Proceedings of the 2001 Winter Simulation Conference
B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, eds.

SENSITIVITY ANALYSIS OF CENSORED OUTPUT THROUGH POLYNOMIAL, LOGISTIC, AND TOBIT REGRESSION META-MODELS: THEORY AND CASE STUDY

Jack P.C. Kleijnen

Department of Information Systems/Center
for Economic Research (CentER)
School of Economics and Business Administration
Tilburg University
5000 LE Tilburg
THE NETHERLANDS

Antonie Vonk Noordegraaf
Mirjam Nielen

Department of Social Sciences
Farm Management Group
Wageningen University
6706 KN Wageningen
THE NETHERLANDS

ABSTRACT

This paper focuses on simulation output that may be censored; that is, the output has a limited range (examples are simulations that have as output the time to occurrence of a specific event - such as a 'rare' event - within a fixed time horizon). For sensitivity analysis of such simulations we discuss three alternatives: (i) traditional polynomial regression models, (ii) logistic or logit regression, and (iii) tobit analysis. The case study concerns the control of a specific animal disease (namely, IBR) in The Netherlands. The simulation experiment has 31 environmental factors or inputs, combined into 64 scenarios - each replicated twice. Traditional polynomial regression gives some estimated main effects with wrong signs. Logit regression correctly predicts whether simulation output is censored or not, for 92% of the scenarios. Tobit analysis does not give effects with wrong signs; it correctly predicts censoring, for 89% of the scenarios.

1 INTRODUCTION

Simulation analysts should always perform *sensitivity analysis*. We define such analysis as the systematic investigation of the reaction of the simulation responses to *extreme* values of the model's input or to *drastic* changes in the model's structure. (For example, what happens to the customers' mean waiting time when their arrival rate doubles; what happens if the priority rule is changed by introducing 'fast lanes'?) Such an analysis helps identify the most important factors in a simulation study. See Kleijnen (2000).

In the simulation literature it is well-known that in sensitivity analysis the gathering of simulation data should be guided by the statistical theory on the *design of experiments* (DOE). Indeed, DOE is a systematic method for specifying inputs for experimentation with the simulation model. See Kleijnen (1998).

Traditionally, DOE uses analysis of variance (ANOVA) or *polynomial regression models* to analyze the input/output (I/O) data of the experiment (be it an experiment with a real or a simulated system). In this paper we examine an important type of simulation models, namely models that generate *censored* outputs; that is, output that has a limited range. Well-known examples are (non-negative) waiting times in queueing simulations, and survival length in rare-event simulations with fixed-time horizons in which the rare event may or may not occur. Outside the simulation field, applications occur in biometrics, econometrics, engineering, etc. Amemiya (1984) states that in 1958 the econometrician Tobin published one of the first analysis of censored data.

Actually, it can be proven that ordinary least squares (OLS) analysis of censored data gives a *biased* estimator; see Amemiya (1984, pp.10-11) and Greene (1997, pp. 956, 963, 966).

Therefore we compare traditional OLS polynomial regression models with two alternatives, namely logistic or logit and tobit regression (tobit analysis has that name in honor of Tobin). These alternatives have never before been applied in simulation - to the best of our knowledge. Through a case study we shall illustrate that these alternatives may indeed be attractive.

Our case study concerns the control of animal diseases (namely, infectious bovine rhinotracheitis or IBR) in The

Netherlands. The current outbreaks of foot-and-mouth disease in Western Europe demonstrate the urgent need for national and international policies on animal health. To support these policies, simulation has already been applied extensively; see Horst et al. (1999). Obviously, these policies often involve risky and costly projects. Details are presented in Vonk Noordegraaf, Nielen, and Kleijnen (2001).

The main result of our case study is that the tobit analysis gives an acceptable metamodel of the underlying simulation model, whereas the traditional polynomial metamodel has some main effects with wrong signs.

We organize the remainder of this paper as follows. In §2 we summarize classic DOE and its concomitant regression analysis through polynomial models. In §3 and §4 we introduce logit and tobit regression respectively. In §5 we summarize our case study. In §6 we present conclusions, and propose future research topics.

