Monte Carlo Sensitivity Analysis of a Complex Ecosystems Model

Hiroyuki HIROOKA

Faculty of Agriculture, Kyoto University,
Sakyo-ku, Kyoto-shi 606

(Received June 12, 1989)

Abstract Monte Carlo sensitivity analysis was applied to an ecosystems model in cattle grazing grassland in order to investigate the contribution of variability on the model parameters to uncertainties in the model predictions. In this method, path coefficient was used as a measure of parameter sensitivity. The results obtained are as follows: (1) The model parameters which enter in the model as exponents were far more important than multiplicative parameters. (2) The highly sensitive parameters to an output are not always included in the equation relating directly to the output. (3) The sensitivity conclusions vary from outputs as well as with time. (4) When compared relative sensitivity coefficient, it was indicated that path coefficient may have advantages that the use allows us to consider interaction effects between model parameters at the same time and to neglect magnitude of perturbation of model parameters.

Jpn. J. Zootech. Sci., 61 (1) : 67-73, 1990

Key words : Monte Carlo method, sensitivity analysis, ecosystems model, path coefficient, model uncertainty

In these three decades, systems simulation has been very often practiced in the research fields of grassland and animal science. Although the approach is quite useful, it is known that there are some limitations. One of them is that the majority of simulation models are deterministic and variability which is not specified deterministically may thus have a significant influence on the model predictions (outputs). If the factors which affect model predictions can be understood and controlled, then the confidence of model predictions can be enhanced.

Generally speaking, a complex simulation model consists of a series of mathematical equations which express the relationships between components of the model and the individual equation is estimated by means of regression and least-square method using the data which were measured in a separate labor-

atory or field experiment. It is therefore important to investigate how estimation error (variability) of a model parameter in each equation affects the uncertainty of model predictions and how such model uncertainty can be reduced.

The sensitivity analysis based on changes of a single model parameter has been the most conventional and widely used method of evaluating the contribution of estimation error in each model parameter to overall model predictions. However, it is indicated that the method has a serious weakness that only one model parameter is usually varied independently and the others can not be taken into consideration.

To overcome the limitation, HIROOKA and YAMADA¹⁾ proposed Monte Carlo sensitivity analysis using path coefficient as a measure of parameter sensitivity. The objective of this

study is to apply the new sensitivity method to the complex ecosystems model for cattle grazing grassland, constructed by OKUBO *et al.*^{2,3)}, and examine the effects of model parameter estimation errors on the model predictions.

Materials and Methods

Description of model

The model used in this study was a compartment model for the dynamics of the grazed grassland ecosystem. Fig. 1 shows the model structure of material flow containing 9 compartments V_i and transfer coefficients of organic matter from V_i to V_i' ($f_{ii'}$). The model consists of nine differential equations concerning each state variables (V_i) (Table 1), and the

transfer coefficients were represented as functions with 66 model parameters (a_k) (Table 2). The more detailed explanation of the model given by OKUBO *et al.*^{2,3)}

Simulation procedure

The Monte Carlo sensitivity analysis proposed by HIROOKA and YAMADA¹⁾ was adopted for quantifying the contribution of variability of model parameter (a_k) to uncertainties of model predictions (V_i ; $i=1\sim 9$). The model simulations were performed by the following steps.

500 values for each model parameter were selected independently from normal distributions specified by means (nominal values shown in Table 3) and standard deviations. A con-

Fig. 1. Flow diagram of ecosystems model in cattle grazing grassland, OKUBO *et al.* (1975).

