JOURNAL OF PLANT PROTECTION RESEARCH **Vol. 52, No. 2 (2012)**

ANALYSIS OF DOWNY MILDEW INFECTION OF FIELD PEA VARIETIES USING THE LOGISTIC MODEL

Ewa Bakinowska[1], Wiesław Pilarczyk[1,2], Agnieszka Osiecka[2], Kazimierz Wiatr[2]*

[1] Department of Mathematical and Statistical Methods, Poznań University of Life Sciences
Wojska Polskiego 28, 60-637 Poznań, Poland
[2] Research Centre for Cultivar Testing in Słupia Wielka, Poland

Received: August 31, 2011
Accepted: January 27, 2012

Abstract: The logistic model is commonly used for analysis of discrete, multinomial data. Such a model was used for the statistical evaluation of data concerning infection of field pea varieties by downy mildew, in two series of field trials. Each series consisted of experiments performed in locations spread over the whole of Poland in the time period from 2002 to 2005. Varieties cultivated on light soils were compared in the first series, and varieties cultivated on rich soils in the second. The most resistant varieties were identified (Sokolik – light soils, Terno – rich soils) and significant differences among varieties were detected. Estimators of model parameters were found using the Fisher scoring method implemented in *logistic glm* procedure of the SAS system.

Key words: downy mildew, field pea, logistic model

INTRODUCTION

Downy mildew is a fungal disease which commonly occurs in Poland (Sadowski *et al*. 1997; Boros and Marcinkowska 2010) as well as all over the world. It occurs on maize populations (Ajala *et al*. 2003), on sunflower (Panković *et al*. 2007), on brassica juncea (Nashaat *et al*. 2004), and in particular on peas (Stegmark 1990; Stegmark 1994). Downy mildew is caused by a pathogen named *Peronospora viciae* (Berk.) Casp. f. sp. *pisi* Sydow. Because of the early time of infection, its consequences can have a considerable influence on yields. It is especially important when the level of infection is severe. High humidity and low temperatures are favorable for the pathogen's appearing in the period before and during blossoming. The type of study we performed, which covered the whole area of Poland, had been not previously carried out. The susceptibility of the varieties before and after their registration had not been identified either. These issues therefore, are worthy of attention.

The main goal of the research was to identify the susceptibility of field pea varieties to downy mildew. This susceptibility was tested in the registration trials. The secondary goal was to identify the most favorable locations (within the studied years) for downy mildew.

MATERIALS AND METHODS

Evaluation of infection by downy mildew was performed in field trials at the blossoming stage. Importantly, the evaluation throughout all years and sites of testing was performed by the same phytopathologist, and therefore the results are free of rater bias. The intensity of downy mildew infection was assessed on each plot using a scale from 0 to 5.

A logistic model was used as the statistical tool for the data analysis. This approach made it possible to compare infection resistance by different varieties of field pea to downy mildew (Bakinowska and Kala 2007; Bocianowski *et al*. 2008; Czerniak *et al*. 2009).

Soil requirement observations were carried out for two groups of field pea varieties. The first group consists of varieties cultivated on rich soils, and the second of varieties cultivated on light soils. The varieties were tested in the years 2002–2005. The trials took place at 12 testing stations in Poland belonging to the Research Centre for Cultivar Testing (Fig. 1). The considered groups of varieties were generally examined at different sites; only in two stations were both groups tested. All trials were conducted in a randomized complete block design with 5 replicates.

Many varieties were examined throughout all the testing years. Only those with three-year results were chosen for the current analysis. Finally, 14 varieties cultivated on rich soils (in all trials) and 7 cultivated on light soils (in all trials) were chosen. Results from 10 sites were taken for the analysis which was independent of the testing years (Figs. 2, 3). The differences in degree of infection among the tested varieties was the main criterion for accepting the results from a particular site. In some trials, the degree of infection did not differentiate between varieties at all (all varieties received the same score) or the differences in the degree of infection was very weak (maximum

*Corresponding address:
ewabak@up.poznan.pl

difference of 1 in the score). In order to have better differentiation among varieties, trials showing such a small difference were excluded from the analyzed set of data. In the analysis, no particular factors describing the experimental environment were introduced. Only the cumulative influence of many factors (type of soil, precipitation level, temperature) within particular sites (among years), are preset in the applied model.

