TWC/14/12
ORIGINAL: English
DATE: June 4, 1996

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Fourteenth Session
Hanover, Germany, June 4 to 6, 1996

THRESHOLD MODELS FOR VISUALLY-OBSERVED DATA

Document prepared by experts from the Netherlands

Threshold models for visually observed data

F.A. van Eeuwijk & J. Jansen
CPRO-DLO
P.O. Box 16
6700 AA Wageningen
The Netherlands
e-mail: f.a.vaneeuwijk@cpro.dlo.nl

Introduction

In DUS testing it is common to distinguish 'visually observed' characteristics from 'measured' characteristics. Measured characteristics are usually assumed to be normally (or at least symmetrically) distributed, which should make their analysis straightforward as standard linear model theory can be used. The analysis of distinctness known as COY-D is an example of the application of standard linear model theory, while the analysis of uniformity as exemplified in the COY-U procedure, though less firmly grounded in the same theory, is at least clearly inspired by that theory. So far, there are for visually observed characteristics no clear analogons to COY-D and COY-U. It is possible to use analysis of variance methods on rankings, but it is not always that clear what should be the interpretation of such analyses. Alternatively, non-parametric methods may be used. However, with these methods it is difficult to arrive at a degree of sophistication comparable to that of the COY-D and COY-U.

Within the visually observed characteristics those of the so-called ordinal type are the most prevalent ones in DUS testing. Ordinal are those data that are observed in categories that can be ranked, but without the distances between categories having a binding interpretation. For example, the categories are ranked from bad to good on a 1 to 9 scale, without the implication that a difference between the categories 1 and 2 has the same meaning as a difference between the categories 5 and 6.

The aim of this paper is the description of a class of models especially developed for ordinal data: *threshold models*. These models allow the same kind of questions to be answered for ordinal data as standard linear models do for measured characteristics. In a sense threshold models are even more general than standard linear models as they provide a very natural means to model expectation and variance simultaneously, thus effectively combining the modelling of distinctness and uniformity. By writing threshold models in the form of so-called generalized linear models, estimation and testing procedures developed in the context of this latter class of models carry over to the class of threshold models. In the remainder of this paer we will first describe the general features of *generalized linear models*, then threshold models will be introduced in their own terms, and subsequently it is indicated how a threshold model can be formulated in terms of a generalized linear model. The use of threshold models will be illustrated in two examples. Some comments about literature and software conclude the paper.

Generalized Linear Models

The class of generalized linear models is an extension of the class of classical linear models. We start with briefly reconsidering the general structure of classical linear models. For standard linear models (ANOVA and regression) there is a vector of independent observation, **y**, containing the elements $y_1, y_2, \ldots, y_n$, that is a realization of the stochastic vector **Y**. The expectation of the element i of the stochastic vector is $E(Y_i) = \mu_i$. The components of **Y** are independently normally distributed with constant variance, σ^2. The structural part of the classical linear model specifies the expectation, μ_i, as a function of a number of regression parameters, $\beta_1, \beta_2, \ldots, \beta_p$, in combination with known covariates $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p$, to give $E(Y_i) = \mu_i = \beta^T x_i$, with β the vector of regresion parameters and $\mathbf{x}_i$ denoting the vector of covariate values for the unit i. The class of generalized linear models can be created from the class of linear models by, firstly, allowing other distributions than the normal for the vector **Y**. These distributions should be members of the exponential family, that is, besides the normal distibution, other distributions are allowed like the binomial distribution (for percentages and fractions) and the Poisson distribution (for counts). Secondly, a so-called link function, g(.), is introduced that connects the linear predictor, $\eta = \beta^T\mathbf{x}$, with the expectation, μ: $g(\mu) = \eta = \beta^T\mathbf{x}$. (For convenience we will often drop the index i). For the classical linear model the link function g(.) is equal to the identity function. For observations in percentages and fractions popular links are the logit, $\eta = \log\{\mu/(1-\mu)\}$, and the probit, $\eta = \Phi^{-1}(\mu)$, with Φ the Normal cumulative distribution function. Use of the logit, or the probit, guarantees that the predicted value for the probability or fraction π_i will be between 0 and 1. For example, if for a certain experiment the probability of success, $P(Y_i=1) = \pi_i$, is modelled with a logit model then $\pi_i = \exp(\beta^T x/(1+\exp(\beta^T x))$. For counts, in contingency tables, the log link is a standard choice, $\eta = \log(\mu)$. Again, predicted values are restricted to the 'natural' range. Counts cannot be negative, and because $\eta = \log(\mu)$, $\mu = \exp(\eta)$, which is always non-negative.

