# Rhynchosporium leaf scald disease incidence: seed source and spatial pattern

Topp, CFE; Hughes, G; Nevison, IM; Butler, Adam; Oxley, S; Havis, ND

*Published in:*
Plant Pathology

*DOI:*
10.1111/ppa.13019

Print publication: 01/08/2019

*Document Version*
Peer reviewed version

# Rhynchosporium leaf scald disease incidence: seed source and spatial pattern

C. F. E. Topp[a], G. Hughes[a*], I. M. Nevison[b], A. Butler[b], S. J. P. Oxley[ac], N. D. Havis[a]

[a] Crop & Soil Systems, SRUC, West Mains Road, Edinburgh, EH9 3JG

[b] Biomathematics and Statistics Scotland, James Clerk Maxwell Building, Peter Guthrie Tait Road, The Kings Buildings, EH9 3FD

[c] Agriculture and Horticulture Development Board, Stoneleigh Park, Warwickshire, CV8 2TL

Running head: Rhynchosporium spatial pattern

*Keywords*: Rhynchosporium leaf scald, *Rhynchosporium commune*, winter barley, disease incidence, binary power law, random coefficients model

* Corresponding author: E-mail gareth.hughes@sruc.ac.uk
C.F.E. Topp (E-mail kairsty.topp@sruc.ac.uk) is the submitting author.

A programme of field trials for the study of the winter barley – *Rhynchosporium commune* pathosystem is reported. The associated seed-borne disease Rhynchosporium leaf scald is regarded as having an important impact on barley yields. The analysis reported here relates to the impact of the seed source (commercial or farm-saved seed) on disease incidence, and to the spatial pattern of Rhynchosporium leaf scald disease incidence. Disease incidence data were calculated from field data recorded as disease severity. Mean disease incidence was higher in the crops grown from farm-saved seed than in the crops grown from commercial seed, although we cannot attach great agronomic significance to this result. The spatial pattern of Rhynchosporium leaf scald disease incidence was characterized in terms of the binary power law (BPL), and was indicative of an aggregated pattern. Programme-wide BPL results were described using a novel phytopathological application of a random coefficients model. These results have application in field sampling for Rhynchosporium leaf scald disease.

## Introduction

Rhynchosporium leaf scald is an important global disease of barley crops in cool temperate countries. Epidemics have been reported as far afield as Northern Europe (Shipton et al, 1974; Avrova & Knogge, 2012; Polley *et al*., 1993), North Africa (Bouajila *et al*., 2007), North and South America (Penner *et al*., 1998; Carmona *et al*., 1997) and Australia (Brown, 1985). The disease, caused by the fungal pathogen *Rhynchosporium commune* (formerly *R. secalis*), appreciably reduces barley yields. Estimates of economic damage can vary but losses of over 60% have been recorded in Africa (Semeane, 1995) and up to 35% in North America (Buchannon & Wallace, 1962; Webster, 1980). From a survey of farmer attitudes covering the period 2011 - 2015, Stetkiewicz *et al*. (2018) reported that most farmers

43 believed foliar diseases of spring-sown barley were important or very important in
44 determining crop yield. Rhynchosporium leaf scald was cited by the majority of farmers
45 questioned as being the most common of the three diseases covered by the survey
46 (Rhynchosporium leaf scald, Ramularia leaf spot caused by *Ramularia collo-cygni* and
47 powdery mildew caused by *Blumeria graminis f. sp. hordei*), as well as having the greatest
48 impact on yield. These survey results are indicative of the contemporary relevance of the
49 analysis of previous studies of the Rhynchosporium leaf scald pathosystem, including those
50 reported by Stetkiewicz (2017) on spring barley, and those reported here on winter barley. An
51 extremely useful review of the Rhynchosporium leaf scald pathosystem (Avrova & Knogge,
52 2012) manages to be both comprehensive and concise, and readers are referred there for full
53 details including the pathogen life cycle, disease symptomatology, fungicidal control and host
54 resistance. Rather than repeat this material en bloc, we will refer to it as required in the
55 context of the work reported here.

56 *R. commune* is a seed-borne pathogen. Thus the initial inoculum may be infected seed,
57 although the pathogen may also survive on debris from previous crops, on stubble of previous
58 crops and on volunteers infected from previous crops. The primary inoculum for *R. commune*
59 is considered to arise from crop debris and seed-borne infection with secondary infection due
60 to the release of rain splash spores from infected lesions (Zhan *et al*., 2008; Fountaine *et al*.,
61 2010). The disease is polycyclic and secondary spread occurs via spore dispersal from
62 infected leaves. The name Rhynchosporium leaf scald refers to the foliar symptoms
63 characteristic of the disease (see Avrova & Knogge, 2012; AHDB 2016). In the work
64 reported here, Rhynchosporium leaf scald was recorded in the field by visual assessment of
65 foliar symptoms on winter-sown barley crops in an experimental programme carried out over
66 a five-year period. Here, we investigate the significance of seed source for the level of
67 disease and present an analysis of the spatial pattern of Rhynchosporium leaf scald. The

latter uses a novel phytopathological application of a random coefficients model to provide a statistical overview of results from programme-wide spatial analysis.

## Materials and methods

### Outline of experimental programme

The basic unit of the winter-sown barley experimental programme was a field trial. Trials took place at four centres in Scotland, UK, over a period of five years (harvest years 2005 – 2009) as follows: Aberdeen (two trial sites), Bush (two trial sites near Edinburgh), Perth/Dundee, and Lanark/Lockerbie (see Table 1 for full details). Initially (2005 – 2006), trials included three treatments: farm-saved seed (no fungicide applied), commercial seed (no fungicide applied), and commercial seed (fungicide applied). The commercial seed (fungicide applied) treatment was discontinued after the first two years. Trials were either first- or second-barley in rotation. In total, there were 26 trials potentially available for analysis (Table 1). A detailed description of the use of the trial data in the statistical analyses of treatment effects and of spatial pattern is given for each analysis below.

The basic unit of a trial was a plot, the area to which treatments were applied. All trials comprised four replicate plots of the treatments applied. Thus trials comprised either 12 plots (2005 – 2006) or 8 plots (2007 – 2009). Plots were 12m × 12m in size. For the purpose of data collection, plots were divided into 576 ‘quadrats’, each 0.5m × 0.5m in size.

## Disease assessment

The quadrat was the basic sampling unit for disease assessment. In each plot of each trial, disease symptoms were visually assessed on a continuous severity scale (0 – 100%) in each quadrat. Here, we consider only a single disease assessment for each trial, specifically the one made at or close to GS (growth stage) 31. Figure 1 illustrates such a % severity disease assessment from a 2005 trial. For data analysis, quadrat-level disease severity values were converted to disease incidence, using a detection threshold of 0.5% severity. Thus for quadrat-level disease severity $\leq 0.5\%$, quadrat-level disease incidence was coded "0", and for quadrat-level severity $> 0.5\%$, quadrat-level incidence was coded as "1".

