# A Study on Seemingly Unrelated Model of Forest Ecological Benefit

Bi Lei[1] , Li Changsheng[2], Xue Yu[3]

[1]School of Municipal and Environmental Engineering, Harbin Institute of Technology, China
[2]School of Humanities and Social Sciences, Harbin Institute of Technology, China
[3]Correspondence author, School of Forestry, Northeast Forestry University, China
Email - xueyuhenry@126.com, bl9999456@126.com

**Abstract:** Based on a modern statistical model, which is seemingly unrelated model, the paper defines the dependent and the independent variable set of forest ecological benefit, and then formulates seemingly unrelated model of forest ecological benefit, and finally devises three methods to solve least square estimation of seemingly unrelated model, GM estimation of parameter when covariance matrix $\sum$ is given, and GM estimation of parameter $\beta$ when covariance matrix $\sum$ is unknown.

[Bi Lei, Li Changsheng, Xue Yu. **A Study on Seemingly Unrelated Model of Forest Ecological Benefit.** *World Rural Observations* 2014; 6(2):111-115]. ISSN: 1944-6543 (Print); ISSN: 1944-6551 (Online). http://www.sciencepub.net/rural. 17

**Key Words:** Forest Ecological Benefit; Seemingly Unrelated Model; Solution

## 1. Seemingly Unrelated Model

Suppose a set of random variable $Y_1, \cdots, Y_q$ in physical quantity of forest ecological benefit meets the linear statistical relation, that is:

$$\begin{cases} Y_{1j} = x_{11}^j \beta_{1j} + x_{12}^j \beta_{2j} + \ldots\ldots + x_{1,pj}^j \beta_{pj,j} + e_{1j} \\ Y_{2j} = x_{21}^j \beta_{1j} + x_{22}^j \beta_{2j} + \ldots\ldots + x_{2,pj}^j \beta_{pj,j} + e_{2j} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ Y_{nj} = x_{n1}^j \beta_{1j} + x_{n2}^j \beta_{2j} + \ldots\ldots + x_{n,pj}^j \beta_{pj,j} + e_{nj} \end{cases}, \quad j=1,2,\ldots\ldots q \qquad (1)$$

which in equation (1) $Y_{ij}$ is the jth ecological benefits and ith samples for observation value of the variable containing observed values with errors. $n_j$ is the sample number and generally each dependent variable sample observed value $Y_{ij}$ is independent.

$x_{11}^j$ is the j'th ecological benefit observed value of the first sample and the first independent variable; $x_{12}^j$ is the j'th ecological benefit observed value of the first sample and the second independent variable;... $x_{1,pj}^j$ is the j'th ecological benefit observed value of the first sample and the $p_{jth}$ independent variable; $x_{n,pj}^j$ is the j'th ecological benefit observed value of the $n_j$th sample and the $p_{jth}$ independent variable. $P_j$ is a variable number.

$\beta_{1j}, \beta_{2j}, \ldots\ldots \beta_{pj,p}$ is jth estimated parameters of ecological benefit.

$e_{1j}, e_{2j}, \ldots\ldots e_{nj}$ is the jth of ecological benefit of observational errors of dependent variable from sample number 1, 2 to $n_j$. Matrix form of which is:

$$Y_j = \underset{1\times p_j}{\mathbf{x}^{(j)}} \underset{p_j\times 1}{\beta^{(j)}} + e_j , \quad 1 \le j \le q \quad \text{(Variables)} \qquad (2)$$

What's worth noticing is that in the above q relations, seemingly they are independent, but the dependent variable set could be different and includes meaning and the amount of variables. In the following two observational data matrix $\mathbf{x}^{(j)}$ and $\mathbf{x}^{(l)}$ of forest ecological benefit, they could be not the same (variables are respectively $p_j$ and $p_l$).

$$\mathbf{x}^{(j)} = \begin{pmatrix} x_{11}^j & x_{12}^j & \cdots & x_{1pj}^j \\ x_{21}^j & x_{22}^j & \cdots & x_{2pj}^j \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}^j & x_{n2}^j & \cdots & x_{npj}^j \end{pmatrix} \text{ and } \mathbf{x}^{(l)} = \begin{pmatrix} x_{11}^j & x_{12}^j & \cdots & x_{1pl}^j \\ x_{21}^j & x_{22}^j & \cdots & x_{2pl}^j \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}^j & x_{n2}^j & \cdots & x_{npl}^j \end{pmatrix}$$

