Theor Appl Genet (2014) 127:1009–1018
DOI 10.1007/s00122-014-2275-1

ORIGINAL PAPER

# Dissecting genetic and non-genetic sources of long-term yield trend in German official variety trials

**Hans-Peter Piepho · Friedrich Laidig · Thomas Drobek · Uwe Meyer**

Received: 15 October 2013 / Accepted: 26 January 2014 / Published online: 20 February 2014


**Abstract**
***Key message*** **Long-term yield trends have genetic and non-genetic components which may be dissected by a linear mixed model with regression terms. Disease-resistance breakdown must be accounted for in the interpretation.**
*Abstract* Long-term yield trends of crop varieties may be studied to identify a genetic trend component due to breeding efforts and a non-genetic trend component due to advances in agronomic practices. Many such studies have been undertaken, and most of these inspect trends either by plotting means against years and/or by some kind of regression analysis based on such plots. Dissection of genetic and non-genetic trend components is a key challenge in such analyses. In the present paper, we consider mixed models with regression components for identifying different sources of trend. We pay particular attention to the effect of disease breakdown, which is shown to be confounded with long-term genetic and non-genetic trends, causing an overestimation of genetic trends based on long-term yield trial data. The models are illustrated using German multi-environment trial data on yield, mildew and Septoria leaf blotch susceptibility for winter wheat and yield, mildew and net blotch susceptibility for spring barley.

This paper is dedicated to Professor Dr. Hanspeter Thöni on the occasion of his 80th birthday.

Communicated by I. Mackay.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00122-014-2275-1) contains supplementary material, which is available to authorized users.

H.-P. Piepho (✉)
Bioinformatics Unit, Institute of Crop Science, University of Hohenheim, Fruwirthstrasse 23, 70599 Stuttgart, Germany
e-mail: piepho@uni-hohenheim.de;
hans-peter.piepho@uni-hohenheim.de

F. Laidig · T. Drobek · U. Meyer
Bundessortenamt, Osterfelddamm 80, 30627 Hannover, Germany

## Introduction

Mean yields of many crops typically show a long-term upward trend. This trend is partly due to genetic improvement of newly released crop varieties. Advancements of agronomic practices and other environmental changes account for the remainder of the increase. There is considerable interest in studies dissecting genetic and environmental causes of yield trend in crops, and several studies with this focus have been published (e.g. Schuster et al. 1977; Silvey 1978, 1981, 1986; Perry and D'Antuono 1989; Schuster 1997; Peltonen-Sainio et al. 2009; Ahlemeyer and Friedt 2011; Mackay et al. 2011; Lopes et al. 2012). Most of these studies use some kind of regression analysis or plots of year and/or genotype means against time to assess time trends.

A particular challenge in analysing time trends is to disentangle genetic trends due to breeding efforts from non-genetic trends due to agronomic progress (farm machinery, weed control, disease and pest control, use of growth regulators, and other contributory causes), climate change, etc. Schuster et al. (1977) and Schuster (1997) fitted a regression of simple year means based on year-wise analyses against time and then determined genetic trend by subtraction of a trend line fitted for a set of standard varieties (checks) that was grown for a longer period, taking the latter trend to represent agronomic advance. Mackay et al. (2011) fitted a linear mixed model with factors genotype, location and year. They used adjusted year means, plotted

against time, to assess non-genetic trend and adjusted genotype means, plotted against the year when a variety entered the trial, to assess genetic trend. We here consider an extension of this useful approach.

Mackay et al. (2011) cautioned that due to genetic and non-genetic trends, a simple mixed model with independent random genotype and year main effects would potentially yield biased results because of the underlying trends in genotype and year effects. This problem may be approached by explicitly fitting fixed regression terms for genetic and non-genetic trend, as well as random deviations from these trend components under a suitable linear mixed model. The main purpose of the present paper is to propose a modelling framework for this purpose and to illustrate its use.

The paper is structured as follows. We first describe the motivating datasets from German registration trials for winter wheat and spring barley. The models are then outlined in detail, followed by a presentation of illustrative results for winter wheat and spring barley. The paper ends with a discussion, emphasizing the inherent difficulties in trying to dissect genetic from non-genetic trends based on long-term variety trial data, and considering the complementation of long-term analyses with yield trials contemporaneously evaluating old and new varieties.