## Statistical analysis of treatment effects

The statistical analysis of treatment effects reported here is a simple comparison of farm-saved seed (no fungicide applied) with commercial seed (no fungicide applied). In two trials, no recorded disease data were available for such a comparison, so the analysis reported here has been carried out on the basis of data from 24 out of the 26 trials (Table 1). Data for a third treatment, commercial seed (fungicide applied), were collected in 10 trials. In five of these trials the fungicide was applied earlier than the date on which the disease assessment under consideration here was made, and in the other five, later. For the five trials in which the plots designated for the commercial seed (fungicide applied) treatment were untreated at the time of the disease assessment, the data are effectively extra replications of data for the commercial seed (no fungicide applied) treatment, and were incorporated in the analysis on that basis. For the five trials in which the plots designated for the commercial seed (fungicide applied) treatment were already treated at the time of the disease assessment, % severity data from the plots for the farm-saved seed (no fungicide applied) treatment and the commercial

seed (no fungicide applied) treatment were still available for the analysis but the data from the commercial seed (fungicide applied) treatment were excluded.

A generalized linear mixed model (GLMM) with the logit link function, binomial error structure and dispersion fixed at unity (Brown & Prescott, 2015) was fitted to the number of quadrats per plot with disease incidence in order to compare treatment effects on mean disease incidence. Individual trials, the trial × treatment interaction and between-plot-within-trial variation were fitted as random effects while the seed source 'treatment' (commercial or farm-saved) was fitted as a fixed effect. Thus the model was of the form

$$\left.\begin{array}{l} y_{jkl} \sim Binomial\left(p_{jkl}, n_{jkl}\right) \\ \ln\left(\dfrac{p_{jkl}}{1-p_{jkl}}\right) = \alpha + \tau_j + \gamma_k + \delta_{jk} + \varepsilon_{jkl} \\ \gamma_k \sim N\left(0, \sigma^2_{trial}\right),\ \delta_{jk} \sim N\left(0, \sigma^2_{trial\times treatment}\right),\ \varepsilon_{jkl} \sim N\left(0, \sigma^2_{individual plot error}\right) \end{array}\right\} \quad (1)$$

in which the subscript $j$ refers to the $j^{th}$ treatment ($j$=1,2), the subscript $k$ refers to the $k^{th}$ trial ($k$=1,…,24), and the subscript $l$ refers to the $l^{th}$ plot ($l$=1,…,12). $y_{jkl}$ is the number of quadrats with disease incidence out of the $n_{jkl}$ quadrats in that plot, and $p_{jkl}$ is the probability of incidence in an individual quadrat in the corresponding plot. $\alpha$ is the grand mean, and $\tau_j$ is the treatment $j$ fixed effect. $\gamma_k$, $\delta_{jk}$ and $\varepsilon_{jkl}$ are the trial, trial × treatment and the individual plot error random effects.

Treatments were applied to entire plots and hence it is appropriate to model plot-level totals without explicitly needing to model spatial pattern within plots. Formal statistical comparisons of treatment effects, using the model (1), must be made on the logit scale and accordingly that is how means and the standard error of difference are presented in the Results section.

To aid understanding, treatment means on the incidence scale have also been presented. In principle, for each of the treatments the mean proportion of disease incidence can be averaged across plots for each trial and then the mean and standard error of these trial means can be computed. An adjustment to this calculation is needed to account for the fact that the number of plots varies between trials and treatments, leading to a lack of balance. We adjust by finding the minimum number of plots per trial × treatment combination ($u$) and then, for each of a large number of simulations ($R$), randomly subsampling exactly $u$ values from each trial × treatment combination in order to simulate a balanced design. For each simulation $r = 1,\ldots,R$ we then calculate the observed mean $m_r$ and associated standard error $s_r$ for each treatment group in the same way that we would calculate these values for a balanced dataset. An overall estimate of the mean for each treatment group, $\overline{m}$, is then given by the mean of the values $m_1,\ldots m_R$, and an overall estimate of the associated standard error by

$$\sqrt{(1/R)\cdot\sum_{r=1}^{R}\left(s_r^2+(m_r-\overline{m})^2\right)}.$$

## Statistical analysis of spatial pattern

The analysis of spatial pattern of *R. commune* presented here is based on the Binary Power Law (BPL) (Hughes & Madden, 1992; Madden *et al*., 2018). The basic unit for the analysis of spatial pattern is the trial (e.g. Figure 1). For each trial, spatial pattern is determined within plots. No reference is made in this analysis to the treatments applied to the plots in a trial. The assumption here is that treatments may affect disease intensity rather in the way that "artificial" methods are used to manipulate levels of disease in experiments designed to study the relationship between crop yield loss and disease intensity (e.g. Sah & MacKenzie, 1987). Thus we are interested, initially, in the variation in disease intensity (recorded as mean

disease incidence) at the plot scale over a trial and, subsequently, the extension of this analysis across trials.

A BPL analysis describes the logarithm of the observed variance of disease incidence as a linear function of the logarithm of the variance of the corresponding random distribution (the binomial distribution). For the analysis described here, we take the variance of the binomial distribution as $p \cdot (1-p)/n$, in which $p$ is estimated by observed mean disease incidence at the plot scale and $n$ is the number of quadrats grouped to form the within-plot sampling unit. Within three trials (2006 Lockerbie, 2009 Aberdeen 1, and 2009 Aberdeen 2, see Table 1) observed mean incidence at the disease assessment reported here was in each case equal to zero in all the trial plots. Within a further three trials (2006 Perth, 2007 Perth, and 2008 Perth, see Table 1), observed mean incidence at the disease assessment reported here was equal to one in all plots for two of the trials, and equal to one in seven out of eight of the trial plots for the third. Data from all these six trials are therefore unsuitable for use in a BPL analysis and have been excluded from further consideration in this context. On that basis, the BPL analysis reported here has been carried out using data from 20 out of the 26 trials (see Table 1).

We motivate our application of BPL methodology by means of an illustration, built on the disease assessment shown in Figure 1. Each of the 12 plots in the trial consists of 24 × 24 quadrats (each 0.5m × 0.5m), for which Figure 1 shows the observations of % disease severity (appropriately binned). As outlined above, these disease severity data were converted to disease incidence for further analysis.