In the above equations, n is just the sample number. Put the variables in (2) into:

$$\mathbf{Y}_j = \begin{pmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{nj} \end{pmatrix}, \quad \mathbf{x}_j = \begin{pmatrix} \mathbf{x}_1^{(j)} \\ \mathbf{x}_2^{(j)} \\ \vdots \\ \mathbf{x}_n^{(j)} \end{pmatrix}, \quad \mathbf{e}_j = \begin{pmatrix} e_{1j} \\ e_{2j} \\ \vdots \\ e_{nj} \end{pmatrix}$$

And we can get (3), which equals to (2).

$$\mathbf{Y}_j = \mathbf{x}_j \beta^{(j)} + \mathbf{e}_j, \quad 1 \le j \le q \qquad (3)$$

Generally, we can suppose the average value of error matrix $\underset{n\times q}{\mathbf{e}} = (\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_q)$ is zero and it is irrelevant (independent) between lines, which has the same covariance matrix, that is:

$$\mathbf{E}(\mathbf{e}) = \mathbf{0}, \quad \mathrm{cov}(\mathbf{e}_{i\bullet}) = \sigma^2 \underset{q\times q}{\Sigma},$$
$$\mathrm{cov}(\mathbf{e}_{i\bullet}, \mathbf{e}_{j\bullet}) = \underset{q\times q}{0}, \quad 1 \le i \ne j \le n \qquad (4)$$

From (3) and (4), we can get the following linear model:

$$\begin{cases} \underset{n\times 1}{\mathbf{Y}_j} = \underset{n\times p_j}{\mathbf{x}_j} \underset{p_j\times 1}{\beta^{(j)}} + \mathbf{e}_{\bullet j} \\ \mathbf{E}(\mathbf{e}_{\bullet j}) = \mathbf{0} \\ \mathrm{cov}(\mathbf{e}_{i\bullet}) = \sigma^2 \underset{q\times q}{\Sigma} \\ \mathrm{cov}(\mathbf{e}_{i\bullet}, \mathbf{e}_{i'\bullet}) = \underset{q\times q}{\mathbf{0}} \end{cases}, \quad 1 \le i \ne i' \le n, \quad 1 \le j \le q \qquad (5)$$

Where: $\mathbf{e}_{i\bullet} = (e_{i1}, \cdots, e_{iq})$, we can call (5) seemingly unrelated linear model (Tang et al, 2009).

## 2. Variable Set

According to the research of Li et al (2006), and Lang et al (2000), dependent and independent variable set of forest ecological benefit were selected.

### 2.1 Dependent Variable of Forest Ecological Benefit (random variable)

y1—Annual ecological benefit of water intercepted by forest canopy ($t/hm^2.a$).

$y_2$—Annual water-retaining capacity of forest litter layer ($t/hm^2.a$)
$y_3$—Annual Ecological benefit of water reserve in non-capillarypore forest soil ($t/hm^2.a$)
$y_4$--Annual forest ecological benefit of fixed soil ($t/hm^2.a$)
$y_5$--Annual forest ecological benefit of conserved fertilizer ($t/hm^2.a$)
$y_6$--Annual forest ecological benefit of absorbed $CO_2$ ($t/hm^2.a$)
$y_7$--Annual forest ecological benefit of releasing $O_2$ ($t/hm^2.a$)
$y_8$--Annual forest ecological benefit of wind-break and sand-fixation ($hm^2/hm^2.a$)

### 2.2 Independent Variable Set of Forest Ecological Benefit

$F_1$ is response value of coniferous forest; $F_2$ is response value of broad leaved forest; $F_3$ is response value of mixed forest. Parameters are $b_{1j}, b_{2j}, b_{3j}$, j=1,2,3.

$Z_1$ is the response value of young age groups; $Z_2$ is the response value of middle age groups; $Z_3$ is the response value of mature age groups. Parameters are $c_{1j}, c_{2j}, c_{3j}$, j=1,2.

