*Listy Biometryczne - Biometrical Letters*
*Vol. 39(2002), No. 1, 11-20*

# The application of logistic curve for the description of the fruitbearing of raspberries using cumulated data

**Małgorzata Szczepanik, Mirosława Wesołowska-Janczarek**

Department of Applied Mathematics,
Agricultural University, Akademicka 13, 20-934 Lublin, Poland

SUMMARY

This paper considers the method of estimation of parameters and testing hypothesis for a specific class of nonlinear growth functions presented by Rasch (1988; Probleme der Angewandten Statistik **24**, 159-191). On the basis of the experimental data the attempt was made to apply this method to the description of the process of raspberry fruitbearing taking cumulated crop into consideration.

KEY WORDS: growth curves, logistic function.

## 1. Introduction

Growth is one of the fundamental biological processes. It is a characteristic of organisms as well as populations. Studies concerning the growth of organisms or populations play an important role in biochemistry, biophysics and genetics.

As Rasch (1988) writes, Malthus (1798) gave the first analytical expression relating the size of the world population to time using the function:

$$f(t) = \theta_1 e^{\theta_2 t}, \quad t > 0,$$

where $\boldsymbol{\theta}^T = (\theta_1, \theta_2)$ is the (transposed) vector of parameters.

Gompertz (1825) investigated population growth for insurance purposes. To make life tables, he used the function

$$f(t) = \theta_1 e^{\theta_2 e^{\theta_3 t}}, \tag{1}$$

where $f(t)$ is the number of individuals at age $t$ in a population of size $f(0) = \theta_1 e^{\theta_2}$ at birth ( $t = 0$) and $\boldsymbol{\theta}^T = (\theta_1, \theta_2, \theta_3)$ is the vector of parameters.

Verhulst (1838, 1845, 1847) defined a system of growth functions $y = f(t)$ by a differential equation

$$\frac{1}{y}\frac{dy}{dt} = 1 - g(y) \tag{2}$$

for some function $g(y)$ and applied it to the human population growth. Assuming that $g(y)$ is a linear function, solving (2) he received the function

$$f(t) = \frac{\theta_1}{1 + \theta_2 e^{\theta_3 t}} \tag{3}$$

and called it the logistic function.

When $g(y) = \alpha(y - \beta)/y$

$$f(t) = \theta_1 + \theta_2 e^{\theta_3 t} \tag{4}$$

is the solution of (2).

The functions (1), (3) and (4) were used to describe the growth of organisms on the basis of data concerning a lot of disciplines, for example Askenasy (1880) studied growth of plants, Wiener (1890) – human growth, Lydtin and Werner (1887) and Minot (1891) – animal growth.

Thus, the expressions presented above are used to describe various processes occuring in time. In this paper, we suggest that the logistic function can be used to describe plants crop in time by means of using cumulated crop.

## 2. The model

Let $\mathbf{y}^T = (y_1, ..., y_h)$ be a vector of random variables depending on time $t$ which can denote $h$ growing characters of animals, plants or other organisms. We assume that our growth model has the form

$$\begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_h \end{pmatrix} = \begin{pmatrix} f_1(t, \theta^{(1)}) \\ \cdot \\ \cdot \\ \cdot \\ f_h(t, \theta^{(h)}) \end{pmatrix} + \begin{pmatrix} e_1(t) \\ \cdot \\ \cdot \\ \cdot \\ e_h(t) \end{pmatrix}, \tag{5}$$

where $f_j(t, \theta^{(j)})$ is called the growth function of character $y_j$ ( $j = 1, ..., h$ ) with parameters $\boldsymbol{\theta}^{(j)}$ and can be given for example by (1), (3) or (4). In our considerations we investigate each component of $\mathbf{y}$ independently and thus set $h = 1$. The above model Rasch (1988) called Model I of growth curves analysis and the parameter vector $\boldsymbol{\theta} = \boldsymbol{\theta}^{(1)}$ is considered as fixed.