Parameters for generalized linear models are estimated by maximum likelihood procedures. A special feature of generalized linear models is that the maximum likelihood estimation can be rewritten to an iterative weighted least squares procedure, so that the calculations can be done by any programme for linear regressions with facilities for the inclusion of weights. The dependent variate in these regressions is the so-called link- adjusted dependent variate **z**, while the weights are just the inverses of the variances of the components of **z**. **z** is a linearized form of the link function to y; $z = g(\mu) + (y-\mu)(\delta\eta/\delta\mu)$, or, $z = \eta + (y-\mu)(\delta\eta/\delta\mu)$. If Z is the stochastic variable corresponding to the realized value z, then the variance of Z, $\mathrm{var}(Z) = \mathrm{var}(Y)(\delta\eta/\delta\mu)^2$, and the weights for the iterative regressions will be 1/var(Z). Each iteration, **z** is calculated using updated values for η and μ, while the weigths are likewise updated. The iterations continue untill some convergence criterion has been met. Standard errors for the parameters can be obtained in the usual way for maximum likelihood procedures from the inverse of the matrix of second derivatives of the likelihood to the parameters. Testing of treatment terms for inclusion in a final model can be done with procedures that are comparable to those used in standard linear model theory. We will use so-called Wald statistics, i.e. quadratic forms in which the sum of squares for a treatment term is divided by the variance of the parameter estimates. These quantities should be approximately Chi-squared distributed with degrees of freedom equal to the number of independent parameters. For testing differences between individual treatments t-tests can be used.

Threshold models for ordinal data

We will think of ordinal data as arranged in two-way tables, with the columns indicating the categories and the rows units to which treatments have been applied. For examples see Table 1 and 3.

Threshold models assume that there is an underlying variable, U_i, with i referring to the treatment, which cannot be observed and that determines in which category an individual observation will lie. The categories are numbered 1 to C, and they are separated by thresholds, θ_1 to θ_{C-1}. The relationship between the value of the underlying variable U_i and the response categories is

Interval		*Category*
$-\infty < u_i \leq \theta_1$	$\rightarrow$	1
$\theta_{c-1} < u_i \leq \theta_c$	$\rightarrow$	$c = 2 \ldots C-1$
$\theta_{C-1} < u_i \leq \infty$	$\rightarrow$	C.

The underlying variable U_i is just a theoretical construct. What is actually observed is the discrete response category, y_i, which takes on the values 1 to C, and what we would like to know is the probability that an individual with the treatment i will show a response in the category c, or, $\pi_{ic} = P(Y_i = c)$. Note that the expected value in the response category c for treatment i is $N_i\pi_{ic}$, with N_i the total number of individuals having received treatment i. The probability π_{ic} is given by $P(\theta_{c-1} < U_i < \theta_c)$. The underlying response, U_i, is thus assumed to follow a linear model, $U_i = \beta^T \mathbf{x}_i + \epsilon_i$, where the vector $\mathbf{x}_i$ contains the values of the explanatory variables corresponding to treatment i. For the error distribution various choices can be made, like the standard normal and logistic distribution. Besides a linear model formulation, with one error term, there exists the possibility of formulating a linear mixed model for U_i, with more than one error term. Taking F(.) to be the distribution of U_i,

$$\pi_{ic} = P(\theta_{c-1} < U_i < \theta_c) = F(\theta_c - \beta^T \mathbf{x}_i) - F(\theta_{c-1} - \beta^T \mathbf{x}_i)$$