Sensitivity Analysis of a Complex Model

Table 1. Differential equations showing flow rate of organic matters in nine compartments (OKUBO *et al* 1975)

$dV_1/dt = f_{01} Q(t) + f_{21} V_2 - (f_{10} + f_{12} + f_{15} + f_{16}) V_1 - F_{17}$
$dV_2/dt = f_{12} V_1 + f_{32} V_3 - (f_{20} + f_{21} + f_{23} + f_{25} + f_{26}) V_2 - F_{27}$
$dV_3/dt = f_{23} V_2 + f_{43} V_4 - (f_{30} + f_{32} + f_{34} + f_{36}) V_3$
$dV_4/dt = f_{04} V_4 + f_{34} V_3 - (f_{40} + f_{43} + f_{46}) V_4$
$dV_5/dt = f_{15} V_1 + f_{25} V_2 - f_{56} V_5$
$dV_6/dt = f_{16} V_1 + f_{26} V_2 + f_{36} V_3 + f_{46} V_4 + f_{56} V_5 - f_{60} V_6$
$dV_7/dt = F_{17} + F_{27} - (f_{78} + f_{79}) V_7$
$dV_8/dt = f_{88} f_{78} V_7 - f_{80} V_8$
$dV_9/dt = f_{79} V_7 - f_{90} V_9$

stant coefficient of variation (*i. e.*, 5% C.V. of the nominal value) was assumed in this study. In practice, two of 66 model parameters (a_6 and a_{63}) were eliminated from this analysis, since $a_6 = 0$ and $a_{63} = a_{33} + a_{34}$. 500 model runs were performed with 500 combinations of the model parameters in the period of 180 days from May 1. The state variables (V_i) at day 120 and day 180 from May 1 (V_i (120) and V_i (180)) were defined as model outputs of interest.

Using the 500 sets of model outputs (V_i (120) or V_i (180)) and the model parameters (a_k), the following linear multiple regression equation was estimated :

$$V_i(120) \text{ or } V_i(180) = b_{i0} + b_{i1} a_1 + b_{i2} a_2 + \dots + b_{ik} a_k \quad (1)$$

Table 2. Expression of transfer coefficients (OKUBO *et al* 1975)

$Q(t) = a_1 + a_2 [\sin\{2(t + a_3)/a_4\}]$	$f_{32} = a_{45} \exp(-a_{46} t) + a_{47} F_{27}$
$L = a_5 (a_6 + a_7 T + a_8 T^2) V_1$	$f_{25} = a_{48} / [1 + \exp\{-a_{49}(t - a_{50})\}]$
$f_{01} = a_9 \{1 - 1/(a_{10} L + 1)\} \{a_{11}/(a_{11} Q + 1)\}$	$f_{26} = a_{51} V_8 + a_{52} V_9$
$f_{21} = a_{12} \exp(-a_{13} t) + a_{14} F_{17}$	$F_{27} = a_{41} V_2 \quad HA < D$
$f_{12} = a_{15} + a_{16} [\sin\{2(t + a_{17})/a_{18}\}]$	$= D(a_{41}/HA) V_2 \quad HA > D$
$f_{10} = a_{19} + a_{20} T$	$f_{34} = a_{53}$
$f_{20} = a_{21} + a_{22} T$	$f_{36} = a_{54}$
$f_{30} = a_{23} + a_{24} T$	$f_{43} = a_{55}$
$f_{40} = a_{25} + a_{26} T$	$f_{46} = a_{56}$
$f_{15} = a_{27} / [1 + \exp\{-a_{28}(t - a_{29})\}]$	$f_{56} = a_{57} \exp(a_{58} t)$
$f_{16} = a_{30} V_8 + a_{31} V_9$	$f_{60} = a_{59} + a_{60} T$
$W = 10^4 V_8 / a_{32}$	$FWNE = a_{61} \ln W^{0.75} + a_{62}$
$D = \{(a_{33} + a_{34}) W^{0.75} + a_{35} W^{0.75} + a_{36}\}$	$f_{78} = a_{38}$
$V_8 / (a_{37} a_{38} a_{39} W^{1.000})$	$f_{88} = a_{37} a_{39} FWNE$
$HA = a_{40} V_1 + a_{41} V_2$	$f_{80} = \{a_{63} W^{0.75} / (1000 * W)\} \cdot FWNE$
$F_{17} = a_{40} V_1 \quad HA < D$	$f_{79} = (1 - a_{38})$
$= D(a_{40}/HA) V_1 \quad HA > D$	$f_{90} = a_{64} + a_{65} T$
$f_{23} = a_{42} / [1 + \exp\{-a_{43}(t - a_{44})\}]$	$f_{04} = a_{66}$

