Sensitivity analysis to evaluate the impact of uncertain factors in a scenario tree model for Classical Swine Fever virus introduction into the Netherlands

De Vos, C.J.[1], Saatkamp, H.W.[1], Nielen, M.[2] & Huirne, R.B.M.[1]
[1] Farm Management Group, Department of Social Sciences, Wageningen University, Hollandseweg 1,
6706 KN Wageningen, The Netherlands.
[2] Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, Yalelaan 7,
3584 CL Utrecht, The Netherlands.

Summary
The main aim of this model is to provide a better understanding of the importance of the different introduction routes for CSFV. This study showed that group screening combined with the statistical techniques of DOE and metamodelling proved to be an effective and efficient screening method to identify those uncertain input parameters in the scenario tree model for CSFV introduction that influenced model results most. Only 128 scenario calculations were needed to specify the final metamodel. It was concluded that four uncertain input parameters change the ranking of risk factors contributing most to the annual P_{CSFV} and thus require a more accurate estimate of their values to make model outcome robust.

Introduction

Introduction of classical swine fever virus (CSFV) is a continuing threat to the pig production sector in the Netherlands. In order to optimally use resources for prevention of CSFV introduction, more quantitative insight into the main risk factors determining the probability of CSFV introduction (P_{CSFV}) is needed. Therefore, a scenario tree model was developed that calculates the annual P_{CSFV} into the Netherlands are (De Vos et al., 2003). The main aim of this model is to provide a better understanding of the importance of the different introduction routes for CSFV.

Only limited data was available to quantify all input parameters required for this scenario tree model. Furthermore, data obtained from experiments or historic CSF epidemics are limited in their use, due to, for example, low frequency of epidemics, differences in virus strains, and changes in preventive measures and control strategies used. Hence, the model contains many uncertain input parameters. In the default calculations, point estimates were used for these input parameters. Using different point estimates for these uncertain input parameters may, however, have large impact on model results and, consequently, the conclusions drawn. The model outcome of most interest for decision support is the ranking of risk factors contributing to the annual $P_{CSFV.}$ If using other values for uncertain input parameters results in a different ranking, model outcome is not robust and might result in different decisions when prioritising preventive measures.

The main goal of the current study was to investigate which of the uncertain input parameters influence model results most and thus require further (empirical) research. For this purpose an extensive sensitivity analysis was performed.

Material and methods

A statistical approach was used based on the techniques of Design of Experiments (DOE) and metamodelling (Kleijnen, 1998). The model contained 257 uncertain input parameters that were all selected as factors (X) for the modelling experiment. Each factor was assigned a low and a high value for the experiment[1]. The response variable (Y) chosen was the annual P_{CSFV} as this output parameter summarises model results. To reduce the number of factors in the modelling experiment, group screening was applied (Watson, 1961). Factors were grouped taking into account the risk factor or country of origin they were linked to and their expected impact on model outcome.

Results of the modelling experiment were used to specify a regression metamodel with the groups as independent variables and the annual P_{CSFV} as dependent variable. The metamodel was specified as the following simple first-order polynomial with k groups:

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

where $\underline{y_i}$ denotes the value of the response variable in scenario *i*, β_0 the intercept, β_h the main effect of group *h*, $x_{i,h}$ the standardised value of group *h* in scenario *i*, and $\underline{e_i}$ the approximation error plus intrinsic noise in scenario *i*. To fit this first-order polynomial to data from the modelling experiment a logit transformation was applied to the response variable, i.e. the annual P_{CSFV}, to project the interval (0,1) at the interval $(-\infty, \infty)$ and enable linear regression (Rothman and Greenland, 1998). Then, ordinary least squares (OLS) regression was used, performing a stepwise selection procedure with $p \leq 0.20$. The fit of the metamodel was evaluated by the R^2_{adj}.

Next, the groups with significant main effects in this metamodel were subdivided into smaller groups, with which a second modelling experiment was conducted. Again a metamodel was specified using OLS ($p \leq 0.20$). These steps were repeated until the number of factors in the groups with significant main effects was sufficiently small to run a modelling experiment with individual factors. The results of this modelling experiment were used to specify the final metamodel.

The factors with significant main effects in the final metamodel were ultimately included in a one-at-the-time (OAT) sensitivity analysis to investigate their impact on the ranking of risk factors. The OAT design contained two scenarios for each factor. Each factor was assigned its low value in one scenario, its high value in a second scenario, and its default value in all other scenarios. The ranking of risk factors in the scenarios of this OAT experiment was compared with the ranking obtained in the original default calculations.

[1] A complete overview of the factors in the experimental design and their default, low, and high values is available on request from the corresponding author.

