\hat{y} = \hat{B}_o + \hat{B}_1 u + \hat{B}_2 v$$
$$= 221.226 + 0.0042\,u - 0.0107\,v$$

حيث $u\,,v$ من الجدول (1).

نحصل على قيم $\hat{y}$ المقدرة، والجدول (3) يضم قيم y الأصلية والمقدرة $\hat{y}$ وبمقارنتها مع بعضها نلاحظ انها قريبة جداً، أي أن سطح الاتجاه الذي حصلنا عليه قريب جداً من السطح الحقيقي.

y	$\hat{y}$
220.04	219.9941
220.54	220.4141
219.56	220.8131
221.26	221.2541
220.28	220.5291
219.81	220.9491
219.3	221.3691
219.92	221.7891
220.45	221.0641
220.96	221.4848
220.87	221.9671
223.04	222.3241
220	219.6216
223.3	222.6966

الجدول (3)

مقارنة بين القيم الحقيقية y والتقديرية $\hat{y}$.

الاستنتاجات والتوصيات

*** الاستنتاجات:**

١. في تحليل سطح الاتجاه افترضنا دالة تغاير معروفة وقمنا بتقدير معلمات هذه الدالة بطريقة الإمكان الأعظم.

٢. اختبار رتبة نموذج سطح الاتجاه تم على أساس الحصول على تقديرات جيدة جدا.

٣. إن سطح الاتجاه للبيانات الحقيقة يتمثل بمستوى موازٍ لمستوي الإحداثيات.

٤. يجب ملاءمة نموذج التغاير من اجل الوصول إلى التقارب في تقدير الإمكان الأعظم.

*** التوصيات:**

١. يمكن اختبار رتبة نموذج سطح الاتجاه بواسطة أسلوب أخر غير الذي اعتمدناه، أنظر Haining (1987).

٢. تطبيق طريقة تحليل سطح الاتجاه على بيانات أخرى مثل التنبؤ عن المعادن والتلوث البيئي والغابات الطبيعية ...الخ.

المصادر

١. عبد الرحمن، نبال صباح (2001): نموذج تغاير مكاني خطي مع التطبيق، رسالة ماجستير غير منشورة، كلية التربية، جامعة الموصل.

٢. عبد الرحمن، نبال صباح (2005): تقدير بيز لمعلمات نموذج الانحلال التربيعي المكاني، بحث مقبول النشر، المجلة العراقية للعلوم الإحصائية.

3. Agrawal, R. and Srikant, R. (1994): Fast Algorithms for Mining Associations Rules. In Proc. of Very Large Databases. Dept of computer science and Engineering. University of Minnesota.

4. Barnett, V. and Lewis, T. (1994): Outliers in Statistical Data, John Wiley, 3rd edition.

5. Brewer, A. C. and Mead, R. (1986): Continuous Second Order Models of Spatial Variation with Application to the Efficiency of Field Crop Experiments. (with discussion), Journal of the Royal Statistical Society. Ser. A. 149, pp. 314-348.

6. Cressie, N. (1993): Statistic for Spatial Data, John Wiley, New York.

7. Haining, R. (1987): Trend – Surface Models with Regional and Local Scales of Variation with an Application to Aerial Survey Data. Technometrics, Vol. 29, No.4, PP. 461-469.

8. Haslett, J. and Rafferty, A. E. (1989): Space – Time Modelling with Long–Memory Dependence: Assessing Ireland’s Wind Power Resource. (with discussion), Applied Statistcis, 38, 1-50.

9. Mardia, K. V. and Marshall, R. J. (1984): Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression, Biometrika, V. 71, 135-146.

10. Mardia, K. V. and Watkins, A. J. (1989): On Multimodality of the Likelihood in the Spatial Linear Model, Biometrika, V. 76, 289-295.

11. Matheron, G. (1963): Principle of Geostatistics.J. Economic Geology.V. 58, pp.1246-1266.

12. Quinlan, J. (1993): Programs for Machine Learning. Morgan Kaufmann Publishers.

13. Ripley, B. D. (1981): Spatial Statistic, John Wiley, New York.

14. Roddick, J. F. and Spiliopoulou, M. (1999): A Bibliography of Temporal, Spatial and Spatio – Temporal Data Mining Research. SIGKDD Explorations 1 (1): 34-38.

15. Shekhar, S. and Chawla, S. (2003): Spatial Databases: A Tour, Prentice Hall (ISBN 0 – 7484 – 0064 –6).

16. Worsley, K. J., Evans, A. C., Strother, S. C. and Tyler, J. L. (1991): A Linear Spatial Correlation Model with Applications to Positron Emission Tomography. J. Am. Statistic. Assoc. V. 86, No. 413, 55-67.