JY represents average annual rainfall with parameters of $d_{1j}$ ,j=1,4,5,6,7,8;

YB represents canopy density with parameters of $d_{2j}$ ,j=1,4,5,6,7,8;
JD represents longitude with parameters of $d_{3j}$ ,j=2,3,4,5,6,7,8;
WD represents latitude with parameters of $d_{4j}$,j=2,3,4,5,6,7,8;
HB represents altitude $d_{53}$。

### 3. Seemingly Unrelated Model of Forest Ecological Benefit

$$
\begin{aligned}
\ln(y_{i1}) &= a_{11}+b_{11}F_{i1} + b_{21}F_{i2} + b_{31}F_{i3} + c_{11}Z_{i1} + c_{21}Z_{i2} + c_{31}Z_{i3} + d_{11}\times\ln(JY_i) + d_{21}YB_i + e_{i1} \\
y_{i3} &= a_{13}+b_{13}F_{i1} + b_{23}F_{i2} + b_{33}F_{i3} + d_{33}JD_i + d_{43}WD_i + d_{53}HB_i + e_{i3} \\
y_{i4} &= a_{14}+c_{14}Z_{i1} + c_{24}Z_{i2} + c_{34}Z_{i3} + d_{14}JY_i + d_{24}YB_i + d_{34}JD_i + d_{44}WD_i + e_{i4} \\
y_{i5} &= a_{15}+c_{15}Z_{i1} + c_{25}Z_{i2} + c_{35}Z_{i3} + d_{15}JY_i + d_{25}YB_i + d_{35}JD_i + d_{45}WD_i + e_{i5} \qquad (6) \\
y_{i6} &= a_{16}+c_{16}Z_{i1} + c_{26}Z_{i2} + c_{36}Z_{i3} + d_{16}JY_i + d_{26}YB_i + d_{36}JD_i + d_{46}WD_i + e_{i6} \\
y_{i7} &= a_{17}+c_{17}Z_{i1} + c_{27}Z_{i2} + c_{37}Z_{i3} + d_{17}JY_i + d_{27}YB_i + d_{37}JD_i + d_{47}WD_i + e_{i7} \\
y_{i8} &= a_{18}+c_{18}Z_{i1} + c_{28}Z_{i2} + c_{38}Z_{i3} + d_{18}JY_i + d_{28}YB_i + d_{38}JD_i + d_{48}WD_i + e_{i8}
\end{aligned}
$$

If it is written by standard form of seeming unrelated model are as follows:

$$
\begin{aligned}
y_{i1} &= \beta_{11}+\beta_{21}X_{i1}+\beta_{31}X_{i2}+\beta_{41}X_{i3}+\beta_{51}X_{i4}+\beta_{61}X_{i5}+\beta_{71}X_{i6}+\beta_{81}X_{i7}+\beta_{91}X_{i8}+e_{i1} \\
y_{i2} &= \beta_{12}+\beta_{22}X_{i1}+\beta_{32}X_{i2}+\beta_{42}X_{i3}+\beta_{52}X_{i4}+\beta_{62}X_{i5}+\beta_{72}X_{i6}+\beta_{10,2}X_{i9}+\beta_{11,2}X_{I,10}+\beta_{12,2}X_{I,11}+e_{i2} \\
y_{i3} &= \beta_{13}+\beta_{23}X_{i1}+\beta_{33}X_{i2}+\beta_{43}X_{i3}+\beta_{10,3}X_{i9}+\beta_{11,3}X_{i,10}+\beta_{12,3}X_{i,11}+e_{i3} \\
y_{i4} &= \beta_{14}+\beta_{54}X_{i4}+\beta_{64}X_{i5}+\beta_{74}X_{i6}+\beta_{8,4}X_{i7}+\beta_{9,4}YB_{i8}+\beta_{10,4}X_{i9}+\beta_{11,4}X_{i,10}+e_{i4} \\
y_{i5} &= \beta_{15}+\beta_{55}X_{i4}+\beta_{65}X_{i5}+\beta_{75}X_{i6}+\beta_{8,5}X_{i7}+\beta_{9,5}YB_{i8}+\beta_{10,5}X_{i9}+\beta_{11,5}X_{i,10}+e_{i5} \\
y_{i6} &= \beta_{16}+\beta_{56}X_{i4}+\beta_{66}X_{i5}+\beta_{76}X_{i6}+\beta_{8,6}X_{i7}+\beta_{9,6}YB_{i8}+\beta_{10,6}X_{i9}+\beta_{11,6}X_{i,10}+e_{i6} \\
y_{i7} &= \beta_{17}+\beta_{57}X_{i4}+\beta_{67}X_{i5}+\beta_{77}X_{i6}+\beta_{8,7}X_{i7}+\beta_{9,7}YB_{i8}+\beta_{10,7}X_{i9}+\beta_{11,7}X_{i,10}+e_{i7} \\
y_{i8} &= \beta_{18}+\beta_{58}X_{i4}+\beta_{68}X_{i5}+\beta_{78}X_{i6}+\beta_{8,8}X_{i7}+\beta_{9,8}YB_{i8}+\beta_{10,8}X_{i9}+\beta_{11,8}X_{i,10}+e_{i8}
\end{aligned}
$$

