Bernd Truberg[1,2], Thilo Hammann[1], Ulrich Darsow[1], Hans-Peter Piepho[3]

Measuring late blight attack of potato foliage in field trials: optimal resource allocation in assessment trials

Erfassung von Krautfäulebefall in Feldversuchen: Optimale Faktorallokation in der Befallserhebung

Abstract

An empirical data set was analysed in order to give recommendations on the optimal resource allocation in a field testing system to measure late blight attack in potato. The data set was derived from an experiment comprising 854 genotypes, three years, two replicates per year, and 16 to 18 scoring dates per year. AUDPC (area under disease progress curve) values were calculated based on percentage of attacked haulm. Artificial inoculation was used to establish late blight in the testing field. Three testing years, two replicates per year, and three scoring dates per year are recommended to be sufficient. The results are based on the assumption, that the quality of data is independent of the frequency of data collection. This assumption is critically discussed.

Key words: *Solanum tuberosum*, *Phytophthora infestans*, AUDPC, field trial, variance components

Zusammenfassung

Ein empirischer Datensatz wurde analysiert, um Empfehlungen zur optimalen Faktorallokation in einem Feldversuchssystem zur Erfassung des Krautfäulebefalls bei Kartoffel zu geben. Der Datensatz stammt von einem Experiment mit 854 Genotypen, drei Jahren, zwei Wiederholungen pro Jahr und 16 bis 18 Boniturterminen pro Jahr. AUDPC (area under disease progress curve) Werte wurden auf der Basis von Prozent Krautbefall berechnet. Zur Etablierung der Krautfäule im Bestand wurde künstliche Inokulation verwendet. Drei Prüfjahre, zwei Wiederholungen pro Jahr, und drei Boniturtermine werden als ausreichend empfohlen. Diese Ergebnisse gehen von der Voraussetzung aus, dass die Qualität der Daten unabhängig ist von der Häufigkeit der Datenerhebung. Diese Voraussetzung wird kritisch diskutiert.

Stichwörter: *Solanum tuberosum*, *Phytophthora infestans*, AUDPC, Feldversuch, Varianzkomponenten

Introduction

Late blight of potato (*Solanum tuberosum*) caused by the oomycete *Phytophthora infestans* is the economically most important disease in commercial potato production worldwide. Breeding for late blight resistance is – apart from the use of agrochemicals – the most promising way to control the disease. To determine the resistance of genotypes in the selection process different methods are in use. Field tests under natural infestation as well as field tests with artificial inoculation in plots with no fungicide treatment are used to characterize genotypes with respect to their late blight resistance reaction. Field trials – as compared to laboratory tests – are highly representative for conditions in commercial production, but are quite

Institution
Julius Kühn Institute, Federal Research Centre for Cultivated Plants, Institute for Breeding Research on Agricultural Crops, Quedlinburg, Germany[1]
since 2010 NORIKA Nordring-Kartoffelzucht- und Vermehrungs-GmbH, Gross Lüsewitz Sanitz, Germany[2]
University of Hohenheim, Department of Crop Production and Grassland Research (340), Bioinformatics (340c), Stuttgart, Germany[3]

Correspondence
Dr. Bernd Truberg, NORIKA Nordring-Kartoffelzucht- und Vermehrungs-GmbH, Parkweg 4, 18190 Sanitz, Germany, E-Mail: truberg@norika.de

Accepted
October 2009

expensive to be conducted. Thorough planning of experiments makes sure that a maximum of information is generated from the money invested into field trials. Informed decision making by the experimenter requires knowledge about the effects associated with the dimensions of a trial. The present study investigates the effects of the dimensions "number of scoring dates per year", "number of replicates", and "number of testing years" on the information generated from trials to determine resistance against late blight of potato genotypes. The results will help experimenters to allocate their resources in an optimal way.

Material and Methods

Experimental Data

The field experiments were conducted at the experimental station in Groß Lüsewitz, Germany, (012° 15' E, 54° 15' N) in the years 2004, 2005, and 2006. The experimental design was a randomized complete block design (RCBD) with two replicates. Trials were artificially inoculated when flowering of the medium early commercial variety "Adretta" reached its end (mid July). A suspension with a concentration of 12,000 zoosporangia/ml was used. It was applied using a hand-held sprayer. One to two leaves close to the ground of only the first plant in each row were inoculated from the bottom side. Inoculation was done with ambient temperature below 16°C and at increasing humidity in the evening. Sprinkler irrigation and wind shield in form of a hemp strip around the field helped in creating a micro climate supportive of late blight infection (Darsow, 2008). Material for the spore suspension was derived from an isolate with high quantitative pathogeneity containing all virulences from V1 to V11. The isolate was maintained on tuber slices.

Late blight infestation was measured as percent of potato tops in the plot at 16 (2004 and 2006) to 19 (2005) scoring dates. 854 clones from four experimental crosses were scored.