Because the observed units (plots) are classified according to an ordinal scale, the logistic model (Agresti 1984; Grizzle *et al*. 1969; McCullagh and Nelder 1989) can be applied for the data analysis. The marginal model can be written in the following form (Miller *et al*. 1993; Bakinowska and Kala 2007; Halekoh *et al*. 2006):

$$\eta_{ji} = \log\frac{\gamma_{ji}}{1-\gamma_{ji}} = \theta_j + \tau_i, \quad j = 1,2,\ldots,k-1, \quad i = 1,2,\ldots,s \qquad (1)$$

where:
θ_j – the border (cutpoint) of the *j*-th category,
k – number of categories,
τ_i – the effect of the *i*-th object (variety or trial),
s – the number of observed objects.

So $\theta_j + \tau_i$ is the cutpoint of the *j*-th category for the *i*-th object and γ_{ji} is the *j*-th cumulative probability corresponding to units of the *i*-th object

$$\gamma_{ji} = \pi_{1i} + \pi_{2i} + \ldots + \pi_{ji} \qquad j = 1,2,\ldots,k-1$$

The results of the classification of the studied units are usually modeled with the use of multinomial distribution, which is determined by probabilities p_{ji}, j = 1,2,...,k, summing to one, $\sum_{j=1}^{k} \pi_{ji} = 1$, and the fixed number of units m_i (m_i – the number of units which are classified to k separate categories). The analysis is aimed at estimating the unknown cumulative probabilities in model (1) based on the experimental data.

The set of equations (1) can be written in compact form as

$$\boldsymbol{\eta} = \mathbf{C}^T \log(\mathbf{L}\boldsymbol{\pi}) = \mathbf{X}\boldsymbol{\beta} \qquad (2)$$

where:
$\boldsymbol{\beta}^T = (\boldsymbol{\theta}^T, \boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_s)$ and $\boldsymbol{\theta}$ denotes the vector of cutpoints (borders of successive categories).

The estimates of unknown parameters in model (2) can be obtained using the maximum likelihood method. The main difficulty is solving the maximum likelihood equation, which is non-linear. The solution can be obtained using iterative methods (McCullagh and Nelder 1989; McCulloch and Searle 2001).

