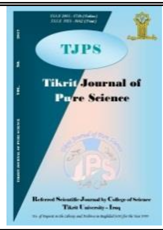
Tikrit Journal of Pure Science 23 (10) 2018 *ISSN: 1813 – 1662 (Print)*
E-ISSN: 2415 – 1726 (On Line)

TJPS

TIKRIT JOURNAL OF PURE SCIENCE

Journal Homepage: http://main.tu-jo.com/ojs/index.php/TJPS/index



APPLICATION TREND SURFACE MODELS WITH ESTIMATION

Shaymaa Riyadh Thanoon
College of Nursing , University of Mosul , Mosul , Iraq

ARTICLE INFO.

Article history:
-Received: 12 / 3 / 2018
-Accepted: 19 / 4 / 2018
-Available online: / / 2018

Keywords: parameters, residuals

Corresponding Author:
Name: **Shaymaa Riyadh**
E-mail:
Shaymaaalnuaimi@yahoo.com
Tel:

Abstract

This research deal with estimation of trend surface analysis and with spatial data with three models to real spatial data represents a rising ground water. The first method in assessment is to estimation trend surface model parameters by (ml), the second method requires decision on the maximum time difference to be calculated (s). while, the third method need a resolution, and the residuals r of dots to take for the conjecture the f(Dij) parameters. The first and second methods require resolution principle of "neighbor" from determines of "W". These three approaches are applied to real data which represent the ground water levels in 47 wells in mountain region in Sin jar district in Nineveh governorate.

Introduction

Trend –surface analysis is used for describing a certain style to represent real desktop data or highlights to represent changes that occur the spatial stochastic process. Base model is usually in the trend surface analysis in a way that ordinary least-squares and that assumes that the residuals are independent but if least-squares method is not efficient the standard errors and test morale be biased. The link between residuum can appear in the graphs, for example of groundwater location in a certain area . Statistical models are often used to represent views in terms of random variables and these models can be used to estimate the based on Probability theory. [1].

attribute actual data trend –surface using of duet dimension polynomial equation An observation at a dot i to enable to writs $Y_i = Y(u_i, v_i)$, where u and v represent a rectangular coordinate system for the area. Trend –surface analysis assume that Y_i consists of two components, a trend component T_i represented by the polynomial equation and a residual component (R_i) so that

$$Y_i = T_i + R_i$$

$$T(u,v) = \beta_{00} + \beta_{10}u + \beta_{01}v + \beta_{20}u^2 + \beta_{11}uv + \beta_{02}v^2 + \cdots + \beta_{pq}\mu^p v^q \quad \ldots\ldots(1)$$

Then term $p+q$ represents the order of the trend surface,first order or linear $p+q=1$ second order or quadratic $p+q=2$, third order or cubic $p+q=3$,and so on.the technique has been documented in a number of textbooks .

To a multidimensional research due to the finder [2], and circulated earth science by [3].

The basic model is often installed by the normal minimum squares (ols) assuming that residues separate, assumption residues are spatially autocorelated, the estimates of the lower squares ineffective ,the standard errors and important test of the biased. Auto correlated can be mistakenly generated by error in determining the order of the form (the auto correlated of the first-order shape residue will be linked real shape quadrature) or measurement error or both. However, of the localized effects of spatial operations that process on a medium scale between the regional trend component and the sit or residual component so that an attempting to force this element into the trend component of the model may be inappropriate or to give rise unnecessary complicated model. The study examines alternative methods for mounting trend-model to accommodate spatial autocorrelation.

Regression analysis

Defines the concept of a general regression mathematical representation for the average relationship between one variable and other variables called independent variables. Specialized Regression analysis to describe the relationship between variables in the form of have from that contains one independent variable is then called (simple linear regression model) write the following formula [4,5]

$$Y = \beta_0 + \beta_1 X + e$$

ISSN: 1813 – 1662 (Print)
E-ISSN: 2415 – 1726 (On Line)

Also, if the from contains several independent variables it is called (multiple linear regression model).General linear model takes a written form and takes more than independent variables and the formula be:

$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_m X_{im} + e_i$

$j = 1,2,3 \ldots, m \; ; \quad Y_i = \beta_0 + \sum_{j=1}^{m} \beta_j X_{ij} + e_i$

$i = 1,2,3 \ldots, n$

Using arrays can formulate the following general form . [6], [5].