Note: In the above equation, the second subscript is the number of random variable and the first is the number of the variable. And the seemingly unrelated model can be changed into generalized one-dimensional linear model:

$$
\underset{nq\times 1}{\mathbf{Y}} = \begin{pmatrix} \underset{n\times 1}{\mathbf{Y}_1} \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_8 \end{pmatrix} \qquad
\mathbf{x} = \begin{pmatrix} \underset{n\times p1}{\mathbf{x}_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \underset{n\times p2}{\mathbf{x}_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \underset{n\times pq}{\mathbf{x}_8} \end{pmatrix}
$$

$$
\mathbf{x}_1 = \begin{pmatrix} 1, X_{1,1}, X_{1,2}, X_{1,3}, X_{1,4}, X_{1,5}, X_{1,6}, X_{1,7}, X_{1,8} \\ 1, X_{2,1}, X_{2,2}, X_{2,3}, X_{2,4}, X_{2,5}, X_{2,6}, X_{2,7}, X_{2,8} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 1, X_{n,1}, X_{n,2}, X_{n,3}, X_{n,4}, X_{n,5}, X_{n,6}, X_{n,7}, X_{n,8} \end{pmatrix}
$$

$$
\mathbf{x}_2 = \begin{pmatrix} 1, X_{1,1}, X_{1,2}, X_{1,3}, X_{1,4}, X_{1,5}, X_{1,6}, X_{1,10}, X_{1,11}, X_{n,12} \\ 1, X_{2,1}, X_{2,2}, X_{2,3}, X_{2,4}, X_{2,5}, X_{2,6}, X_{2,10}, X_{2,11}, X_{n,12} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 1, X_{n,1}, X_{n,2}, X_{n,3}, X_{n,4}, X_{n,5}, X_{n,6}, X_{n,10}, X_{n,11}, X_{n,12} \end{pmatrix}
$$

$$\mathbf{x}_3 = \begin{pmatrix} 1, X_{1,1}, X_{1,2}, X_{1,3}, X_{1,9}, X_{1,10}, X_{1,11} \\ 1, X_{2,1}, X_{2,2}, X_{2,3}, X_{2,9}, X_{2,10}, X_{2,11} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 1, X_{n,1}, X_{n,2}, X_{n,3}, X_{n,9}, X_{n,10}, X_{n,11} \end{pmatrix}$$

$$\mathbf{x}_4 = \begin{pmatrix} 1, X_{1,4}, X_{1,5}, X_{1,6}, X_{1,7}, X_{1,8}, X_{1,9}, X_{1,10} \\ 1, X_{1,4}, X_{2,5}, X_{2,6}, X_{2,7}, X_{2,8}, X_{2,9}, X_{2,10} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 1, X_{n,4}, X_{n,5}, X_{n,6}, X_{n,7}, X_{n,8}, X_{n,9}, X_{n,10} \end{pmatrix},$$

Where: $x_5,x_6,x_7,x_8$ is the same as $x_4$.

$$\beta = \begin{pmatrix} \beta^{(1)} \\ \beta^{(2)} \\ \vdots \\ \beta^{(8)} \end{pmatrix} \qquad e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_8 \end{pmatrix} \qquad (7)$$

And can be put as the generalized one-dimensional linear model standard form of (8).

$$\begin{cases} \underset{(nq)\times 1}{\mathbf{Y}} = \underset{(nq)\times pq}{\mathbf{x}} \underset{p\times 1}{\beta} + \underset{nq\times 1}{e} \\ \mathbf{E}(\mathbf{e}) = \mathbf{0} \\ \operatorname{cov}(\mathbf{e}) = \sigma^2 \underset{q\times q}{\Sigma} \otimes I_n \end{cases} \qquad (8)$$

The number of columns of design matrix $x$ is $p = \sum_{j=1}^{q} p_j$ (equivalent parameter vector dimension of $\beta$), namely $p$ is the sum of the dimensions of $q$ 'th parameter vector $\beta^{(j)}$, $1 \le j \le q$.

