A MODEL OF DEFORESTATION WITH SPATIAL DEPENDENCY BY HUMAN POPULATION GROWTH*

SHOJIRO TANAKA† AND RYUEI NISHII‡

Abstract. Deforestation is a result of complex causality chains in most cases. But identification of limited number of factors shall provide comprehensive general understanding of the vital phenomenon at a broad scale, as well as projection for the future. Only two factors – human population and relief energy (difference of minimum altitude from the maximum in a sampled area) – were found to give sufficient elucidation of deforestation by a model, whose functional form was verified by linear combinations of dummy variables. Likelihood with spatial dependency was derived and applied to real data of one-kilometer spatial resolution, with which our model showed the best relative appropriateness.

Key words. AIC, environment model, relief energy, remotely-sensed information

1. Introduction.

Terabyte-scale information from multispectral sensors and microwave radiometers is rapidly increasing through remotely-sensed measurements, such as ozone concentration, water vapor, soil moisture as well as traditional land-cover and surface temperature. Yet interrelations of such information with other data are insufficiently studied in terms of mathematical models of socio-economic causality to terrestrial surfaces [4].

Increases of cultivated land for cash crops, grazing of cattle, shifting cultivation, logging, and fuel-wood requirements in developing countries as well as airborne pollutants and acid rain in developed nations should be major driving force for deforestation [5]. That means almost all the causes are strongly related to human activities [4].

In this research, we incorporated the typical human factor of population into forest coverage ratio by applying grid-cell data. We assume that the process of deforestation has strong dependency on human population in the same area.

We consider one-kilometer square area on the earth. Let N be the population, and R be the relief energy. Here, R denotes difference of minimum altitude from the maximum. Also let $F \equiv F(N, R)$ be the forest areal rate which includes open forest ($0 \leq F \leq 1$). In [8], we discussed the relation between F and N observed at areas with small R not greater than 20 meters in Japan. Based on 107 samples, we selected the following non-linear regression model:

$$F(N, R) = \gamma \exp\left(-\alpha N^2\right) + e, \quad (0 \leq R < 20) \tag{1.1}$$

where e is an independent error, α is a positive deforestation coefficient, and γ is a positive constant. See Appendix for the theoretical derivation of (1.1) by a differential equation. Forest relative reduction rate is assumed to increase in a certain proportion to the human population growth in an area.

Our previous research, however, has at least two drawbacks:

- Since the model (1.1) is valid only for regions with small relief energy R, an extended model applicable for any values R is required. It is known that

*日本語タイトル：「人口増による森林減少の空間モデル」, 田中 章司郎・西井 龍映

†Department of Mathematics and Computer Science, Faculty of Science and Engineering, Shimane University, 1060 Nishikawatsu, Matsue, 690-8504, Japan (tanaka@cis.shimane-u.ac.jp).

‡Department of Mathematics and Information Science, Faculty of Integrated Arts and Sciences, Hiroshima University, 1-7-1 Kagamiyama, Higashi-Hiroshima, 739-8521, Japan (nishii@mis.hiroshima-u.ac.jp).

Table 2.1
Basic statistics of two datasets.

Variable	Mean	Minimum value	Maximum value	Standard deviation
Dataset I ($0 \leq F \leq 1$, sample size: 8697)				
F	0.7730	0.000	1.000	0.2643
N	334.0	0.000	55,050	1,367
R	154.5	0.000	580.0	86.11
Dataset II ($0 < F \leq 1$, sample size: 8538)				
F	0.7874	1.138×10^{-4}	1.000	0.2446
N	288.8	0.000	22,070	1,091
R	157.1	0.000	580.0	84.52

geological features such as elevation and slope of a region are more important than climate variables [3].

- Data were assumed to be spatially independent. Spatial correlation should be incorporated in the extended model.

For discussing these points, we suppose that the forest areal rate $F(N, R)$ is expressed by the following additive model:

$$F(N, R) = g(N) + h(R) + e \tag{1.2}$$

where the errors e are spatially correlated. The spatial model (1.2) is an extension of the model (1.1).