When we consider growth functions in animal or plant studies, we will meet some specific problems. If we select individuals of the population for growth, we must remember that every individual has a different growth function. We will assume that growth follows the same type of growth function $f(t, \boldsymbol{\theta}_i)$ for each individual of a population with respect to some particular trait but parameters $\boldsymbol{\theta}_i$ may vary from one individual to another. Realisations $\boldsymbol{\theta}_i$ for different individuals can be identified with values of a random variable (Leech and Healy, 1959; Krauss et al., 1967).

Growth curves for individuals can be fitted if more than $p$ measurements are available for every individual, where $p$ is the number of parameters of the fitted growth function. Later we will analyse $n$ measurements $y_i = y(t_i)$ , $i = 1, ..., n,\ \ n > p$ for more than $p$ different time points.

We consider the model:

$$\mathbf{y} = f(t_i, \boldsymbol{\theta}) + \mathbf{e}, \tag{6}$$

where $i = 1, ..., n,\quad n > p,\quad \boldsymbol{\theta} \in \Omega,\ \dim(\Omega) = p,\ \ t_i \in [x_1, x_u],\ x_1 < x_u,\ \ E(\mathbf{e}) = 0,\ Var(\mathbf{e}) = \sigma^2 I_n,\ \mathbf{e}^T = (e_1, ..., e_n)$ and $\boldsymbol{\theta}^T = (\theta_1, ..., \theta_p)$ is the vector of unknown parameters of growth function $f$.

## 3. The estimation of parameters and hypothesis testing

The elements of the vector $\boldsymbol{\theta}$ from (6) are usually estimated using the least squares method. Thus the estimator $\widehat{\boldsymbol{\theta}}$ is the vector for which the expression

$$R(\boldsymbol{\theta}) = \sum_{i=1}^{n} [y_i - f(t_i, \boldsymbol{\theta})]^2 \tag{7}$$

reaches minimum. The solution is the result of an iteration. Therefore the distribution of $\boldsymbol{\theta}$ is not known (Rasch, 1988).

For $n > p$ the mostly used estimator of $\sigma^2$ from (6) is

$$s^2 = \frac{R(\widehat{\boldsymbol{\theta}})}{n - p}. \tag{8}$$

Denoting the values of partial derivatives of the growth function for parameters at points $t_i$ by

$$f_j(t_i, \boldsymbol{\theta}) = \frac{\partial f(t_i, \boldsymbol{\theta})}{\partial \theta_j}, \quad i = 1, ..., n, \quad j = 1, ..., p,$$

and setting them in the matrix

$$\mathbf{F}_i = \mathbf{F}_i(\boldsymbol{\theta}) = (f_1(t_i, \boldsymbol{\theta}), ..., f_p(t_i, \boldsymbol{\theta})),$$

calculating

$$\left[\sum_{i=1}^{n} \mathbf{F}_i^T(\boldsymbol{\theta})\mathbf{F}_i(\boldsymbol{\theta})\right]^{-1} = (v_{jk})\,, \quad j=1,...,p, \;\; k=1,...,p \tag{9}$$

and estimating $\sigma^2$ by $s^2$ according to (8), the estimated asymptotic covariance matrix of the parameter vector $\widehat{\boldsymbol{\theta}}^T = \left(\widehat{\theta}_1,...,\widehat{\theta}_p\right)$ is obtained:

$$\left[s^2\sum_{i=1}^{n} \mathbf{F}_i^T(\boldsymbol{\theta})\mathbf{F}_i(\boldsymbol{\theta})\right]^{-1} = \left(s^{-2}v_{jk}\right). \tag{10}$$

To test the null hypothesis

$$H_0^j:\; \theta_j = \theta_{j0} \quad \text{against} \quad H_1^j:\; \theta_j \neq \theta_{j0}$$

for parameters $\theta_j$, $j=1,...,p$, in accordance with Rasch and Schimke's (1984, 1985) papers for the group of growth function, including (1), (3), (4) and other functions, the $t$-Student test is useful:

$$t_j = \frac{\widehat{\theta}_j - \theta_{j0}}{s\sqrt{\hat{v}_{jj}}}, \quad j=1,...,p, \tag{11}$$

where $s$ is calculated from (8), and $\hat{v}_{jj}$ are estimated diagonal elements of matrix (9). For $|t_j| > t_{\alpha,n-p}$ hypothesis $H_0^j$ should be rejected at the significance level $\alpha$ ( $t_{\alpha,n-p}$ is the critical value of Student's distribution).