Thus, for finding the response category probabilities, π_{ic}, using a threshold model, it is necessary to choose an adequate distribution F(.) for the underlying variable U_i, and then to estimate the thresholds (cutpoints), θ_c, and the regression parameters, β. A reasonable choice for F(.) is often the normal distribution, implying that $U_i - \beta^T \mathbf{x}_i$ has a standard normal distribution. Every treatment i receives its own normal distribution whose location on the underlying scale is determined by its mean, $\beta^T \mathbf{x}_i$. The width (dispersion) of the distributions can be chosen to be independent or dependent on the treatments, according to the structure of the data. Location and width of a treatment distribution determine together with the positions of the cutpoints the probabilities of the response categories for a treatment i. Or, differences in the 'response pattern' between treatments are attributed to differences in location and width of the underlying distribution. These ideas are graphically illustrated in Figure 1. More formally this is expressed by

Category	*Probability*
1	$\Phi\left[\dfrac{\theta_1-\mu_i}{\sigma_i}\right]$
$c = 2 \ldots C-1$	$\Phi\left[\dfrac{\theta_c-\mu_i}{\sigma_i}\right] - \Phi\left[\dfrac{\theta_{c-1}-\mu_i}{\sigma_i}\right]$
C	$1 - \Phi\left[\dfrac{\theta_{C-1}-\mu_i}{\sigma_i}\right]$,

where Φ represents the cumulative normal distribution function.

Threshold models do not automatically fall within the generalized linear model framework. However, with some effort they can be brought within that family of models, after which the theory on generalized linear models becomes applicable.

Threshold models as generalized linear models

In standard generalized linear models there is a direct, one-to-one correspondence between an observation y_i and the linear predictor η_i. The threshold model as formulated above in a sense relates every observation to a difference between *two* linear predictors, as is evident from a reconsideration of $\pi_{ic} = P(\theta_{c-1} < U_i < \theta_c) = F(\theta_c - \beta^T x_i) - F(\theta_{c-1} - \beta^T x_i)$. Two ways have been proposed to bring threshold models within the class of generalized linear models. The simplest solution to this problem is not to consider the category probabilities, π_{ic}, but the cumulative probabilities γ_{ic}, where $\gamma_{ic} = P(Y_i \leq c)$. A popular model for the cumulative probabilities is the proportional-odds model

$$\log[\gamma_{ic}(x_i)/(1 - \gamma_{ic}(x_i))] = \theta_c - \beta^T x_i, \text{ for } c = 1,\ldots,C-1,$$

with $\gamma_{ic}(x_i) = P(Y_i \leq c \mid x_i)$. The proportional-odds model derives its name from the fact that the ratio of the odds of the event $Y_i \leq c$ for the pair of treatments $\mathbf{x}_1$ and $\mathbf{x}_2$ is independent of the category:

$$[\gamma_{ic}(\mathbf{x}_1)/(1 - \gamma_{ic}(\mathbf{x}_1))] \;/\; [\gamma_{ic}(\mathbf{x}_2)/(1 - \gamma_{ic}(\mathbf{x}_2))] = \exp(\beta^T(\mathbf{x}_1 - \mathbf{x}_2)).$$

The proportional-odds model has a logit link for the cumulative response probability. The distribution for this response can be derived from that of the multinomial distribution, which is the appropriate distribution for the response probabilities π_{ic}. In principle the stage is then set for the application of the general theory for generalized linear models.