The vector of observation, denoted by Y and matrix notation the model is

$Y = X\beta + e \quad (2)$

the matrix X is

$$X = \begin{bmatrix} 1 & u_1 & v_1 & u_1^2 & v_1^2 & u_1 v_1 & \ldots & u_1^P v_1^q \\ 1 & u_2 & v_2 & u_2^2 & v_2^2 & u_2 v_2 & \ldots & u_2^P v_2^q \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & u_n & v_n & u_n^2 & v_n^2 & u_n v_n & \ldots & u_n^P v_n^q \end{bmatrix}$$

and

$\beta' = [\beta_{00}, \beta_{10}, \beta_{01}, \beta_{20}, \beta_{02}, \beta_{11}, \ldots, \beta_{pq}]$

That $E(e) = 0$ and $E(ee') = \Omega$, That Ω matrix of nxn identical , [7].

In most practical applications Ω be unknown data shall be the standard case appreciation when independent data, to find a linear model parameters estimation indicated in equation (2) using(ols) follow: [8]

$\hat{\beta} = (X'X)^{-1}X'Y \quad (3)$

If there is a Spatial Autocorrelation between subjective random errors u_i because of random error in each site depends on nieghbouring sites and the error in this case we

$E(a_i a_j) \neq 0 \qquad \forall\, i \neq j$

From the simplest *Autocorrelation is* that appears in the *first order* autoregressive scheme . [8]

$a_i = \rho a_{i-1} + e_i$

That ρ autocorrelation between a_i and a_{i-1} and random error for this model takes the same assumptions used in some way (OLS) and $e_i \sim N(0, \sigma_e^2)$

The relationship is found by (4) $\quad \rho(h) = \frac{C(h)}{\sigma(i)\sigma(j)}$

That j=2,3,4,…….,n and i=1,2,3,…...,n-1

Not that $\sigma(i)$ and $\sigma(j)$ represent the standard deviation of the data and $C(h)$ represents the heterogeneity and which can be found from the relationship:

$$C(h) = \frac{1}{|N(h)|} \sum_{i,j=1}^{N(h)} \left(Y(s_i) - \bar{Y}(s_i)\right)\left(Y(s_j) - \bar{Y}(s_j)\right) \qquad (5)$$

$N(h) = \{(s_i, s_j): s_i - s_j = h \;\; ; \;\; i,j = 1,2,\cdots,n\}$

Application

application has on real data represent groundwater levels for 47 borehole in sin jar in Nineveh. As each watch its coordinates by location (S_i) and $Y(S_i)$ because staying i=1,2,3,…,47 in the exploration area and the data shown in table.

Appropriate height table Y meters and v, u coordinates for a region spend Sinjarin Nineveh

u_i	v_i	$Y(u_i, v_i)$	u_i	v_i	$Y(u_i, v_i)$
1.9	3.7	6.2	1.5	3.1	6.1
3.8	4.3	7.4	1.3	5.1	5.8
1.5	3.7	6.2	1.1	5.4	5.6
1.4	3.9	6.2	1.2	4.4	5.4
4.5	4.4	7.3	1.4	4.2	5.3
1.8	3.8	6.1	3.2	3.9	7.1
1.5	4	6.1	1.6	3.5	6.2
3	4	3.6	1.8	3.1	6.1
3	5.1	3.4	1.8	3.3	6.2
3.3	4.9	3.4	1.8	2.9	6.1
4.9	4.2	7.4	1.1	5.1	5.8
1.7	3.5	6.2	1.5	4.9	6.6
4.7	4	7.3	1.4	5.4	5.2
1.8	3.4	6.2	1.4	4.7	5.2
1.7	2.9	6.2	1.7	2.8	6
0.8	3.1	3.9	1.5	3	6
0.8	3.3	3.5	2	3.3	6.1
0.8	3	3.5	0.8	4.5	5.9
3.2	3.6	3.2	1.3	4.9	5.6
3	3.9	7.2	1.2	5.2	5.5
1.7	3.8	6.2	1	4.4	5.2
1.5	3.9	6.2	1	4.2	5.1
4.5	4.2	7.3	0.7	3.1	3.9
4	4.5	7.5			