AUDPC values

AUDPC stands for "Area Under Disease Progress Curve". The AUDPC value was calculated according to the formula used by Shaner and Finney (1977) with a slight modification:

$$(1)\qquad AUDPC = \sum_{i=1}^{n-1}\left(\left(\frac{x_{i+1}+x_i}{2}\right)(t_{i+1}-t_i)\right)$$

with

i – index for scoring date
x_i – percentage of foliage infestation at scoring date i
t_i – scoring date i expressed in days after scoring date 1
n – total number of scoring dates in the trial

AUDPC values are a measure for the level of late blight attack (Colon, 1994; Andrivon et al., 2006). The positive genetic correlation between the level of late blight attack and late maturity, the latter being an undesirable trait, has lead to the development of a range of measures for maturity-corrected resistance (Bormann et al., 2004; Darsow and Hansen, 2004; Emrich et al., 2008; Truberg et al., 2009). These measures are all based on AUDPC and maturity data. In this paper all analyses are based on AUDPC values, although determination of foliage blight resistance requires a separation of the maturity effect from the true resistance. To eliminate the error connected with the measurement of maturity, AUDPC values were used as reference instead of maturity-corrected resistance, which commonly is used as a measure in selection experiments.

Heritability

For each of the scoring dates the heritability on a trial mean basis for percent infested foliage was calculated using the formula

$$(2)\qquad h^2 = \frac{VarG}{VarG + (VarErr/r)}$$

with

h^2 – heritability
$VarG$ – genetic variance component
$VarErr$ – residual variance component
r – number of replicates

The heritabilities on a trial mean basis for each of the single scoring dates were used to describe the information content of each of the single scoring dates.

The heritability on a mean-across-years basis for the AUDPC values was calculated using the formula

$$(3)\qquad h^2 = \frac{VarG}{VarG + (VarGY/y) + (VarErr/(y \times r))}$$

with

h^2 – heritability
$VarG$ – genetic variance component
$VarGY$ – genotype-year interaction variance component
$VarErr$ – residual variance component
y – number of years
r – number of replicates

The heritabilities on a mean-across-years basis for each of the AUDPC values were used to describe the information content of each of the single scoring dates in respect to their usefulness to differentiate between tested genotypes.

Correlations of percent foliage infestation values at single scoring dates with final AUDPC values

Correlations of percent foliage infestation values at single scoring dates with final AUDPC values were calculated as Spearman's product-moment correlations. The calculated correlations were used to describe the information content of each of the single scoring dates with respect to their usefulness to predict the AUDPC value of the whole season.

Estimation of variance components

Using the syntax of PIEPHO et al. (2003) the following model was used for the estimation of variance components for the percent foliage infestation values at a single scoring date:

Y = R : GENOTYPE + R • GENOTYPE

The underscored term corresponds to the residual error.

Estimates from this model were used in formula (2) to calculate the heritabilities on a trial mean basis for percent infested foliage.

Using the syntax of PIEPHO et al. (2003) the following model was used for the estimation of variance components for the AUDPC values (all effects random):

Y = GENOTYPE + YEAR + YEAR • R + GENOTYPE • YEAR + GENOTYPE • YEAR • R

The underscored term corresponds to the residual error.

Estimates from this model were used in formula (3) to calculate the heritabilities on a mean-across-years basis for the AUDPC values.

Simulation of data sets by random sampling

To investigate how many scoring dates are needed for the calculation of reliable AUDPC values, subsamples with a reduced number of scoring dates were drawn from the original data set. Correlations of AUDPC values calculated from the subsamples with AUDPC values calculated from the original data set were determined.

To simulate data sets with different numbers of scoring dates, random samples were drawn from the complete data set. 1,000 data sets for each level tested were generated by random sampling of scoring dates without replacement from the original data set. Levels were increased from two to eight scoring dates per data set. So in total, 7,000 data sets were generated from the original data set.

For each of the generated data sets the correlation of the AUDPC values with the AUDPC values from the original data set was calculated. The mean value, maximum value, and the minimum value of these correlations for each level of scoring dates in a single year are reported in "Results". The minimum value of the correlations represents a worst-case scenario, i.e. choosing the scoring dates with the lowest possible information content. The maximum value of the correlations represents a best-case scenario, i.e. choosing the scoring dates with the highest possible information content.

Simulation of data sets by systematic sampling

To imitate the approach that would be taken in a practical breeding programme to reduce the number of scoring dates, equal intervals between the single scoring dates were chosen as a restriction in the generation of the simulated data sets. This way, data sets were generated by choosing only every second scoring date from the original data set or every third scoring date etc. Always starting with the first scoring date from the original data set, for choosing every second scoring date two different data sets were generated, for choosing every third data set three different data sets were generated etc. The practice to choose every second scoring date from the original data set is referred to as sampling type 2, the practice to choose every third scoring date is referred to as sampling type 3 etc.