A confidence interval for parameters $\theta_j$ is given by:

$$\left[\widehat{\theta}_j - st_{\alpha,n-p}\sqrt{\hat{v}_{jj}} \;\; ; \;\; \widehat{\theta}_j + st_{\alpha,n-p}\sqrt{\hat{v}_{jj}}\right]. \tag{12}$$

## 4. Application

The estimators and the test, which were presented in the Section 3, are going to be used now to analyse some data obtained from the experiment carried out by the Departament of Fruitgrowing, Agricultural University, Lublin. The experiment, which was performed in four blocks, referred to the fruitbearing of raspberries. The data obtained in 1988 and 1989 were taken into consideration. It defined the cumulated crop (in kilogrammes), which was gathered on each field in blocks. Fruits of two cultivars of raspberries (Mailing Promise – cultivar 1 and Mailing Seedling – cultivar 2) were collected at 11 time points in 1988 ($n$ = 11) and 17 time points in 1989 ($n$ = 17). Each type of the raspberries was cultivated in 4 fields (one field in each block), which had 15 square meters.

A logistic function (3) with $p = 3$ parameters was chosen as a growth function. The values of parameters of functions which describe the process of fruitbearing (taking into consideration cumulated crops) were estimated for each field. In addition to this, the same calculations were made for average cumulated crops (average of cumulated each cultivar's crops from fields from 4 blocks). The results of estimation of parameters of a logistic function were made by the use of formulas (7), (8), (9), (12). The above results are illustrated in Tables 1, 2. Graphs, which show the process of fruitbearing, are presented in Figures 1a, 1b, 1c, 1d. The graphs of logistic curves, which describe the average cumulated crop of two cultivars of the fruit in 1988 and 1989 are presented in Figures 2a and 2b. The above curves are presented to show the differences in fruitbearing graphically.

On the basis of the values given in the Tables 1, 2, we can see that the longest (95%) confidence intervals were obtained in each case for the parameter $\theta_2$ and the shortest confidence intervals were obtained for $\theta_3$. The values of estimated parameters for particular cultivars and fields in blocks are strongly differentiated in two consecutive years. Thus, it confirms that in 1988 and in 1989 there is a different dynamic of fruitbearing of the cultivars and fields in blocks as well. In these years we took into consideration a different length of time of fruitbearing ( 26 days in 1988 and 39 days in 1989).

Checking analytic properties and graphs of the logistic functions for obtained estimators of parameters (e.g. $\widehat{\theta}_1 > 0$, $\widehat{\theta}_2 > 0$ and $\widehat{\theta}_3 < 0$) we can claim that $\theta_1$ is the biggest value of the function $f(t) = f(t_n)$ ($t_n$ is the last time point). Thus, from the

**Table 1.** Estimators of logistic functions' parameters describing cumulated crop from individual fields (1, 2, 3, 4 – block number) and average cumulated crop of two cultivars of raspberries in 1988