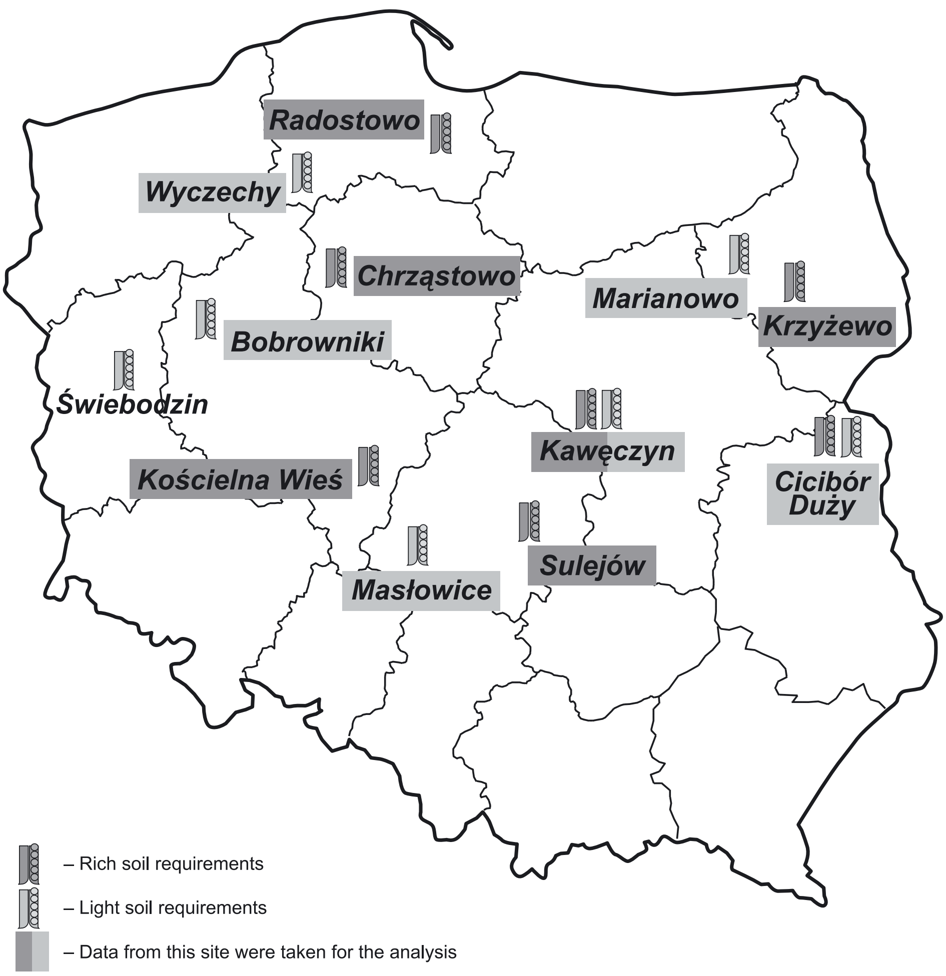


Fig. 1. Testing stations where trials on field pea cultivars were carried out

One of these methods is known as the Newton-Raphson method (N-R) and is based on the Hessian matrix of likelihood function. This method is simplified if the matrix of second derivatives of the likelihood function is replaced by its expectation, which leads to the Fisher information matrix. This approach is known as the Fisher scoring method (FS) (McCulloch and Searle 2001; Bakinowska 2004).

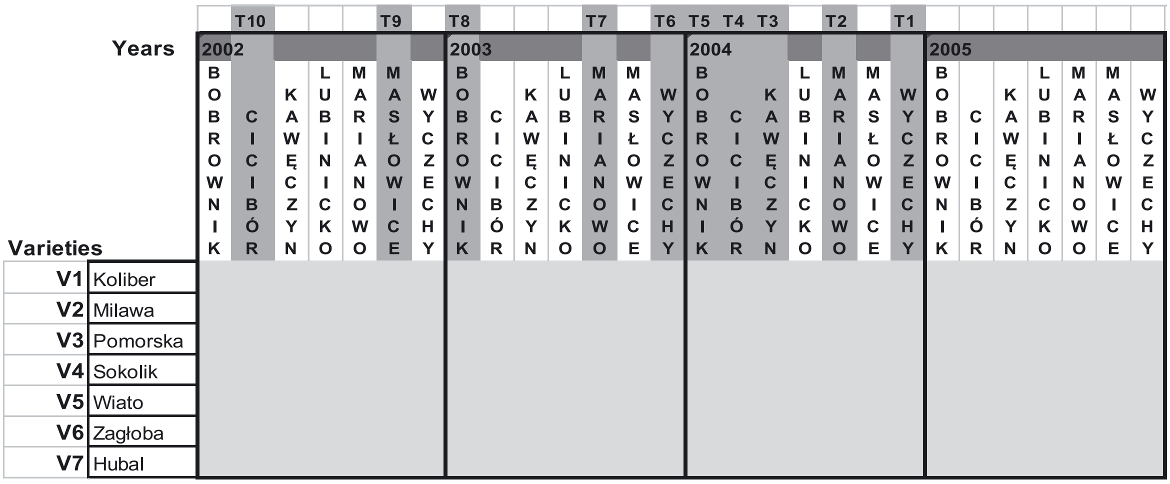
To test the hypothesis

$$\mathbf{H}_0 : \beta = \beta_*$$

Wald's statistic can be used. Wald's test statistic has the following form

$$(\hat{\beta} - \beta_*)^T \mathbf{F}(\hat{\beta})(\hat{\beta} - \beta_*)$$

and has approximate χ^2_p distribution (p being the order of β_*), where $\hat{\beta}$ is an estimate of β, and $\mathbf{F}(\hat{\beta})$ is the information matrix of $\hat{\beta}$ (Agresti 1984; McCulloch and Searle 2001).