2 CLASSIC DOE AND POLYNOMIAL REGRESSION

In DOE applied to simulation, a factor can be an input parameter, a variable, or a structural assumption (which implies a qualitative factor). A factor has at least two values or levels in the simulation experiment (set of simulation runs). A factor combination is the specific *scenario* that defines the input of a simulation run, which yields the output of that run. That output usually consists of multiple response types. In this paper, however, we focus on a single response type.

We do not discuss the various types of designs, but refer to Kleijnen (1998) and Kleijnen and Sargent (2000). Suffice it to say that in our case study we have 31 factors, each at two levels, together forming a set of $2^{31-25} = 64$ scenarios. Here the focus is on how to analyze these I/O data.

We denote the simulation's I/O data by (X, w) where X is an $N \times k$ matrix when there are k factors and N simulation runs. Actually some runs may use identical factor combinations x_i but m_i different (pseudo)random numbers: $N = \sum_{i=1}^{n} m_i$ where n denotes the number of different scenarios replicated m_i times. Hence, $x_i = (x_{i;1}, \ldots, x_{i;k})$ occurs m_i times in $\boldsymbol{X}$. The output is $w = (w_1, \ldots, w_N)'$.

The classic analysis in DOE uses the following *polynomial model* (ANOVA with fixed effects):

$$y_i = \beta_0 + \sum_{h=1}^{k} \beta_h x_{i;h} + \sum_{h=1}^{k} \beta_{h;h'} x_{i;h} x_{i;h'} + \ldots + e_i = \boldsymbol{x}'_i \boldsymbol{\beta} + e_i \quad (1)$$

where y_i denotes the metamodel predictor of $E(w_i)$, the expected simulation output; β consists of the grand or overall mean β_0; the main effect of factor h, namely β_h; the two-factor interaction between the factors h and h' namely $\beta_{h;h'}$; the dots denote higher-order interactions (which play an important role in classic ANOVA; we ignore these interactions because they are hard to interpret); e denotes *white noise*; that is, e is normally, independently and identically distributed (NIID) with zero mean. We denote the variance of e by σ^2. Note that e captures both the intrinsic simulation noise (caused by the use of random numbers) plus the lack of fit (approximation error) of the regression metamodel.

Because the noise is IID, the *best linear unbiased estimator* (BLUE) of the regression parameters β in Equation (1) is given by OLS:

$$\hat{\beta} = (X'X)^{-1} X'w \quad . \quad (2)$$

The OLS estimator in Equation (2) has the following covariance matrix:

$$cov(\hat{\beta}) = (X'X)^{-1} \sigma^2 \quad . \quad (3)$$

There are two methods for estimating the variance σ^2 in Equation (3). If $m_i > 1$ (as is the case in most simulations, including our case study), then we may use the *pooled variance estimator*

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} \hat{\sigma}_i^2}{n} = \frac{\sum_{i=1}^{n} (w_{i;r} - \bar{w}_i)^2}{n(m-1)} \quad (4)$$

where $\bar{w}_i = \sum_{r=1}^{m} w_{i;r}/m$ and we assume that the number of replicates is a constant m, for simplicity of presentation (in the case study, $m = 2$). Note that in our case study, some scenarios give zero estimated variances, but this does not make the pooled estimator zero.

Often, however, practitioners - especially when using standard statistical software (as we do: see SPSS 1999) - use the *mean squared residuals* (MSR):

$$MSR = \frac{\sum_{i=1}^{n} (\bar{w}_i - \hat{y}_i)^2}{n - q} \quad (5)$$

where q denotes the number of regression parameters in $\hat{\beta}$ and $n > q$. Note that this MSR has expected value $var(\bar{w})$

$= var(w)/m$ if and only if the regression model has no lack of fit. If the regression model, however, is not specified correctly, then MSR *overestimates* the variance, so relatively important factor effects have a higher probability of being declared non-significant.

To test the effect's *significance*, practitioners assume normally distributed simulation output *w*. This yields the well-known *t* statistic:

$$t_{\nu} = \frac{\hat{\beta}_j}{\sqrt{\hat{var}(\hat{\beta}_j)}} \quad (j = 1, \ldots, q) \tag{6}$$

where $\hat{\beta}_j$ follows from Equation (2) and $\hat{var}(\hat{\beta}_j)$ follows from Equation (3) combined with either Equation (4) or Equation (5), which implies $\nu = n(m - 1)$ and $\nu = n - q$ respectively.

