Neth. J. Pl. Path. 92 (1986) 197-206

Measures of disease intensity in powdery mildew (Erysiphe graminis) of winter wheat. 1. Errors in estimating pustule number

R.A. DAAMEN

Research Institute for Plant Protection (IPO), P.O. Box 9060, 6700 GW Wageningen, the Netherlands

Accepted 1 May 1986

Abstract

Assessments of pustule number and severity of powdery mildew on winter wheat in the Netherlands were made in commercial fields and in experimental plots. The sample variance (s^2) of the number of pustules per leaf (m) was fairly constant over years, varieties, growth stages and leaf postitions, but depended strongly on the average pustule number: $s^2 = 2 \cdot 2\overline{m}^{1.5}$. The effect of sample size on the precision of the estimate is discussed and it is concluded that it is difficult to estimate low disease intensities accurately. Estimates are given for the detection level of pustule counts in relation to sample size.

Mildew intensity on the lower surface of leaves can be estimated from the intensity on the upper surface. This method reduces the duration of the observation, but introduces an additional error. At low disease intensities and small sample sizes this method is more efficient than sampling mildew on both surfaces of leaves. The common practice of assessments of the upper surface of leaves only may not be the most efficient method.

Additional keywords: Triticum aestivum, epidemiology, sampling method, detection level.

Introduction

The choice of a disease assessment method is a difficult problem. No widely accepted method for foliar diseases of wheat is yet available, and the choice of a method is usually a compromise between the objectives and the resources available. A clear distinction can be made between the resources in an experiment and monitoring a disease in commercial fields. This paper considers the precision of counting pustules as a method to assess mildew *(Erysiphe graminis)* intensity. In a following paper, the precision of incidence counts as method to assess mildew intensity is discussed. To avoid confusion, intensity will be used in this paper to refer to the quantity of disease, irrespective of the assessment method; incidence is used when the intensity is expressed in proportion of leaves diseased and severity when intensity is in proportion of leaf surface diseased.

In studies of yield loss caused by powdery mildew in winter wheat in the Netherlands, assessments of disease were made by counting the number of pustules on leaves. There were three reasons for using this method. First, the aim was to estimate yield loss at low disease intensities; pustule counts are then not time-

consuming. Second, the assessment method should be defined exactly, so that it can be used independent of observer bias. Third, pustule counts can be transformed easily to the more common percentages, using disease assessment keys.

Rouse et al. (1980, 1981) studied the distribution of mildew pustules on wheat leaves. They found that the distribution on artificially inoculated seedling leaves fitted the negative binomial distribution, but on spontaneously infected leaves at the end of tillering the distribution did not fit any of the tested distributions. Koch (1978) described the precision of barley mildew assessments. He concluded that assessments of the percentage leaf surface diseased could not be used to estimate low disease intensities.

The objective of this paper is to examine the precision of mildew pustule counts on full grown wheat plants as a method of disease assessment at low disease intensities. Diseases in wheat are usually assessed only on the upper surface of leaves, although mildew can occur also on the under surface. Attention is given to the error introduced by observing the upper surface alone.

Materials and methods

Data. This paper is based on two data sets, one from commercial fields and one from field experiments, both in the Netherlands.

Commercial fields. In 1980 and 1981, 31 and 11 commercial fields, respectively, of different winter wheat varieties were sampled during May-June. A sample of 40 tillers per field was taken and the development stage determined (Zadoks et al., 1974). The top full grown leaf (flag leaf of adult plant) was designated leaf position 1. For each leaf position the number of mildew pustules was counted on the upper and lower surface of the leaves. In the same samples, leaves without disease were counted in the three upper positions. The number of disease-free leaves was counted like farmers will normally do in the field. All observations were made with the naked eye by experienced observers. Fields were sampled whether or not fungicides had been used.

Field experiments. From 11 field experiments (1980 to 1983), 31 sets of mildew pustules counts (upper leaf surface) in untreated plots were available. In each assessment at least 60 and usually 90 tillers were sampled. Average pustule number and sample variance were computed per leaf position for each sample. From the same pustule counts the number of leaves without disease was determined. The mildew assessments were were made at a speed of roughly 100 tillers per hour, at low disease intensities. The duration of the observation increased at higher disease intensities, and pustule numbers above 20 were estimated instead of counted.

Statistics. To obtain an estimate of the mildew severity, key 1.1.2 (Ubels and Van der Vliet, 1979) which is in common use in the Netherlands, was used to transform the average pustule number into percentage values. According to this key, for fully grown winter wheat, 5 pustules correspond to 1% severity.