Estimating the usual method model parameters

Model (1), rate first of the case where the directional orderly arrangement of the trend surface is known . [9] proposed a regression pattern for the regression models with the remaining spatial autocorrelation by following these steps; [10]

First Step. Appreciation β in (1) (call this $\hat{\beta}$ by (ols) assuming that the

Residuals is in depended

$\hat{\beta} = (X'X)^{-1}X'Y$

Second Step. Gain the (ols) residuals , $\hat{e} = Y - X\hat{\beta}$

Third Step. Use the residuals to achieve $\hat{\Omega}$ (or $\hat{\Omega}^{-1}$,if this can be gained immediately.

Fourth Step. Appreciation β by generalized (gls).

$\tilde{\boldsymbol{\beta}} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y \qquad (6)$

Fifth Step . get the (gls) residuals $\hat{e} = Y - X\hat{\beta}$,and return to third step. Iterate until convergence.

The critical step is three, because the specification of Ω could be approached in several different methods. I contrast a maximum likelihood (ml) approach (method 1) with two other approaches, direct estimation of the residual autocovariances (way 2) and estimation of an autocovariance function for the residue (way3).

Throughout, I shall write *e* for residuals, it being understood that at the first cycles the (ols) residuals $\hat{e}$ are used, but in subsequent cycles (gls) residuals $\hat{e}$ are used.

Method 1 : estimation maximum likelihood

[11]studied the ML estimators for spatial regression model parameters where errors are automatically linked. trend-surface shape trend is a special case of regression modeling so that a natural start. point is to study this approach.

The auto covariance among Y values is stated by a parametric model function, a position two values (usually only dependent space between sites) and a set of unknown parameters (θ). The autocovariances matrix must be positive, clear, and twice latent in relation to unknown parameters. Assuming a Gaussian process (ml) of β ,θ are achieved by maximize – likelihood

$$f(Y, L(\beta), L(\vartheta)) = \frac{|\Omega|^{-\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}(Y-X\beta)'\Omega^{-1}(Y-X\beta)}$$

$$L(\beta, \theta, Y) = -\frac{n}{2}\ln 2\pi - \frac{1}{2}\ln|\Omega| - \frac{1}{2}(Y - X\beta)'\Omega^{-1}(Y - X\beta)$$

Where $|\Omega|$ denotes determinat of . the second model is the mulataneous autoregressive scheme can be written as:

$e = \rho We + a$, where $a \sim N(0, \sigma^2 I)$ and W is an $n \times n$ connectivity or describing linkage matrix the proximity or neighbor relationship among the sites

i j is an neighbor of

$W_{ij} = 1$ if

if $i = j$ $W_{ij} = 0$

$W_{ij} = 0$ otherwise

The next step is to create a matrix Ω .

$$\Omega = \sigma^2[(I - \rho W')(I - \rho W)]^{-1}$$

$$\tilde{\boldsymbol{\beta}} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}Y$$

Then find the estimated values trend surface model $\tilde{Y}$

That, $\tilde{Y} = X\tilde{\boldsymbol{\beta}}$

Comparative selection coefficient adopted R^2 as well as maen square error MSE

$MSE = \frac{SSe}{n-k}$, that SSe= $Y'Y - (\tilde{\boldsymbol{\beta}})'X'Y$

$$R^2 = \frac{SSR}{SST}$$

$R^2 = 0.606$

$MSE = 1.9908$

To improve the previous models suggested following methods.