Results

Information content of single scoring dates

Heritabilities of percent foliage infestation values at single scoring dates in the years 2004, 2005, and 2006 are shown in Fig. 1. For all of the three years the following can be said: early in the season heritabilities are relatively low. For the month of August heritabilities show consistently high values. The end of the season shows no clear tendency.

Correlations of percent foliage infestation values at single scoring dates with final AUDPC values in the years 2004, 2005, and 2006 are shown in Fig. 2. For each of the years the correlations show a bell-shaped curve reaching its maximum value at the middle of August.

Heritabilties of AUDPC values with varying numbers of replicates and years

Estimation of variance components for the AUDPC values yielded the following results:

VarG = 615,738
VarY = 57,202
VarGY = 89,587
VarErr = 168,075

Using formula (3) heritabilities were calculated for different numbers of replicates (see Fig. 3) and years (see Fig. 4). Heritabilities show higher values for increasing numbers of replicates with the gain substantially levelling off for numbers greater than two. Heritabilities also show higher values for increasing numbers of years with the gain substantially levelling off for more than three years.

Simulation of data sets by random sampling

The correlations between the AUDPC values from the dataset with different numbers of scoring dates and the AUDPC values from the original data set are shown in Fig. 5. Mean values, maximum values, and minimum values of correlations of AUDPC values calculated from data sets containing different numbers of scoring dates with AUDPC values calculated from the original data set are reported there. The data sets containing different numbers of scoring dates were generated from the original data set by random sampling without replacement.

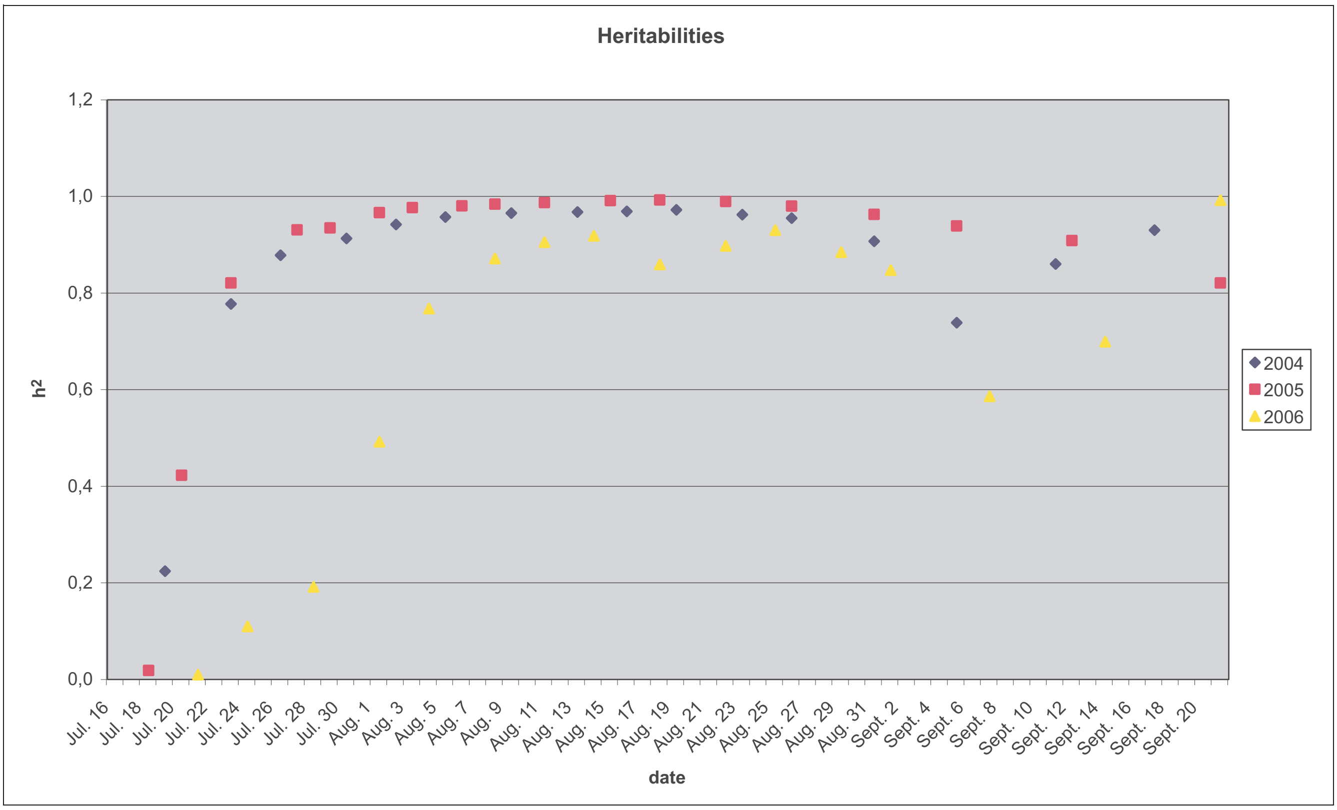

Fig. 1. Heritabilities (h²) of percent foliage infestation values at single scoring dates (date) in the years 2004, 2005, and 2006.