2. Identification of Trend Functions with Independent Errors.

Figure 2.1 illustrates a location of our test field, and it also shows spatial maps of the variables. We use two datasets. One named Dataset I has all range of F, whereas the other called Dataset II omits data with $F = 0$. The reason for this is that areas with no forests in the test field are either very urbanized, or large agricultural fields including cattle ranches. The land-use types are known to cause anomaly in analysis since the development patterns are different in terms of deforestation [7]. Table 2.1 gives the basic statistics of the two datasets. For the preparation of them, see also [7].

2.1. Estimation of Functional Forms.

The regression model (1.1) holds only for data with small R. Our datasets include the data with large R, so the estimation procedure of parameters in the model (1.1) hardly converges. This implies that the effect due to R in the model (1.2) is not negligible. Thus we need to specify the functional form of $h(R)$. We first approximate $h(R)$ by a step function, say $h_0(R)$, over 16 intervals $r_0 = 0 < r_1 = 20 < r_2 = 40 < r_3 = 60 < \cdots < r_{15} = 300 < r_{16} = \infty$, i.e.,

$$h_0(R) = \omega_i \quad \text{if} \quad r_{i-1} \leq R < r_i \quad \text{for} \quad i = 1, 2, \cdots, 16. \tag{2.1}$$

Here, ω_1 is set to zero as a baseline. By using the regression model:

$$F(N, R) = \gamma \exp\left(-\alpha N^{\beta}\right) + h_0(R) + e \tag{2.2}$$

for Dataset I, we get a preliminary functional form of $h(R)$ shown in Fig.2.2. The figure strongly indicates that forest rate increases by a logarithmic function of relief

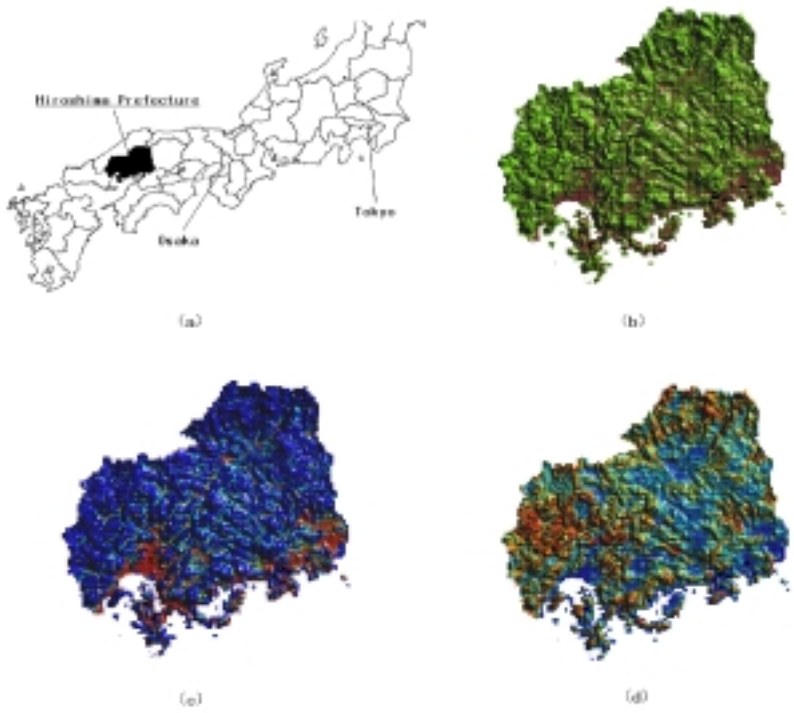

Fig. 2.1. Location of test field and maps of variables: (a) test field, (b) forest coverage ratio F, (c) population density N, (d) relief energy R. In each map from (b) to (d), altitude is overlaid with corresponding variables. (Originally in color)

energy R. Hence, we fit the parametric functions for $h(R)$:

$$h_1(R) = \begin{cases} \delta \log(R - \theta + 1) & \text{if } R > \theta \\ 0 & \text{if } 0 \le R \le \theta \end{cases} \tag{2.3}$$

$$h_2(R) = \begin{cases} \delta \log(R/\theta) & \text{if } R > \theta \\ 0 & \text{if } 0 \le R \le \theta \end{cases} \tag{2.4}$$