Now consider a single plot. The following description adopts the notation of Madden *et al*. (2018). Suppose we merge groups of $n = 4$ adjacent quadrats (in a 2 × 2 arrangement); in such a group, $X_i = 0, 1, 2, 3$ or 4 represents the number of diseased units (quadrats coded

"1") and $x_i$ (= $X_i/n$) is the proportion of diseased units on a 5-point scale between zero and one. The entire plot comprises $N$ = 144 such groups. We can calculate the proportion of diseased units in a plot (mean disease incidence) as

$$\hat{p} = \frac{\sum X_i}{n \cdot N} = \frac{\sum x_i}{N} \qquad (i = 0, 1, \ldots, N). \tag{2}$$

where the notation indicates that mean disease incidence calculated at the plot scale is an estimate of the probability of disease incidence at that scale, $p$. Then the plot-scale *observed variance* of disease incidence is estimated as

$$\hat{v} = \frac{\sum (x_i - \hat{p})^2}{N - 1}. \tag{3}$$

Suppose that the disease status of a group (of $n$ = 4 adjacent quadrats in a 2 × 2 arrangement) is independent of the disease status of other groups in the same plot; then $x$ has a binomial distribution with variance $v_{bin} = p \cdot (1 - p) / n$. Then the plot-scale *binomial variance* of disease incidence is estimated as

$$\hat{v}_{bin} = \left[\hat{p} \cdot (1 - \hat{p})\right] / n. \tag{4}$$

The binomial distribution is the random distribution for proportions, so for the purpose of an analysis of spatial pattern, a comparison of $\hat{v}$ and $\hat{v}_{bin}$ is of interest. Aggregation (extra-binomial variation, overdispersion) is indicated by $\hat{v} > \hat{v}_{bin}$.

Now, for the trial illustrated in Figure 1, there are 12 plots, each of which yields a value of $\hat{v}$ and $\hat{v}_{bin}$. The BPL can be characterized by $v = A_p \cdot \left[p \cdot (1 - p) / n\right]^b$, in which in practice $v$ and $p$ are replaced by their estimates and the resulting graphical plot has logarithmic scales on both axes (Hughes & Madden, 1992; Madden *et al.*, 2018). Then, as

199 illustrated in Figure 2A, there is typically a linear relationship from which the parameters *b*
200 (slope) and ln($A_p$) (intercept) may be estimated by linear regression based on

201 $$\ln(\hat{v}) = \ln(A_p) + b \cdot \ln(\hat{v}_{bin}). \quad (5)$$

202 We have used natural (i.e. base *e*) logs in BPL equations, although base 10 [i.e. $\log_{10}(\bullet)$] or
203 any other base can be used. The choice of log base has no effect on the slope parameter, but it
204 does affect the intercept.

205 Figure 2 shows four different versions of the analysis described above, resulting from
206 four different ways of internally dividing up each of the 12 plots in the trial. These are
207 described by *n* (the number of adjacent quadrats merged to form a group, always in a square
208 arrangement) and *N* (the number of groups, each of *n* quadrats), such that $n \times N = 576$ (the
209 total number of quadrats in a plot). Thus we have, in order of increasing group size, BPL
210 analyses denoted: G4Q (*n*=4, *N*=144, groups are 1m × 1m), G9Q (*n*=9, *N*=64, 1.5m × 1.5m),
211 G16Q (*n*=16, *N*=36, 2m × 2m), G36Q (*n*=36, *N*=16, 3m × 3m) (Figure 2).

212 When $b = 1$ and $A_p = 1$ (i.e. $\ln(A_p) = 0$), then the BPL reduces to $\hat{v} = \hat{v}_{bin}$; that is, the
213 observed variance equals the binomial variance. The "binomial line" representing this
214 situation is shown in Figure 2 for reference. We now have a basis for interpreting graphical
215 plots such as Figure 2 in relation to spatial randomness of disease incidence (and,
216 particularly, deviations from spatial randomness in the direction of aggregation) for a single
217 trial at the within-plot scale. Estimates of $b = 1$ and $A_p > 1$ (i.e. $\ln(A_p) > 0$) are indicative of
218 aggregation that does not depend on the level of disease incidence. Estimates of $b > 1$ are
219 indicative of aggregation that systematically varies with the level of disease incidence.
220 Typically, most of the observed values are above the reference line. Sometimes observed
221 values may be close to, or even below, the reference line when mean disease incidence is

close to 0 or 1 (i.e. towards the left end of the horizontal axis). At such values, the spatial pattern of disease incidence is indistinguishable from random.

What we require now is a method of extending such analysis across the series of trials available for spatial analysis (Table 1). At each of the four quadrat groupings (G4Q, G9Q, G16Q and G36Q) a random coefficients model (Brown & Prescott, 2015) was fitted across trials for the relationship (after natural log transformation) between the individual plot-level observed variances and the corresponding theoretical variances based on the binomial distribution. This assumes that for each trial that the intercept and slope of the relationship follows a bivariate normal distribution. Thus the model takes into account not only the lack of fit within individual trials but also that the relationship may vary from trial to trial. The random coefficients model is written as

$$\left.\begin{aligned} &\ln\left(\hat{v}_{kl}\right) = \ln\left(A_p\right) + \ln\left(A_{p,k}\right) + \left(b + b_k\right)\cdot \ln\left(\hat{v}_{bin,kl}\right) + e_{kl} \\ &\begin{pmatrix} A_{p,k} \\ b_k \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{A_{p,k}} & \sigma_{A_{p,k}b_k} \\ \sigma_{A_{p,k}b_k} & \sigma^2_{b_k} \end{pmatrix}\right) \end{aligned}\right\} \qquad (6)$$

where ln($A_p$) and $b$ are the fixed effect population intercept and slope, while ln($A_{p,k}$) and $b_k$ are the corresponding intercept and slope random effects for the $k^{th}$ trial. The trial-specific intercepts and slopes respectively are derived by summing their population fixed effect and trial-specific random effect.

## Results

### Statistical analysis of treatment effects

Based on the fitted GLMM, mean disease incidence was significantly different in the crops grown from farm-saved seed compared with the crops grown from commercial seed (P =

243 0.012) (see Table 2). The finding that Rhynchosporium leaf scald mean disease incidence
244 was higher in the crops grown from farm-saved seed than in the crops grown from
245 commercial seed (Table 2) is consistent with what might be expected for a disease caused by
246 a pathogen for which seed-borne inoculum is an important component of the life-cycle, but
247 we cannot attach too much agronomic significance to this result. There are no records of the
248 initial level of seed infection for the two treatments, so in effect the result merely bolsters an
249 assumption about the comparative levels of seed hygiene in farm-saved and commercial seed
250 lots, rather than providing firm evidence.

251

252 Statistical analysis of spatial pattern

253 To facilitate illustration of the application of the BPL at four different quadrat groupings we
254 first considered disease incidence data derived from the disease assessment shown in Figure
255 1. Working at the plot scale, for each grouping, the natural logarithms of the observed and
256 binomial variances were plotted against each other in Figure 2. The fitted relationships
257 between log-transformed variances are shown along with a reference line representing the
258 situation in which the observed variance equals the binomial variance. For the selected trial it
259 is clear from slopes $b > 1$ and intercepts $\ln(A_p) > 0$ that aggregation varied with mean disease
260 incidence (see Madden *et al*., 2018). There was also an apparent relationship between the
261 quadrat size and both the slope and intercept, both increasing as quadrat groupings became
262 larger.