Under the normality assumption, the OLS estimator is also the maximum likelihood (ML) estimator. We shall return to ML below.

Actually most simulation practitioners (including us) use *common random numbers* (CRN), so the white noise assumption is violated. In fact, the *n* responses are correlated (hopefully, positively correlated to decrease the variances of the estimated main effects). Obviously CRN implies $m_i = m$. If there are enough replicates ($m > n$), then we could estimate these correlations, and replace OLS by generalized LS (or GLS) to obtain BLUE. In our case study, however, we have only two replicates, because each replicate requires much computer time.

Obviously, we should *validate* the assumed metamodel. So we check whether the polynomial model in Equation (1) is an adequate predictor of $E(w)$. To save computer time, we can use *cross-validation*, as follows.

We temporarily remove the i^{th} I/O combination (we delete all *m* replicates of that combination), and estimate the polynomial model from the remaining I/O data, which yields (say) $\hat{\beta}_{-i}$ with $i = 1, \ldots, n$. This estimate we use to compute the predictor $\hat{y}_i = x'_i \hat{\boldsymbol{\beta}}_{-i}$. This we repeat for each of the *n* input combinations. Finally, we make a scatter plot of these *n* predictors $\hat{y}_i$ versus the *n* corresponding average simulation responses $\overline{w}_i$. This plot should show an estimated Pearson linear correlation coefficient (say) ρ close to one.

3 LOGIT REGRESSION

The 'rare event' literature focuses on methods for improving the statistical accuracy when estimating the probability of a specific rare event happening within a certain time frame. We, however, emphasize that - once such a probability is estimated - the simulation analysts should try to identify the most important factors that affect that probability. In this section we propose logit regression models for such a sensitivity analysis (in §5 we shall present a case study that concerns a 'not so rare' event).

For logit regression, the original simulation output *w* (defined in §2) is changed into the *binary* variable w^*. In our case study, *w* denotes the time it takes for a specific event to occur. We transform *w* to 1 if for *w* the censoring event does occur, and to 0 if not:

$$\begin{aligned} w^* &= 1 \ \textit{if} \ w = c \\ w^* &= 0 \ \textit{if} \ w < c \end{aligned} \tag{7}$$

where in the case study we set $c = 1{,}000$. Logit regression models uses the regression dependent variable *y* to predict $P(w^* = 1) = E(w^*)$:

$$y = \frac{e^{\beta' x}}{1 + e^{\beta' x}} = \frac{1}{1 + e^{-\beta' x}} \tag{8}$$

so that $0 \le y \le 1$; see Greene (1997, p. 874), Hosmer and Lemeshow (1989, p. 6), Long (1996, p. 49), and SPSS (1996, p. 37).

Estimation of the effects β in Equation (8) uses ML, instead of OLS. Unfortunately, no explicit formula for the ML estimator of the factor effects are available: ML requires computerized iterative search; see Greene (1997, pp. 173-219) and Long (1997, pp. 54-61).

However, it is well-known that in general, ML estimators have asymptotic normal distributions with mean β and covariance matrix

$$cov(\hat{\boldsymbol{\beta}}) = -E[H(\boldsymbol{\beta})]^{-1} \tag{9}$$

where *H* denotes the Hessian matrix with the second-order derivatives of the log-likelihood function $\delta^2 lnL(\boldsymbol{\beta})/\delta\boldsymbol{\beta}\delta\boldsymbol{\beta}'$; see Long (1997, p. 32, 58) and also Amemiya (1984, p. 17) and Greene (1997, p. 966). To compute Equation (9) in the case study, we shall use SPSS (1999)'s binary logit regression procedure (other software is mentioned by Hosmer and Lemeshow 1989).

Combining Hosmer and Lemeshow (1989, pp. 82-89) and SPSS (1999), we decide to use the following model building procedure.

(i) Start with univariate analysis of each independent variable *x*. Only variables with a *p* value below 0.25 are selected for the multivariate logit model; that *p* is based on Wald's statistic, which has a chi-square distribution (with degrees of freedom equal to the number of constraints tested; also see Long, 1997, pp. 87-93).

(ii) Next, perform backwards elimination of main effects. Follow up by testing those interactions thought to be relevant; that is, those interactions suggested by knowledge of the real system being

simulated. Significance testing with $p < 0.05$ is based on the change in the -2log-likelihood statistic, which has a chi-square distribution (also see Long, 1997, p. 109).