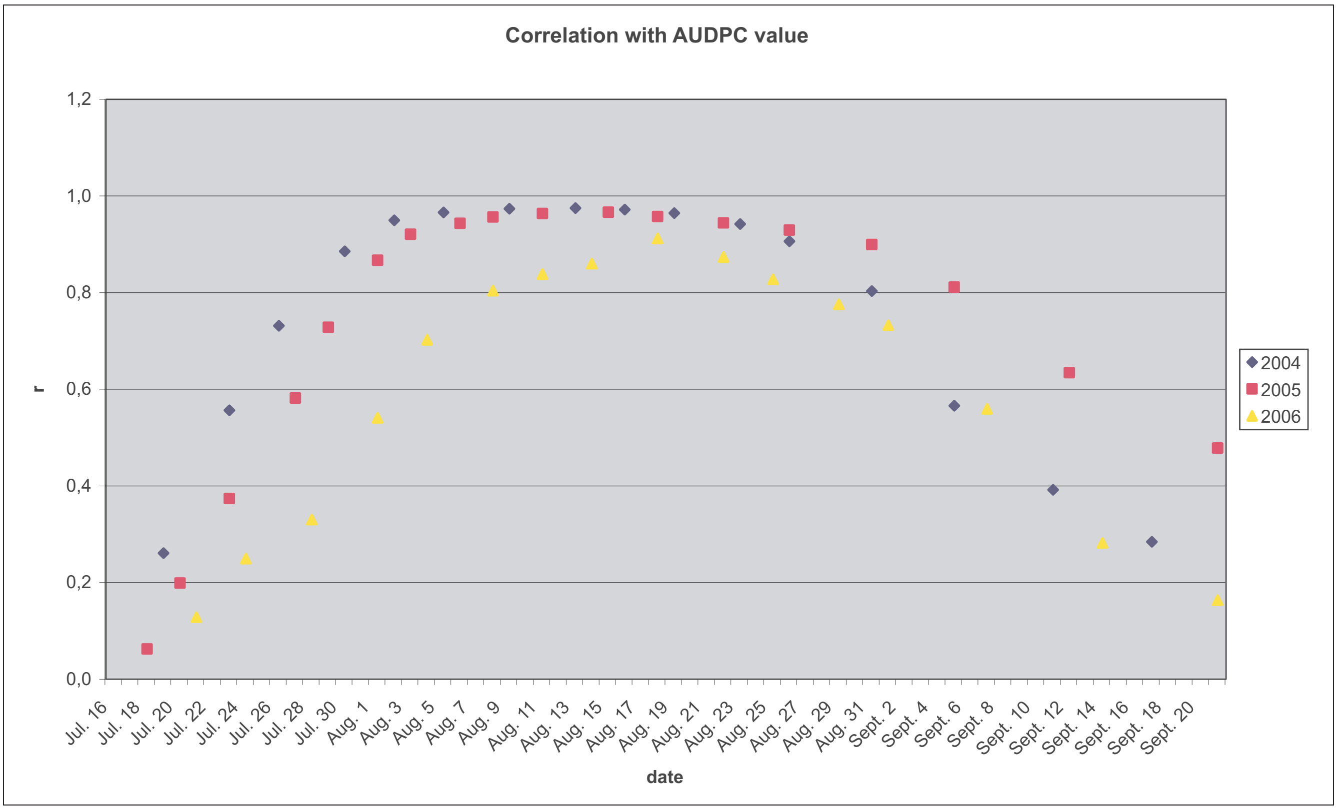

Fig. 2. Correlations (r) of percent foliage infestation values at single scoring dates (date) with final AUDPC values in the years 2004, 2005, and 2006.