263 While Figure 2 serves to illustrate typical BPL relationships between observed and
264 binomial variances, it does so only for a single trial. Applying the principle of estimating
265 parameters of the BPL to the relationship between observed and binomial variances but now

266 considering all 20 trials available for analysis (see Table 1), the fitted random coefficients
267 model (6) at each quadrat grouping gave estimates of the population averages (fixed effect
268 estimates) of $b$ and $\ln(A_p)$. These are shown in Table 3, from which a relationship between
269 quadrat size and both the slope and intercept is again apparent. In all four cases it can be seen
270 that the estimates of the slope parameter $b$ were statistically significantly greater than 1. It is
271 also evident that the parameter estimates for $b$ increased as quadrat groupings increased in
272 size.

273 For the four different quadrat groupings, each individual trial has its own estimated
274 slopes and intercepts, which are a combination of the population average (fixed effects) and
275 trial-specific deviations (random effects). Key aspects of the respective distributions of the
276 trial-specific slopes and trial-specific intercepts are depicted graphically in Figure 3.

277

278 Discussion

279 Large-scale field experimentation of the kind on which the analysis reported here is based is
280 highly resource intensive. It is therefore advantageous if the results of such experimentation
281 can be applied to aspects of routine crop management that are themselves resource intensive,
282 of which disease surveillance is a prime example. Looking again at Figure 1, it represents a
283 disease assessment recorded by intensive mapping. All the Rhynchosporium leaf scald
284 disease assessments that contributed to the analysis reported here were similarly recorded by
285 intensive mapping. This is feasible for a research project, but not a practical proposition for
286 disease assessments made in the context of crop protection decision making, where data will
287 typically be collected by some kind of sparse sampling.

For the study reported here, Rhynchosporium leaf scald disease intensity was measured in terms of disease severity, but converted to disease incidence for the purpose of statistical analysis of treatment effects and of spatial pattern. As noted by Paul *et al*. (2005), lack of knowledge of the statistical distribution of disease severity as a continuous random variable limits its analytical usefulness, by comparison with disease incidence. The conversion from severity to incidence was carried out at the level of the individual units of disease assessment, so a relationship between mean severity and mean incidence (e.g. Seem (1984), McRoberts *et al*. (2003)) is not required here. In Germany, Rhynchosporium leaf scald intensity is measured as disease incidence in the context of crop protection decision making (Institut für Pflanzenschutz, 2018). In passing, we note that the GS31 (or near) disease assessments reported here show a wide range of disease, whether on a severity scale (Figure 1) or an incidence scale (Table 1). The implication is that GS31 assessments would likely be too late in the winter barley growing season for use in a decision process relating to management of Rhynchosporium leaf scald, at least for cultivars lacking good resistance.

Since it was introduced by Hughes & Madden (1992), the Binary Power Law has been used to characterize aggregation of disease incidence in a large number of pathosystems (Madden *et al*., 2018), although not previously for Rhynchosporium leaf scald of barley. Typically, BPL analyses are illustrated by a graphical plot of the observed variance of disease incidence against the corresponding theoretical binomial variance, with logarithmic scales on both axes. Aggregation varying with mean disease incidence is indicated for slope $b > 1$ and intercept $\ln(A_p) > 0$. Here, this is shown for a single trial in Figure 2, where there is a relationship between quadrat size and both the slope and intercept of the BPL regression.

When it comes to the experimental programme as a whole (comprising 20 trials available for spatial analysis), we used a random coefficients model to obtain an estimated BPL slope and intercept for each trial based on the sum of a population fixed effect and a

trial-specific random effect. To the best of our knowledge, the deployment of a random coefficients model in the context of a programme-wide BPL analysis is a novel phytopathological application. The results (summarized in Figure 3) further illustrate the apparent relationship between quadrat size and both the slope and intercept of the BPL regression, and thus, the correlation between $b$ and $\ln(A_p)$.

The apparent relationship between quadrat size and both the slope and intercept of the BPL regression has also been noted in previous field studies (e.g. Bassanezi *et al*., 2002; Dallot *et al*., 2003; Humeau *et al*., 2006; Batista *et al*., 2008). Although these results are clear, field studies that provide a basis for characterizing observed patterns of disease do not necessarily elucidate the process(es) underlying those patterns. In this respect, a simulation study by Xu & Ridout (2000) is enlightening. The spatiotemporal spread of plant diseases was simulated using a stochastic model to study the effects of initial conditions (number of plants initially infected and their spatial pattern), spore dispersal gradient, and the dimensions of sampling quadrats, on spatial summary statistics (including BPL parameter estimates) for simulated epidemics. Such simulations show BPL parameter estimates increasing with quadrat size (see, for example, Figure 10 in Madden *et al*., 2018). The effects of the size of the sampling unit $n$ arise as a result of the relationship of the BPL parameters $b$ and $A_p$ to the index of dispersion and the intracluster correlation coefficient (measures of spatial aggregation for incidence data), as illustrated in Figure 4 of Madden *et al*. (2018).

The BPL analysis reported here contributes to the design of sampling in two ways. We consider the intensive mapping in terms of $N$ groups of $n$ sampling units (quadrats) each, and use this as a basis to estimate parameters of the BPL describing aggregation of disease incidence. Such analysis allows the specification of cluster sampling designs in which disease incidence may be estimated with a pre-specified level of precision (Madden & Hughes, 1999a). Further, we have considered a range of sizes of sampling unit in our analysis, which

together constitute a spatial hierarchy. Such hierarchies have application in epidemiological research (e.g. Hughes *et al*., 1997; Turechek & McRoberts, 2013; Madden *et al*., 2018), and also in practical disease assessment for crop protection decision making, where hierarchical sampling of groups of sampling units saves on resources devoted to sampling while taking account of the spatial pattern of disease incidence (Madden & Hughes, 1999a;b; Arnold *et al*., 2017). Thus we have an analytical basis for field sampling of barley crops for Rhynchosporium leaf scald disease incidence that can contribute to a process for disease management decision making.

Acknowledgements

The field assessments of disease were carried out by Alison Laing, Susan Moore, Charlie Kasperek and John Swaney. The authors thank Alasdair Sykes for preparing the graphic shown in Figure 1. The authors acknowledge funding support for this work from the Scottish Government's Rural and Environment Science and Analytical Services (RESAS) Division's Strategic Research Programme.

References

AHDB, 2016. *Barley Disease Management Guide*. Stoneleigh Park, UK: Agriculture and Horticulture Development Board.

Arnold K, Golino DA, McRoberts N. 2017. A synoptic analysis of the temporal and spatial aspects of grapevine leafroll disease in a historic Napa vineyard and experimental vine blocks. *Phytopathology* **107,** 418−26.

361 Avrova A, Knogge W, 2012. *Rhynchosporium commune*: a persistent threat to barley
362 cultivation. *Molecular Plant Pathology* **13,** 986–97.