| Block | $\hat{\theta}_1$ | $\hat{\theta}_2$ | $\hat{\theta}_3$ | $s$ | $\hat{v}_{11}$ | $\hat{v}_{22}$ | $\hat{v}_{33}$ | Confidence interval for $\theta_1$ | | Confidence interval for $\theta_2$ | | Confidence interval for $\theta_3$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cultivar 2 | | | | | | | | | | | | | |
| 1 | 12.88 | 10.58 | −0.22 | 0.65 | 1.11 | 15.85 | 0.002 | 11.31 | 14.45 | 4.65 | 16.52 | −0.28 | −0.15 |
| 2 | 13.18 | 10.98 | −0.20 | 0.34 | 1.51 | 16.10 | 0.002 | 12.21 | 14.15 | 7.81 | 14.16 | −0.24 | −0.17 |
| 3 | 11.46 | 9.39 | −0.23 | 0.59 | 0.78 | 15.36 | 0.003 | 10.26 | 12.65 | 4.06 | 14.71 | −0.30 | −0.16 |
| 4 | 15.39 | 17.39 | −0.25 | 0.44 | 0.85 | 44.80 | 0.007 | 14.45 | 16.33 | 10.56 | 24.21 | −0.29 | −0.21 |
| average | 13.25 | 11.64 | −0.22 | 0.43 | 1.05 | 19.65 | 0.002 | 12.24 | 14.26 | 7.26 | 16.02 | −0.27 | −0.18 |
| Cultivar 1 | | | | | | | | | | | | | |
| 1 | 3.05 | 12.39 | −0.19 | 0.13 | 2.22 | 402.54 | 0.034 | 2.57 | 3.53 | 5.94 | 18.86 | −0.25 | −0.14 |
| 2 | 10.41 | 37.99 | −0.28 | 0.27 | 0.87 | 812.12 | 0.005 | 9.83 | 10.99 | 20.03 | 55.96 | −0.33 | −0.24 |
| 3 | 9.09 | 15.32 | −0.21 | 0.38 | 1.78 | 84.48 | 0.004 | 7.93 | 10.27 | 7.25 | 23.38 | −0.27 | −0.16 |
| 4 | 6.97 | 9.42 | −0.25 | 0.58 | 0.62 | 44.01 | 0.008 | 5.91 | 8.03 | 0.49 | 18.35 | −0.37 | −0.13 |
| average | 7.43 | 15.92 | −0.23 | 0.29 | 1.20 | 145.52 | 0.006 | 6.69 | 8.18 | 7.73 | 24.08 | −0.28 | −0.18 |

**Table 2.** Estimators of logistic functions' parameters describing cumulated crop from individual fields (1, 2, 3, 4 – block number) and average cumulated crop of two cultivars of raspberries in 1989

| Block | $\hat{\theta}_1$ | $\hat{\theta}_2$ | $\hat{\theta}_3$ | $s$ | $\hat{v}_{11}$ | $\hat{v}_{22}$ | $\hat{v}_{33}$ | Confidence interval for $\theta_1$ | | Confidence interval for $\theta_2$ | | Confidence interval for $\theta_3$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cultivar 1 | | | | | | | | | | | | | |
| 1 | 23.51 | 31.68 | -0.29 | 0.69 | 0.18 | 81.71 | 0.001 | 22.89 | 24.14 | 18.26 | 45.09 | -0.32 | -0.25 |
| 2 | 21.68 | 27.99 | -0.26 | 0.78 | 0.20 | 64.53 | 0.001 | 20.92 | 22.43 | 14.57 | 41.43 | -0.29 | -0.22 |
| 3 | 20.73 | 39.59 | -0.31 | 0.54 | 0.16 | 198.29 | 0.001 | 20.26 | 21.20 | 23.19 | 55.99 | -0.34 | -0.27 |
| 4 | 16.72 | 24.14 | -0.26 | 0.51 | 0.19 | 74.54 | 0.001 | 16.24 | 17.19 | 14.64 | 33.64 | -0.29 | -0.23 |
| average | 20.66 | 29.95 | -0.28 | 0.55 | 0.18 | 89.23 | 0.001 | 20.16 | 21.16 | 18.84 | 41.06 | -0.31 | -0.25 |
| Cultivar 2 | | | | | | | | | | | | | |
| 1 | 9.85 | 34.42 | -0.17 | 0.46 | 0.89 | 453.41 | 0.001 | 8.93 | 10.78 | 13.57 | 55.28 | -0.21 | -0.13 |
| 2 | 9.13 | 52.66 | -0.24 | 0.25 | 0.31 | 1768.24 | 0.003 | 8.84 | 9.43 | 30.08 | 75.25 | -0.27 | -0.21 |
| 3 | 7.00 | 143.33 | -0.26 | 0.16 | 0.37 | 38966.71 | 0.006 | 6.79 | 7.21 | 76.67 | 209.98 | -0.28 | -0.23 |
| 4 | 5.62 | 63.36 | -0.26 | 0.22 | 0.26 | 7809.03 | 0.009 | 5.38 | 5.86 | 21.42 | 105.30 | -0.31 | -0.22 |
| average | 7.81 | 55.73 | -0.22 | 0.21 | 0.39 | 2723.05 | 0.003 | 7.53 | 8.09 | 32.39 | 79.07 | -0.25 | -0.19 |