An alternative way to bring threshold models within the class of generalized linear models uses the theory on composite link functions. Composite link functions allow the expectation, μ, to be connected to a function of more than one link function. To repeat, in standard generalized linear models $g(\mu) = \eta$, the link function g(.) connects the expectation, μ, with the linear predictor, $\eta = \beta^T x$. The inverse of the link, h(.), is defined by $\mu = h(\eta)$. Assume now that there are more linear predictors, for example two, η_1 and η_2, and define $\kappa_1 = h_1(\eta_1)$ and $\kappa_2 = h_2(\eta_2)$. Then we can imagine the expectation, μ, to be a function of κ_1 and κ_2. For example, $\mu = \kappa_1 - \kappa_2$. In the framework of the threshold model we suggest

$$\pi_{ic} = \kappa_{i,c} - \kappa_{i,c-1} = \Phi(\theta_c - \beta^T x_i) - \Phi(\theta_{c-1} - \beta^T x_i).$$

Using this result a working variate and weight function can be derived that form the basis for an iterative reweighted linear regression, and once again threshold models are brought within the class of generalized linear models.

Example 1; Number of internodes in maize

In document TWC/13/3 some methods are presented for the analysis of ordinal data. Table 2 of that document contains for 12 maize varieties the number of plants with a specified number of internodes, where the total number of plants per variety is either 15 or 30. These data are reproduced in the body of Table 1. The definition of the classes can be read from the top of the table. In TWC/13/3 these data were analysed to illustrate a proposed measure for the concentration (homogeneity) of the observations around a modal class(es). These data have been reanalysed using a threshold model. The underlying distribution was chosen to be the normal. A conclusion in TWC/13/3 was that considerable differences in concentration were present. Therefore, it seemed logical to start with a threshold model with a separate mean and dispersion for every variety. The adequacy of such a model should be checked. In general it is a good procedure to start with the most elaborate model for the dispersion and mean, and then first try to reduce the model for the dispersion. Subsequently, conditional on the model for the dispersion one can try to reduce the model for the mean.

For the analysis we used the Genstat procedure CLASS (see section on Literature and software). Within that procedure two identification constraints are used: (1) the first cutpoint is chosen to be equal to zero; (2) the dispersion for the first treatment (variety) is chosen equal to unity. The required directive is essentially;

```
CLASS [UDISTRIBUTION = Normal;  \
       FIXED         = Variety; \
       FDISPERSION   = Variety] data.
```

One of the first things to check after having fitted the model is whether the categories were well separated. Looking at the estimated cutpoints at the top of Table 1 it is clear that the choice of categories was alright, the distances between the cutpoints were significant. It can also be seen that the earlier categories were further apart than the later categories. This is important information, because it makes the application of an ANOVA disputable. The estimated values for the dispersions showed some significant differences (t-tests were done on the logarithms of the dispersion, only results are given, + and -), supporting the inclusion of a model for the dispersion (dependence on variety). It may be expected that there should be a relationship between the coefficient of concentration of TWC/13/3 and the dispersion as estimated in the threshold model. The correlation between both quantities was 0.73. Thus, the facility for modelling separate dispersions in the threshold model provides a convenient method for quantifying differences in concentration (homogeneity/uniformity). Given the model for the dispersion the differences between the means were tested. The Wald statistic for the variety main effects amounted to 27.7 on 11 degrees of freedom, which was significant. The differences between variety means can be assessed from the column for the estimated means and the standard error of a difference. Table 2 gives the fitted values from the threshold model. Comparison of Table 1 and 2 illustrates the quality of the fit. In addition, Table 2 exhibits the character of the threshold model: shifted (normal) distributions on an underlying scale, with additional differences in dispersion.

The usefulness of the threshold model in this example is illustrated by the natural way in which distinctness and uniformity can be modelled for the previously notoriously difficult to analyse visually observed characteristics of the ordinal type.

Example 2; Soil coverage in sugar beet

This example concerns visual assessments of soil coverage for 9 varieties of sugar beet over 3 years. The observations were made on a 1 to 9 scale (Table 3). As a first step in the analysis the model for the dispersion was investigated conditional on a complete model for the mean, i.e., main effects for year and variety plus the interaction of variety by year. The underlying distribution was chosen to be the normal distribution. No heterogeneity of variance was assessed and so the analysis was continued with homoscedasticity. Again the complete model for the mean was fitted. The Genstat directive was

```
CLASS [UDISTRIBUTION = Normal; \
       FIXED = Year+Variety+Variety.Year] data
```