Foot note T : Atmosphere temperature, Q(t) : Inpnt of solar energ.

$$(k = 1 \sim 66, k \neq 6, 63)$$

and all partial regression coefficients (b_{ik}) were standardized as :

$$\beta_{ik} = \sigma_k / \sigma_{V_i} \cdot b_{ik} \quad (2)$$

where σ_k and σ_{V_i} denote the standard deviations of a_k and V_i respectively. By definition, these standard partial regression coefficients are really path coefficients⁴⁾. In the present study, the path coefficient (β_{ik}) was adopted as a measure of representing sensitivity of a model parameter (a_k) on model outputs (V_i (120) and V_i (180)). This measure can be interpreted as the change in V_i , as a fraction of one standard deviation of V_i , produced by one standard deviation in a_k .

The path coefficient (β_{ik}) can be numerically represented as follows :

$$\beta_{ik} = (V_{ik}^* - V_{ik}^0) / (a_k^* - a_k^0) \cdot \sigma_k / \sigma_{V_i} \quad (3)$$

where a_k^0 is the nominal values of the k^{th} model parameter (non-perturbed state) ; a_k^* is the value of the k^{th} model parameter in its perturbed state ; V_{ik}^* is the value of i^{th} state variable at a_k^* ; V_{ik}^0 is the value of state variable at a_k^0 .

Since $\sigma_k = a_k^0 u_k / 100$ and $a_k^* = a_k^0 (1 + u_k / 100)$, equation (3) can be rewritten as :

$$\beta_{ik} = (V_{ik}^* - V_{ik}^0) / \sigma_{V_i} \quad (4)$$

Table 3. Values of model parameters used in the model. All values were derived from Okubo *et al.* and they were regarded as best estimates of individual parameters

Model parameters	values	model parameters	values	model parameters	values
a_1	3100	a_{23}	0.0001	a_{45}	0.005
a_2	1100	a_{23}	0.00028	a_{46}	0.06
a_3	50	a_{24}	0.0028	a_{47}	0.05
a_4	365	a_{25}	0.00083	a_{48}	1.0
a_5	0.025	a_{26}	1.0	a_{49}	0.055
a_6	0	a_{27}	0.065	a_{50}	180
a_7	0.08	a_{28}	180	a_{51}	0.0001
a_8	-0.002	a_{29}	0.001	a_{52}	0.0001
a_9	15	a_{30}	0.001	a_{53}	0.008
a_{10}	2	a_{31}	750	a_{54}	0.001
a_{11}	0.0015	a_{32}	41.0	a_{55}	0.003
a_{12}	0.04	a_{33}	42.0	a_{56}	0.0015
a_{13}	0.06	a_{34}	41.0	a_{57}	0.0004
a_{14}	0.05	a_{35}	1730	a_{58}	0.0015
a_{15}	0.02	a_{36}	4.0	a_{59}	0.0004
a_{16}	0.015	a_{37}	0.53	a_{60}	0.0013
a_{17}	-12.5	a_{38}	0.5	a_{61}	-0.3875
a_{18}	250	a_{39}	0.75	a_{62}	1.8732
a_{19}	0.0055	a_{40}	0.75	a_{63}	$a_{33} + a_{34}$
a_{20}	0.0016	a_{41}	1.0	a_{64}	0.002
a_{21}	0.0028	a_{42}	0.04	a_{65}	0.005
a_{22}	0.00083	a_{43}	170	a_{66}	0.0002

where u_k is the size of the perturbation.

Results and Discussion

Table 4 shows the top 5 ranked model parameters with regard to 9 outputs (state variables) at day 120 and day 180, when ranked by the absolute values of path coefficient. The higher ranked parameters appears to be more important, because the uncertainty of model outputs increases as the estimation error of a model parameter becomes larger.