363 Bassanezi RB, Bergamin Filho A, Amorim L, Gimenes-Fernandes N, Gottwald TR, Bove
364 JM, 2003. Spatial and temporal analyses of citrus sudden death as a tool to generate
365 hypotheses concerning its etiology. *Phytopathology* **93,** 502–12.

366 Batista L, Bassanezi RB, Laranjeira FF. 2008. Comparative epidemiology of citrus tristeza in
367 Cuba and citrus sudden death in Brazil. *Tropical Plant Pathology* **33,** 348–55.

368 Bouajila A, Abang MM, Haouas S, Udupa S, Rezgui S, Baum M, Yahyaoui A (2007).
369 Genetic diversity of *Rhynchosporium secalis* in Tunisia as revealed by pathotype,
370 AFLP, and microsatellite analyses. *Mycopathologia* **163,** 281-94.
371 [https://doi.org/10.1007/s11046-007-9012-0] Accessed 6 January 2019.

372 Brown H, Prescott R, 2015. *Applied Mixed Models in Medicine*, 3rd edn. Chichester, UK:
373 John Wiley & Sons Ltd.

374 Brown JS, 1985. Pathogenic variation among isolates of *Rhynchosporium secalis* from
375 cultivated barley growing in Victoria, Australia. *Euphytica* **34,** 129–33.

376 Buchannon KW, Wallace HAH, 1962. Note on the effect of leaf disease on yield, bushel
377 weight and thousand-kernel weight of Parkland barley. *Canadian Journal of Plant*
378 *Science* **42**, 534–6.

379 Carmona MA, Moschini RC, Conti HA, 1997. Meteorological factors influencing the
380 incidence of barley scald and its spatial distribution over the Argentine pampas
381 region. *Journal of Plant Pathology* **79**,203–9.

382 Dallot S, Gottwald T, Labonne G, Quiot JB, 2003. Spatial pattern analysis of sharka disease
383 (*Plum pox virus* strain M) in peach orchards of southern France. *Phytopathology* **93,**
384 1543−52.

385 Fountaine JM, Shaw MW, Ward E, Fraaije BA, 2010. The role of seeds and airborne
386 inoculum in the initiation of leaf blotch (*Rhynchosporium secalis*) epidemics in winter
387 barley. *Plant Pathology* **59,** 330–37.

388 Hughes G, Madden LV, 1992. Aggregation and incidence of disease. *Plant Pathology* **41,**
389 657−60.

390 Hughes G, McRoberts N, Madden LV, Gottwald TR. 1997. Relationships between disease
391 incidence at two levels in a spatial hierarchy. *Phytopathology* 87, 542–50.

392 Humeau L, Roumagnac P, Picard Y, Robene-Soustrade I, Chiroleu F, Gagnevin L, Pruvost O,
393 2006. Quantitative and molecular epidemiology of bacterial blight of onion in seed
394 production fields. *Phytopathology* **96,** 1345−54.

395 Institut für Pflanzenschutz, 2018. *Arbeitsanleitung zum gezielten Fungizideinsatz in*
396 *Praxisbetrieben nach dem erweiterten Gerstenmodell Bayern*. Bayerische
397 Landesanstalt für Landwirtschaft.
398 [https://www.lfl.bayern.de/mam/cms07/ips/dateien/arbeitsanleitung_gerstenmodell_b
399 ayern_praxisbetriebe.pdf] Accessed 28 December 2018.

400 Madden LV, Hughes G, 1999a. Sampling for plant disease incidence. *Phytopathology* **89,**
401 1088−103.

402 Madden LV, Hughes G, 1999b. An effective sample size for predicting plant disease
403 incidence in a spatial hierarchy. *Phytopathology* **89,** 770–81.

404 Madden LV, Hughes G, Bucker Moraes W, Xu X-M, Turechek WW. 2018. Twenty-five
405 years of the binary power law for characterizing heterogeneity of disease incidence.
406 *Phytopathology* **108,** 656–80.

McRoberts N, Hughes G, Madden LV, 2003. The theoretical basis and practical application of relationships between different disease intensity measures in plants. *Annals of Applied Biology* **142,** 191−211.

Paul PA, El-Allaf SM, Lipps PE, Madden LV, 2005. Relationships between incidence and severity of Fusarium head blight on winter wheat in Ohio. *Phytopathology* **95,** 1049−60.

Penner GA, Legge, WG, Tekauz A, 1998. Identification of isolate specific sources of scald resistance in Turkish barley (*Hordeum vulgare*) accessions. *Euphytica* **99,** 111–4.

Polley RW, Thomas MR, Slough JE, Bradshaw NJ, 1993. Surveys of diseases of winter barley in England and Wales, 1981–91. *Annals of Applied Biology* **123,** 287–307.

Sah DN, MacKenzie DR, 1987. Methods of generating different levels of disease epidemics on loss experiments. In: Teng PS, ed. *Crop Loss Assessment and Pest Management*. St Paul, USA: APS Press, 90–6.

Seem RC, 1984. Disease incidence and severity relationships. *Annual Review of Phytopathology* **22,** 137−50.

Semeane Y, 1995. Importance and control of barley leaf blights in Ethiopia, *Rachis* **14,** 83–9.

Shipton WA, Boyd WJR, Ali SM, 1974. Scald of barley. *Review of Plant Pathology* **53,** 839–61.

Stetkiewicz SS, 2017. *An interdisciplinary assessment of the potential for improving integrated pest management practice in Scottish spring barley*. Edinburgh, UK: The University of Edinburgh.

Stetkiewicz SS, Bruce A, Burnett FJ, Ennos RA, Topp CFE, 2018. Perception vs practice: farmer attitudes towards and uptake of IPM in Scottish spring barley. *Crop Protection* **112,** 96–102.

Turechek WW, McRoberts N, 2013. Considerations of scale in the analysis of spatial pattern of plant disease epidemics. *Annual Review of Phytopathology* **51,** 453−72.

Webster RKL, Jackson F, Schaller CW, 1980. Sources of resistance in barley to *Rhynchosporium secalis*. *Plant Disease* **64,** 88–90.

Xu X-M, Ridout MS, 2000. Effects of quadrat size and shape, initial epidemic conditions, and spore dispersal gradient on spatial statistics of plant disease epidemics. *Phytopathology* **90,** 738−50.

Zhan J, Fitt BDL, Pinnschmidt HO, Oxley SJP, Newton AC, 2008. Resistance, epidemiology and sustainable management of *Rhynchosporium secalis* populations on barley. *Plant Pathology* **57,** 1–14.

442 FIGURE LEGENDS

443

444 Figure 1 The diagram shows the outcome of the Rhynchosporium leaf scald disease
445 assessment on 8 April for the 2005 Bush (2) trial (for further details see Table 1). The key
446 indicates % disease severity in 0.5×0.5m quadrats. Graphic prepared by Alasdair Sykes.