(iii) The fit of the resulting logit model is evaluated through Nagelkerke's R^2 statistic and the fraction of scenarios classified correctly. That R^2 is defined through the likelihood function; it has the same interpretation as the regular R^2; see Nagelkerke (1991) and SPSS (1999, p.46) (also see Long, 1997, pp. 102-109 and Hosmer and Lemeshow 1989, pp. 135-175). Logistic regression classifies a predicted simulation output as censored when $\hat{y} > 0.50$, for a particular scenario; see SPSS (1999, p. 39); we might call 0.50 the watershed value. Next we consider the actual simulation w, and check if it is correctly classified: $w = c$? (See equation 7.)

We shall illustrate this procedure through our case study in §5.

4 TOBIT ANALYSIS

Logit regression throws away much information when transforming the original simulation output w into a binary variable w^* through Equation (7). Tobit analysis, however, uses the following transformation of the so-called *latent* variable w^*:

$$\begin{aligned} w &= w^* \text{ if } w^* < c \\ w &= c \text{ if } w^* \geq c \; . \end{aligned} \tag{10}$$

Note that the precise value of the latent variable cannot be observed when censoring occurs. This latent variable w^* is predicted by (say) $y^* = \beta' \boldsymbol{x} + e$, which equals Equation (1) if y is replaced by y^*; see Greene (1997, p. 962) and Long (1997, pp. 196, 211).

To compute the ML estimator of β in the case study, we use LIMDEP 7.0; see Long (1997, pp. 204-206) and also Greene (1997, p. 191).

The model building procedure applied for tobit regression is similar to the one for traditional polynomial regression, namely stepwise selection of main effects and subsequent testing for interactions; see Long (1997, pp. 206-208).

5 CASE STUDY: NATIONAL ANIMAL DISEASE CONTROL

Our case study concerns the simulation of a national program for IBR eradication, which should lead to Dutch cattle farms *free* of IBR.

We simulate outbreaks and control of infections, per week. Inputs include vaccination parameters; outputs include costs and epidemiogical results such as number of outbreaks.

In the sensitivity analysis of this simulation we distinguish 31 environmental factors related to the spread of the IBR virus within and among farms. Five of these factors are qualitative (they concern the distribution function type). Each factor has two extreme levels, standardized as 0 and 1 respectively so the relative importance of the factors is quantified by the regression effects β; see Bettonvil and Kleijnen (1990) for polynomial regression and Long (1997, pp. 61-82) for logit analysis. These 31 factors and their levels are detailed by Vonk Noordegraaf et al. (2000).

As the design for these $k = 31$ factors we select a 2^{31-25} (so-called resolution-4) design, so we simulate $n = 64$ scenarios. Even though we simulate only $m = 2$ replications, the total computer time is almost two weeks while we use five PCs - with 533 MHZ clock speed - in parallel.

Whereas the simulation model generates multiple outputs, we focus on a single output, namely the number of weeks needed to reach a *prevalence level* of 5% in the national dairy cattle population. The simulation run terminates whenever that level drops below 5%, or whenever the simulated period reaches 1,000 weeks. In other words, the output w is *censored* at 1,000 weeks.

For each scenario, we take the average output of two replications as the output. This gives the same OLS estimate as taking the individual outputs; see Kleijnen (1987, p. 195).

Now we present the results of this case study for the three alternative regression metamodels.

5.1 Polynomial Regression

As we explained above, we start with a first-order regression model; that is, an ANOVA model without interactions. We compute the OLS estimates of these effects, and stepwise eliminate those factors that have no significant main effects; see Equation (6): backwards elimination. Next we decide on the addition of two-factor interactions between the factors that remain after the backwards elimination. We use SPSS software, and a significance level of 5%. We evaluate the resulting polynomial through R^2 adjusted for the number of effects, and $\hat{\rho}$ based on cross-validation. Our results are as follows.