For the effect $g(N)$ due to population N, we will confirm the relation (1.1) for the Dataset I. We similarly approximate $g(N)$ by a step function, say $g_0(N)$, over 9 intervals $0 = n_0 < n_1 < n_2 < \cdots < n_9 = \infty$, i.e.,

$$g_0(N) = \psi_i \quad \text{if} \quad n_{i-1} \le N < n_i \quad \text{for} \quad i = 1, 2, \cdots, 9 \quad \text{with} \quad \psi_1 = 0. \tag{2.5}$$

Actually, eight terminal points n_i are determined as follows: $\log(n_1 + 1)$, $\log(n_2 + 1)$, $\ldots$, $\log(n_8 + 1)$ are respectively set by 3.0, 4.0, 4.5, 5.0, 5.5, 6.0, 7.0, 8.0 for taking balances of sample sizes in the intervals into account.

By fitting the regression model:

$$F(N, R) = g_0(N) + h_2(R) + e \tag{2.6}$$

to Dataset I, $g(N)$ is roughly estimated by the step function shown in Fig.2.3. This figure shows that the functional form of $g(N)$ is close to $g_2(N)$ which was already

selected in regions with small R.

Thus, the following candidate models for $g(N)$ are examined:

$$(2.7)\quad g_1(N) = \gamma \exp(-\alpha N),\ g_2(N) = \gamma \exp\left(-\alpha N^2\right),\ g_\beta(N) = \gamma \exp\left(-\alpha N^\beta\right).$$

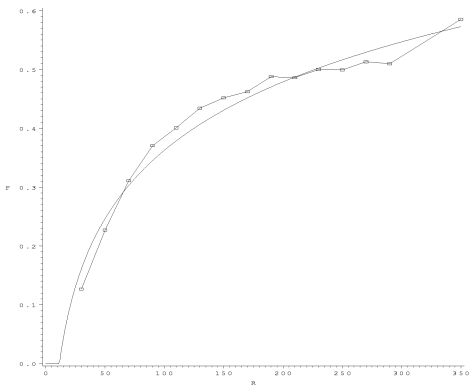

Fig.2.2. Estimated step function $h_\theta(R)$ for Dataset I with $(R, h_\theta(R))$. Estimated values are marked as square, and smooth line indicates the estimated function $h_2(R)$ based on the model $E[F] = g_\beta(N) + h_2(R)$.

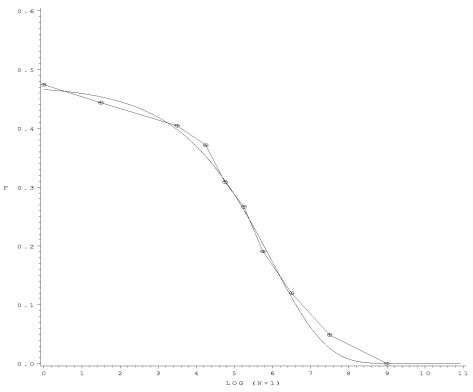

Fig.2.3. Estimated step function $g_\theta(N)$ for Dataset I with $(\log(N+1), g_\theta(N))$. Estimated values are marked as cross in a circle, and smooth line indicates the estimated function $g_\beta(N)$ based on the model $E[F] = g_\beta(N) + h_2(R)$.

2.2. Model Selection under Independent Assumption. Ordinary non-linear parameter estimations are employed first. The initial values are carefully determined by grid search.

Many criteria to select the best model in regression are proposed in the literature, and it is known that the cross-validation, FPE and AIC have the same asymptotic properties, see e.g., [6]. Among candidate statistical models, Akaike's Information Criterion (AIC) prefers the model which minimizes the value. In the normal regression model with k explanatory variables, it is known that AIC is reduced to $n \log \widehat{\sigma}^2 + 2(k+1)$, where n is a sample size and $\widehat{\sigma}^2$ is the maximum likelihood estimate of the variance σ^2 [1].