447

448 Figure 2 Relationships between observed and binomial variances of disease incidence on
449 natural logarithm axes (equation 1) at the plot scale, from the disease assessment on 8 April
450 for the 2005 Bush (2) trial (see also Figure 1). Based on Equation 5, ordinary least squares
451 (OLS) regression lines fitted to the data are shown (as solid lines) for four different quadrat
452 groupings within plots: (a) G4Q (144 groups of 4 quadrats), $\ln(A_p) = 1.233$ (SE = 0.312), $b$ =
453 1.177 (0.0755); (b) G9Q (64 groups of 9 quadrats), $\ln(A_p) = 1.687$ (0.456), $b = 1.207$
454 (0.0927); (c) G16Q (36 groups of 16 quadrats), $\ln(A_p) = 2.335$ (0.679), $b = 1.278$ (0.124); and
455 (d) G36Q (16 groups of 36 quadrats), $\ln(A_p) = 3.457$ (0.991), $b = 1.374$ (0.158). On each
456 graph, the dashed line is the reference line with $\ln(A_p) = 0$, $b = 1$ which represents the
457 situation in which the observed variance equals the binomial variance.

458

459

460 Figure 3 Boxplots summarizing (a) intercepts ($\ln[A_p]$) and (b) slopes ($b$) of binary power law
461 (BPL) relationships between ln(observed variance) and ln(binomial variance) of disease
462 incidence for 20 individual trials (see Table 1) calculated via random coefficients regression.
463 The solid horizontal line within a box represents the median value, and the top and bottom of
464 a box represent the 75$^{th}$ percentile (Q3) and the 25$^{th}$ percentile (Q1), respectively. The upper
465 vertical line from a box extends to the highest data value within Q3 + 1.5·(Q3 – Q1), the

lower vertical line from a box extends to the lowest data value within Q1 – 1.5·(Q3 – Q1), and individual data values falling beyond the range delimited by the upper and lower vertical lines are indicated by filled circles. Results are shown for four different quadrat groupings within plots: G4Q (144 groups of 4 quadrats), G9Q (64 groups of 9 quadrats), G16Q (36 groups of 16 quadrats) and G36Q (16 groups of 36 quadrats).

Table 1 Winter barley experimental programme details.

| Year | Centre (trial site) | Ordnance survey grid reference | Cultivar (resistance rating)[a] | Disease assessment date | Growth stage | Minimum recorded incidence[b] | Maximum recorded incidence[c] | Treatments analysis (● = yes)[d] | Spatial analysis (● = yes)[e] |
|---|---|---|---|---|---|---|---|---|---|
| 2005 | Bush (1) | NT 253 652 | Sumo (5) | 22-Mar[f] | 24 – 29 | 0.0156 | 0.5278 | ● | ● |
| 2005 | Bush (2)[g] | NT 246 650 | Sumo (5) | 08-Apr[f] | 31 | 0.0035 | 0.9167 | ● | ● |
| 2005 | Lockerbie | NY 115 806 | Sumo (5) | 13-Apr[f] | 31 | 0.9965 | 1.0000 | ● | ● |
| 2005 | Perth | NO 082 179 | Sumo (5) | 12-Apr[f] | 31 | 0.8819 | 1.0000 | ● | ● |
| 2006 | Aberdeen (1) | NJ 904 252 | Sumo (5) | 09-May[h] | 31 | 0.2309 | 0.2674 | X | ● |
| 2006 | Aberdeen (2) | NJ 775 275 | Sumo (5) | 17-May[h] | 32 | 0.0625 | 0.5816 | X | ● |
| 2006 | Bush (1) | NT 243 649 | Sumo (5) | 25-Apr[h] | 31 | 0.9965 | 1.0000 | ● | ● |
| 2006 | Bush (2) | NT 247 653 | Sumo (5) | 25-Apr[h] | 31 | 0.0087 | 0.7951 | ● | ● |
| 2006 | Lockerbie | NY 113 799 | Sumo (5) | 18-Apr[f] | 26 | 0.0000 | 0.0000 | ● | X |
| 2006 | Perth | NO 055 235 | Sumo (5) | 24-Apr[h] | 31 | 1.0000 | 1.0000 | ● | X |
| 2007 | Bush (1) | NT 250 659 | Haka (5) | 24-Apr | 32 | 0.3986 | 0.9130 | ● | ● |
| 2007 | Bush (2) | NT 251 660 | Haka (5) | 24-Apr | 32 | 0.8021 | 0.9462 | ● | ● |
| 2007 | Lanark | NS 907 384 | Haka (5) | 16-Apr | 31 | 0.9844 | 1.0000 | ● | ● |
| 2007 | Perth | NO 045 238 | Haka (5) | 12-Apr | 31 | 1.0000 | 1.0000 | ● | X |
| 2008 | Aberdeen (1) | NJ 874 107 | Haka (5) | 06-Mar | 26 | 0.7413 | 1.0000 | ● | ● |
| 2008 | Aberdeen (2) | NJ 874 107 | Haka (5) | 06-Mar | 26 | 0.8837 | 0.9965 | ● | ● |
| 2008 | Bush (1) | NT 253 656 | Haka (5) | 23-Apr | 30 | 0.0000 | 0.0139 | ● | ● |
| 2008 | Bush (2) | NT 253 656 | Haka (5) | 01-May | 31 | 0.9740 | 1.0000 | ● | ● |
| 2008 | Lanark | NS 904 383 | Haka (5) | 22-Apr | 29 | 0.9774 | 1.0000 | ● | ● |
| 2008 | Perth | NO 048 235 | Haka (5) | 21-Apr | 30 | 0.9983 | 1.0000 | ● | X |
| 2009 | Aberdeen (1) | NJ 874 107 | Saffron (6) | 11-May | 31 | 0.0000 | 0.0000 | ● | X |
| 2009 | Aberdeen (2) | NJ 874 107 | Saffron (6) | 12-May | 31 | 0.0000 | 0.0000 | ● | X |
| 2009 | Bush (1) | NT 246 650 | Saffron (6) | 30-Apr | 30 | 0.0017 | 0.7205 | ● | ● |
| 2009 | Bush (2) | NT 246 649 | Saffron (6) | 20-Apr | 31 | 0.8941 | 0.9983 | ● | ● |
| 2009 | Lanark | NS 906 381 | Saffron (6) | 15-Apr | 26 | 0.0417 | 0.4983 | ● | ● |
| 2009 | Dundee | NO 303 332 | Saffron (6) | 24-Apr | 31 | 0.0035 | 0.0816 | ● | ● |

[a] See the Recommended List archive at https://cereals.ahdb.org.uk/varieties/ahdb-recommended-lists.aspx. The varieties Sumo, Haka and Saffron were sown in 2005-2006, 2007-2008, and 2009, respectively. All three varieties had similar disease resistance ratings.
[b] Minimum mean disease incidence at the plot scale (based on the smallest number of quadrats coded "1" out of 576).
[c] Maximum mean disease incidence at the plot scale (based on the largest number of quadrats coded "1" out of 576).
[d] In this column, X indicates a trial in which only plots for the commercial seed (fungicide applied) treatment were assessed, and that fungicide had been applied prior to the disease assessment. These trials were excluded from the analysis of treatment effects.
[e] In this column, X indicates a trial in which the range of mean disease incidence at the plot scale was insufficient for spatial analysis.
[f] Trial in which the plots designated for the commercial seed (fungicide applied) treatment were fungicide treated after the date of the disease assessment. These plots were included in the analysis of treatment effects as extra replicates of the commercial seed (no fungicide applied) treatment.
[g] See Figure 1.
[h] Trial in which the plots designated for the commercial seed (fungicide applied) treatment were fungicide treated before the date of the disease assessment. These plots were excluded from the analysis of treatment effects.