Of the 31 factors, 11 factors give significant main effects (for details see Vonk Noordegraaf et al. 2001). Moreover, 3 two-factor interactions are significant too. These interactions increase the adjusted R^2 from 0.72 to 0.82. *Unfortunately, two main effects have signs that conflict with prior expert knowledge.*

Cross-validation gives a scatter plot with $\hat{\rho} = 0.97$. However, the simulation gives 23 out of 64 outputs that are censored at 1,000 weeks (so w =1000), whereas *the*

polynomial predicts outputs that exceed this censoring limit for these scenarios (so $\hat{y} > 1000$)

So the polynomial metamodel does not adequately approximate the behavior of the underlying simulation model. The explanation of this undesirable result may be that OLS gives biased estimators for censored data. (In this case-study, the metamodel may adequately approximate the behavior of the real system: because of computer constraints the simulation stops after 1,000 weeks, whereas the real system may reach the 5% prevalence level after more than 1,000 weeks. Kleijnen and Sargent (2000) discusses the validation of metamodels against both the simulated and the real systems.)

5.2 Logit Analysis

The logit model predicts the probability of censoring the simulation output; see Equations (7) and (8). In our case study we find that this model has only six significant main effects and no interactions. These six effects form a subset of the eleven main effects in the polynomial model, which also had three significant interactions.

The fit of the resulting model is evaluated through Nagelkerke's R^2 statistic - which turns out to be 0.81 - and the fraction of correctly classified scenarios. That fraction turns out to be 92.2%. More precisely: of the 23 censored scenarios, 21 are classified correctly. Of the 41 uncensored scenarios, 38 are classified correctly. (Note that each scenario is replicated twice. The 23 'censored' scenarios give replicated outputs that both are 1,000. The 41 'uncensored' scenarios include two scenarios that give one censored and one non-censored output, so its average is smaller than 1000 and the scenario is not considered censored.)

5.3 Tobit Analysis

For our case study we find that the tobit model in Equation (10) has the same significant main effects and interactions as the polynomial had, *except* for the two main effects with *wrong signs* in the traditional polynomial (so the tobit model has nine instead of eleven main effects).

To validate this metamodel, we do not use cross-validation: the software does not provide this facility. (Of course, manual calculations would have been possible, but tedious.) Instead, we plot the simulated versus the predicted values, which gives $\hat{\rho} = 0.91$ or $R^2 = 0.83$ (R^2 was 0.82 for the polynomial model).

Analogous to logit regression, tobit analysis shows the probability of each scenario being censored: If the tobit output $\hat{y}^*$ exceeds the threshold $c = 1{,}000$, then we predict censoring. (Actually, the LIMDEP software gives both this $\hat{y}^*$ and the estimated probability of censoring; the latter probability may be compared with 0.5 to classify the scenario. We prefer the first approach, since it does not require a watershed value. Fortunately, both approaches give identical results.) The overall fraction of scenarios correctly classified is 89.1%. Of the 23 censored scenarios, 16 are classified correctly. All 41 uncensored scenarios are classified correctly!

6 CONCLUSIONS AND FUTURE RESEARCH

Our case study with censored simulation output showed that polynomial regression may give significant main effects with the *wrong signs* - even though the cross-validation's $\hat{\rho}$ and the classic R^2 are acceptable. Our explanation is that OLS gives biased estimators in case of censored data. And indeed, censoring occurred for 23 of the 64 simulated scenarios.

Regression techniques more suitable for censored data are logit and tobit regression.

Logit regression gives information on those factors that significantly impact the probability of a specific event - in our case study that event is censoring a simulation output. The fraction of correctly classified scenarios was high, namely 92%.

This logit regression is appropriate if we are interested only in a *binary probability* such as censoring or non-censoring (or a rare event happening or not happening). Tobit regression, however, gives more information: it also estimates factor effects on the *non-censored continuous* simulation output.

In our case study, tobit analysis gave a correctly predicted fraction nearly as high as logit analysis gave, namely 89%. Compared with OLS, tobit regression did not contain the two factors with wrong signs. Altogether we consider the tobit model to be the valid metamodel of our underlying simulation model with censored continuous output.

In *future research* the following issues may be addressed.

Rare event simulations have ignored sensitivity analysis. Logit analysis deserves further research.

Because waiting times cannot be negative, tobit analysis of queueing simulations deserve further exploration.

Further, nonnormality of the output gives a biased ML estimator; see Amemiya (1984, p. 25) and Greene (1997, p. 971). In our case study we take the average of two replicates, so possible nonnormality is reduced. Nevertheless, in general this problem may be further investigated.

