SENSITIVITY ANALYSIS BY EXPERIMENTAL DESIGN AND METAMODELLING FOR INTERIBR-ENDEMIC

A. VONK NOORDEGRAAF*, M. NIELEN* J.P.C. KLEIJNEN **

SUMMARY

In many scientific studies, sensitivity analysis of simulation models is performed by changing only one factor at a time. Such an approach results in less accurate estimates of factor effects, and does not allow for estimation of interaction between factors. Experimental design and metamodelling (Kleijnen & Sargent, 2000) supports a structural approach to sensitivity analysis, and is more effective and efficient in estimating factor effects, including interactions. This paper applies these techniques to the simulation model InterIBR-endemic, which simulates the spread and control of BHV1 within and between cattle farms (Vonk Noordegraaf et al., 2000). Linear (OLS) and non-linear (logistic and tobit regression) regression metamodels were fitted to the input-output data of the simulation experiments. When dealing with a censored outcome variable, tobit regression is considered more appropriate than OLS. Future field studies should focus on getting better estimates of factors to which the simulation model is most sensitive.

INTRODUCTION

To support decision makers in the national BHV1-eradication program in The Netherlands, the spatial, dynamic and stochastic simulation model InterIBR-endemic was developed (Vonk Noordegraaf et al., 2000). InterIBR-endemic simulates the spread and control of BHV1 within and between cattle farms in The Netherlands. This model contains many uncertain input factors and as part of verification and validation, it is important to evaluate the sensitivity of model-outcome to these factors. Sensitivity analysis allows for identification of parameters that have most impact on model outcome. In many scientific studies, sensitivity analysis is performed by changing only one factor at a time (OAT designs). This results in less accurate estimates of factor effects, and does not allow for estimation of interaction between factors (Kleijnen, 1998). The techniques of Design of Experiments (DOE) and metamodelling (Kleijnen and Sargent, 2000) support a structural approach to sensitivity analysis, and are more effective and efficient in estimating factor effects, including interactions.

In a simulation context, DOE can be defined as selecting, from the great number of possible combinations of factor levels, the set that actually needs to be simulated in an experiment with the simulation model, in order to quantify factor effects (Hunter & Naylor, 1970). The

* Department of Social Sciences, Farm Management Group, Wageningen University, Hollandseweg 1, 6706 KN Wageningen, The Netherlands

** Department of Information Systems, Center for Economic Research, Tilburg University, 5000 LE Tilburg, The Netherlands

simulation model is run for this set of combinations, and resulting input-output data are analysed by regression analysis to derive conclusions about the importance (sensitivity) of the factors. This analysis is based on a metamodel, which is defined as a model of the simulation model. Figure 1 shows the relationships among metamodel, simulation model and problem entity.

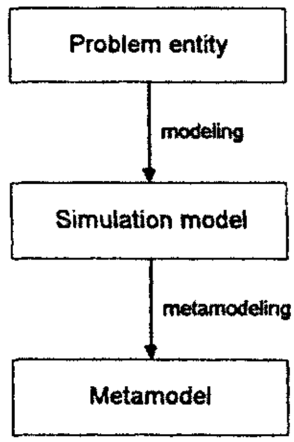

Fig 1. Metamodel, simulation model and problem entity .

The concept of metamodels can be explained by viewing the simulation model as a function that turns input factors into output performance measures. The explicit form of this function is unknown, but by experimentation with the model this function is approximated with a metamodel. The metamodel then treats the simulation model as a black box (Bettonvil & Kleijnen, 1996). The purpose of a metamodel is to estimate the response surface; the metamodel can then be used, instead of the actual simulation program, to learn about how the response surface would behave over various regions of the input-factor space (Law & Kelton, 2000).

The main goal of this paper is to show how the techniques of experimental design and metamodelling can be applied in the sensitivity analysis of complex simulation models. Furthermore, the application of three regression techniques fitting the simulation input–output transformation is demonstrated: Ordinary Least Squares (OLS), logistic and tobit regression. A theoretical discussion on the application of these regression techniques in veterinary epidemiology has been presented at the SVEPM in 1998 (Carpenter, 1998).