Method 2: lag order (autocovariance estimates).

residuals Autocovriance the makes direct estimated . Let C(s) indicate residuals Autocovriance at lag s. state

$$C(0) = \sum_{i=1}^{n} \frac{u_i^2}{n}$$

$$C(s) = \sum_{i,j} \frac{u_i u_j}{n(s)} \quad s = 1,2 \ldots, s$$

$\hat{\omega}_{ij} = c(s)$ if j is an s lag neighbor of i , $s = 1,2 \ldots, s$

$\hat{\omega}_{ij} = c(0)$ if i =, j

$\hat{\omega}_{ij} = 0$ otherwise

$$\widehat{\Omega_*} = \sigma^2[(I - \rho\hat{\omega})(I - \rho\hat{\omega})]^{-1}$$

$$\tilde{\boldsymbol{\beta}}_* = \left(X'\widehat{\Omega_*}^{-1}X\right)^{-1} X'\widehat{\Omega_*}^{-1}Y$$

[10] estimated several approaches of autocovariances.

Comparative selection coefficient adopted R^2 as well as mean square error MSE .

$MSE = \frac{SSe}{n-k}$, that SSe= $Y'Y - (\tilde{\boldsymbol{\beta}}_*)'X'Y$

$$R^2 = \frac{SSR}{SST}$$

$R^2 = 0.755$

$MSE = 1.001$

Method 3: autocovariance function

discussed [12] the Auto covariance function are parameters are estimated of a continuous in a different context He assumed that the structure of dependence among the residuals .

the function can be described by $u_i = F(D_{ij})u_i + e_i$, $F(D_{ij})$ refer to function of the distance among j ,i . It is more a summed , $E[e_i] = 0$

$$u_i u_j = F(D_{ij})u_j^2 + e_i e_j$$

providing u_i and e_i are independent, that

$$E[F(D_{ij})] = \frac{E[u_i u_j]}{E[u_j^2]} \equiv \rho_{ij}$$

Where (ρ_{ij}) is the autocorrelation separated among two sites by a distance $\mathbf{D_{ij}}$. Although this seems like a approach a reasonable, e and u processes cannot be independent , and in general the expected value will be an underestimate of the true autocorrelation function.

Suggested approximating the distance function by the expression, $F_A(D_{ij}) = A + BD_{ij} + CD_{ij}^2$ and discuss an (ols) procedure for estimated the parameters A,B, and C . f A=1, then

$F_A(D_{ij}) = 1$ if $D_{ij} = 0$

Procedure so that $F_1(D_{ij})$ can be interpreted as an autocorrelation function.

depends the way to define constant radius (r) around each point within which pairs of location captivated for the interpret of A, B, and C. past interim autocorrelation is equal 0.

r < If D_{ij} $\hat{F}(D_{ij})$ $\hat{w}_{ij} =$

0 otherwise.= $\hat{w}_{ij}$

The way, the previous way unlike, allows the values in Ω to be adgusted for distance betwwen location, it may be suitable for the case of irregular site distributions. For this reason. For this reason.

I representing evaluating. When you make up $\hat{w}_{ij}$ in equation

$$\sigma^2[(I - \rho\hat{w})(I - \rho\hat{w})]^{-1} = \Omega_{**}$$

estimate (gls) get by

$$\tilde{\boldsymbol{\beta}}_{**} = \left(X'\,\widehat{\Omega_{**}}^{-1}X\right)^{-1} X'\widehat{\Omega_{**}}^{-1}Y$$

Comparative selection coefficient adopted R^2 as well as mean square error MSE .

$MSE = \frac{SSe}{n-k}$, that SSe= $Y'Y - (\tilde{\boldsymbol{\beta}}_{**})'X'Y$

$$R^2 = \frac{SSR}{SST}$$

$R^2 = 0.9222$

$MSE = 0.8921$

methods comparing the used for error rate previously roads note that I said error rate significantly this proof that estimate of variance components and distance function approximation contributed significantly to reduce error and improve the model.

The following table shows the comparison between some values of truth and estimated values.