Results from this analysis are given in Table 3 (cutpoints and means) and Table 4 (fitted values). The categories were well separated. Notable is that the categories were more or less equidistant, except for the last categories. An ANOVA for these data would be less inappropriate than for the maize data. The Wald statistics showed significance for all treatment terms:

Source	Wald Statistic	Degrees of Freedom
Year	12.7	2
Variety	122.4	8
Variety.Year	48.0	16

There was marked interaction between variety and year. Inspection of the variety by year means in Table 3 learns that this interaction was largely due to the varieties Hilton and Univers. Comparison of Table 3 and Table 4 shows that the threshold model with equal dispersion and a full model for the mean fitted adequately. From Table 4 the nature of the threshold model can be observed as a collection of shifted distributions with equal width. Although it is interesting to know that there was variety by year interaction, and one should try to find an explanation for that interaction, in DUS testing the mean over the three year period is still of most importance. These means are given in Table 5. Over the years Cordelia (3.56) had the highest soil coverage and Hilde the lowest (1.34), a range of 2.22. The standard error of a difference indicates that differences on the underlying scale should exceed 0.50 to be considered significant. Comparing this figure with the range, the threshold model can be said to be sufficiently distinctive for these data.

More analyses are possible, for example, by choosing the variety by year interaction to be a random term. We fitted such a model by

```
CLASS [UDISTRIBUTION = Normal;       \
       FIXED         = Year+Variety; \
       RANDOM        = Variety.Year] data
```

The variance component on the underlying scale for the variety by year interaction was 0.195 (+/- 0.101). The Wald statistics were

Source	Wald Statistic	Degrees of Freedom
Year	6.0	2
Variety	52.2	8

No qualitative differences appeared between the mixed model analysis and the above fixed model analysis. The Wald statistics are lower when the variety by year interaction is random, because the variances of the parameters increase due to the extra variance component. The standard error of a difference between variety means over the three years was 0.44 for the mixed model, while the differences between the variety means were hardly different from those for the fixed model (see Table 5).

This example shows how threshold models allow model formulations for ordinal data (on the underlying scale) that are completely analogous to those for measured characteristics, another illustration of the wide practical potential of these models.

Literature and software

The standard text for generalized linear models is: P. McCullagh and J.A. Nelder, 1989, Generalized linear models, 2nd edn., Chapman and Hall. A classical paper on ordinal data analysis is: P. McCullagh, 1980, Regression models for ordinal data, J. R. Statist. Soc. B, 42: 109-142. An elaborate treatise on threshold models can be found in: J. Jansen, 1993, Generalized linear mixed models and their application in plant breeding research, Ph. D. thesis Technical University Eindhoven. Further recommended reading: J. Jansen, 1992, Statistical analysis of threshold data from experiments with nested errors, Computational Statistics and Data Analysis, 13: 319-330; and A. Keen and B. Engel, Analysis of a mixed model for ordinal data by iterative re-weighted REML, Statistica Neerlandica, in press.

Due to the authors' bias towards Genstat (Genstat 5 Committee, 1993, Genstat 5 release 3 reference manual, Oxford University Publications) only reference was made to the facilities for ordinal data analysis in Genstat. In the present release there is a facility for ordinal regression within the standard generalized linear regression set up. In addition there is the Genstat procedure CLASS written by A. Keen (DLO - Agricultural Mathematics Group, Wageningen) that has wider facilities than the standard ordinal regression directives. CLASS offers the possibility for defining mixed models for the linear predictor. Estimation is based on restricted maximum likelihood (REML). Alternatively, there are Genstat programmes written by the second author, that allow a limited set of mixed model formulations. These latter programmes are based on maximum likelihood.

Table 1. Observed number of internodes for 12 maize varieties (reproduction of Table 2 of TWC/13/3), cutpoints, location and scale parameters. Average standard error of a difference between two consecutive cutpoints was 0.43. Average standard error of a difference between variety means was 0.58.