MATERIALS AND METHODS

The ten steps suggested by Kleijnen and Sargent (2000) were adopted for development of a metamodel: (1) Determine the goal of the metamodel; (2) Identify the inputs and their characteristics; (3) Specify the domain of applicability; (4) Identify the output variables and their characteristics; (5) Specify the accuracy required of the metamodel; (6) Specify the metamodel's validity measures and their required values; (7) Specify the metamodel, and review this specification; (8) Specify a design including tactical issues, and review the DOE; (9) Fit the metamodel and (10) Determine the validity of the fitted metamodel. In this paper only a few of these steps are highlighted, with the main emphasis on step 9, regression techniques to fit the metamodel.

Metamodel variables

In steps 2 and 3, factors and their levels were identified. Just as in OAT designs, there was no general prescription for which factors to select and what levels to assign to each factor; this depends on the goal of the study. A total of 31 factors used in the simulation model InterIBR-endemic was selected. These factors were all related to disease spread, and had in common that they were uncontrollable by decision makers and their estimation contained uncertainty. The control program was considered fixed. Sensitivity analysis requires that each factor has at least two levels, and therefore a lower and upper level were determined for each factor. Values assigned to each level reflected uncertainty of factor values in real life, based on data if available, or expert opinion otherwise. These levels also determined the experimental frame for which the metamodel was valid. Factor levels were standardised to 0 and 1, to enable comparison of factor effects by relative importance.

In step 4, simulation outputs of interest in the sensitivity analysis were selected. These were used as dependent variables in regression analysis, and for each of these outputs a metamodel was specified. In this paper the focus is on only one simulation outcome; mean number of weeks necessary to reduce the prevalence level to 5% in the dairy cattle population, applying the national control programme. Because simulation stopped when this prevalence level was not reached within 1000 simulated weeks, data were considered to be censored.

Metamodel definition and analysis

Specification of the form of the metamodel was required in step 7 of the metamodelling process. Initially, the metamodel was specified as a simple first-order polynomial, in which the independent variables (X) were standardised at either 0 or 1:

$$y_i = \beta_0 + \sum_{h=1}^{k} \beta_h x_{i,h} + e_i$$

In this additive metamodel, y_i denoted the simulation response of factor combination i, β_0 the overall mean, β_h the main effect of factor h, $x_{i,h}$ the value of the standardised factor h in combination i, and e_i represented approximation error. Later, this metamodel was extended with effect modifiers (interactions).

To allow efficient estimation of the coefficients (factor effects) in this metamodel, an experimental design was constructed (step 8). In this design, each scenario represented a combination of factor levels. Dealing with 31 factors and two levels for each factor, a total of 2^{31} scenarios could be constructed. However, to give unique estimates of the 31 main effects and overall mean of the metamodel, a minimum of 32 factor scenarios would suffice. Because estimation of certain two-factor interactions requires more scenarios, a design with 64 scenarios was constructed, by applying the Foldover principle to a 2^{31-26} fractional factorial design (Kleijnen, 1998). The resulting design matrix was orthogonal, thereby minimising the variance of the estimated factor effects. In total, 64 simulation experiments were performed with the simulation model, each experiment replicated twice. Using 5 computers, (Pentium III, 600 Mhz), total calculation time was about 2 weeks.

Fitting the specified metamodel to the resulting input-output data (step 9), classic DOE uses Ordinary Least Squares (OLS). When the 5% prevalence level was not reached within 1000 weeks (threshold), simulation output was set to 1000 weeks, although the true value could have been much higher. This is called upper censoring. With OLS regression, censored observations will result in underestimation of the factor effects, and therefore produce inconsistent estimates. Dealing with censored data, a censored regression model or tobit model may be more appropriate (Long, 1997; Greene, 1997; Carpenter, 1998). The tobit model includes information about the censoring, and thereby provides consistent estimates of factor effects (Long, 1997). The form of the underlying tobit metamodel was similar to the OLS metamodel, with the difference that the dependent variable y was now a latent variable. Observations were never seen above the threshold value of 1000 weeks. Tobit regression is based on maximum likelihood estimation, where the log likelihood of the censored regression model consists of two parts; one corresponding to the classical regression for the non-limit observations and one corresponding to the probabilities for the limit observations (Greene, 1997). Using tobit regression, the expected value of an upper censored variable equals (Long, 1997):