T1–T10 – Trials

Fig. 2. Varieties, sites and years of the light soils, and also shown are the shaded trials which were included in the analysis

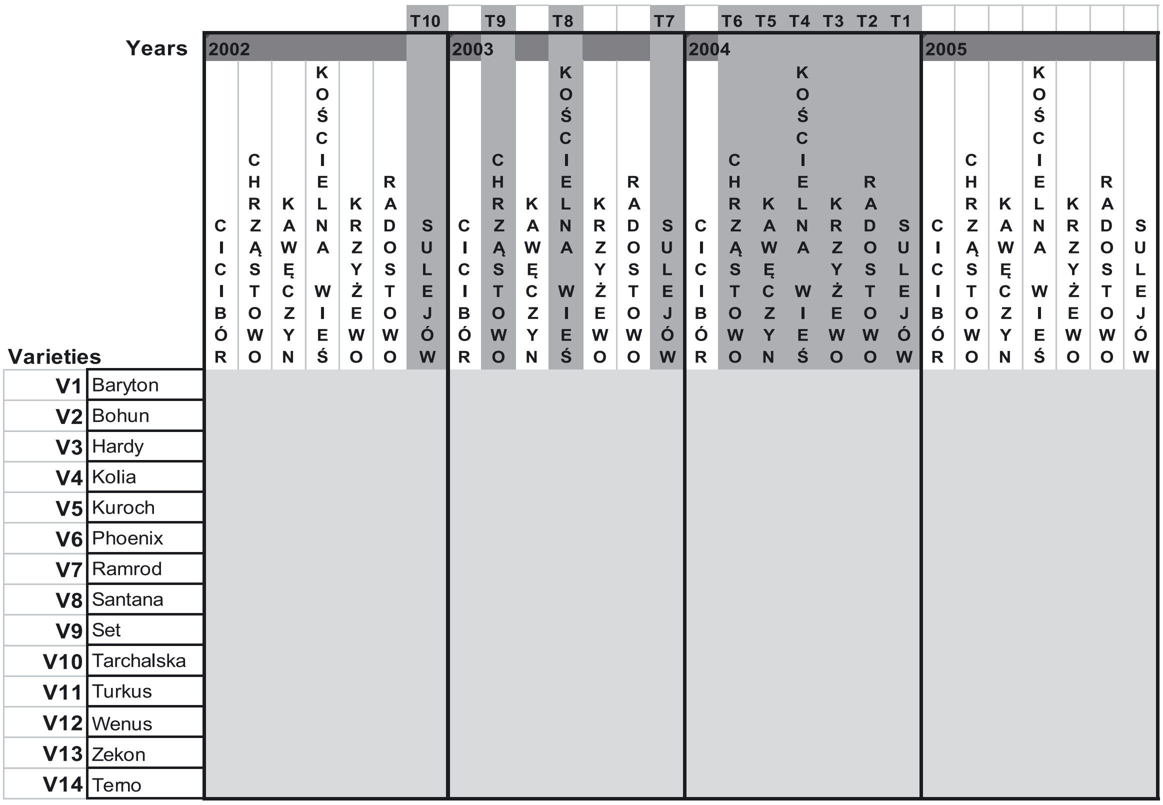


Fig. 3. Varieties, sites and years of the rich soils, and also shown are the shaded trials which were included in the analysis

All analyses were performed using SAS software (Statistical Analysis System, SAS Inst. 1997). The estimates were found using the Fisher scoring procedure. The results from 10 sites in three years were selected for analysis (7 varieties cultivated on light soils and 14 cultivated on rich soils were included).

RESULTS

Light Soils

The first analysis entailed the comparison of 6 varieties with the variety Hubal, which was indicated as the standard variety by crop experts. The obtained estimates of unknown parameters and values of Wald's statistic are presented in table 1. In the last column of this table the *p*-values are given. Based on the results obtained, we can draw the conclusion that only the variety Koliber does not differ significantly (at $\alpha = 0.01$) from Hubal with respect to level of infection by downy mildew.

