



HAL
open science

Constrained estimators of treatment parameters in semiparametric models

Marcin Przystalski, Pawel Krajewski

► **To cite this version:**

Marcin Przystalski, Pawel Krajewski. Constrained estimators of treatment parameters in semiparametric models. *Statistics and Probability Letters*, 2009, 77 (9), pp.914. 10.1016/j.spl.2006.12.013 . hal-00565444

HAL Id: hal-00565444

<https://hal.science/hal-00565444>

Submitted on 13 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Author's Accepted Manuscript

Constrained estimators of treatment parameters in semiparametric models

Marcin Przystalski, Paweł Krajewski

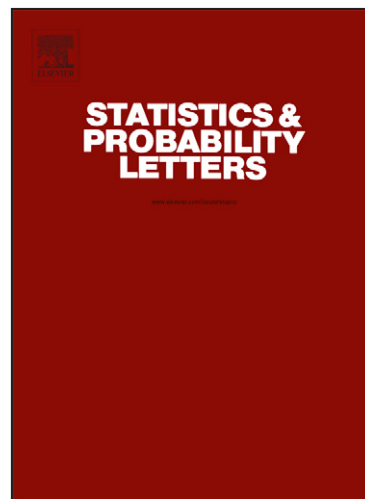
PII: S0167-7152(07)00030-2
DOI: doi:10.1016/j.spl.2006.12.013
Reference: STAPRO 4563

To appear in: *Statistics & Probability Letters*

Received date: 13 September 2006
Accepted date: 19 December 2006

Cite this article as: Marcin Przystalski and Paweł Krajewski, Constrained estimators of treatment parameters in semiparametric models, *Statistics & Probability Letters* (2007), doi:10.1016/j.spl.2006.12.013

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



www.elsevier.com/locate/stapro

Constrained Estimators of Treatment Parameters in Semiparametric Models

Short Title

Constrained Estimators in Semiparametric Models

Marcin Przystalski,¹ Paweł Krajewski

Biometry Laboratory, Institute of Plant Genetics, Polish Academy of Sciences

Strzeszyńska 34, 60 – 479 Poznań, Poland

Abstract

Semiparametric models are generalizations of parametric regression models. We present a method of estimation of treatment effects in a semiparametric model with one smoothing term under additional conditions on their linear functions and its application to hypothesis testing.

KEY WORDS: cubic smoothing splines, estimation, hypothesis testing, loess.

¹Corresponding author: Marcin Przystalski, tel.:(48)0616550254, email: mprz@igr.poznan.pl

1 Introduction

Semiparametric models are one of the possible generalizations of parametric regression models. A model with one smooth term can be expressed in the form (Durban et al., 1999)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{f} + \mathbf{e}, \quad (1)$$

where $\mathbf{y} = (y_1, \dots, y_n)'$ is a vector of observed values, \mathbf{X} is an $n \times p$ design matrix, with $\mathbf{1}$, the column vector of 1's, included in its column space, $\boldsymbol{\beta}$ is a $p \times 1$ vector of treatment effects, \mathbf{f} is a smooth term, that is, a function of a covariate \mathbf{z} , \mathbf{e} is a vector of i.i.d. errors with expected value $\mathbf{0}$ and covariance matrix $\sigma^2\mathbf{I}_n$.

The term \mathbf{f} in model (1) is called a smoother. Hastie and Tibshirani (1989) described several smoothers, e.g. cubic smoothing splines and locally weighted running line smoother (loess).

Durban et al. (1999) proposed a test for linear functions of treatment effects in semiparametric models with any number of smoothing terms based on approximate standard errors. To avoid using such approximations the following problem has to be considered: how to estimate the treatment effects in semiparametric models under some additional conditions on their linear functions, in particular under conditions forming a set of contrasts of treatment effects, and secondly how to apply these estimates in hypothesis testing.

The purpose of this paper is to give the answer to these two questions. In section 2 we introduce basic notions related to estimation in semiparametric models. In section 3 we discuss construction of the estimates under additional conditions. In section 4 we describe a testing procedure. Section 5 contains an example based on real data. A few concluding remarks are given in the final section.