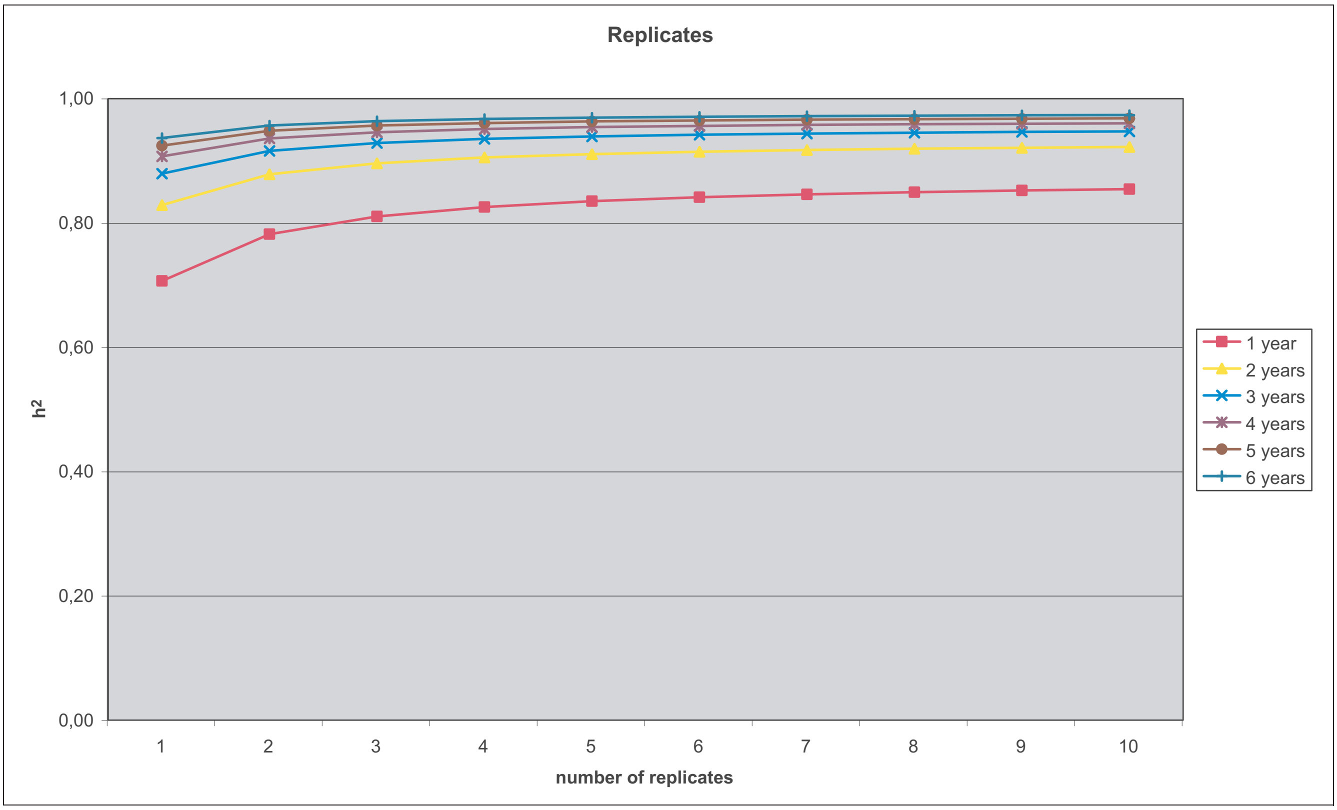

Fig. 3. Heritabilities (h^2) of final AUDPC values for varying numbers of replicates per year calculated using formula (3) with the variance components estimated from experimental data.