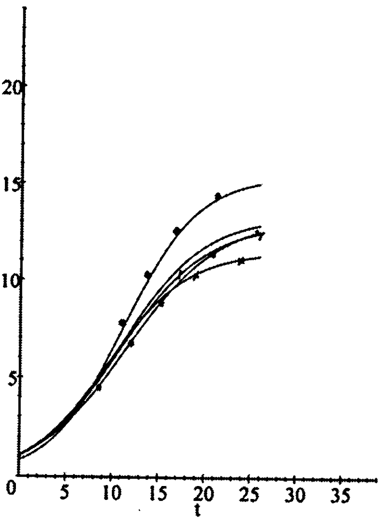
equation $f(0) = \frac{\theta_1}{1+\theta_2}$ for the fixed value $\theta_1$, we can come to a conclusion that the larger $\theta_2$ the later the begining of fruitbearing and the graph crosses the vertical axis closer to zero. On the other hand, if we take into consideration

$$f''(t) = -\frac{\theta_1\theta_2\theta_3^2\left(1-\theta_2 e^{t\theta_3}\right)}{\left(1+\theta_2 e^{t\theta_3}\right)^3},$$

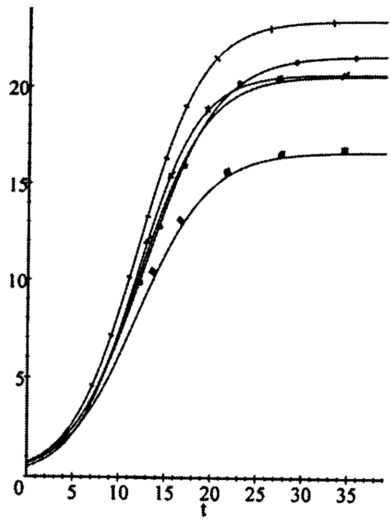
then the inflection point of $f(t)$ is

$$\left(-\frac{1}{\theta_3}\ln\theta_2, \frac{\theta_1}{1+\theta_2^{(1-t)}}\right).$$

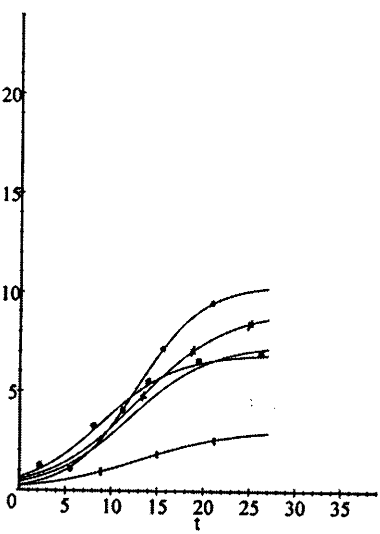
Parameters for average cumulated crops have to be taken if we want to prove whether it is possible to take common function, which describes the course of fruitbearing of a given cultivar in a given year for all four fields. Hypotheses referred to particular parameters were tested. In order to check hypotheses referring to the values of consecutive parameters $\theta_1$, $\theta_2$, $\theta_3$ of the logistic functions which describe the fruitbearing of cultivars in both years, the values of the test function (11) were calculated. The values $\widehat{\theta}_j$, $s$ and $\widehat{v}_{jj}$ were taken from the Tables 1 or 2. As $\theta_{j0}$ suitable numbers were taken from the rows which were called 'average'. For example, suitable estimators from row 1 of Table 1 were used to test the hypothesis $H_0^{j=1}$ : $\theta_1 = 13.25$, for the cultivar 1 and the field in block number 1 for data from 1988. Obtained values of the test functions were given in Table 3. The critical values were $t_{0.05,8} = 2.306$ (in 1988)