Table 2 Comparison of treatment effects on disease incidence.

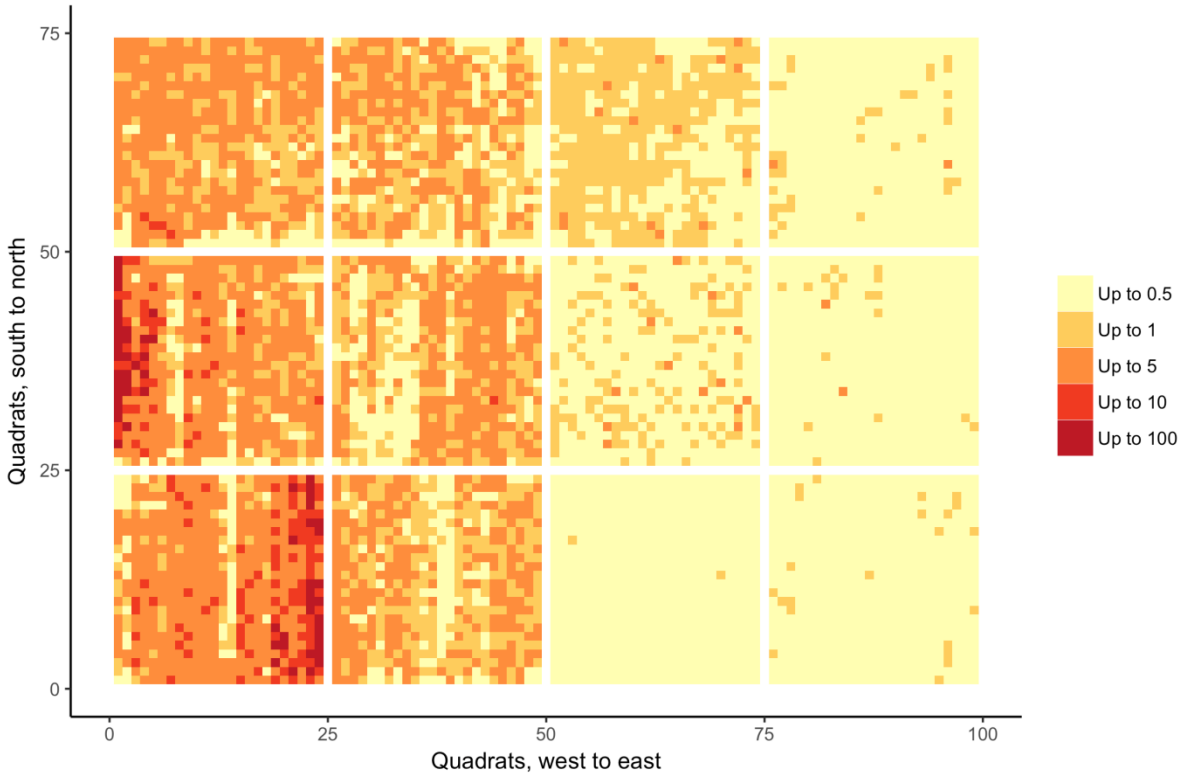
| | Seed source | |
|---|---|---|
| | Commercial | Farm-saved |
| Mean (logit scale) | 1.086 | 1.988 |
| | SED[a] = 0.322 | |
| Mean (incidence scale) | 0.615 | 0.649 |
| SEM[b] | 0.0881 | 0.0840 |

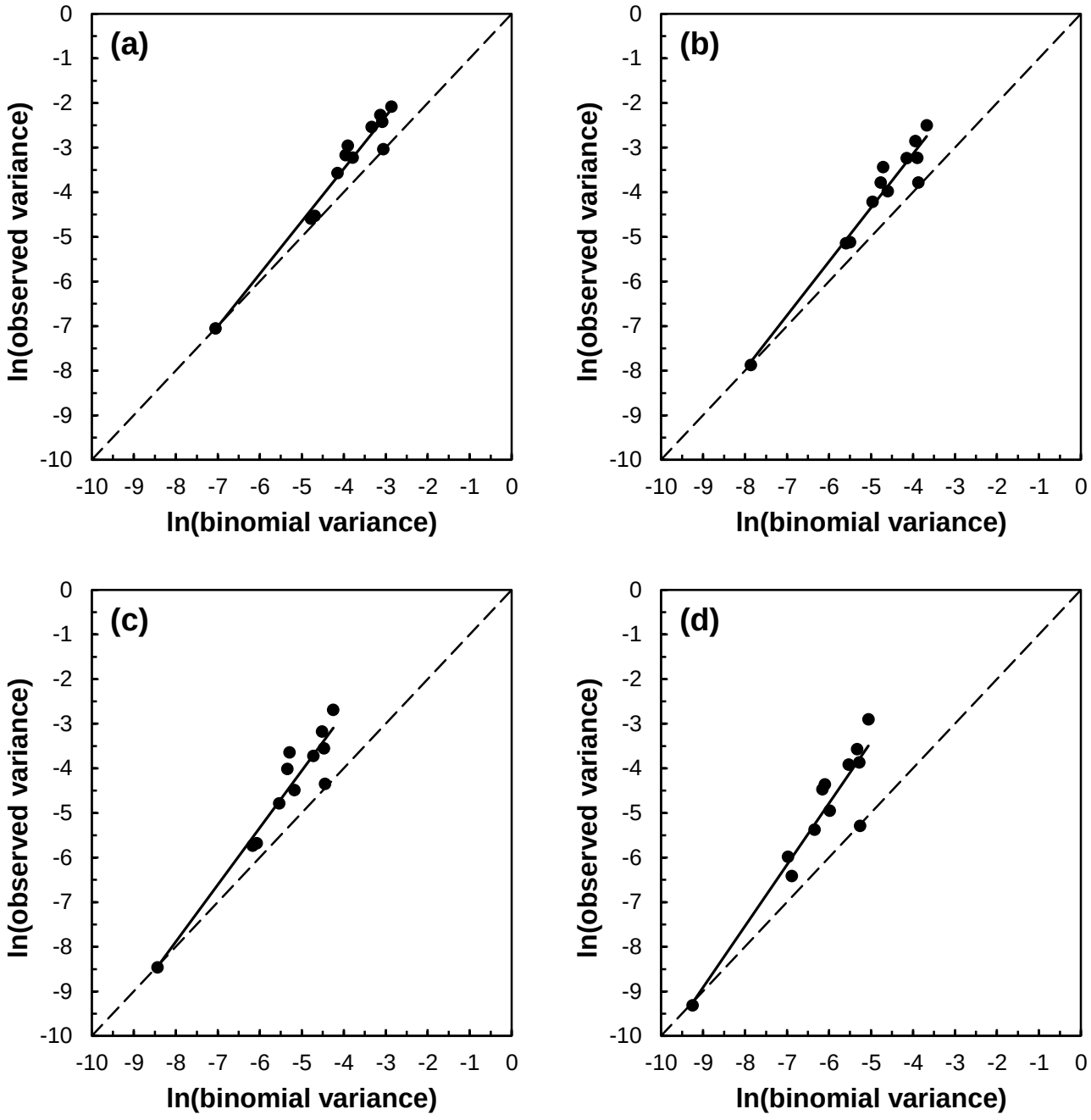
[a] standard error of the difference, 17 d.f.