2 Estimation in semiparametric models

Usually, semiparametric models are fitted by the backfitting algorithm. For model (1) the estimates of β and f can be written explicitly as

$$\begin{aligned}\widehat{\beta} &= \mathbf{A} \mathbf{y}, \\ \widehat{f} &= \mathbf{S} (\mathbf{y} - \mathbf{X} \widehat{\beta}) = \mathbf{S} (\mathbf{I} - \mathbf{X} \mathbf{A}) \mathbf{y},\end{aligned}\tag{2}$$

where $\mathbf{A} = \left\{ \mathbf{X}' (\mathbf{I} - \mathbf{S}) \mathbf{X} \right\}^{-1} \mathbf{X}' (\mathbf{I} - \mathbf{S})$ and \mathbf{S} is a smoother matrix which depends on smoothing parameter δ , provided that $\mathbf{X}' (\mathbf{I} - \mathbf{S}) \mathbf{X}$ is invertible. In this paper we do not consider the problem of choosing the smoothing parameter. Some methods of doing this may be found in Durban et al. (2003), Eubank et al. (1998), and Schimek (2000).

When, instead of the backfitting algorithm, we consider a penalized least squares criterion

$$\|\mathbf{y} - \mathbf{X}\beta - \mathbf{f}\|^2 + \mathbf{f}' (\mathbf{S}^{-1} - \mathbf{I}) \mathbf{f},\tag{3}$$

we derive the same estimates as (2) (see Buja et al., 1989, page 500). However this criterion can be used only for smoothers with a symmetric smoother matrix \mathbf{S} .

If we consider approach proposed by Speckman (1988), which minimizes

$$\|\mathbf{y} - \mathbf{X}\beta - \mathbf{S} (\mathbf{y} - \mathbf{X}\beta)\|^2,\tag{4}$$

where \mathbf{S} is an arbitrary smoother matrix, we obtain a solution different from (2), namely

$$\begin{aligned}\widetilde{\beta} &= \mathbf{A} \mathbf{y}, \\ \widetilde{f} &= \mathbf{S} (\mathbf{y} - \mathbf{X} \widetilde{\beta}),\end{aligned}\tag{5}$$

where $\mathbf{A} = \left\{ \mathbf{X}' (\mathbf{I} - \mathbf{S})' (\mathbf{I} - \mathbf{S}) \mathbf{X} \right\}^{-1} \mathbf{X}' (\mathbf{I} - \mathbf{S})' (\mathbf{I} - \mathbf{S})$. The Speckman approach was considered by Buja et al. (1989), Eubank et al. (1998), Schimek (2000).

For both discussed cases, the variance of the estimator of $\boldsymbol{\beta}$ is $\sigma^2 \mathbf{A} \mathbf{A}'$. If $\mathbf{H} = \mathbf{S} + (\mathbf{I} - \mathbf{S}) \mathbf{X} \mathbf{A}$ is the "hat" matrix for (1), then $\hat{\mathbf{y}} = \mathbf{H} \mathbf{y} = \mathbf{X} \hat{\boldsymbol{\beta}} + \hat{\mathbf{f}}$ and, following Durban et al. (1999), σ^2 can be estimated as $\hat{\sigma}^2 = \frac{1}{n - \text{tr} \mathbf{H}} \sum_{i=1}^n (y_i - \hat{y}_i)^2$.

3 Constrained estimators in semiparametric models

Let us assume that we have an additional condition for the vector of treatment parameters of the form

$$\mathbf{R} \boldsymbol{\beta} = \mathbf{r}. \quad (6)$$

We would like to find estimate \mathbf{b} of $\boldsymbol{\beta}$ in model (1) such that condition (6) is satisfied. For this reason we consider two cases: when the smoother matrix \mathbf{S} is symmetric and a second case when \mathbf{S} is arbitrary. The following theorem gives the solution of this problem for the first case.