Table 2.2 shows the comparative result by AIC. Model 1 with a large number of parameters is chosen among the independent models from Model 1 to Model 9.

Nonetheless, from the viewpoint to obtain tractable functional forms, we selected Models 8 and 9 as viable candidate for further investigations by spatial model (Models 10 and 11 will be discussed in Section 3.3). We note that if the estimated mean value $E[F(N,R)]$ exceeds 1.0, then it is truncated to 1.0 for all models.

3. Parameter Estimation with Spatially-Dependent Errors.

3.1. Spatial Model for Rectangular Region. Let F_{ij} be forest coverage rate at image coordinate $(i,j), i = 1,2,\cdots,m; j = 1,2,\cdots,n$, and let $\mu_{ij} = \mu_{ij}(\boldsymbol{\beta})$ be the expected values of F_{ij} specified by a parameter vector $\boldsymbol{\beta}$. In the previous section, we fit six continuous regression models other than the models with step functions. If we take a model $\mathrm{E}\left[F(N_{ij}, R_{ij})\right] = \mu_{ij}(\boldsymbol{\beta}) = g_2(N_{ij}) + h_1(R_{ij})$, the unknown parameter vector $\boldsymbol{\beta}$ is given by $(\alpha, \gamma, \delta, \theta)'$, see (2.3) and (2.7) as an example of the elements of $\boldsymbol{\beta}$.

Table 2.2
Comparison of models by AIC (34000 is added to AIC values below).
Numerals with asterisk(s) denote the best three ranking.

Model ID	Regression Models	♯ of param.	AIC Dataset I	AIC Dataset II
1	$g_\beta(N)+h_0(R)$	18	657.0*	-11.0*
2	$g_0(N)+h_1(R)$	12	807.2	68.7
3	$g_0(N)+h_2(R)$	12	763.8	39.6
4	$g_1(N)+h_1(R)$	4	824.0	448.3
5	$g_1(N)+h_2(R)$	4	777.0	343.2
6	$g_2(N)+h_1(R)$	4	1172.1	992.7
7	$g_2(N)+h_2(R)$	4	1111.0	880.8
8	$g_\beta(N)+h_1(R)$	5	772.4	97.0
9	$g_\beta(N)+h_2(R)$	5	722.8	35.7
10	$g_\beta(N)+h_1(R)$	6	285.3**	-204.8**
11	$g_\beta(N)+h_2(R)$	6	196.6***	-497.3***

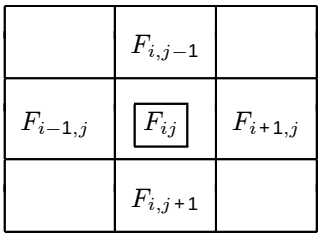

Fig.3.1. Adjacency in grid cells.

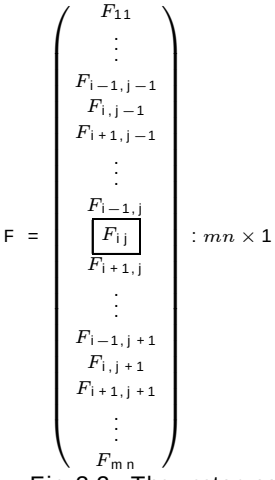

Fig.3.2. The vector notation.

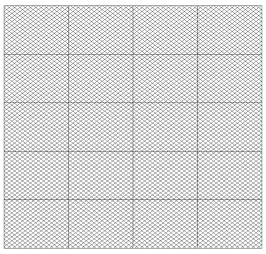

Fig.3.3. An example of rectangular data.

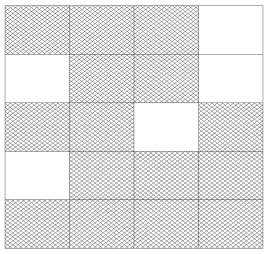

Fig.3.4. Data with missing values.