$$E(y \mid x_i) = [\Pr(uncensored \mid x_i) \times E(y \mid y < \tau) + [\Pr(censored \mid x_i) \times E(y \mid y = \tau)]$$

where $\Pr(\text{censored}|x_i)$ is the probability of a scenario with factor combination x_i being censored and τ the threshold value. Long (1997) shows that $E(y|x_i)$ is non-linear in x. To identify which factors significantly contributed to the event that the simulation outcome censored at 1000 weeks, logistic regression was performed. For the logistic model, the dependent variable was made dichotomous by transforming simulation output to 1 if censored (y=1000), and to 0 if not censored (y<1000). Logistic regression uses a log linear model in which the probability of the simulation outcome being censored (y=1) is modelled as (Hosmer & Lemeshow, 1989):

$$E(y \mid x_i) = \pi(x_i) = \frac{e^{\beta_0 + \beta_1 x_1 + ... + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + ... + \beta_k x_k}}$$

Estimation of factor effects by logistic regression is based on a maximum likelihood procedure, using the logit transformation of $\pi(x)$ (Hosmer & Lemeshow, 1989). For all regression models, factors were excluded from the model with a backward elimination procedure. Possible interaction terms were investigated and added to the regression model with a forward conditional selection procedure.

<u>Metamodel validation</u>

To validate the metamodel with respect to the simulation model (step 10), new scenarios could be run, and simulation output compared with metamodel output. Because the simulation model required a lot of computer time, the technique of cross-validation, which requires no new simulation runs, was applied. Cross-validation means that factor input combinations (scenarios) are eliminated one by one, the regression model re-estimated, and the resulting metamodel used to predict simulation realisation for the combination eliminated. These predictions are then compared with the corresponding simulation responses (Van Groenendaal & Kleijnen, 1997; Kleijnen & Sargent, 2000). Cross-validation was applied to the metamodel estimated by OLS, deleting all 64 scenarios one by one, re-estimating coefficients on 63 scenarios, and predicting simulation realisation for the deleted scenario.

RESULTS

From the 64 scenarios simulated, 23 did not reach a 5% prevalence level in the dairy cattle population within 1000 weeks. Table 1 shows the fitted metamodel applying OLS and tobit regression of y on x for all observations, with the censored observations included as y=1000.

Table 1. Factor estimates and individual p-values of fitted metamodel for simulation outcome 'weeks to 5% prevalence', using OLS and tobit regression.

	OLS regression		Tobit regression	
Factor	Estimate	P-value	Estimate	P-value
Intercept	217.6	0.002	-18.0	0.840
Local spread	175.9	0.005	251.5	0.004
Reactivation rate transport	158.5	0.000	220.8	0.000
Yearly reactivation rate	226.1	0.000	297.1	0.000
Professional contact	92.4	0.011	139.7	0.005
R_0 non-vaccinated	91.4	0.070	109.6	0.093
R_0 killed vaccine	125.5	0.014	200.0	0.004
Weeks young stock infected	-89.6[a]	0.013	n.s.[b]	
Hygiene certified farm	-52.4	0.294	-64.0	0.325
Bulk threshold prevalence	-98.9[a]	0.007	n.s.[b]	
Sero sensitivity	-118.7	0.001	-117.2	0.020
Vaccine type used	103.7	0.041	173.1	0.011
Interactions				
Vaccine type x R_0 killed	188.0	0.010	212.6	0.042
Local x Hygiene	-201.2	0.006	-297.1	0.004
Local x R_0 non	280.8	0.000	445.3	0.000

[a] Sign of factor estimate opposite to expectation

[b] Main effect of factor not significant ($p<0.05$) in backward elimination procedure and therefore not included in final metamodel

The adjusted R^2 of the linear regression model using OLS was 0.82. Estimates for each factor in Table 1 reflect the expected change of the outcome variable when changing a factor from its low (0) to high (1) level. For example, changing the yearly reactivation rate in the simulation model from its low to high value increased the number of weeks required to reach the 5% prevalence level in dairy cattle to 226 weeks according to the metamodel fitted with OLS. In general, using upper censoring, tobit regression resulted in increased estimates compared to OLS regression. Most factors had a positive estimate due to increased risk of virus transmission. However, increasing hygiene on certified farms and sensitivity of serological tests, reduced the value of the outcome variable both in the OLS and tobit model, reflecting preventive effects. In the OLS metamodel, two factors had negative signs not in keeping with prior expectation, but these factors were not significant using tobit regression. In both models, three interactions had a significant effect on simulation outcome.