From Table 4, the following interesting points were noted. First, the model parameters used as exponents in the model (*e. g.*, a_{28} , a_{29} , a_{43} , a_{44} , a_{49} and a_{50}) tended to be relatively more important than multiplicative parameters. This may be because small changes in these parameters result in relative larger changes in the value of the derivatives.

Secondly, although the highly sensitive parameters to an output were generally included in the equations relating directly to the output,

there were several exceptions. For example, the importance of model parameters, a_{61} and a_{62} , were high ranking with respect to V_1 (leaf biomass) at day 120, although the parameters are included in other equation concerning factors of conversion from net energy to body weight in cattle (FWNE). This may be due to the dynamic nature of the present model; increase of cattle gain (V_6) caused by increasing a_{61} and a_{62} may lead to increase of herbage intake by cattle (V_7) and thereby decrease of leaf mass (V_1).

Table 5 shows the Spearman rank correlation comparing the sensitivity conclusions between the different state variables at day 120 (above diagonal), between the different state variables at day 180 (below diagonal) and between the same state variables at day 120 and day 180 (diagonal). It is considered that the correlation represents the degree of similarity between sensitivity conclusions from the rank ordering of parameter-output pairs.

Sensitivity Analysis of a Complex Model

Table 4. Top 5 ranked model parameters and their path coefficients for nine state variables at day 120 and day 180

Day 120 Rank	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9
1	a_{29} (0.669)	a_{44} (0.730)	a_{44} (-0.727)	a_{44} (-0.691)	a_{29} (-0.707)	a_{60} (-0.553)	a_{62} (0.507)	a_{62} (0.735)	a_{38} (-0.586)
2	a_9 (0.413)	a_{50} (0.404)	a_{43} (-0.270)	a_{26} (-0.506)	a_{50} (-0.456)	a_{62} (0.449)	a_{61} (0.466)	a_{61} (0.666)	a_{62} (0.382)
3	a_{62} (-0.242)	a_9 (0.167)	a_9 (0.227)	a_{43} (-0.337)	a_{28} (-0.261)	a_9 (0.438)	a_{39} (-0.358)	a_{32} (0.183)	a_{61} (0.348)
4	a_{61} (-0.234)	a_{18} (0.147)	a_{62} (-0.218)	a_{53} (0.211)	a_{49} (-0.207)	a_{61} (0.391)	a_{37} (-0.350)	a_{35} (0.062)	a_{39} (-0.287)
5	a_{28} (0.209)	a_{29} (0.146)	a_{50} (0.210)	a_9 (0.171)	a_{44} (0.192)	a_{29} (0.198)	a_{38} (-0.336)	a_{36} (0.039)	a_{37} (-0.278)
Day 180 Rank	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9
1	a_{29} (0.485)	a_{50} (0.403)	a_{29} (0.657)	a_{44} (-0.617)	a_9 (0.513)	a_{29} (0.797)	a_{29} (0.382)	a_{62} (0.731)	a_{38} (-0.488)
2	a_{62} (-0.098)	a_{44} (-0.189)	a_{44} (-0.369)	a_{29} (0.333)	a_{50} (-0.356)	a_{60} (-0.354)	a_{50} (0.326)	a_{61} (0.679)	a_{61} (0.369)
3	a_{31} (-0.097)	a_{54} (-0.167)	a_{62} (-0.217)	a_{50} (0.255)	a_{62} (-0.346)	a_9 (0.330)	a_{18} (0.184)	a_{32} (0.169)	a_{62} (0.362)
4	a_{34} (0.074)	a_{30} (-0.138)	a_{61} (-0.216)	a_{26} (-0.233)	a_{61} (-0.312)	a_{62} (0.231)	a_{44} (-0.167)	a_{35} (0.055)	a_{37} (0.246)
5	a_{27} (-0.069)	a_{15} (-0.132)	a_{50} (0.201)	a_{43} (-0.233)	a_{44} (0.261)	a_{61} (0.205)	a_{66} (0.125)	a_{36} (0.031)	a_{39} (0.201)