Let $\boldsymbol{F} : mn \times 1$ be a vector of forest rates observed at the rectangular regions with size $m \times n$, see Fig. 3.2. Also, let H be an $mn \times mn$ adjacency matrix whose rows and columns are corresponding to the indices of the column vector $\boldsymbol{F}$. Namely, entries of H are defined by 1 if two cells are adjacent, otherwise they are defined by 0. Hence, H is symmetric, and the sum of row is equal to four if the row is corresponding to the inner points of the rectangular. For example, a region like Fig. 3.3, its adjacency

matrix H shall be given by (3.1).

$$
\text{(3.1)} \qquad \mathsf{H} = \begin{pmatrix}
0100\,1000\,0000\,0000\,0000 \\
1010\,0100\,0000\,0000\,0000 \\
0101\,0010\,0000\,0000\,0000 \\
0010\,0001\,0000\,0000\,0000 \\
1000\,0100\,1000\,0000\,0000 \\
0100\,1010\,0100\,0000\,0000 \\
0010\,0101\,0010\,0000\,0000 \\
0001\,0010\,0001\,0000\,0000 \\
0000\,1000\,0100\,1000\,0000 \\
0000\,0100\,1010\,0100\,0000 \\
0000\,0010\,0101\,0010\,0000 \\
0000\,0001\,0010\,0001\,0000 \\
0000\,0000\,1000\,0100\,1000 \\
0000\,0000\,0100\,1010\,0100 \\
0000\,0000\,0010\,0101\,0010 \\
0000\,0000\,0001\,0010\,0001 \\
0000\,0000\,0000\,1000\,0100 \\
0000\,0000\,0000\,0100\,1010 \\
0000\,0000\,0000\,0010\,0101 \\
0000\,0000\,0000\,0001\,0010
\end{pmatrix}, \quad \mathsf{A}_4 = \begin{pmatrix}
0100 \\ 1010 \\ 0101 \\ 0010
\end{pmatrix}, \quad \mathsf{A}_5 = \begin{pmatrix}
01000 \\ 10100 \\ 01010 \\ 00101 \\ 00010
\end{pmatrix}
$$

The adjacency matrix of rectangular data like this case can be simply expressed by using Kronecker product as $H = A_5 \otimes I + I \otimes A_4$, see (3.1).

Taking the influence of the neighborhood into account, we assume that the conditional distribution of F_{ij} with values for given adjacent cells is expressed by the following normal distribution:

$$
\begin{aligned}
&[\, F_{ij} \mid F_{i-1,j}, F_{i+1,j},\ F_{i,j-1},\ F_{i,j+1}] \\
&\sim N\left(\mu_{ij} + \phi(\eta_{i+1,j}+\eta_{i-1,j}+\eta_{i,j+1}+\eta_{i,j-1}),\ \sigma^2\right)
\end{aligned} \qquad \text{(3.2)}
$$

where $\eta_{ij} = F_{ij} - \mu_{ij}$ and so on (see Fig. 3.1). Parameter ϕ captures the spatial dependency. It should be noted that the spatial model (3.2) is reduced to ordinary regression model considered in the preceding section when $\phi = 0$.

Now, by the conditional distribution (3.2), we have the joint distribution of the forest coverage vector $\boldsymbol{F}$, see Chapter 6 of [2]. The joint distribution is an mn-variate normal with the mean vector $\boldsymbol{\mu} \equiv \boldsymbol{\mu}(\boldsymbol{\beta}) : mn \times 1$ and the variance-covariance matrix $\sigma^2(I - \phi H)^{-1}$, where $\boldsymbol{\mu}$ is constituted by the means μ_{ij}. The likelihood function, say $L(\boldsymbol{\beta}, \sigma^2, \phi)$, is therefore given by

$$
L(\boldsymbol{\beta}, \sigma^2, \phi) = (2\pi\sigma^2)^{-mn/2} |I - \phi H|^{1/2} \exp\{-Q(\boldsymbol{\beta}, \phi)/(2\sigma^2)\} \qquad \text{(3.3)}
$$

where

$$
Q(\boldsymbol{\beta}, \phi) = \{\boldsymbol{F} - \boldsymbol{\mu}(\boldsymbol{\beta})\}' (I - \phi H) \{\boldsymbol{F} - \boldsymbol{\mu}(\boldsymbol{\beta})\}. \qquad \text{(3.4)}
$$

The maximum likelihood estimators (MLEs) can be obtained as follows.