Statistical analyses were performed by means of Genstat V, release 4.04 A and B (Alvey et al., 1983). Analyses of variance were carried out using multiple regression analyses with dummy variables for main effects only. For example the simple relation

between two variables, x and y, is described as:

$$y = a + b \cdot x + \text{error1} \quad (1)$$

and can be analysed by means of ordinary least squares regression to estimate a and b and the error with $n - 2$ degrees of freedom. If the relation between the variables depends on other factors, for example years with three levels ($i = 1, 2, 3$) and varieties with two levels ($j = 1, 2$), the relation for main effects only will have the form:

$$y(\text{year, variety}) = a(\text{year}) + a(\text{variety}) + [b(\text{year}) + b(\text{variety})] \cdot x + \text{error2} \quad (2)$$

The combined effect of years and varieties in this example can be evaluated by calculating the F-statistic over the sums of squares (SS) of models 1 and 2:

$$F = \frac{(\text{SSerror1} - \text{SSerror2}) / [2(i - 1) + 2(j - 1)]}{\text{SSerror2} / [n - 2 - 2(i - 1) - 2(j - 1)]} \quad (3)$$

by which the significance of factors (as in analyses of variance) can be evaluated in covariance designs. Since the factors may be partially confounded, this analysis is followed by a test on each individual factor, adjusted for the other factors. Chatterjee and Price (1977) gave a clear treatment of these tests.

To summarise the fit of a regression model, the ordinary R-squared statistic (R^2) will be used, computed as:

(SStotal − SSerror) / SStotal.

Results

Sample variance of pustule number per leaf. The sample variance of the population density of an organism is often variable, and dependent on the mean density. Taylor (1961) proposed a power function to describe the sample variance (s^2) in relation to population density (M):

$$s^2 = a \cdot M^b \quad (4)$$

in which a and b are parameters. This function is used to describe the relation between the sample variance [Var(m)] of the pustule counts from the upper surface of leaves and the mean number of pustules ($\overline{m}$) in the sample. Taking logs is a convenient way to stabilise variance and to estimate the parameters a and b:

$$\ln[\text{Var}(m)] = \ln(a) + b \cdot \ln(\overline{m}) \quad (5)$$

First, it was tested whether the relation was independent of years, varieties and growth stages, for each leaf position.

Analyses of variance on main effects (see Methods) were carried out, using the data set of the field experiments (Table 1). The results show no strong evidence for deviations due to years, varieties or growth stages. For the top leaves, these factors had a weak significant effect ($p < 0.05$). Further analyses showed that this deviation was due mainly to a year effect; in 1981 parameters a and b deviated (top leaf). In 1981 the attack of mildew was relatively severe and it is possible that the attention of the observers was weakened and the low disease intensities on the top leaf not properly observed. No data are available to analyse this deviation further. Since this deviation was slight and occurred on the top leaf in one year only, the parameters for each individual leaf

Table 1. Sample variance of mildew pustule counts. Analyses of variance to test for deviations from common a and b in ln[Var(m)] = ln(a) + b·ln($\bar{m}$) due to years (1980 to 1983), growth stage of the crop (DC $\leqslant$59, DC >59) and varieties (Arminda, Okapi, Caribo). Experimental plots (n = 32), forth leaf position n = 17.

Source of variation	df	Top leaf SS	Leaf 2 SS	Leaf 3 SS	Leaf 4	
					df	SS
1. common a and b	1	95.1	146.3	169.2	1	101.1
2. different a and b	12	2.3	2.1	2.2	12	0.8
2a. years	6	1.3	0.8	0.7	6	0.2
2b. growth stage	2	0.2	0.1	0.3	2	0.1
2c. varieties	4	0.4	1.2	0.2	4	0.4
3. residual	18	1.1	2.2	2.2	3	0.5
4. total	31	98.5	150.7	173.5	16	102.3
All factors	F 12/18	3.2*[1]	1.4ns[1]	1.5ns	F 12/3	0.4ns
Years	F 6/18	3.7*	1.1ns	1.0ns	F 6/3	0.2ns
Growth stage	F 2/18	0.2ns	0.3ns	1.5ns	F 2/3	0.4ns
Varieties	F 4/18	1.5ns	2.5ns	0.4ns	F 4/3	0.7ns

[1] ns = not significant ($p > 0.05$); * = significant at $p = 0.05$.

position are assumed constant over the years, varieties and growth stages. Estimates of the parameters and R^2 are given in Table 2. The fit of the descriptive model is good ($R^2 > 0.96$). Parameters a and b have the same magnitude for the four leaf positions and do not differ significantly between leaf positions ('a posteriori testing'). The conclusion is that average values of a and b can be used, irrespective of the leaf position. Thus, the sample variance of pustule counts [Var(m)] can be described by:

$$\text{Var}(m) = 2.2\,\bar{m}^{1.55} \qquad (6)$$

To illustrate the relation, the data for the third leaf are plotted in Fig. 1.