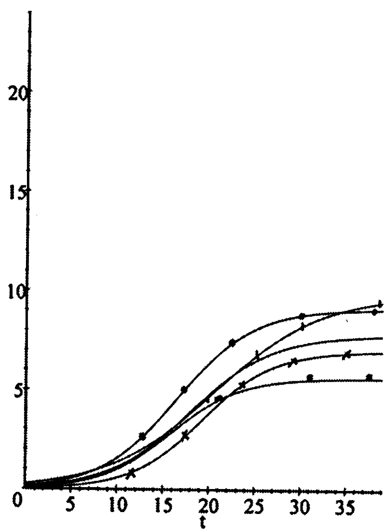

a) Cultivar 1, 1988




c) Cultivar 1, 1989

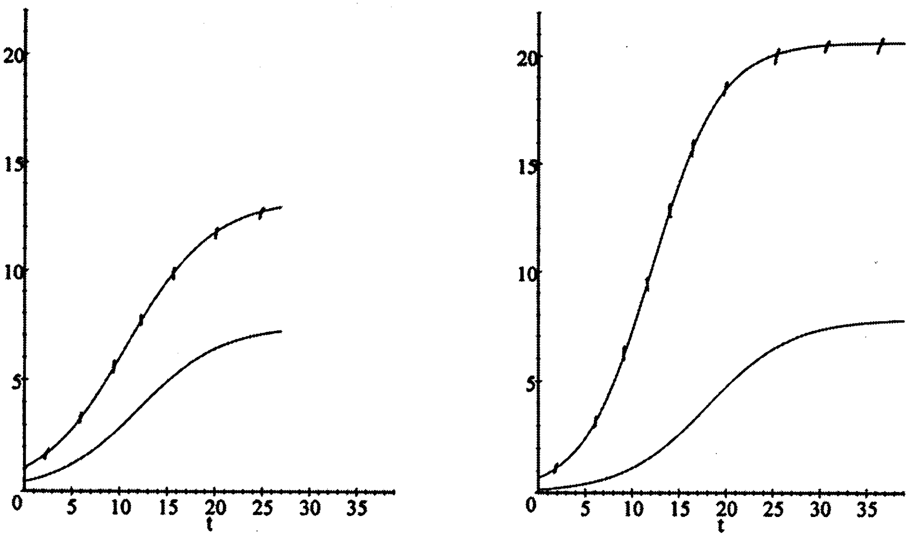



b) Cultivar 2, 1988




d) Cultivar 2, 1989

**Figure 1.** Fitted logistic functions describing cumulated crop (in klogrammes) for individual fields in blocks (—+— in block 1, —●— in block 2, —×— in block 3, —▪— in block4) and for average cumulated crops (—— ) of the cultivars of raspberries in two consecutive years




a) year 1988 b) year 1989

**Figure 1.** Fitted logistic functions describing avearge cumulated crop (in kilogrammes) of the cultivars of raspberries ( —+— cultivar 1, —— cultivar 2) in two consecutive years

**Table 3.** Values of the test function (11) for data from 1988 and 1989 (significances in bold type)

| Hypothesis for the parameter | 1988 | | | 1989 | | |
|---|---|---|---|---|---|---|
| | $\theta_1$ | $\theta_2$ | $\theta_3$ | $\theta_1$ | $\theta_2$ | $\theta_3$ |
| Cultivar 1 | | | | | | |
| 1 | -0.549 | -0.413 | 0.242 | **9.834** | 0.277 | -0.517 |
| 2 | -0.176 | 2.524 | 1.304 | **2.890** | -0.312 | 1.091 |
| 3 | **-3.457** | -0.979 | -0.279 | 0.332 | 1.261 | -1.780 |
| 4 | **5.242** | 1.940 | -0.757 | **-17.537** | -1.313 | 1.039 |
| Cultivar 2 | | | | | | |
| 1 | **-21.163** | -1.258 | 1.366 | **4.739** | **-2.189** | **3.156** |
| 2 | **11.733** | **2.839** | **-2.849** | **9.545** | -0.292 | -1.255 |
| 3 | **3.283** | -0.169 | 0.749 | **-8.528** | **2.826** | **-2.819** |
| 4 | -1.004 | **4.113** | -0.364 | **-19.585** | 0.391 | -1.903 |

and $t_{0.05,14} = 2.145$ (in 1989). By analysing the results, we can claim that the cultivar 1 and data from 1988 had the least significant differences. The parameter $\theta_1$ decides on the differentiation of the process of the function for the cultivar 1. As far as the second cultivar of the fruit is concerned, we have to say that significant differences were also discovered for the remaining parameters in both years. Hence investigated common growth functions of both cultivars do not describe exactly fruitbearing in particular fields from blocks.