Theorem 3.1 *If the smoother matrix \mathbf{S} is symmetric then the estimate of $\boldsymbol{\beta}$ in model (1) obtained by using the penalized least squares criterion (3), satisfying condition (6), is equal to*

$$\mathbf{b} = \hat{\boldsymbol{\beta}} + \mathbf{D}^{-1} \mathbf{R}' \left\{ \mathbf{R} \mathbf{D}^{-1} \mathbf{R}' \right\}^{-1} (\mathbf{r} - \mathbf{R} \hat{\boldsymbol{\beta}})$$

and its covariance matrix to $\text{Var}(\mathbf{b}) = \sigma^2 \mathbf{C} \mathbf{C}'$, where $\mathbf{D} = \mathbf{X}' (\mathbf{I} - \mathbf{S}) \mathbf{X}$

and $\mathbf{C} = \left(\mathbf{I} - \mathbf{D}^{-1} \mathbf{R}' \left\{ \mathbf{R} \mathbf{D}^{-1} \mathbf{R}' \right\}^{-1} \mathbf{R} \right) \mathbf{D}^{-1} \mathbf{X}' (\mathbf{I} - \mathbf{S})$.

Proof: Let us assume that the smoother matrix \mathbf{S} is symmetric and we want to find an estimate of $\boldsymbol{\beta}$ in model (1) satisfying condition (6) using the criterion (3). For this purpose we construct the function

$$F(\boldsymbol{\beta}, \mathbf{f}, \boldsymbol{\lambda}) = \|\mathbf{y} - \mathbf{X} \boldsymbol{\beta} - \mathbf{f}\|^2 + \mathbf{f}' (\mathbf{S}^{-1} - \mathbf{I}) \mathbf{f} - 2 \boldsymbol{\lambda}' (\mathbf{R} \boldsymbol{\beta} - \mathbf{r}).$$

By differentiating function F with respect to β , f and λ we obtain the following system of vector equations

$$\begin{aligned}\frac{\partial F(\beta, f, \lambda)}{\partial \beta} &= -2\mathbf{X}'(\mathbf{y} - \mathbf{X}\beta - f) - 2\mathbf{R}'\lambda = \mathbf{0} \\ \frac{\partial F(\beta, f, \lambda)}{\partial f} &= -2(\mathbf{y} - \mathbf{X}\beta - f) + 2(\mathbf{S}^{-1} - \mathbf{I})f = \mathbf{0} \\ \frac{\partial F(\beta, f, \lambda)}{\partial \lambda} &= \mathbf{R}\beta - \mathbf{r} = \mathbf{0}.\end{aligned}$$

If we get f from the second equation as

$$f = \mathbf{S}(\mathbf{y} - \mathbf{X}\beta)$$

and we put it into the first equation, we obtain

$$\mathbf{X}'[\mathbf{y} - \mathbf{X}\beta - \mathbf{S}(\mathbf{y} - \mathbf{X}\beta)] + \mathbf{R}'\lambda = \mathbf{0}.$$

Solving this equation with respect to β we get that

$$\beta = \left\{ \mathbf{X}'(\mathbf{I} - \mathbf{S})\mathbf{X} \right\}^{-1} \left(\mathbf{X}'(\mathbf{I} - \mathbf{S})\mathbf{y} - \mathbf{R}'\lambda \right).$$

Such β we put into the third equation and we have

$$\mathbf{R} \left[\left\{ \mathbf{X}'(\mathbf{I} - \mathbf{S})\mathbf{X} \right\}^{-1} \left(\mathbf{X}'(\mathbf{I} - \mathbf{S})\mathbf{y} - \mathbf{R}'\lambda \right) \right] = \mathbf{r}.$$

If we put $\mathbf{D} = \mathbf{X}'(\mathbf{I} - \mathbf{S})\mathbf{X}$, and notice that $\left\{ \mathbf{X}'(\mathbf{I} - \mathbf{S})\mathbf{X} \right\}^{-1} \mathbf{X}'(\mathbf{I} - \mathbf{S})\mathbf{y}$ is equal to $\hat{\beta}$ in (2), we get the following equality

$$\mathbf{R}\hat{\beta} - \mathbf{R}\mathbf{D}^{-1}\mathbf{R}'\lambda = \mathbf{r}.$$

Because $\mathbf{R}\mathbf{D}^{-1}\mathbf{R}' > 0$, we may write the optimal λ as

$$\hat{\lambda} = \left\{ \mathbf{R}\mathbf{D}^{-1}\mathbf{R}' \right\}^{-1} \left(\mathbf{R}\hat{\beta} - \mathbf{r} \right).$$