In table 2, the values of probabilities and cumulative probabilities are presented. In the upper part of that table, for example, the entry 0.307 (for the variety in column V3 – Pomorska – and row γ_1) denotes that the probability that this variety will receive a score not larger than 1 is 0.307. Similarly, the value 0.847 (in row γ_2) for the same variety means that the probability of receiving a score not larger than 2 is 0.847, and so on. In the lower part of the table, the probabilities of the assignment of each variety to particular categories are presented. For example, the entry 0.556 (variety V6 – Zagłoba – in row π_2) denotes that the average probability of the assignment of this variety to category 3 is 0.556.

Comparing the probabilities and cumulative probabilities in table 2, we can conclude that the variety Sokolik is the best in terms of resistance to infection by downy mildew. The variety Pomorska is also resistant. On the other hand, Koliber and Hubal belong to the most susceptible varieties.

The next analysis concerns the comparison of trials. The results are given in tables 3 and 4. The level of infection in trials T1, T4 and T5 differs significantly from that in trial T10. The highest infection appeared in trial T1, the lowest in trial T7.

The detailed interpretation of the results in these tables is similar to the interpretation of the data in table 1 and 2. For example, the value 0.620 in column T1 and row π_3 of table 4 means that 0.620 is the probability that this location will receive an average score of 3.

Table 1. Comparison of varieties – levels of infection: 0, 1, 2, 3, 4

		Estimate	Wald's Statistic	p
Cutpoint 0	θ_0	-6.644	74.722	<.0001
Cutpoint 1	θ_1	-2.717	73.182	<.0001
Cutpoint 2	θ_2	-0.194	0.487	0.4852
Cutpoint 3	θ_3	3.021	42.397	<.0001
V1-Koliber	τ_1	0.367	0.901	0.3425
V2-Milawa	τ_2	1.705	18.233	<.0001
V3-Pomorska	τ_3	1.903	22.449	<.0001
V4-Sokolik	τ_4	2.376	30.481	<.0001
V5-Wiato	τ_5	1.214	9.533	0.002
V6-Zagłoba	τ_6	1.627	16.687	<.0001
V7-Hubal	τ_7	0	–	–

Table 3. Comparison of trials – levels of infection: 0, 1, 2, 3, 4

		Estimate	Wald's Statistic	p
Cutpoint 0	θ_0	-4.714	36.626	<.0001
Cutpoint 1	θ_1	-0.792	5.011	0.025
Cutpoint 2	θ_2	1.677	20.889	<.0001
Cutpoint 3	θ_3	4.945	80.483	<.0001
T1	τ_1	-2.536	22.612	<.0001
T2	τ_2	-0.767	2.564	0.109
T3	τ_3	-0.162	0.116	0.734
T4	τ_4	-1.063	4.905	0.027
T5	τ_5	-1.099	5.238	0.022
T6	τ_6	-0.287	0.361	0.548
T7	τ_7	0.608	1.624	0.203
T8	τ_8	-0.756	2.495	0.114
T9	τ_9	-0.525	1.121	0.290
T10	τ_{10}	0	–	–

Table 2. Comparison of varieties – probabilities and cumulative probabilities

	V1	V2	V3	V4	V5	V6	V7
$\gamma_0=\pi_0$	0.002	0.007	0.009	0.014	0.004	0.007	0.001
$\gamma_1=\pi_0+\pi_1$	0.087	0.267	0.307	0.416	0.182	0.252	0.062
$\gamma_2=\pi_0+\pi_1+\pi_2$	0.543	0.819	0.847	0.899	0.735	0.807	0.452
$\gamma_3=\pi_0+\pi_1+\pi_2+\pi_3$	0.967	0.991	0.993	0.995	0.986	0.991	0.954
π_0	0.002	0.007	0.009	0.014	0.004	0.007	0.001
π_1	0.085	0.260	0.298	0.402	0.178	0.245	0.061
π_2	0.456	0.553	0.540	0.483	0.553	0.556	0.390
π_3	0.424	0.172	0.146	0.097	0.251	0.183	0.502
π_4	0.033	0.009	0.007	0.005	0.014	0.009	0.046