Sample size and precision of estimation. The variance of the mean pustule number

Table 2. Sample variance of pustule counts. Estimates of ln(a) and b in Var(m) = ln(a) + b·ln($\bar{m}$) and their standard deviations; and an estimate of a and its 95% confidence limits. Experimental plots n = 32, fourth leaf n = 17.

Leaf position	ln(a)		b		R^2	a	
top leaf	0.87	(0.09)	1.49	(0.05)	0.97	2.4	(2.9-2.0)
2nd leaf	0.78	(0.07)	1.56	(0.05)	0.97	2.2	(2.5-1.9)
3rd leaf	0.79	(0.09)	1.55	(0.05)	0.97	2.2	(2.6-1.8)
4th leaf	0.64	(0.08)	1.61	(0.05)	0.99	1.9	(2.3-1.6)
Average	0.81	(0.03)	1.55	(0.02)	0.98	2.2	(2.4-2.1)

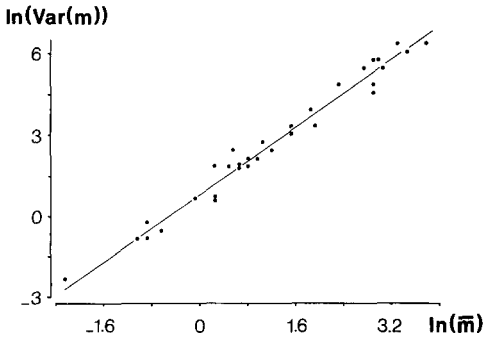

Fig. 1. Sample variance, log-transformed (ln[Var(m)]) in relation to log mean pustule number (ln($\overline{m}$)), upper surface third leaf position. Regression: ln[Var(m)] = 0.789 + 1.55 ln($\overline{m}$); R^2 = 0.97.

[Var($\overline{m}$)] on the upper surface of a leaf from a sample of size n will be estimated as usual by:

$$\text{Var}(\overline{m}) = \text{Var}(m)/n$$

substitution of (6) gives:

$$\text{Var}(\overline{m}) = \frac{2.2\,\overline{m}^{1.55}}{n} \tag{7}$$

Approximate confidence limits for mean pustule number can be computed by:

$$\overline{m} \pm u_p\sqrt{\text{Var}(\overline{m})} \tag{8}$$

in which u_p is the (100 − 0.5p) percentile of the normal distribution; p was arbitrarily chosen as 5%. The approximate confidence limits for sample means at low disease intensities at fixed sample sizes (15 and 30), were computed from (7) and (8) (Table 3). The width of the 95% confidence limits depends strongly on the disease intensity.

To estimate the sample mean with a certain precision, a particular minimum sample size is required. The coefficient of variation (CV) can be used as a measure of precision:

$$CV = \frac{\sqrt{\text{Var}(\overline{m})}}{\overline{m}} \tag{9}$$

As the sample mean ($\overline{m}$) is an estimate of the population mean (M), (7) and (9) can be used to estimate the required sample size n at different population densities. Substitution of (7) into (9) gives Equation 10, for estimation of the required sample size, n:

$$n = \frac{2.2M^{-0.45}}{CV^2} \tag{10}$$

In Fig. 2 the relation between the required sample size and the population mean is given for fixed coefficients of variation. If M < 20, a coefficient of variation of the mean of 0.1 is not attainable with n < 50. Apparently, in wheat powdery mildew the sample error is large, especially at low disease intensities. The detection level of pustule counts is therefore of special interest.

The detection level is defined as the disease intensity M_d (in pustules per leaf), at

Table 3. Estimated 95% confidence limits for mean pustule number ($\overline{m}$) on the upper surface of leaves, for two sample sizes (n).