$Y(s)$	Method 1	Method 2	Method 3
6.2	3.8	6.1	6.5
7.4	7.9	5.8	7.9
6.2	4.2	5.6	6.1
6.2	5.2	5.4	6.4
7.3	7.7	5.3	6
6.1	6.1	7.1	6.2
6.1	5	6.2	6.2
3.6	2.6	6.1	3.1
3.4	1.4	6.2	3.2
3.4	2.4	6.1	6.1
7.4	6.1	5.8	7.1
6.2	4	6.6	6.9
7.3	5.3	5.2	7.2
6.2	4.2	5.2	6.2
6.2	6.9	6	6.1
3.9	3.1	6	4.2
3.5	2.5	6.1	3
3.5	3.1	5.9	3
3.2	3.9	5.6	3.6
7.2	4.2	5.5	6.7
6.2	6.9	5.2	6.2
6.2	6.1	5.1	6.1
7.3	5.3	3.9	7.1
7.5	8.5		

The table represents a standard comparison of methods by the coefficient of determination and mean square error.

ESTIMATION	METHOD 1	METHOD 2	METHOD 3
R^2	0.606	0.755	0.9222
MSE	1.9908	1.001	0.8921

Conclusion

The present study has been considered as a repeat pattern to estimating the parameters of the trend surface models with residuals autocorrelation. The procedures discussed here would seem appropriate in in those situations in which three levels of variation are suspects in the data, but they may also be useful in avoiding order misspecification trend-surface analysis. The importance of fine tests does not seem to be available, but sufficient information can be taken from the variance of the common contrast matrix to design approximate tests. Of conclusion supply, the (ml) way to be the best, Method 3 works well, and special suitable irregular Method 2 requires the evaluation of multiple autocorrelation requests and shows the progress of the method 1 in the current context, however, perhaps to a lot of autocorrelation in the data sites, from all this we can conclude that it is better on the third method way depending on the comparison criteria.

References

1- Donald, W.; Meals, J. S., Steven, A. Dressing, and Jon, B. (2011). Statistical analysis for monotonic trends, Tech Notes 6, November 2011. Developed for U.S. *Environmental Protection Agency by Tetra Tech, Inc., Fairfax*, **23**.

2- Journel, A.G. and Huijbregts, C.J. (1978). Mining Geostatistics. *New York Academic Press*, **600**.

3- Krumbein, W. C. (1956). Regional and local components in facies maps. *Bulletin of the American Association of Petroleum Geologists*, **40(9)**: 2163-2194.

4- Chatterjee, S. and Hadi, A. S. (2015). Regression analysis by example. John Wiley & Sons.

5- Draper, N. R., & Smith, H. (1981). Applied Regression Analysis 2nd ed Wiley. New York, NY.

6- Freedman, D. A. (2009). Statistical models: theory and practice. cambridge university press.

7- Webster, R., & Oliver, M. A. (2007). Geostatistics for environmental scientists. John Wiley & Sons.

8- Haining, R. P. and Haining, R. (2003). Spatial data analysis: theory and practice. Cambridge University Press.

9- Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, **70(349)**:120-126.

10-Ripley, B. D. (1981). Spatial Statistics. Wiley, New York.

11- Mardia, K. V. and Marshall, R. J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. Biometrika, **71(1)**: 135-146.

12-Agterberg, F.P. (1970), Autocorrelation Function in geology, in geostatistic. A colloquium, ed. D.F Merriam, new york: plenum press, 113-141.

تطبيق نماذج سطح الاتجاه مع التقدير

شيماء رياض ذنون

كلية التمريض ، جامعة الموصل ، الموصل ، العراق

الملخص

يتناول البحث تحليل سطح الاتجاه للبيانات المكانية مع تطبيق ثلاث نماذج على بيانات مكانية حقيقية تمثل ارتفاع مناسيب المياه الجوفية . تتمثل الطريقة الاولى لتقدير معلمات نموذج سطح الاتجاه بواسطة (ml), الطريقة الثانية تتطلب اتخاذ قرار بشان الفارق الزمني الاعلى الذي يتم حسابه ب (s). وفي نفس الوقت تحتاج الطريقة الثالثة الى حل وحساب البواقي r التي تاخذ لتقدير معلمات الدالة f(Dij). تتطلب الطريقة الاولى والثانية مبدأ "الجوار" الذي تمثل ب "W" . تم التطبيق على بيانات حقيقية تمثل ارتفاع مناسيب المياه الجوفية في 47 بئرا مع احداثيات مواقعها في قضاء سنجار في محافظة نينوى.