Different scenarios may give different variances - not only different means. Such heteroscedasticity is discussed in Greene (1997, p. 967). Also remember our discussion (in §2) of the two estimators of the variance (namely, the pooled and the MSR estimators).

We ignore possible effects of CRN in our analysis of polynomial, logit, and tobit analyses. Obviously, CRN does affect the likelihood function.

There are more alternatives besides logit and tobit regression models; for example, survival analysis and the

generalized linear model (GLM); see Greene (1997, pp. 984-999) and Long (1997, p. 257).

A final general issue is the model building procedure: regression modeling is an art. The final regression model should be statistically 'optimal' (for example, give minimum prediction errors) and - more important - should be acceptable to the users (industrial engineers, economists, managers, etc.).

REFERENCES

Amemiya, T. (1984), Tobit models: a survey. *Journal of Econometrics* 24: 3-61

Bettonvil, B. and J.PC. Kleijnen (1990), Measurement scales and resolution IV designs. *American Journal of Mathematical and Management Sciences* 10: 309-322

Greene, W.H. (1997), *Econometric analysis; third edition*. Upper Saddle River, New Jersey: Prentice-Hall International

Horst, H.S., A.A. Dijkhuizen, R.B.M. Huirne, and M.P.M. Meuwissen (1999), Monte Carlo simulation of virus introduction into the Netherlands. *Preventive Veterinary Medicine* 41, 209-229

Hosmer, D.W. and S. Lemeshow (1989), *Applied logistic regression*. New York: Wiley

Kleijnen, J.P.C. (2000), Strategic directions in verification, validation, and accreditation research: a personal view. In *Proceedings of the 2000 Winter Simulation Conference*, ed. J.A. Joines, R.R. Barton, K Kang, and P.A. Fishwick, 909-916, Piscataway, New Jersey, Institute of Electrical and Electronic Engineers

--- (1998), Experimental design for sensitivity analysis, optimization, and validation of simulation models. In: *Handbook of Simulation*, ed. J. Banks, New York: Wiley

--- (1987), *Statistical tools for simulation practitioners*. New York: Marcel Dekker

--- and R.G. Sargent (2000), A methodology for the fitting and validation of metamodels in simulation. *European Journal of Operational Research*, 120: 14-29

Long, J.S. (1997), *Regression models for categorical and limited dependent variables*. Thousand Oaks: Sage

Nagelkerke, N.J.D. (1991), A note of general definition of the coefficient of determination, *Biometrika*, 78: 691-692

SPSS (1999), *Regression Model TM 10.0.5*, Chicago: SPSS

Vonk Noordegraaf, A., M. Nielen, and J.P.C. Kleijnen (2001), Sensitivity analysis by experimental design and metamodeling: case study on simulation in national animal disease control. Working Paper submitted for publication

AUTHOR BIOGRAPHIES

JACK P.C. KLEIJNEN is a Professor of Simulation and Information Systems. His research concerns simulation, mathematical statistics, information systems, and logistics; this research resulted in six books and nearly 160 articles. He has been a consultant for several organizations in the USA and Europe, and has served on many international editorial boards and scientific committees. He spent several years in the USA, at both universities and companies, and received a number of international fellowships and awards. His e-mail and web address are <kleijnen@kub.nl> and <http://www.tilburguniversity.nl/faculties/few/im/staff/kleijnen/>.

ANTONIE VONK NOORDEGRAAF finished his Master's in Animal Science at the Wageningen Agricultural University in 1998. He is working now at the Farm Management Group of Wageningen University, on a Ph.D. project entitled 'Simulation modeling to support policy making in the control of BHV1 in The Netherlands'. His e-mail address and web address are <Antonie.VonkNoordegraaf@Alg.ABE.WAU.NL> and <www.sls.wageningen-ur.nl/abe/>.

MIRJAM NIELEN is a veterinarian by training, who works as a veterinary epidemiologist within the Animal Health Economics unit of the Farm Management Group of Wageningen University. She is involved in various projects on decision support for the control of diseases in farm animals; these diseases may cause losses at the levels of the individual farm, the industry, or the country. Her recent research focuses on diseases with control at the national level, such as classical swine fever and foot and mouth disease. Her e-mail address and web address are <Mirjam.Nielen@Alg.ABE.WAU.NL> and <www.sls.wageningen-ur.nl/abe/>.