Finally, the estimate of β may be written as

$$\mathbf{b} = \hat{\beta} + \mathbf{D}^{-1}\mathbf{R}' \left\{ \mathbf{R}\mathbf{D}^{-1}\mathbf{R}' \right\}^{-1} \left(\mathbf{r} - \mathbf{R}\hat{\beta} \right).$$

□

Proposition 3.2 *If \mathbf{S} is an arbitrary smoother matrix then the estimate of $\boldsymbol{\beta}$ in model (1) under criterion (4) such that condition (6) is satisfied, is equal to*

$$\mathbf{b} = \tilde{\boldsymbol{\beta}} + \mathbf{Z}^{-1}\mathbf{R}' \left\{ \mathbf{R}\mathbf{Z}^{-1}\mathbf{R}' \right\}^{-1} (\mathbf{r} - \mathbf{R}\tilde{\boldsymbol{\beta}})$$

and its covariance matrix to $\text{Var}(\mathbf{b}) = \sigma^2\mathbf{B}\mathbf{B}'$, where $\mathbf{Z} = \mathbf{X}'(\mathbf{I} - \mathbf{S})'(\mathbf{I} - \mathbf{S})\mathbf{X}$,

$$\text{and } \mathbf{B} = \left(\mathbf{I} - \mathbf{Z}^{-1}\mathbf{R}' \left\{ \mathbf{R}\mathbf{Z}^{-1}\mathbf{R}' \right\}^{-1} \mathbf{R} \right) \mathbf{Z}^{-1}\mathbf{X}'(\mathbf{I} - \mathbf{S})'(\mathbf{I} - \mathbf{S}).$$

Proof: Transforming \mathbf{y} and \mathbf{X} in model (1), to $\tilde{\mathbf{y}} = (\mathbf{I} - \mathbf{S})\mathbf{y}$ and $\tilde{\mathbf{X}} = (\mathbf{I} - \mathbf{S})\mathbf{X}$ (see Speckman 1988, Eubank et al. 1998, Schimek 2000), applying RLSE for linear models and (5) we get the thesis.

4 Hypothesis testing

Let us assume that we tested a hypothesis $H_0 : \boldsymbol{\beta} = \mathbf{0}$ against $H_1 : \boldsymbol{\beta} \neq \mathbf{0}$. If we denote by $\mathbf{y}'\mathbf{G}_i\mathbf{y}$, $i = 1, 2$, $\mathbf{G}_1 = (\mathbf{I} - \mathbf{Q})'(\mathbf{I} - \mathbf{Q})$ and $\mathbf{G}_2 = (\mathbf{I} - \mathbf{H})'(\mathbf{I} - \mathbf{H})$, the residual sum of squares for model under H_0 and the non restricted model, respectively, then the test of null hypothesis against the alternative is based on the variance ratio

$$F = \frac{(\mathbf{y}'\mathbf{G}_1\mathbf{y} - \mathbf{y}'\mathbf{G}_2\mathbf{y})/\nu_1}{\mathbf{y}'\mathbf{G}_2\mathbf{y}/\delta_1} = \frac{\mathbf{y}'(\mathbf{G}_1 - \mathbf{G}_2)\mathbf{y}/\nu_1}{\mathbf{y}'\mathbf{G}_2\mathbf{y}/\delta_1}. \quad (7)$$

According to Cleveland and Devlin (1988), this test statistic has an approximate F distribution with $\gamma_1 = \nu_1^2/\nu_2$ and $\gamma_2 = \delta_1^2/\delta_2$ degrees of freedom, where $\nu_1 = \text{tr}(\mathbf{G}_1 - \mathbf{G}_2)$, $\nu_2 = \text{tr}[(\mathbf{G}_1 - \mathbf{G}_2)^2]$, $\delta_1 = \text{tr}(\mathbf{G}_2)$ and $\delta_2 = \text{tr}(\mathbf{G}_2^2)$.