## 4. Solutions

Seemingly unrelated model is a special generalized one-dimensional liner model, so its parameter estimation should be respectively used least squares estimate and GM estimate or proposed GM estimate.

### 4.1 Least Squares Estimate of Seemingly Unrelated Model

Under the condition of simplified one-dimensional linear model (when ∑=I), the estimate of generalized one-dimensional liner model was called ordinary least squares estimate. Apparently, it is a unbiased and accordant estimate (If adding the observation frequency n of index vectors, we can advance the precision of the parameter estimation.), however, it is still not the optimal estimate for not considering ∑.

$$\hat{\beta} = (\mathbf{x}'\mathbf{x})^{+}\mathbf{x}'\mathbf{Y} \qquad (9)$$

Or we could use

$$\hat{\beta}^{(j)} = (\mathbf{x}'_j\mathbf{x}'_j)^{+}\mathbf{x}_j\mathbf{Y}_j \qquad (10)$$

as estimation of $\beta$. The equation (10) equals to a mathematical regression.

### 4.2 When Covariance Matrix $\Sigma$ is Known, the GM Estimate of Parameter $\beta$

According to theories of generalized linear models, if we use $\Sigma \otimes I_n$ to replace $\mathbf{V}$, the GM estimate of parameter $\beta$ in seemingly unrelated model could be got.

$$\tilde{\beta} = \left(\mathbf{x}'\left(\Sigma^{-1} \otimes I_n\right)\mathbf{x}\right)^{+}\mathbf{x}'\left(\Sigma^{-1} \otimes I_n\right)\mathbf{Y} \qquad (11)$$

When design matrix $\mathbf{x}$ is full rank of column, the GM estimate of parameter $\tilde{\beta}$ is best linear unbiased estimator of parameter $\beta$ and it is still mutually compatible. In the same way, the mean value of the GM estimate of parameter $\tilde{\beta}$ and covariance matrix could be got.

### 4.3 When Covariance Matrix $\Sigma$ is Unknown, the GM Estimate of Parameter $\beta$

When covariance matrix $\Sigma$ is unknown, firstly we need to estimate covariance matrix and then can get the proposed GM estimate of parameter $\tilde{\beta}$ from $\beta$. The two-step regression method is as followed:

1. Calculate the ordinary least squares estimate $\hat{\beta}^{(j)}$ of $\beta^{(j)}$.

2. Calculate $\hat{\sigma}_{ij} = \frac{\left(\mathbf{Y}_i - \mathbf{x}_i\hat{\beta}^{(i)}\right)'\left(\mathbf{Y}_j - \mathbf{x}_j\hat{\beta}^{(j)}\right)}{n - \overline{p}}$ and $\overline{p} = \frac{1}{q}\sum_{i=1}^{q} p_i$ is the dimension of variable X.

3. Mark $\hat{\Sigma} = \left(\hat{\sigma}_{ij}\right)_{q\times q}$, and then ordinary least squares estimate of error structure matrix of $\Sigma \otimes I_n$ is $\hat{\Sigma} \otimes I_n$.

4. The proposed GM estimate of parameter $\beta$ is $\tilde{\tilde{\beta}} = \left(\mathbf{x}'\left(\hat{\Sigma}^{-1} \otimes I_n\right)\mathbf{x}\right)^{+}\mathbf{x}'\left(\hat{\Sigma}^{-1} \otimes I_n\right)\mathbf{Y}$.

Where: $\tilde{\tilde{\beta}}$ is seemingly unrelated estimator and can be calculated by iterative method.

The above estimate parameters can be calculated by two-step curve fitting method, and be realized by Mat lab program.

6/22/2014

**Reference**

[1] Tang Shouzheng, Lang Kuijian, Li Haikui. 2009. *Computation of Statistics and Biological Mathematics Model* [M]. Beijing: Science Pres, p197-200

[2] Li Junqing, Niu Shukui. 2006. *Forest Ecology*[M]. Beijing: Higher Education Press, p468-469

[3] Lang Kuijian, Li Changsheng, Yin You, et al. 2000. *The Measurement Theory and method of 10 Forest Ecological Benefits for Forestry Ecological Engineering* [J]. Journal of Northeast Forestry University, (1): p1-8.