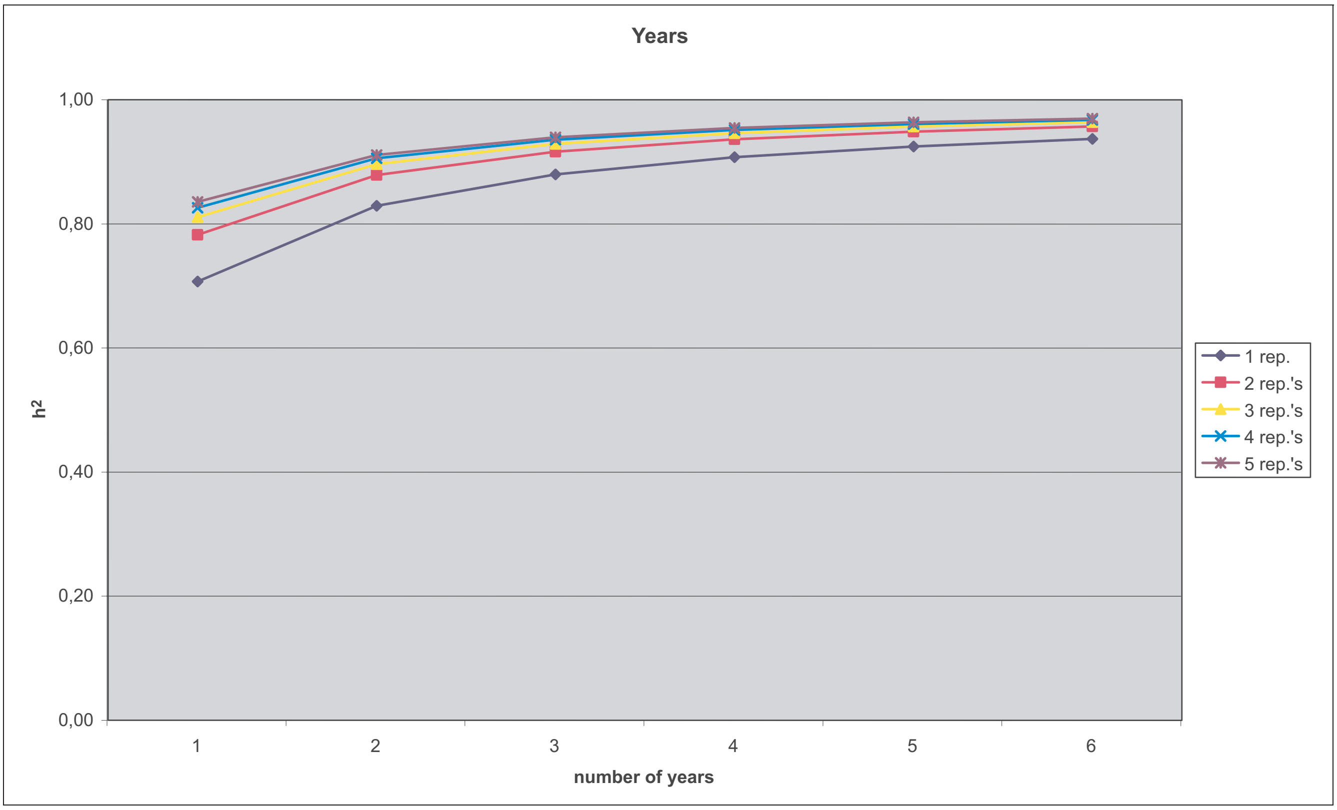

Fig. 4. Heritabilities (h^2) of final AUDPC values for varying numbers of years calculated using formula (3) with the variance components estimated from experimental data.