If the general hypothesis $H_0 : \boldsymbol{\beta} = \mathbf{0}$ is rejected, then we may be interested in testing hypotheses of the type $H_{0,\mathbf{R}} : \mathbf{R}\boldsymbol{\beta} = \mathbf{0}$ against $H_{1,\mathbf{R}} : \mathbf{R}\boldsymbol{\beta} \neq \mathbf{0}$, with an appropriate \mathbf{R} .

From theorem 3.1 or proposition 3.2 with $\mathbf{r} = \mathbf{0}$ we know how to obtain the estimate of $\boldsymbol{\beta}$ under this null hypothesis. Denoting again by $\mathbf{y}'\mathbf{G}_i\mathbf{y}$, $i = 1, 2$, the residual sum of squares for the model with restriction and for the model without restriction, and using test statistic (7), we reject the null hypothesis when $F > F_\alpha$.

5 Example

To illustrate the method we use the yield data from the spring barley experiment used by Durban et al. (2003). The data concern 272 varieties. Among them, first 12 were established as controls and 260 were tested varieties. The trial was sown in a row and column design with 2 replicates. In each replicate there were 8 rows and 34 columns. Plot yields were provided in tonnes per hectare.

We take as a smoother in our semiparametric model (1) the two dimensional loess smoother across rows and columns. As a smoothing parameter we take $\delta = 125/544 \approx 0.2298$. We compare the thirty-first variety with the controls. We can write the null hypothesis in the form

$$H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{0},$$

where the i -th row of \mathbf{R} has -1 in the i -th position and 1 in the 31st. Using proposition 3.2 with $\mathbf{r} = \mathbf{0}$ and (7) we get the values of the residual sums of squares and the test statistic as in Table 1. The degrees of freedom are $\gamma_1 = 27.94$ and $\gamma_2 = 543.7$

Table 1

Now we compare the thirty-first variety with each control variety separately, i.e. we want to test hypotheses

$$H_0 : \mathbf{r}'_i \boldsymbol{\beta} = 0 \quad i = 1, \dots, 12,$$

$$H_1 : \mathbf{r}'_i \boldsymbol{\beta} \neq 0,$$

where \mathbf{r}'_i is the i -th row of \mathbf{R} . Applying proposition 3.2 with $\mathbf{r} = \mathbf{0}$ and (7) we get that the difference between thirty-first variety and first, second, third, fifth, and eighth is not significant. Contrasts between thirty-first variety and fourth, sixth, seventh, ninth, tenth,

eleventh, twelfth are significant.

Table 2

If we apply the proposition of calculating standard errors of contrasts described by Durban et al. (1999) to this example and use as the test statistic the ratio

$$F = \frac{(\mathbf{r}'\hat{\boldsymbol{\beta}})^2}{\widehat{\text{Var}}(\mathbf{r}'\hat{\boldsymbol{\beta}})}$$

with 1 and 272 degrees of freedom, we get that there are no significant differences between thirty first variety and first, second, third, fourth, fifth, eighth, tenth. The rest of the contrasts is significant.

Table 3

If for the same data set we consider the same null hypothesis, but as a smoother we use the cubic smoothing spline over row number and as δ we take 0.2, then using theorem 3.1 with $\mathbf{r} = \mathbf{0}$ we get that the test statistic has $\gamma_1 = 72.48$ and $\gamma_2 = 500.9$ degrees of freedom, and we can not reject the null hypothesis as shown in Table 4.

Table 4

6 Discussion

When the general hypothesis concerning treatment effects is rejected, usually more than one treatment comparison is of interest. The proposed method allows us to test a set of linear functions of treatment effects in a semiparametric model, not just only one function like in the method based on the approximate standard error. Both methods used in the example are approximate: ours because we use approximation of the F distribution, and the second because it uses approximation of standard errors described by Durban et al. (1999).

Further research is needed to find the estimates of treatment effects under restriction using backfitting algorithm. Also, it is interesting to see how the estimates will look like under additional conditions when we consider more than one smoothing term in the model.

Acknowledgments

The authors would like to thank referees for careful reading and Christine Hackett for help in obtaining the barley data set. This work was partially supported by the Polish Ministry of Education and Science grant SPB/COST/KN/DWM 74/2005-2008.