Table 4. Comparison of trials – probabilities and cumulative probabilities

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
$\gamma_0=\pi_0$	0.001	0.004	0.008	0.003	0.003	0.007	0.016	0.004	0.005	0.009
$\gamma_1=\pi_0+\pi_1$	0.035	0.174	0.278	0.135	0.131	0.254	0.454	0.175	0.211	0.312
$\gamma_2=\pi_0+\pi_1+\pi_2$	0.297	0.713	0.820	0.649	0.641	0.801	0.908	0.715	0.760	0.843
$\gamma_3=\pi_0+\pi_1+\pi_2+\pi_3$	0.917	0.985	0.992	0.980	0.979	0.991	0.996	0.985	0.988	0.993
π_0	0.001	0.004	0.008	0.003	0.003	0.007	0.016	0.004	0.005	0.009
π_1	0.034	0.170	0.270	0.132	0.128	0.247	0.438	0.171	0.206	0.303
π_2	0.263	0.539	0.542	0.514	0.510	0.547	0.454	0.540	0.549	0.531
π_3	0.620	0.272	0.172	0.331	0.338	0.190	0.089	0.270	0.228	0.150
π_4	0.083	0.015	0.008	0.020	0.021	0.009	0.004	0.015	0.012	0.007

Rich Soils

The second analysis concerns the comparison of 13 varieties with the variety Terno, which was indicated as the standard variety by crop experts. The obtained estimates of unknown parameters and values of Wald's statistic are presented in table 5. In the last column of this table the *p*-values are presented. Based on the results obtained, we can draw the conclusion that all tested varieties differ significantly (at $\alpha = 0.01$) from the variety Terno with respect to level of infection by downy mildew.

In table 6, the values of probabilities and cumulative probabilities are presented. In the upper part of that table, for example, the entry 0.543 (for the variety in column V4 – Kolia –and row γ_2) denotes that the probability that this variety will receive a score not larger than 2 is 0.543. Similarly, the value 0.942 (in row γ_3) for the same variety means that the probability of receiving a score not larger than 3 is 0.942, and so on.

In the lower part of this table, the probabilities of the assignment of each variety to particular categories are presented. For example, the entry 0.274 (variety V9 – Set – in row π_3 denotes that the average probability of assignment of this variety to category 4 is 0.274.

Comparing the probabilities and cumulative probabilities in table 6 we can conclude that the varieties Terno and Phoenix are the best in terms of resistance to infection by downy mildew. The most susceptible were varieties Ramrod and Turkus.

The next analysis concerns the comparison of trials. The results are given in tables 7 and 8. The level of infection in trials T2, T3 and T6 did not differ significantly from that in trial T10. The highest infection appeared in trials T1 and T2, the lowest in trials T7 and T9.

The detailed interpretation of the results in these tables is similar to the interpretation of data in tables 5, 6 and 7, and tables 3–4 for light soils.

Table 5. Comparison of varieties – levels of infection: 0, 1, 2, 3, 4

		Estimate	Wald's Statistic	*p*
Cutpoint 0	θ_0	–1.840	27.745	<.0001
Cutpoint 1	θ_1	1.094	11.687	0.0006
Cutpoint 2	θ_2	3.315	95.590	<.0001
Cutpoint 3	θ_3	5.932	224.739	<.0001
V1-Baryton	τ_1	–2.156	25.901	<.0001
V2-Bohun	τ_2	–1.539	13.458	0.0002
V3-Hardy	τ_3	–2.239	26.541	<.0001
V4-Kolia	τ_4	–3.144	53.453	<.0001
V5-Kuroch	τ_5	–2.870	44.953	<.0001
V6-Phoenix	τ_6	–0.937	5.026	0.025
V7-Ramrod	τ_7	–3.765	74.328	<.0001
V8-Santana	τ_8	–2.270	28.625	<.0001
V9-Set	τ_9	–2.493	34.291	<.0001
V10-Tarchalska	τ_{10}	–2.756	41.581	<.0001
V11-Turkus	τ_{11}	–3.225	56.070	<.0001
V12-Wenus	τ_{12}	–3.106	52.252	<.0001
V13-Zekon	τ_{13}	–1.279	9.337	0.0022
V14-Terno	t_{14}	0	–	–