Results

Three experiments were needed to obtain the final metamodel containing only 6 out of the 257 uncertain input parameters (Table 1). The R^2_{adj} of this regression metamodel was 0.757. The factor estimates (ß) in Table 1 reflect the expected effect on the logit of the annual P_{CSFV} when changing a factor from its low to its high value.

Factor	Description	ß	S.E.	p-value
X_0	Intercept	-3.332	0.137	0.000
X_1	Expected number of CSF epidemics per year in Germany	0.700	0.137	0.000
X_{16}	Expected number of CSF epidemics per year in Belgium	0.426	0.137	0.005
X_{31}	Expected number of CSF epidemics per year in the United Kingdom	0.986	0.137	0.000
X_{245}	Probability of an infective dose of CSFV being transmitted from a contaminated livestock truck to a susceptible pig	0.252	0.137	0.078
X_{246}	Probability that CSFV survives in an empty livestock truck travelling over a distance of 0-900 km	0.387	0.137	0.009
X_{247}	Probability that CSFV survives in an empty livestock truck travelling over a distance of 901-1800 km	0.269	0.137	0.061

Table 1. Significant factor effects ($p \leq 0.20$) in the final metamodel; dependent variable is the logit of the annual $P_{CSFV.}$

The OAT sensitivity analysis was performed with the six factors that had significant main effects in the final metamodel and hence consisted of 12 scenarios that were compared with the default. In Fig. 1, for each scenario the relative contribution of risk factors to the annual P_{CSFV} into the Netherlands is shown. The ranking of risk factors was especially changed by a high expected number of CSF epidemics per year in the United Kingdom (X_{31} high) and a low probability that CSFV survives in an empty livestock truck travelling over a distance of 0-900 km (X_{246} low). The input parameters concerning the expected number of CSF epidemics per year in Germany and Belgium also influenced the ranking of risk factors, but to a lesser extent (X_1 and X_{16}).

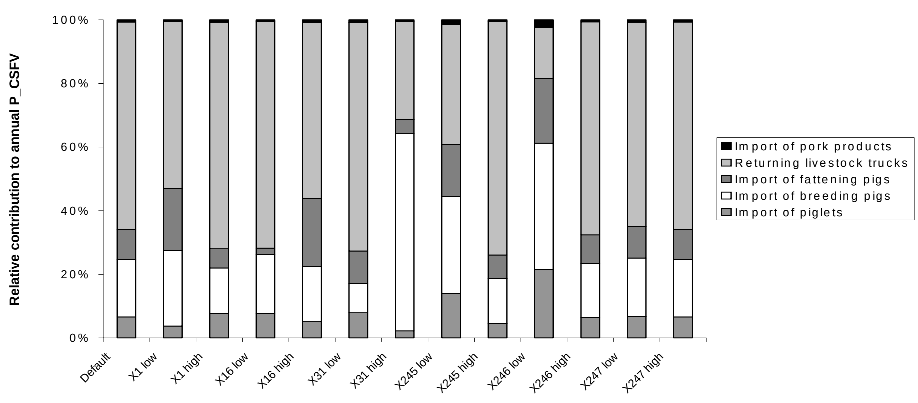

Fig. 1. Relative contribution of risk factors to the annual P_{CSFV} in the default calculations and the scenarios of the OAT design.

Discussion and conclusions

This study showed that group screening combined with the statistical techniques of DOE and metamodelling proved to be an effective and efficient screening method to identify those uncertain input parameters in the scenario tree model for CSFV introduction that influenced model results most. Only 128 scenario calculations were needed to specify the final metamodel. It was concluded that four uncertain input parameters change the ranking of risk factors contributing most to the annual P_{CSFV} and thus require a more accurate estimate of their values to make model outcome robust. For one of them, i.e. the probability of CSFV survival in an empty livestock truck travelling over a distance of 0-900 km, experiments can be conducted to estimate its value more precisely. For the others, i.e. the expected number of CSF epidemics per year in Germany, Belgium, and the United Kingdom, a more precise estimate can only be obtained by observing a longer time period.

References

De Vos, C.J., Saatkamp, H.W., Nielen, M., Huirne, R.B.M., 2003. Scenario tree modelling to analyse the probability of classical swine fever virus introduction into member states of the European Union. Risk Analysis, in press.

Kleijnen, J.P.C., 1998. Experimental design for sensitivity analysis, optimisation, and validation of simulation models. In: Banks, J. (Ed.), Handbook of Simulation. John Wiley & Sons, New York, pp. 173–223.

Rothman, K.J., Greenland, S., 1998. Modern epidemiology, Second edition, Lippincott-Raven Publishers, Philadelphia, 737 pp.

Watson, G.S., 1961. A study of the group screening method. Technometrics 3, 371-388.