$\overline{m}$	n = 15		n = 30	
	lower	upper	lower	upper
1	0.2	1.8	0.4	1.6
3	1.1	4.9	1.7	4.3
5	2.1	7.9	3.1	6.9
7	3.3	11	4.5	9.5
9	4.5	14	6.0	12
11	5.7	16	7.5	15
13	7.0	19	9.0	17
15	8.3	22	10	20
17	9.6	24	12	22
19	11	27	13	24

which there is a 97.5% probability of presence of the disease on a sample of n leaves. So at least 1 pustule should be present in a sample of n leaves, thus at average 1/n pustules per leaf. Assuming approximate normality of the mean pustule number, the detection level M_d can be estimated by:

$$\frac{1}{n} = M_d - u_p\sqrt{Var(M_d)}$$

where u_p is the 97.5 percentile of the normal distribution ($u_p = 1.96$) and $Var(M_d)$ can be estimated by (7). Hence:

$$\frac{1}{n} = M_d - 1.96\,\frac{\sqrt{2.2\,M_d^{1.55}}}{n} \qquad (11)$$

Equation 11 is solved by iteration. A sample size of 15 leaves gives a detection level of $M_d = 0.5$ pustules per leaf, and for 30 leaves $M_d = 0.2$ pustules per leaf. So mildew intensities below 0.5 pustules per leaf cannot be properly estimated from samples of 15 leaves or fewer.

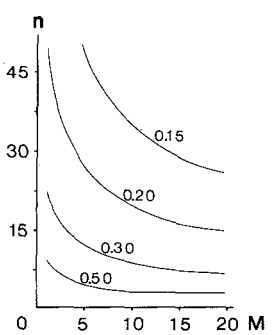

Fig. 2. Minimum sample size (n, leaves) required to estimate the pustule number (M) per leaf with a certain precision. The values plotted are the different coefficients of variation of the mean.

Sampling powdery mildew on the upper and lower surface of leaves. It is common practice in cereals to assess diseases on the upper surface of leaves only, although mildew can occur on both surfaces. If mildew on both surfaces of the leaves is of interest (e.g. causing damage), the assessment of upper surface only causes a systematic underestimation of the mildew population of interest. If there exist a relation between the mildew intensity on the upper and under surface of leaves, then it is possible to predict the mildew intensity on both surfaces from the intensity on the upper surface of leaves. Such a prediction introduces an additional error. On the other hand, let us assume that this procedure is twice as quick as the direct estimate. In other words, it allows for a doubling of the sample size and a gain in precision due to a reduction of the sampling variation. The question is now whether the error due to sampling a limited part of the population (upper surface leaves only) can be neglected when compared with the reduction of the error due to sampling variation. This question can be answered by comparison of two variance estimates. One is when pustules are counted at both surfaces of n leaves. The other is the variance of the predicted pustule number on both surfaces from the pustule number on the upper surfaces of 2n leaves.

A direct estimate of the sample variance of the total pustule number (m_t) on both surfaces of the leaf was obtained from 46 samples of the data set of commercial fields (using Equation 5). Estimates were: ln(a) = 1.17 (sd. 0.084; a = 3.2); b = 1.51 (sd. 0.048); R^2 = 0.96. The b-value in (4) can thus be assumed to describe the sample variance of total pustule number (m_t) and pustule number on upper surface (m_u, see Table 2), but the a-values cannot. So the variance of the mean total pustule number $\overline{m}_t$ is described by:

$$\mathrm{Var}(\overline{m}_t) = \frac{a_t \cdot \overline{m}_t^{\,b}}{n_t} \; ; \; a_t = 3.2; \; b = 1.53 \qquad (12)$$

The next step is to describe the variance of the predicted total pustule number ($\overline{m}_t$) from the pustule number at the upper surface only. Suppose that the ratio between mean pustule number on lower surface ($\overline{m}_l$) and upper surface ($\overline{m}_u$) of leaves is constant:

$$\frac{\overline{m}_l}{\overline{m}_u} = c \qquad (13)$$

For each sample of the data set of commercial fields, the ratio (c) of pustule number at lower and upper surface was calculated. Analyses of variance were carried out for each leaf postition, to check for constancy of the parameter c for the different years, varieties and growth stages. The results of these analyses showed no significant deviations from a common parameter c. The estimates of c for each leaf postition have about the same magnitude (Table 4); that for the third leaf is lower than those for the other leave positions, but this difference is not significant ('a posteriori testing'). The ratio, c, of the pustule numbers is thus independent of the leaf position. The estimate of a common c for the top four leaves is: c = 0.52 (sd. 0.04); the residual variance equals 0.23. For illustration, parameter c is plotted in Fig. 3 in relation to the pustule number on the upper surface for the third leaf position.