Table 2 shows the metamodel based on logistic regression, where the event of interest was the simulation not reaching a 5% prevalence level within 1000 weeks. Factors in this model also appeared significant in the OLS and tobit regression model.

Table 2. Factor estimates and individual p-values of fitted metamodel using logistic regression where event was the simulation outcome being censored at 1000 weeks.

Factor	Estimate	St. error	P-value
Constant	-10.6	3.2	0.001
Local spread	4.2	1.5	0.013
Yearly reactivation rate	3.4	1.4	0.004
R_0 non-vaccinated	4.1	1.4	0.004
R_0 killed vaccine	4.1	1.2	0.013
Hygiene certified farm	-2.9	1.4	0.004
Vaccine type used	4.1	3.2	0.001

From this metamodel, the probability of the outcome value being censored was calculated for each scenario, and compared to the simulation outcome. Using a cut-off value of 0.5, the overall fraction of correctly classified scenarios by the metamodel was 92.2%. From the scenarios being censored, 21 out of 23 were classified correctly by the logistic metamodel, and from the uncensored scenarios, 38 out of 41 were classified correctly.

Figure 2 shows a scatter plot of the results from cross-validating the metamodel based on OLS regression. The correlation coefficient between predicted and true outcome was 0.97.

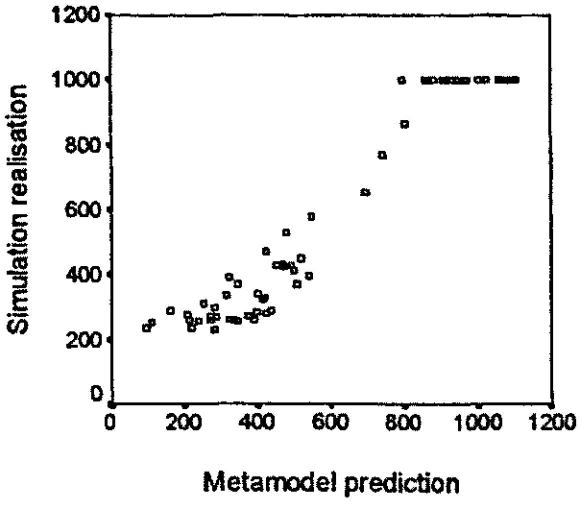

Fig 2. Scatter plot of OLS regression prediction and simulation realisation, where metamodel prediction was based on cross-validation procedure.

DISCUSSION

The main goal of this paper was to present the application of experimental design and metamodelling as part of verification and validation of a complex simulation model. Whereas changing one factor at a time is common practice for sensitivity analysis of simulation models in the field of veterinary epidemiology and economics, it does not meet the statistical requirements that are obtained by experimental design and metamodelling in estimating and testing for significance of factor effects and interactions between factors. An exception is the study of Stärk and Pfeiffer (1999), although their focus was on five factors only.

Another goal of this paper was to show the application of three different regression techniques to the fitting of simulation input-output data obtained from this experiment. Whereas OLS and logistic regression are well known in veterinary epidemiology, tobit regression has only been applied recently (Ekstrand & Carpenter, 1998). Based on the adjusted R^2 and cross-validation results, model fit using OLS was quite good and the metamodel appeared to be valid with respect to the simulation model. Two factors that entered significantly into the metamodel, however, had an estimated sign opposite to prior belief. Programming code for these factors was verified and tested, but no errors were found. These factors did not appear significant in the metamodel using tobit regression. In general, dealing with censored data, OLS will produce inconsistent estimates of factor effects, whereas tobit regression takes into account information obtained from censored data (Greene, 1997). Logistic regression uses the information less efficiently than tobit regression, because continuous output is transformed into binary data. In this study it did provide additional information on factors for which the simulation model was most sensitive.