Firstly, fix ϕ. The regression parameter $\boldsymbol{\beta}$ is estimated by minimizing the quadratic form $Q(\boldsymbol{\beta}, \phi)$ of (3.4), say $\widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}}(\phi)$. Then, σ^2 is estimated by

$$
\widehat{\sigma}^2(\phi) = Q(\widehat{\boldsymbol{\beta}}, \phi)/(mn).
$$

Secondly, ϕ is estimated by maximizing the profile likelihood. This is equivalent to minimize the following function of ϕ:

$$-2\log L(\widehat{\boldsymbol{\beta}}(\phi),\widehat{\sigma}^2(\phi),\phi) = n\log(2\pi e) + n\log\widehat{\sigma}^2(\phi) - \log|I-\phi H|, \tag{3.5}$$

say $\widehat{\phi}$. Finally, we have the MLE for the parameters by $\widehat{\boldsymbol{\beta}}(\widehat{\phi})$ and $\widehat{\sigma}(\widehat{\phi})$ by using MLE $\widehat{\phi}$. Hence,

$$\text{AIC} = mn\log(2\pi e) + mn\log\widehat{\sigma}^2(\widehat{\phi}) - \log|I-\widehat{\phi}H| + 2p \tag{3.6}$$

where p denotes number of unknown parameters under consideration.

3.2. Actual Calculations for Non-Rectangular Data Assembly. Most of municipal boundaries as well as the inside holes like lakes do not always have a rectangular shape. One method for the irregularity is to remove the data on the edge (as if to peel off an outer shell), and to use the cells only that have four neighbors. But this strategy loses much information. For effective use of the data, we will illustrate how to find the adjacency matrix by using Figs. 3.3 and 3.4.

$$P=\begin{pmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1
\end{pmatrix}:20\times 15,\quad
H^*=\begin{pmatrix}
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&1&0&1&0&0&0&0&0&0&0&0&0\\
0&0&1&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&1&0&0&0&0&0&0&0\\
0&0&0&0&1&0&1&0&1&0&0&0&0&0&0\\
0&0&0&0&0&1&0&1&0&1&0&0&0&0&0\\
0&0&0&0&0&0&1&0&1&0&1&0&0&0&0\\
0&0&0&0&0&0&0&1&0&1&0&1&0&0&0\\
0&0&0&0&0&0&0&0&1&0&1&0&1&0&0\\
0&0&0&0&0&0&0&0&0&1&0&1&0&1&0\\
0&0&0&0&0&0&0&0&0&0&1&0&1&0&1\\
0&0&0&0&0&0&0&0&0&0&0&1&0&1&0
\end{pmatrix}:15\times 15 \tag{3.7}$$

Suppose that data are available only at the hatched cells in Fig. 3.4. An adjacency matrix H of the complete data at which Fig. 3.3 is embedded is given by (3.1). Then, the matrix P shown in (3.7) denotes a projection matrix from the complete data to the data with missing values. Then, we can find the adjacency matrix H^* of Fig. 3.4 is given by

$$H^* = P'HP. \tag{3.8}$$

Thus, the adjacency matrix of data with missing values can be obtained with less memory requirements because that of the rectangular data is easily found by Kronecker product. Our dataset I with $n = 8697$ is embedded in the rectangular of size 130×115, and the adjacency matrix is derived by this process.