Table 5. Spearman rank correlations comparing the sensitivity conclusions between the different state variables at day 120 (above diagonal), between the different state variables at day 120 (below diagonal) and between the same state variables at day 120 and day 180 (diagonal)

State variables	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9
V_1	0.002	0.651**	0.592**	0.300*	0.437**	0.683**	0.098	0.155	0.047
V_2	0.038	0.277*	0.751**	0.506**	0.448**	0.567**	0.195	0.144	-0.052
V_3	-0.121	0.231*	0.850**	0.799**	0.426**	0.612**	0.399**	0.227	0.081
V_4	-0.106	0.157	0.849**	0.871**	0.308*	0.434**	0.322**	0.071	-0.004
V_5	-0.132	0.203	0.751**	0.658**	0.507**	0.252*	0.085	0.146	-0.095
V_6	-0.076	0.037	0.373**	0.314*	0.526**	0.694**	0.346**	0.062	0.379**
V_7	-0.032	0.338**	0.374	0.295*	0.346**	0.179	0.057	0.307*	0.590**
V_8	-0.126	0.026	0.438**	0.354**	0.462**	0.309*	0.124	0.475**	0.144
V_9	-0.001	0.208	0.434**	0.399**	0.410**	0.048	0.233	0.401**	0.435**

* Significant at $p < 0.05$

** Significant at $p < 0.01$

When compared different variables (above and below diagonal), only correlations between V_3 (rhizome, which is organ for matter reserves) and four variables (V_4 , V_5 , V_6 and V_7) were commonly significant at $P < 0.01$ at both day 120 and day 180. Further the correlations of the same state variable at day 120 and 180 were not significant at $P < 0.05$ for V_1 , and V_7 .

These results indicate that not only sensitivity conclusions may vary from state variables, but also the conclusions for a given state variable may vary with time.

In general, sensitivity analysis provides information about which of the model parameters are important in determining the behavior of model outputs. For this reason, this analy-

sis may help researchers understand which model parameters should be estimated with high accuracy. SUMMERS and MCKELLAR⁵⁾ stated that new research is most needed to identify the relationships which are high sensitive and which are least understood.

The sensitivity of the model outputs (V_{ik}) to a change in model parameter (a_k) is defined as the partial derivative with respect to a_k . In the most simulation studies, however, an iterative method (*i. e.*, conventional sensitivity analysis), in which single parameter is varied systematically and the resulting effects on model outputs is recorded, has been taken, because it is difficult to derive the partial derivatives analytically in a complex simulation model. The sensitivity coefficient is defined as the partial derivative of V_{ik} with respect to a_k and numerically approximated by :

$$S_{ik} = \partial V_{ik} / \partial a_k \\ = (V_{ik}^* - V_{ik}^o) / (a_k^* - a_k^o) \quad (5)$$

It is suggested that a dimensionless relative sensitivity coefficient (RS_{ik}) may be more adequate as a measure of sensitivity than the sensitivity coefficient (Eq. 5)^{6,7)}. It can be numerically represented as follows :

$$RS_{ik} = (\partial V_{ik} / \partial a_k) / (V_{ik}^o / a_k^o) \\ = [(V_{ik}^* - V_{ik}^o) / V_{ik}^o] \cdot [a_k^o / (a_k^* - a_k^o)] \quad (6)$$

This measure is defined as the percentage change in V_{ik} resulting from 1% increase in a_k . Since $a_k^* = a_k^o (1 + u_k/100)$, it can be rewritten :

$$RS_{ik} = (V_{ik}^* - V_{ik}^o) / V_{ik}^o \cdot 100 / u_k \quad (7)$$

By definition, both of the relative sensitivity and the path coefficients can be commonly regarded as relative importance of each model parameter. In fact, HIROOKA and YAMADA¹⁾ showed that the two coefficients yield almost the same sensitivity conclusions in a simple model.