Table 7. Comparison of trials – levels of infection: 0, 1, 2, 3, 4

		Estimate	Wald's Statistic	p
Cutpoint 0	θ_0	–5.341	206.934	<.0001
Cutpoint 1	θ_1	–2.365	75.532	<.0001
Cutpoint 2	θ_2	0.169	0.470	0.4929
Cutpoint 3	θ_3	2.953	86.856	<.0001
T1	τ_1	–0.854	6.360	0.0117
T2	τ_2	–0.537	2.570	0.1089
T3	τ_3	0.646	3.689	0.0548
T4	τ_4	1.725	25.095	<.0001
T5	τ_5	1.505	19.277	<.0001
T6	τ_6	–0.043	0.013	0.9094
T7	τ_7	3.250	71.865	<.0001
T8	τ_8	1.399	16.719	<.0001
T9	τ_9	2.680	57.477	<.0001
T10	τ_{10}	0	–	–

Table 6. Comparison of varieties – probabilities and cumulative probabilities

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
$\gamma_0=\pi_0$	0.018	0.033	0.017	0.007	0.009	0.059	0.004	0.016	0.013	0.010	0.006	0.007	0.042	0.137
$\gamma_1=\pi_0+\pi_1$	0.257	0.391	0.241	0.114	0.145	0.539	0.065	0.236	0.198	0.159	0.106	0.118	0.454	0.749
$\gamma_2=\pi_0+\pi_1+\pi_2$	0.761	0.855	0.746	0.543	0.609	0.915	0.389	0.740	0.695	0.636	0.523	0.552	0.885	0.965
$\gamma_3=\pi_0+\pi_1+\pi_2+\pi_3$	0.978	0.988	0.976	0.942	0.955	0.993	0.897	0.975	0.969	0.960	0.937	0.944	0.991	0.997
π_0	0.018	0.033	0.017	0.007	0.009	0.059	0.004	0.016	0.013	0.010	0.006	0.007	0.042	0.137
π_1	0.239	0.358	0.225	0.107	0.136	0.481	0.061	0.220	0.185	0.150	0.100	0.111	0.412	0.612
π_2	0.504	0.465	0.504	0.429	0.465	0.376	0.325	0.504	0.497	0.477	0.416	0.434	0.431	0.216
π_3	0.216	0.133	0.230	0.399	0.346	0.078	0.508	0.235	0.274	0.324	0.415	0.392	0.106	0.032
π_4	0.022	0.012	0.024	0.058	0.045	0.007	0.103	0.025	0.031	0.040	0.063	0.056	0.009	0.003

Table 8. Comparison of trials – probabilities and cumulative probabilities

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
$\gamma_0=\pi_0$	0.002	0.003	0.009	0.026	0.021	0.005	0.110	0.019	0.065	0.005
$\gamma_1=\pi_0+\pi_1$	0.038	0.052	0.152	0.345	0.297	0.083	0.708	0.276	0.578	0.086
$\gamma_2=\pi_0+\pi_1+\pi_2$	0.335	0.409	0.693	0.869	0.842	0.531	0.968	0.827	0.945	0.542
$\gamma_3=\pi_0+\pi_1+\pi_2+\pi_3$	0.891	0.918	0.973	0.991	0.989	0.948	0.998	0.987	0.996	0.950
π_0	0.002	0.003	0.009	0.026	0.021	0.005	0.110	0.019	0.065	0.005
π_1	0.036	0.049	0.143	0.319	0.276	0.078	0.598	0.257	0.513	0.081
π_2	0.297	0.357	0.541	0.524	0.545	0.449	0.260	0.552	0.367	0.456
π_3	0.556	0.509	0.280	0.122	0.146	0.417	0.030	0.160	0.051	0.408
π_4	0.109	0.082	0.027	0.009	0.011	0.052	0.002	0.013	0.004	0.050

CONCLUSIONS

Not all available data were used in the analysis. The chosen subset of the data was constructed in a way that made it possible to highlight the differences among varieties in their resistance to the downy mildew pathogen. However, because the effects of locations and years were treated as fixed, the conclusions are limited to circumstances (environments) covered by the locations and years used in the analysis. To draw more general conclusions, it is necessary to treat locations (and years) as random factors. This will be the next stage of our research. Nevertheless, under this rather simple model, some differences in the resistance of varieties to downy mildew were detected. Obtained results are, to wide extent, in accordance with the results published by The Research Centre for Cultivar Testing (Osiecka COBORU 2011). Nevertheless, a direct comparison of our results with the results published by COBORU is impossible, as their results are from the years 2008 to 2010. Another difference is that they used a nine degree scale – from 9 to 1, while in our research the phytopatological scale from 0 to 5 was applied. But the ranking of the common subset of varieties is quite similar. Hence, the applied method proved its usefulness. When interest is focused on the occurrence of downy mildew across the country, the whole set of data is more suitable than the subset used here. Such an analysis is planned for the near future.