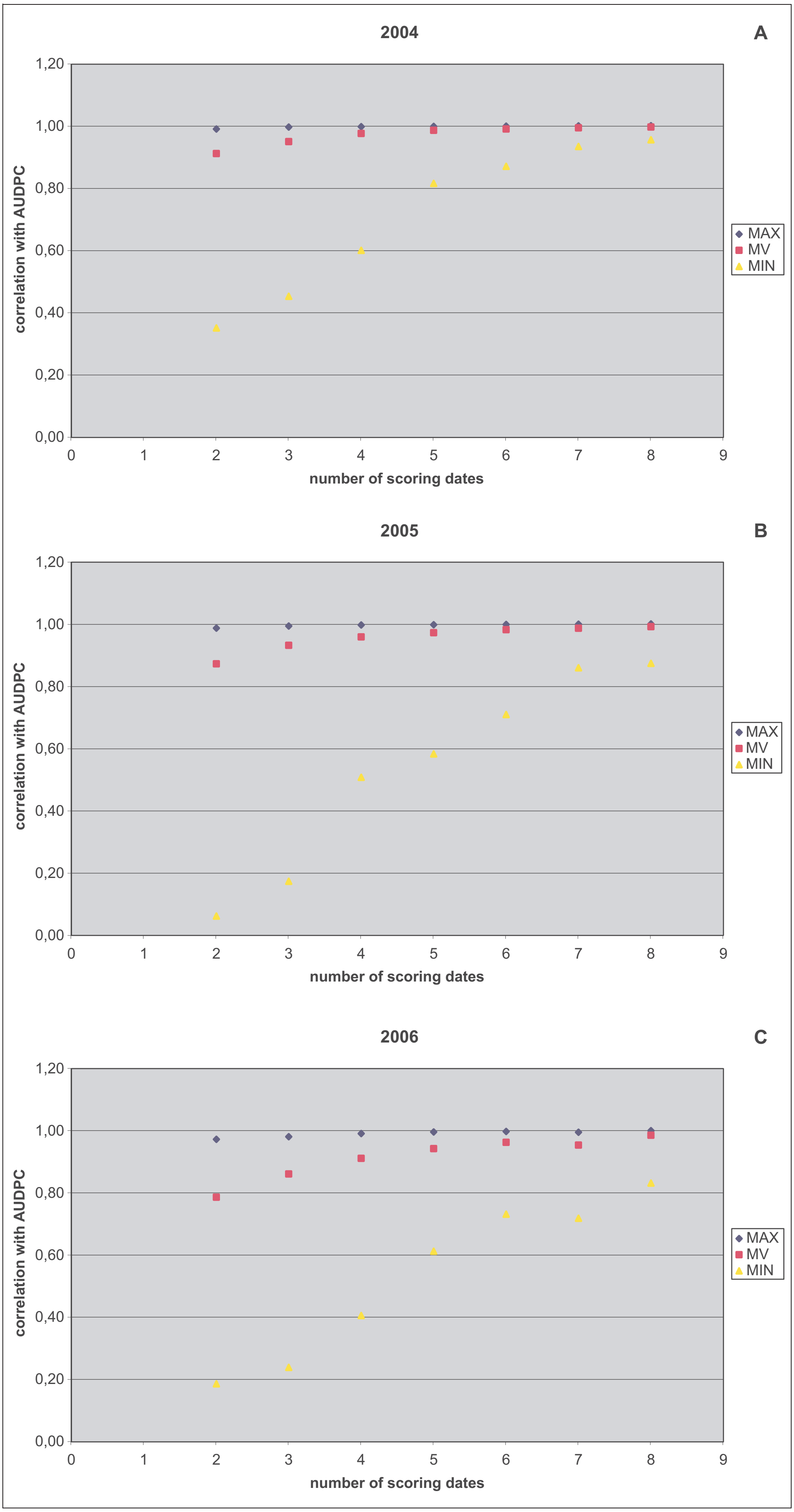

Fig. 5. Mean values, maximum values, and minimum values of correlations of AUDPC values calculated from data sets containing different numbers of scoring dates with AUDPC values calculated from the original data set for the years 2004 (A), 2005 (B), and 2006 (C). The data sets containing different numbers of scoring dates were generated from the original data set by random sampling without replacement. 1,000 samples were drawn for each number of scoring dates.

1,000 samples were drawn for each number of scoring dates. The minimum values of the correlations show a steep increase with higher numbers of scoring dates. Mean values and maximum values show a high level from the beginning with a slight increase for higher numbers of scoring dates.

Simulation of data sets by systematic sampling
The correlations between the AUDPC values from the dataset resulting from different sampling types and the AUDPC values from the original data set are shown in Fig. 6. Mean values of correlations of AUDPC values calculated from data sets resulting from different sampling types with AUDPC values calculated from the original data set are reported there. Only a slight decline in correlations is found with more unfavourable sampling types up to level 5. The lowest value, which is the one for sample type 5 and the year 2006, still lies above 0.95.

Discussion

The results are representative for an experimental set-up with artificial inoculation only. Given this experimental set up, more than three years only lead to a very limited gain in information on the tested genotypes.

Numbers of replicates beyond two hardly give any additional information.

No statement can be made on an optimal number of locations since the data are based on just one location and no variance components for the main effect and interactions can be calculated.

For the number of scoring dates one can see from the maximum values reached at the simulation of data sets by random sampling, that as few as two scoring dates will capture almost all the information for the whole season. This low number of necessary scoring dates found for our data is in good accordance with the findings of JEGER and VILJANEN-ROLLINSON (2001) for the measurement of stripe rust caused by *Puccinia striiformis* f.sp. *tritici* on wheat (*Triticum aestivum*). The high correlations of even single scoring dates with the final AUDPC value further confirm the low number of necessary scoring dates. The simulation of data sets by systematic sampling shows that sampling type 5, which basically leaves three scoring dates in the data set (one in the beginning, one in the middle, and one at the end of the season), still leads to all correlations showing values above 0.95. Choosing the first and the last scoring date more to the centre of the season, even would improve the correlation as can be seen from the results for the single scoring dates. Therefore, three reasonably determined scoring dates can be expected to capture most of the entire information for the whole season from a statistical point of view.

In this paper, the quality of scoring data is assumed to be independent of the number of scoring dates in the