Summing up, we can say that the method which was used shows that the cultivar 2 of raspberries (Mailing Seedling) had a bigger variability of its fruitbearing than the cultivar 1 (Mailing Promise). Besides, as we look at the graphs we can see that the cultivar 2 started its fruitbearing a little later than the cultivar 1 (graphs cross the vertical axis closer to the beginning of the co-ordinate system). The Figure 1b also denotes that not only the course of fruitbearing of the cultivar 2 was differentiated, but also its begining was differentiated on fields from different blocks.

## References


Askenasy, E. (1880). Über eine neue Methode, um die Verteilung der Wachstumstintensitat in wachsenden Teilen zu bestimmen. *Verh. Natur. Med., Ver. Heidelberg*, p. 70-173.

Gompertz, B. (1825). On the nature of the Function Expressive of the Law Of Human Mortality, and on an New Method of Determining the Value of Life Contingencies. *Phil. Trans. Roy. Soc.*, 513- 585.

Krauss, G. F., Siegel, F. B. and D. C. Hurst (1967). A probability structure for growth curves. *Biometrics* **23**, 217- 225.

Leech, R. B. and Healy, M. J. R. (1959). The analysis of experiments on growth rate. *Biometrics* **15**, 98- 106.

Lydtin, A. und Werner, H. (1887). Das deutsche Rind. Beschreibung der in Deutschland heimischen Rinderschlage. *Arbeiten der DLG* **4**, 41.

Malthus T. R. (1798). *Essay on the principle of population.* London.

Minot, C. S. (1891). Senescence and Rejuvenation. *J. Physiol.* **12**, 97- 153.

Rasch D. (1988). Growth curve analysis – a review and recent results. *Probleme der Angewandten Statistik* **24**, 159- 191.

Rasch, D., Schimke, E. (1984a). A test for exponential regression and its robustness. Berlin: *VEB Dt. Verlag d. Wissenschaft*, 104-112.

Rasch, D., Schimke, E. (1984b). Konfidenzschatzugen und Tests Fűr Parameter von Wachstumsfunktionen. *Arch. Tierz.* **27**, 553- 562.

Rasch, D., Schimke, E. (1985a). A test for exponential regression and its robustness. *Reidel Publ. Co.*

Rasch, D., Schimke, E. (1985b). Die Robustheit von Konfidenzschatzungen und Tests Fur Parameter der exponentiellen Regression. *Probleme der Angewandten Statistik* **13**, 40-92.

Verhulst, P. F. (1838). Notice sur la loi que la population suit dans son acroissement. *Corr. Math. et Phys. Publ. par A. Quetelet* **20**, 113- 121.

Verhulst, P. F. (1845). Recherches mathematiques sur la loi d'accroissement de la population. *Nouv. Mem. Acad. Roy. Sci. et Belles- Lettr.* Bruxelles **18**, 1- 38.

Verhulst, P. F. (1847). Deuxie'me memoire sur la loi d'accroissement de la population. *Nouv. Mem. Acad. Roy. Sci. et Belles- Lettr.* Bruxelles **20**, 1- 32.

Wiener, F. (1890). Das Wachstum des menschlichen Korpers. *Verh. naturw. Ver. Karlsruhe* **98**.

*Received 20 December 2001; revised 21 March 2002*

## Zastosowanie krzywej logistycznej do opisu przebiegu owocowania malin z wykorzystaniem skumulowanych danych

STRESZCZENIE

W pracy rozważa się metodę estymacji parametrów i testowania hipotez dla pewnej klasy nieliniowych funkcji wzrostu prezentowaną przez Rascha (1988). Podjęto też próbę zastosowania tej metody do opisu przebiegu plonowania malin poprzez uwzględnienie plonu skumulowanego wykorzystując do tego celu dane uzyskane z eksperymentu.

SŁOWA KLUCZOWE: krzywe wzrostu, funkcja logistyczna.