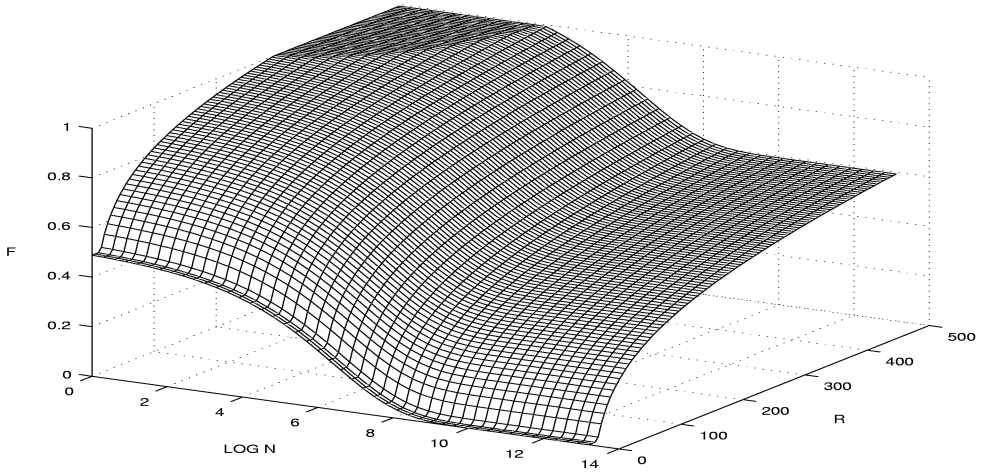

Fig. 4.1. The estimated mean surface of Model 11 for Dataset I. Flat surface is on the truncated values to 1.0 when $E[F(N,R)] > 1.0$, see Section 2.2.

3.3. Parameter Estimates and Model Selection. Values of AIC of spatial models are tabulated in Table 2.2 as Models 10 and 11. Other models are derived under independent assumptions. It can be seen that the spatial models improve AIC drastically in both of the Datasets.

In the spatially independent cases, it should be natural to select $g_\beta(N)$ for the first term of population. But it is subtle to judge which of the two forms shall be chosen for the second term of relief, due to small difference of the relative appropriateness between Models 8 and 9.

In the spatially dependent cases, $h_2(R)$ gives nonetheless apparently better result. Therefore the following model (Model 11) is to be duly selected.

$$E\left[F(N,R)\right] = \gamma \exp\left(-\alpha N^{\beta}\right) + \begin{cases} \delta \log(\frac{R}{\theta}) & \text{if } R > \theta \\ 0 & \text{if } 0 \le R \le \theta \end{cases} \tag{3.9}$$

4. Suggestions from Empirical Studies.

4.1. Areas without Forest Cover, Cell Resolution, and Spatial Dependency. Orange groves and warehouses in the coastal region along the sea are identified as areas without forests, as well as big agricultural lands such as cattle ranches and paddy fields in the hinterland [7]. Those areas give anomaly to data sampled from small grid-cells. But consideration to spatial dependence of data should provide stability, which can be seen in the large reduction of AIC values in models 10 and 11, compared to the independent cases. This indicates original data with small resolution can be well utilized by spatial model.

It is reported that adjacency to developed land, and proximity to transportation networks and major human settlements, are important factors that determine regional patterns of land development [3]. In these cases, spatial data dependency is strongly

Table 4.1
Estimated parameters.

Parameter	Model 8	Model 9	Model 10	Model 11
Dataset I				
α	0.01408	0.01483	0.01838	0.02189
β	0.7264	0.7033	0.6480	0.5950
γ	0.7789	0.7483	0.4731	0.4992
δ	0.09281	0.1678	0.08846	0.1667
θ	27.77	11.52	27.85	13.60
ϕ	–	–	0.1559	0.1609
Dataset II				
α	0.02727	0.02721	0.02184	0.02580
β	0.5448	0.5300	0.5876	0.5163
γ	0.5453	0.4967	0.5549	0.6243
δ	0.07083	0.1611	0.07310	0.1594
θ	39.08	23.16	29.17	25.84
ϕ	–	–	0.1510	0.1610

assumed, and should be applied to the analysis since fragmentation and dispersion of forests can be taken into account.