ACKNOWLEDGEMENTS

We are grateful to the Referee of the paper for the constructive comments and remarks.

REFERENCES

Agresti A. 1984. Analysis of Ordinal Categorical Data. Wiley, New York, 287 pp.

Ajala S.O., Kling J.G., Kim S.K., Obajimi A.O. 2003. Improvement of maize populations for resistance to downy mildew. Plant Breeding 122: 328–333.

Bakinowska E. 2004. Analyzing experiments by generalized linear models. Biometrical Let. 41 (2): 37–49.

Bakinowska E., Kala R. 2007. An application of logistic models for comparison of varieties of seed pea with respect to lodging. Biometrical Let. 44 (2): 143–154.

Bocianowski J., Bakinowska E., Rybiński W. 2008. Analysis of selected grasspea mutants by generalized linear model. Colloquium Biometricum 38: 161–171.

Boros L., Marcinkowska J. 2010. Assessment of selected pea genotypes reaction to ascochyta blight under field conditions and the impact of disease severity on yield components. J. Agric. Sci. 2 (3): 84–91.

Czerniak A., Górna M., Bakinowska E., Kayzer D. 2009. Functionality of the animal crossing built over trunk road No. 5 in the Wielkopolski National Park using a logistic model. Polish J. Environ. Studies 18 (3A): 59–64.

Grizzle J.E., Starmer C.F., Koch G.G. 1969. Analysis of categorical data by linear models. Biometrics 25: 489–504.

Halekoh U., Hojsgaard S., Yan J. 2006. The R package geepack for generalized estimating equations. J. Statistical Software 15 (2): 1-11.

Marcinkowska J. 1997. Healthiness of peas grown for dry seeds. [in Polish]. Biul. IHAR 201: 279–287.

McCullagh P., Nelder J.A. 1989. Generalized Linear Models. 2nd ed. Chapman and Hall, London, 511 pp.

McCulloch Ch.E., Searle S.R. 2001. Generalized, Linear, and Mixed Models. Wiley, New York, 325 pp.

Miller M.E., Davis Ch.S., Landis J.R. 1993. The analysis of longitudinal polytomous data: generalized estimating equations and connections with weighted least squares. Biometrics 49: 1033–1044.

Nashaat N.I., Heran A., Awasthi R.P., Kolte S.J. 2004. Differential response and genes for resistance to *Peronospora parasitica* (downy mildew) in *Brassica juncea* (mustard). Plant Breeding 123: 512–515.

Osiecka A. 2011. Descriptive list of agricultural plant varieties. The research Centre for Cultivar Testing at Słupia Wielka (COBORU), Agricultural Cultivars, part 2, 151 pp.

Panković D., Radovanović N., Jocić S., Satovic Z., ŠKorić D. 2007. Development of co-dominant amplified polymorphic sequence markers for resistance of sunflower to downy mildew race 730. Plant Breeding 126: 440–444.

Sadowski C., Skinder Z., Wilczewski E. 1997. The occurrence of downy mildew (*Peronospora viciae* (Berk.) Casp. f. sp. *pisi*) in the winter form of pea. Adv. Agric. Sci. Problem 446: 451–454.

SAS Institute. 1997. SAS/STAT software: Changes and enhancements through release 6.12. SAS Inst., Cary, NC.

Stegmark R. 1994. Downy mildew on peas (*Peronospora viciae* f. sp. *pisi*). Agron. Sustain. Dev. 14 (10): 641–647.

Stegmark R. 1990. Selection for partial resistance to downy mildew in peas by means of greenhouse tests. Euphytica 53 (2): 87–95.