Category	≤8	9	10	11	12	≥13	Mean	Disp.	
Cutpoint		0.00	2.07	4.24	5.68	7.14	in model	in model	
Vi	0	0	4	14	11	1	5.34	1.00	
SM1	0	2	14	10	1	3	4.21	1.68	+
KX	0	0	0	14	16	0	5.66	0.34	-
SM2	0	0	2	2	11	0	5.77	0.96	
Gr	1	4	13	10	1	1	3.59	1.73	+
Ju	0	0	0	11	19	0	5.74	0.32	-
Mo	0	4	21	5	0	0	3.17	1.00	
Sa	1	10	14	5	0	0	2.58	1.45	
XD	16	14	0	0	0	0	-0.09	0.42	-
Ca	0	2	6	3	3	1	4.21	1.92	+
Ag	0	0	1	12	2	0	5.02	0.55	-
Le	0	4	25	1	0	0	2.84	0.74	

+: dispersion significantly larger than the standard posed by variety Vi (1.00).
-: dispersion significantly smaller then the standard.

Table 2. Fitted values for the data of Table 1.

Category	8	9	10	11	12	13	Mean	Disp.
Cutpoint		0.00	2.07	4.24	5.68	7.14	in model	in model
Vi	0.0	0.0	3.7	14.7	10.4	1.2	5.34	1.00
SM1	0.2	2.7	12.0	9.2	4.7	1.3	4.21	1.68
KX	0.0	0.0	0.0	14.0	16.0	0.0	5.66	0.34
SM2	0.0	0.0	0.7	5.9	7.1	1.2	5.77	0.96
Gr	0.5	4.9	13.6	7.4	3.0	0.7	3.59	1.73
Ju	0.0	0.0	0.0	11.0	19.0	0.0	5.74	0.32
Mo	0.0	3.7	21.6	4.4	0.2	0.0	3.17	1.00
Sa	1.1	9.4	15.5	3.5	0.5	0.0	2.58	1.45
XD	16.0	14.0	0.0	0.0	0.0	0.0	-0.09	0.42
Ca	0.2	1.7	5.5	4.1	2.5	1.0	4.21	1.92
Ag	0.0	0.0	1.0	12.0	2.0	0.0	5.02	0.55
Le	0.0	4.0	25.0	1.0	0.0	0.0	2.84	0.74

Table 3. Visual assessments of soil coverage by 9 varieties of sugar beets over a period of 3 years. Given are the total number of plants over 3 replicates that belonged to a certain category. Average standard error for the difference between two cutpoints was 0.15. The average standard error for a difference between two variety by year means was 0.43, for a difference between two year means it was 0.14.

Category Cutpoints in model	2	3	4	5	6	7	8	9	Mean in model
	0.00	0.89	1.82	2.83	3.73	4.64	6.19		
Year 1									
Cordelia	0	0	1	1	4	4	3	0	3.79
Evita	0	0	1	4	5	2	1	0	3.13
Fatima	0	0	0	4	4	2	3	0	3.58
Hilde	0	4	6	3	0	0	0	0	1.29
Hilton	3	2	4	3	0	1	0	0	1.20
KWS335	0	0	0	4	2	3	3	1	3.92
KWS337	0	0	2	4	3	3	1	0	3.05
Univers	0	1	5	4	3	0	0	0	2.03
Winner	0	0	1	4	6	2	0	0	2.98
									2.77
Year 2									
Cordelia	0	0	0	3	4	3	0	0	3.27
Evita	0	0	2	3	0	3	2	0	3.27
Fatima	0	0	1	3	3	3	0	0	3.07
Hilde	0	2	5	1	2	0	0	0	1.65
Hilton	0	1	1	4	1	2	1	0	2.79
KWS335	1	1	0	0	4	3	1	0	3.08
KWS337	0	0	0	2	4	2	2	0	3.66
Univers	0	0	2	6	1	0	1	0	2.51
Winner	0	0	2	3	4	1	0	0	2.70
									2.89
Year 3									
Cordelia	0	0	0	3	4	4	2	0	3.64
Evita	0	0	0	0	7	2	4	0	4.03
Fatima	0	0	0	1	2	5	4	1	4.44
Hilde	2	5	2	3	1	0	0	0	1.07
Hilton	0	2	5	4	2	0	0	0	1.81
KWS335	0	0	0	0	4	5	4	0	4.24
KWS337	0	0	0	3	5	4	1	0	3.49
Univers	0	0	0	1	5	5	2	0	3.86
Winner	0	0	0	0	2	9	2	0	4.21
									3.42

Table 4. Fitted values for data of Table 3.