4.2. Interpretations of Estimated Parameters. A specific activity such as "collecting, compiling and regularly updating and distributing information on land classification and land use, including data on forest cover, areas suitable for afforestation, endangered species, ecological values, traditional/indigenous land use values, biomass and productivity, correlating demographic, socio-economic and forest resources information at the micro- and macro-levels, and undertaking periodic analyses of forest programmes," was adopted to develop in the Agenda 21 [9]. It is suggested by this study that the typical socio-economic factor to deforestation can be measured and compared by region with relation (3.9).

We regard the followings as worth being scrutinized through systematic transdisciplinary studies: the coefficient α to N in relation (3.9) represent a land-cover pattern reflecting the regional productive structure, and the exponent β represent extent of population pressure to forests. Demographic cohort composition in the community would give strong influence to the pressure degree. Forests are well preserved when both α and β are small compared to population size of the region. We would like to name α as 'base trend' coefficient and β as 'cohort spurt' coefficient in terms of deforestation.

Boundaries to determine the regions are to be defined, but for large-scale analysis it would be sufficient to use typical climate and apparent economic classification such as developed region i in temperate zone, developing region j in tropical zone.

5. Conclusions. Drastic improvement of relative appropriateness could be seen in spatial models. For the term of population, $g(N) = \gamma \exp\left(-\alpha N^{\beta}\right)$ is selected

among many possible candidates. We regard the coefficients α and β could be plain environmental indicators if we compare α_i and β_i by region. Namely if the both coefficients are estimated by region with its suffix i regularly year-by-year (t), then $\alpha = \alpha_i(t),\ \beta = \beta_i(t)$:

- cross-sectional comparative comprehension can be visualized at $t = t_m$ with scatter plot, $\beta_i(t_m)$ vs. $\alpha_i(t_m)$, with $i = 1, 2, \cdots, n$ where n is a total number of surveyed regions,
- time-series trend can be observed for a region k by over-plotting $\beta_k(t_1)$ vs. $\alpha_k(t_1)$, $\beta_k(t_2)$ vs. $\alpha_k(t_2)$, ... with tracing arrows.

If we obtain smaller α and β chronological values in a region, then the applied environmental policy is successful. Problems however remain in the second term of $h(R)$ with its theoretical aspect. Further verification is necessary applying the relation to many other forested regions.

REFERENCES

[1] H. Akaike, Information theory and an extension of the maximum likelihood principle, 2nd International Symposium on Information Theory (1973) Akadémiai Kiado, Budapest, pp. 267–281.
[2] N. A. C. Cressie, Statistics for Spatial Data, John Wiley & Sons, New York, 1991.
[3] C. A. S. Hall, et al., Modelling spatial and temporal patterns of tropical land use change, Journal of Biogeography, 22 (1995), pp. 753–757.
[4] E. F. Lambin, Modelling and monitoring land-cover change processes in tropical regions, Progress in Physical Geography, 21 (1997), pp. 375–393.
[5] N. Myers, The world's forests and human populations: the environmental interconnections, Pop. Dev. Rev., 16 (supplement) (1990), pp. 1–15.
[6] R. Nishii, Asymptotic properties of criteria for selection of variables in multiple regression, Ann. Statist., 12 (1984), pp. 758–765.
[7] S. Tanaka, A quantitative aspect on man-land interrelations: case study of deforestation in Japan, Ecological Engineering, 4 (1995), pp. 163–172.
[8] S. Tanaka and R. Nishii, A model of deforestation by human population interactions, Environmental and Ecological Statistics, 4 (1997), pp. 83–91.
[9] United Nations, Agenda 21 Forest Principles (draft Rio Declaration), United Nations Publications, New York, 1992.

Appendix: Theoretical Modeling. Suppose that relative rate of forest reduction increases with human population size by the following differential equation:

$$-F^{-1}dF/dN = \nu N^{\xi} \qquad (F > 0).$$

where ν is a positive coefficient. The solution is easily obtained by

$$F = \gamma \exp\left(-\alpha N^{\beta}\right) \quad \text{with} \quad \alpha = \nu/(\xi + 1),\ \beta = \xi + 1.$$

If the deforestation arithmetically (equivalent expression to linearly) increases with human population size , $\xi = 1$, then $\beta = 2$.