Although the relative sensitivity coefficient has been conventionally used as a measure of parameter sensitivity, it is known that the coefficient has an essential disadvantage ; since

model parameters are usually varied independently, the effects of other parameters are virtually ignored. If we want to consider interaction effects between model parameters, it is necessary to use factorial designs by computers ; *i. e.*, the model should be run such that all levels of a model parameter are combined with all levels of every other parameter. For example, when operated using set of m model parameters with n levels, n^m runs are needed to identify the interaction effects of model parameters. Even if a fractional factorial design is adopted, the computer costs of repeated runs may become excessive when this method is applied to a large and complex model¹⁾.

On the other hand, the Monte Carlo sensitivity analysis used in this study makes it possible to consider the effects of all model parameters at the same time, because path coefficient can be obtained after the linear effects of other parameter have been removed.

In addition, there are other advantages in this method as follows. First, it is conceptually and technically simple to generate the normal distributions on up-to-date digital computers, and the options for the purpose are supplied in some simulation packages. Secondly, the path coefficient allows us to neglect the effect of magnitude of model parameters. It was assumed in this study that all model parameters have equal relative variability (*i. e.*, C.V. = 5%), although this assumption may be unrealistic in the actual ecosystems. However, as shown in equation (4), since path coefficient is free from the magnitude of perturbation of model parameter (*i. e.* u_k value), it is useful in such conditions.

From the above results, it is suggested that the Monte Carlo sensitivity analysis using the path coefficient as a measure of parameter sensitivity will be a rather useful method and it may be widely applicable to complex simulation models. However, needless to say, some problems still remain unsolved in this study.

For example, the method assumes only linear relationships between model parameters and model predictions, despite the fact that most of simulation models are non-linear. Consequently, further research in this field will be required to apply non-linear statistics which can account for most of the high-order non-linear effects.

Acknowledgements

The author is grateful to Professor Y. YAMADA and Professor A. MIYAZAKI for their helpful suggestions and encouragement throughout this study. The author also wishes to thank Professor T. OKUBO, University of Tokyo, for his helpful suggestions and correc-

tion of this paper.

References

- 1) HIROOKA, H. and Y. YAMADA. J. JASS., 5 : 62-71. 1989. (in Japanese)
- 2) OKUBO T., S. HIROSAKI and T. OKUBO. In: Ecological Studies in Japanese Grassland (Numata, N, eds.) : 268-275. Univ. Tokyo Press. 1975.
- 3) OKUBO, T., S. TAKAHASI and T. AKIYAMA. World Rev. Anim. Prod., 13 : 45-50. 1977.
- 4) LI, C.C. Path Analysis-a prime. Boxwood Press, Pacific Grove, California. 1975.
- 5) SUMMERS, J.K. and H.N. McKELLAR. Ecological Modelling., 13 : 283-301. 1981.
- 6) ANDERSON, J.R. Rev. Mktg. Agr. Econ., 42 : 3-55. 1974.
- 7) DOWING, D.J., R.H. GARDNER and F.O. HOFFMAN. Technometrics, 27 : 151-163. 1985.

複雑な生態系モデルのモンテカルロ感度分析

広岡博之

京都大学農学部, 京都市 606

モンテカルロ感度分析を放牧牛と草地に関する生態系モデルに適用して、モデルの予測の不確かさに対するモデルパラメーターの分散の寄与を調べた。この方法においては、径路係数がパラメーターの感度の指標として用いられた。得られた結果は次に示すとおりである。(1) モデルにおいて指数として用いられているモデルパラメーターは、乗数として用いられているものよりずっと重要であった。(2) モデルの出力に対して、感度の高いパラメーターが必ずしもその出力と直接関連する式に含まれているとはかぎらなかった。(3) 感度分析の結果は、出力のみならず時間によってさまざまであった。(4) 相対的感度係数と比べて、径路係数はモデルパラメーター間の相互関係を同時に考慮でき、さらにモデルパラメーターの変化の大きさを無視できる利点のあることが示された。

日畜会報, 61 (1) : 67-73, 1990