This paper only showed the metamodel for one output of the simulation model. Other outcome variables were investigated, such as the total disease control costs and number of outbreaks on certified farms. For each output a separate metamodel was developed. Because response variables were correlated, multivariate regression was also applied.

The goal of this study was to identify which uncertain factors had greatest impact on model outcomes of interest. Factors included in the final metamodel had most impact on outcome of the simulation model, changing factor level from low to high. It is essential to realize that the importance is based on the low and high level assigned to each factor (i.e., experimental frame). Low and high values chosen in this study, were supposed to reflect uncertainty of these factors in the real world. If the model is a good representation of the real system, a sensitive region established in the model can, by association, be considered to be so in the real system. With this assumption, it can be concluded that field studies must focus on getting better estimates of factors included in the metamodels. These may include; local spread, reactivation rate at transport and on farm, professional contact, R_0 for non-vaccinated herds and for herds vaccinated with killed vaccine. Also, some factors found to be important can be used to support advice given to farmers in the current eradication programme, such as the importance of hygiene on certified farms and preference for live vaccine. Three interactions between factors were found to be significant. If factor 'vaccine type used' was at its high level (all farmers use killed vaccine), the level of R_0 for killed vaccine was found to be very important on model outcome. Also, interactions between 'local spread' and 'hygiene certified farms' and between 'local spread' and 'R_0 non-vaccinated herds' were found to be significant. Most interactions were related to the risk of introduction of virus on a farm (local spread), and the consequent virus circulation (hygiene, R_0 killed vaccine and R_0 non-vaccinated herd). If a sensitivity analysis had been performed with one factor at a time, these interaction effects could not have been estimated.

REFERENCES

Bettonvil, B. and Kleijnen, J.P.C. (1996). Searching for important factors in simulation models with many factors: Sequential bifurcation. Eur. J. Oper. Res. 96, 180-194

Carpenter, T.E. (1998). Logistic regression, an illogical choice? In: Thrusfield, M.V., Goodall, E.A. (Eds.), Proceedings of the SVEPM, 25-27 March, Ennis, Ireland.

Ekstrand, C. and Carpenter, T.E. (1998). Using a tobit regression model to analyse risk factors for foot-pad dermatitis in commercially grown broilers. Prev. Vet. Med. 37, 219-228

Hosmer, D.W. and Lemeshow, S. (1989). Applied logistic regression. Wiley and Sons, New York, 307 p.

Hunter, J.S. and Naylor, T.H. (1970). Experimental design for computer simulation experiments. Manag. Science 16, 422-434

Kleijnen, J.P.C. and Sargent, R.G. (2000). A methodology for fitting and validating metamodels in simulation. Eur. J. Oper. Res. 120, 14-29

Kleijnen, J.P.C. (1998). Experimental design for sensitivity analysis, optimization, and validation of simulation models. In: Banks, J. (Ed.). Handbook of simulation: principles, methodology, advances, applications, and practice. New York, Wiley, 173-223

Law, A.M. and Kelton, W.D. (2000). Simulation Modelling and Analysis. McGraw-Hill, Boston, 759 p

Long, J.S. (1997). Regression models for categorical and limited dependent variables. Sage Publications, Inc., 297 p

Stärk, K.D.C. and D.U. Pfeiffer, (1999). The use of experimental design for sensitivity analysis of a computer simulation experiment. In: Thrusfield, M.V., Goodall, E.A. (Eds.), Proceedings of the SVEPM 24-26 March, Bristol, England.

Van Groenendaal, W.J.H. and Kleijnen, J.P.C. (1997). On the assessment of economic risk: Factorial design versus Monte Carlo methods. J. Rel. Eng. Sys. Safety, 57, 91-102

Vonk Noordegraaf, A., Nielen, M., De Jong, M.C.M., Frankena, K., Lam, T.J.G.M. Franken, P. & Dijkhuizen, A.A., 2000. Simulation and monitoring to support the BHV1 eradication program in The Netherlands. In: Salman, M.D., Morley, P.S., Ruch-Gallie, R. (Eds.), Proceeding of the 9th ISVEE Symposium, Colorado, 514-516.