Category Cutpoints in model	2	3	4	5	6	7	8	9	Mean in model
	0.00	0.89	1.82	2.83	3.73	4.64	6.19		
Year 1									
Cordelia	0.0	0.0	0.3	1.9	4.0	4.3	2.5	0.1	3.79
Evita	0.0	0.2	1.1	3.7	4.5	2.7	0.8	0.0	3.13
Fatima	0.0	0.0	0.5	2.4	4.3	3.9	1.8	0.1	3.58
Hilde	1.3	3.2	4.7	3.1	0.7	0.1	0.0	0.0	1.29
Hilton	1.5	3.4	4.6	2.8	0.6	0.1	0.0	0.0	1.20
KWS335	0.0	0.0	0.2	1.5	3.7	4.4	2.9	0.1	3.92
KWS337	0.0	0.2	1.2	3.9	4.4	2.5	0.7	0.0	3.05
Univers	0.3	1.4	3.8	4.8	2.2	0.5	0.1	0.0	2.03
Winner	0.0	0.2	1.4	4.1	4.3	2.3	0.6	0.0	2.98
Year 2									
Cordelia	0.0	0.1	0.7	2.5	3.5	2.4	0.8	0.0	3.27
Evita	0.0	0.1	0.7	2.5	3.5	2.4	0.8	0.0	3.27
Fatima	0.0	0.1	0.9	3.0	3.4	2.0	0.6	0.0	3.07
Hilde	0.5	1.7	3.5	3.1	1.0	0.2	0.0	0.0	1.65
Hilton	0.0	0.3	1.4	3.5	3.1	1.4	0.3	0.0	2.79
KWS335	0.0	0.1	0.9	3.0	3.4	2.0	0.6	0.0	3.08
KWS337	0.0	0.0	0.3	1.7	3.2	3.1	1.6	0.1	3.66
Univers	0.1	0.5	1.9	3.8	2.7	1.0	0.2	0.0	2.51
Winner	0.0	0.3	1.6	3.6	3.0	1.3	0.3	0.0	2.70
Year 3									
Cordelia	0.0	0.0	0.4	2.2	4.3	4.0	2.0	0.1	3.64
Evita	0.0	0.0	0.2	1.3	3.5	4.5	3.3	0.2	4.03
Fatima	0.0	0.0	0.1	0.6	2.4	4.4	5.0	0.5	4.44
Hilde	1.9	3.7	4.5	2.4	0.5	0.0	0.0	0.0	1.07
Hilton	0.5	1.9	4.3	4.4	1.7	0.3	0.0	0.0	1.81
KWS335	0.0	0.0	0.1	0.9	2.9	4.6	4.1	0.3	4.24
KWS337	0.0	0.1	0.6	2.7	4.4	3.7	1.6	0.0	3.49
Univers	0.0	0.0	0.3	1.7	3.9	4.4	2.7	0.1	3.86
Winner	0.0	0.0	0.1	1.0	3.0	4.6	4.0	0.3	4.21

Table 5. Sugar beet variety means with average standard error of a difference (SED) according to a completely fixed model and a mixed model with random variety by year interaction.

	Interaction	
	Fixed	Random
Cordelia	3.56	3.46
Evita	3.47	3.37
Fatima	3.69	3.60
Hilde	1.34	1.28
Hilton	1.93	1.84
KWS335	3.74	3.65
KWS337	3.40	3.28
Univers	2.80	2.71
Winner	3.30	3.21
Constant	3.03	2.93
Average SED	0.25	0.44

[End of document]