References

- Buja A., Hastie T., Tibshirani R. (1989). Linear smoothers and additive models. *Ann. Stat.* **17**, 453–555.
- Cleveland W. S., Develin S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *J. Amer. Stat. Assoc.* **83**, 596–610.
- Durban M., Hackett C. A., Currie I. D. (1999). Approximate standard errors in semiparametric models. *Biometrics* **55**, 699–703.
- Durban M., Hackett C. A., McNicol J. W., Newton A. C., Thomas W. T. B., Currie I. D., (2003). The practical use of semiparametric models in field trials. *J. Agr. Biol. Envir. Stat.* **8**, 48–66.
- Eubank R. L., Kambour E. L., Kim J. T., Klipple K., Reese C. S., Schimek M. (1998). Estimation in partially linear models. *Compu. Stat. Data An.* **29**, 27–34.
- Schimek M. (2000). Estimation and inference in partially linear models with smoothing splines. *J. Stat. Plan. and Infer.* **91**, 525–540.
- Speckman P. (1988). Kernel smoothing in partial linear models. *J. Roy. Stat. Soc. Ser. B* **50**, 413–436.

Table 1: Results of testing simultaneous comparison of variety no. 31 with the controls using loess smoother.

Model	Residual SS	Variance ratio	P-value
(a) without restriction	19.13	-	-
(b) with restriction	22.91	4.260	< .001

Accepted manuscript

Table 2: Hypothesis testing to investigate the significance of the individual contrasts using loess smoother.

Contrast	Residual SS	Degrees of freedom		Variance ratio	P-value
		γ_1	γ_2		
$\beta_{31} - \beta_1$	19.14	4.253	543.7	0.1057	0.9840
$\beta_{31} - \beta_2$	19.20	4.543	543.7	0.9989	0.4135
$\beta_{31} - \beta_3$	19.25	4.434	543.7	1.593	0.1683
$\beta_{31} - \beta_4$	19.33	4.566	543.7	2.695	0.0241
$\beta_{31} - \beta_5$	19.24	4.606	543.7	1.549	0.1783
$\beta_{31} - \beta_6$	19.56	4.662	543.7	5.883	< 0.001
$\beta_{31} - \beta_7$	19.59	4.567	543.7	6.296	< 0.001
$\beta_{31} - \beta_8$	19.13	4.528	543.7	0.0004	1
$\beta_{31} - \beta_9$	20.05	4.407	543.7	12.67	< 0.001
$\beta_{31} - \beta_{10}$	19.44	4.409	543.7	4.255	< 0.001
$\beta_{31} - \beta_{11}$	19.51	4.596	543.7	5.190	< 0.001
$\beta_{31} - \beta_{12}$	19.40	4.600	543.7	3.655	0.0039

Table 3: Hypothesis testing to investigate the significance of the contrasts using approximate standard errors.

Contrast	$r'\hat{\beta}$	Standard error	F statistic	P-value
$\beta_{31} - \beta_1$	0.0590	0.2803	0.04	0.8334
$\beta_{31} - \beta_2$	0.3107	0.2839	1.20	0.2747
$\beta_{31} - \beta_3$	0.3913	0.2835	1.90	0.1687
$\beta_{31} - \beta_4$	0.4677	0.2839	2.71	0.1006
$\beta_{31} - \beta_5$	0.3959	0.2845	1.94	0.1652
$\beta_{31} - \beta_6$	0.7155	0.2861	6.30	0.0127
$\beta_{31} - \beta_7$	0.7347	0.2843	6.68	0.0103
$\beta_{31} - \beta_8$	-0.0028	0.2839	0.00	0.9922
$\beta_{31} - \beta_9$	0.9954	0.2817	12.49	< 0.001
$\beta_{31} - \beta_{10}$	-0.5281	0.2830	3.48	0.0631
$\beta_{31} - \beta_{11}$	0.6566	0.2837	5.36	0.0214
$\beta_{31} - \beta_{12}$	0.6063	0.2883	4.52	0.0345

Table 4: Results of testing simultaneous comparison of variety no. 31 with the controls using cubic smoothing spline. The variance ratio is calculated with (a) as the denominator.

Model	Residual SS	Variance ratio	P-value
(a) without restriction	2.103	-	-
(b) with restriction	3.248	0.3843	1

Accepted manuscript