Table 4. Relation of mean pustule number on lower ($\overline{m}_l$) and upper surface ($\overline{m}_u$) of leaves. Fitted values for c (= $\overline{m}_l/\overline{m}_u$) and its standard deviation. Commercial fields n = 42, fourth leaf n = 41. Top leaf and second leaf 15 and 3 observations, respectively, with c = 0/0 were excluded.

Leaf position	c	
top leaf	0.58	(0.16)
2nd leaf	0.56	(0.05)
3rd leaf	0.47	(0.05)
4th leaf	0.52	(0.06)
Average	0.52	(0.04)

The total pustule number ($\overline{m}_t$) can be predicted from the pustule number on the upper surface of leaves ($\overline{m}_u$) using (13):

$$\hat{\overline{m}}_t = (1 + c)\cdot\overline{m}_u \tag{14}$$

The variance of this new estimate may be calculated by adding the average conditional variance and the variance of conditional average (Rao, 1965, p. 97):

$$\begin{aligned}\mathrm{Var}(\hat{\overline{m}}_t) &= \mathrm{E}\cdot\mathrm{Var}(\hat{\overline{m}}_t)\,|\,\overline{m}_u) + \mathrm{Var}[\mathrm{E}(\hat{\overline{m}}_t\,|\,\overline{m}_u)] \\ &\approx \mathrm{Var}(1 + c)\cdot[\mathrm{Var}(\overline{m}_u) + \overline{m}_u^2] + (1 + c)^2\cdot\mathrm{Var}(\overline{m}_u)\end{aligned} \tag{15}$$

The variance of $\overline{m}_u$ is described by (7) and the variance of c was 0.23. Assume that sampling the upper surface alone is twice as quick as sampling the whole leaf. Sampling the upper surface alone on 2n leaves is a better strategy than sampling n whole leaves if the variance from equation 15 is less than or equal to the variance of equation 12:

$$\frac{a_u\cdot\overline{m}_u^b}{2n}\cdot[(1 + c)^2 + \mathrm{Var}(c)] + \overline{m}_u^2\cdot\mathrm{Var}(c) \leqslant \frac{a_t\cdot(1 + c)^b\cdot\overline{m}_u^b}{n}$$

or when

$$n \leqslant 14.25\,\overline{m}_u^{-0.47} \tag{16}$$

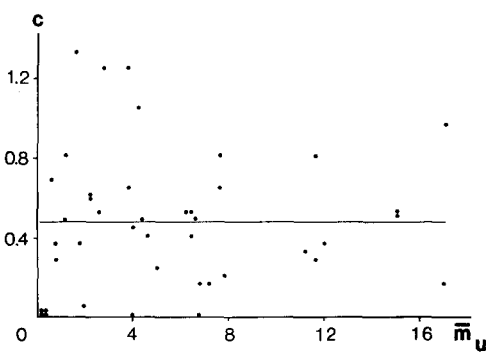

Fig. 3. Ratio (c) of mean pustule number on lower ($\overline{m}_l$) and upper surface ($\overline{m}_u$) of the third leaf, plotted against mean pustule number on the upper surface. Ratio: $\overline{m}_l/\overline{m}_u$ = 0.47 (sd. = 0.05).

As the sample mean $\overline{m}_u$ is an estimate of the population mean M_u, (16) may be used to estimate the breakpoint sample size at different mildew intensities. The breakpoint (n) indicates the sample size at which both methods are equal in precision, at lower sample sizes it is more efficient to double the sample size and to inspect the upper surface of the leaves only. For a range of M_u from 1 to 20 pustules per leaf, n was calculated using (16):

M_u	1	5	10	15	20
$n \leqslant$	14	6	4	4	3

Thus only for low disease intensities and small sample sizes is it more efficient to double the sample size and to inspect the upper surface only. For example; inspection of the upper surface of 12 leaves instead of both surfaces of 6 leaves is efficient at disease intensities of up to 5 pustules per (upper surface) leaf.

Discussion

No statistical frequency distribution has yet been found to describe the dispersion of powdery mildew pustules on leaves. Rouse et al. (1981) found that the negative binomial gave the closest fit, though it still deviated significantly, and it has been shown that only the dispersion of pustules among artificially inoculated seedling leaves fits the negative binomial (Rouse et al. 1980). Nevertheless, the sampling error of pustule counts was adequately described using Taylor's (1961) power function. This description does not yield more insight into the underlying dispersal process, but that was not the aim of this paper.