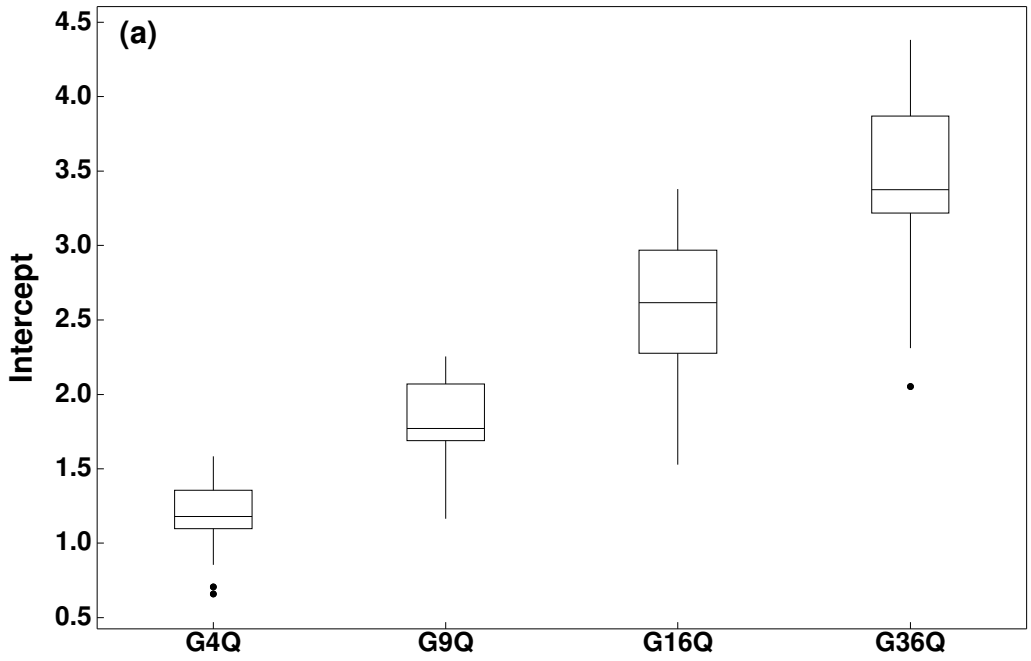
[b] standard error of the mean

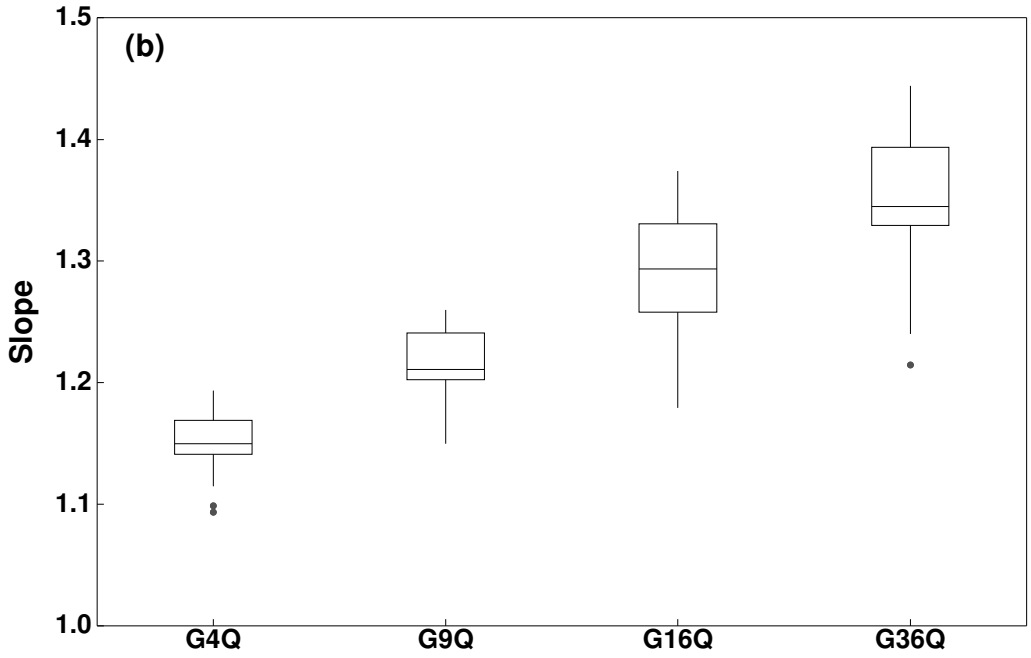
Table 3 Binary power law coefficients from random coefficients model fitted across trials.

| | Coefficient | |
|---|---|---|
| | Intercept | Slope |
| Quadrat grouping | Estimate (standard error) | Estimate (standard error) |
| G4Q | 1.166 (0.102) | 1.148 (0.016) |
| G9Q | 1.818 (0.141) | 1.215 (0.021) |
| G16Q | 2.583 (0.213) | 1.290 (0.028) |
| G36Q | 3.433 (0.293) | 1.350 (0.036) |


75
50
25
0
Quadrats, south to north
0
25
50
75
100
Quadrats, west to east
Up to 0.5
Up to 1
Up to 5
Up to 10
Up to 100


(a)
(b)
(c)
(d)
ln(observed variance)
ln(binomial variance)
0
-1
-2
-3
-4
-5
-6
-7
-8
-9
-10


(a)
Intercept
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
G4Q
G9Q
G16Q
G36Q


(b)
Slope
1.5
1.4
1.3
1.2
1.1
1.0
G4Q
G9Q
G16Q
G36Q