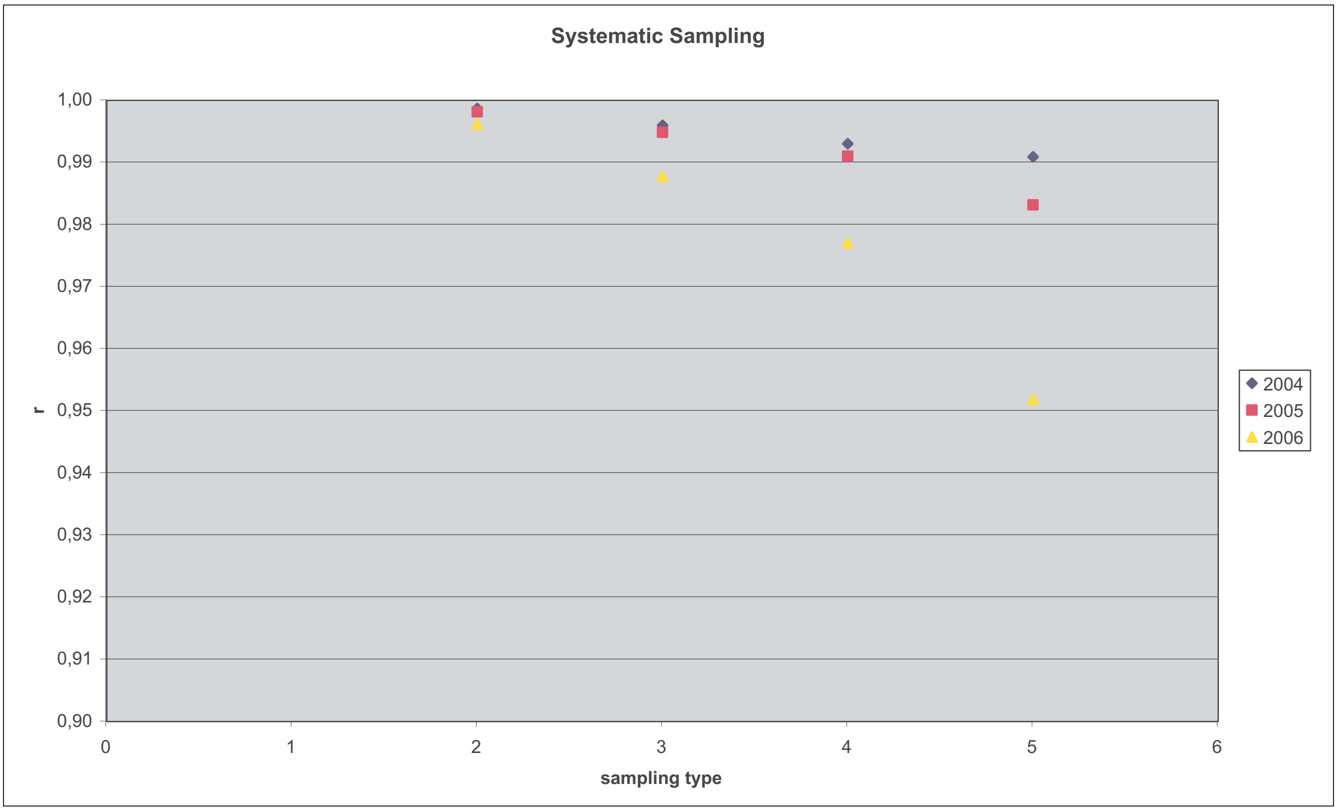

Fig. 6. Mean values of correlations (r) between results from different sampling types and original AUDPC values in the years 2004, 2005, and 2006. The practice to choose every second scoring date from the original data set is referred to as sampling type 2, the practice to choose every third scoring date is referred to as sampling type 3 etc.

experiment. A reduction of the number of scoring dates might have an adverse effect on the ability of the experimenter to reproducibly score late blight and to distinguish haulm damage caused by *Phytophthora infestans* from damage caused by other agents. This possible negative effect of a reduction of the number of scoring dates on the precision of the data is not accounted for in the present analyses.

Acknowledgement

This study is funded by the German Federal Ministry of Economics and Technology (BMWi) by resolution of the German Bundestag.

References

ANDRIVON, D., R. PELLE, D. ELLISSECHE, 2006: Assessing resistance types and levels to epidemic diseases from the analysis of disease progress curves : Principles and application to potato late blight. American Journal of Potato Research **83**, 455-461.

BORMANN, C., A. RICKERT, R. RUIZ, J. PAAL, J. LUEBECK, J. STRAHWALD, K. BUHR, C. GEBHARDT, 2004: Tagging quantitative trait loci for maturity-corrected late blight resistance in tetraploid potato with PCR-based candidate gene markers. Molecular Plant Microbe Interactions **17** (10), 1126-1138.

COLON, L., 1994: Resistance to *Phytophthora infestans* in *Solanum tuberosum* and wild *Solanum* species (1994). Ph.D. Thesis Wageningen, 159 p.

DARSOW, U., 2008: Vorlaufzüchtung der Kartoffel auf quantitative *Phytophthora*-Resistenz im ILK Groß Lüsewitz in der Ressortforschung des BMELV – Stand der Forschung und Züchtung – Pre-breeding for quantitative resistance of potato to late blight at the Insitute of Agricultural Crops Groß Lüsewitz in the departmental reseach of BMELV. -state of research and breeding-. Mitteilungen aus dem Julius Kühn-Institut **415**, 128 pp.

DARSOW, U., J. HANSEN, 2004: Reliability of different parameters to estimate relative foliage blight resistance and its relation to maturity in potato. Plant Breeding and Seed Science **50**, 81-93.

EMRICH, K., F. WILDE, T. MIEDANER, H.P. PIEPHO, 2008: REML estimation for adjusting the Fusarium head blight rating to a phenological date in selection experiments of wheat with artificial infection. Theoretical and Applied Genetics **117**, 65-73.

JEGER, M.J., S.L.H. VILJANEN-ROLLINSON, 2001: The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. Theoretical and Applied Genetics **102**, 32-40.

PIEPHO, H.P., A. BÜCHSE, K. EMRICH, 2003: A hitchhiker's guide to the mixed model analysis of randomized experiments. Journal of Agronomy and Crop Science **189**, 310-322.

SHANER, G., R.E. FINNEY, 1977: The effect of nitrogen fertilization on the expression of slow-mildewing resistance in Knox wheat. Phytopathology **67**, 1051-1056.

TRUBERG, B., T. HAMMANN, U. DARSOW, H.-P. PIEPHO, 2009: Empirischer Vergleich verschiedener Methoden zur Reife-Korrektur von Daten zum Befall mit Krautfäule (*Phytophthora infestans* (Mont.) de Bary) in Selektionsexperimenten bei der Kartoffel (*Solanum tuberosum* subsp. *tuberosum*). Journal für Kulturpflanzen **61** (3), 77-81.