Koch (1978) found that in barley powdery mildew, the sampling error of assessments of the percentage leaf surface diseased became unacceptably large at disease severities below 2%. The sampling error of wheat powdery mildew pustule counts depends also on the disease intensity. The error is large at intensities below 5 pustules per leaf (1% severity). Even using pustule counts it is difficult to determine low disease intensities, but at least the error can be estimated.

The error introduced by sampling only the upper surface of leaves and prediction of mildew intensity on the lower surface instead of direct assessments on both surfaces has been estimated, and suggests that at low disease intensities (< 5 pustules per leaf) and small sample sizes (< 12 leaves) it is more efficient to sample the upper surface only instead of both surfaces. In other situations it is better to sample both surfaces and to halve the size of the sample.

In this paper the occurrence of mildew on the leaf sheaths was not accounted for. Estimates based on examination of only the upper surface of leaves will become very inefficient if the whole mildew population on a plant or crop is of interest. Assessment of the total pustule number per tiller may be far more accurate than the sampling of the upper surface of leaves which is currently the common practice.

Acknowledgements

The author wishes to thank Drs J. de Bree and professors R. Rabbinge and J.C. Zadoks for the valuable discussions and the critical reviewing of the manuscript. The

author is grateful to Dr S.A. Ward for the discussion and corrections of the English, and for the assistance of Mr G. van der Vliet and Mr A.K. van Harxen and Mr W.A. Wietsma during the field work. The field work was financed by the Netherlands Grain Centre (NGC).

Samenvatting

Meting van meeldauwaantastingen (Erysiphe graminis) in wintertarwe. 1. Schattingsfouten van het aantal puistjes

Aantallen puistjes meeldauw per blad werden geteld in praktijkpercelen en veldproeven met wintertarwe. De steekproefvariantie van het aantal puistjes was tamelijk constant in de jaren, rassen, gewasstadia en bladposities, maar was sterk afhankelijk van het gemiddeld aantal puistjes ($\overline{m}$): $s^2 = 2{,}2\ \overline{m}^{1{,}5}$. Het effect van de steekproefgrootte op de nauwkeurigheid van de schatting wordt besproken en het blijkt dat het moeilijk is om lichte aantastingen nauwkeurig te schatten. Er worden schattingen gegeven van de detectiegrens in afhankelijkheid van de steekproefgrootte.

Meeldauwaantastingen aan de onderkant van het blad, kunnen worden geschat uit de aantasting op de bovenkant van het blad. Deze methode levert een tijdsbesparing op, maar ook een extra onnauwkeurigheid. Alleen bij lichte aantastingen en kleine steekproeven is deze methode efficiënter dan een directe tweezijdige bemonstering. Het schatten van meeldauw op de bovenkant van bladeren is, hoewel algemeen gebruikelijk, waarschijnlijk niet de meest efficiënte methode.

References

Alvey, N.G. et al., 1983. Genstat. A general statistical program. Rothamsted Experimental Station, NAG, Oxford.

Chatterjee, S. & Price, B., 1977. Regression analysis by example. John Wiley & Sons, p. 74-100.

Koch, H., 1978. Die verteilung von *Erysiphe graminis* DC. f. sp. *Hordei Marchal* in Gerstenbeständen und ihr Einfluss auf die Stichprobenauswahl für Befallserhebungen. Dissertation, Giessen, 120 pp.

Rao, C.R., 1965. Linear statistical inference and its application. 2nd edition. John Wiley & Sons, (p. 96-98).

Rouse, D.I., MacKenzie, D.R., Nelson, R.R. & Elliot, V.J., 1981. Distribution of wheat powdery mildew incidence in field plots and relationship to disease severity. Phytopathology 71: 1015-1020.

Rouse, D.I., Nelson, R.R. & MacKenzie, D.R., 1980. A stochastic model of horizontal resistance based on frequency distributions. Phytopathology 70: 951-954.

Taylor, L.R., 1961. Aggregation, variance and the mean. Nature (London) 189: 732-735.

Ubbels, E. & Vliet, G. van der, 1979. Beoordeling van tarwe op aantasting door schimmelziekten. IPO Mededeling Nr. 798, 72 pp.

Zadoks, J.C., Chang, T.T. & Konzak, C.F., 1974. A decimal code for the growth stages of cereals. Bulletin Eucarpia 7: 42-52.