## Materials and methods

### Datasets

We use data from official variety trials conducted by the Bundessortenamt (Hannover) in Germany for the crops winter wheat and spring barley. The traits considered in detail are yield and an ordinal mildew susceptibility score visually assessed on a 1–9 scale (1 = not diseased, 9 = fully diseased). We also looked at susceptibility scores for Septoria leaf blotch in winter wheat and net blotch in spring barley. Results are presented in the supplemental material. All trials were laid out as a split-plot design with two replications. Two different intensities were used. Intensity 1 (I1) comprises best agronomic practices with equal fertilizer rates as under intensity 2 (I2) for spring barley in all years and winter wheat from 2005 onwards and without application of crop protection measures. Before 2005 the fertilizer rates for winter wheat were reduced by at least 40 kg per ha compared to intensity 2. Under intensity 2, crop protection against prevailing diseases and pests was applied. At each location, the two intensities were laid out in main plots and varieties in subplots, which were completely randomized within main plots. For the subsequent analyses we use genotype-by-intensity means for location-year combinations, which were available from routine analyses conducted by the Bundessortenamt. We only included genotypes that had undergone at least 3 years of testing (Mackay et al. 2011). Further information on the datasets is given in Table 1.

### Approach for dissecting genetic and non-genetic sources of trend

#### *The basic idea*

For a given intensity, one may develop a model for trend in multi-year variety trial data based on a standard three-way model with factors genotype, location and year (Laidig et al. 2008) as

$$y_{ijk} = \mu + G_i + L_j + Y_k + (LY)_{jk} + (GL)_{ij} + (GY)_{ik} + (GLY)_{ijk}, \quad (1)$$

where $y_{ijk}$ is the mean yield of the $i$th genotype in the $j$th location and $k$th year, $\mu$ is the overall mean, $G_i$ is the main effect of the $i$th genotype, $L_j$ is the main effect of the $j$th location, $Y_k$ is the main effect of the $k$th year, $(LY)_{jk}$ is the $jk$th location × year interaction effect, $(GL)_{ij}$ is the $ij$th genotype × location interaction effect, $(GY)_{ik}$ is the $ik$th genotype × year interaction effect, and $(GLY)_{ijk}$ is a residual comprising both genotype × location × year interaction as well as the error of a mean. This model assumes that locations are crossed with years, i.e. at least some locations are

**Table 1** Basic information on the yield trial data for winter wheat and spring barley

| Crop | Trait | No. of observations | No. of genotypes | Average age of genotype | No. of years | No. of locations | Percentage of genotype-year-location-combinations |
|---|---|---|---|---|---|---|---|
| Winter wheat | Grain yield (dt/ha) | 22,820 | 286 | 3.51 | 30 | 115 | 2.13 |
| | Mildew susceptibility (1–9) | 17,174 | 286 | 3.51 | 30 | 108 | 1.70 |
| | Septoria leaf blotch susceptibility (1–9) | 17,341 | 286 | 3.51 | 30 | 105 | 1.90 |
| Spring barley | Grain yield (dt/ha) | 15,871 | 176 | 3.78 | 30 | 113 | 2.37 |
| | Mildew susceptibility (1–9) | 10,533 | 176 | 3.78 | 30 | 108 | 1.68 |
| | Net blotch Susceptibility (1–9) | 7,874 | 175 | 3.56 | 29 | 88 | 1.62 |

used across several years. If all effects except $\mu$, $G_i$ and $Y_k$ are assumed to be random and independent with constant variance for each effect, genetic and non-genetic time trend may be studied by inspecting the adjusted means for genotypes and years (Mackay et al. 2011). We here carry this approach one step further by explicitly integrating regression terms for time trends.

Time trends will need to be incorporated in an explicit model for main effects of year and genotype. If there is a linear genetic time trend, then $G_i$ should reflect this trend when plotted against the year that a genotype first entered the trials. Thus, we may model $G_i$ explicitly as a function of year of first testing, $r_i$, i.e.

$$G_i = \beta r_i + H_i, \tag{2}$$

where $\beta$ is a fixed regression coefficient for genetic trend and $H_i$ models random deviation of $G_i$ from the genetic trend line. We assume that $H_i$ follows a normal distribution with zero mean and variance $\sigma^2_H$. Similarly, if there is a linear non-genetic time trend, then the year main effect can be modelled as

$$Y_k = \gamma t_k + Z_k, \tag{3}$$

where $\gamma$ is a fixed regression coefficient for non-genetic trend, $t_k$ is the continuous covariate for the calendar year and $Z_k$ is a random residual following a normal distribution with zero mean and variance $\sigma^2_Z$. Genetic and non-genetic trend can be quantified by the regression coefficients $\beta$ and $\gamma$, respectively, directly giving the yield increase per year, measured in the same units as $y_{ijk}$. If graphical inspection of the data, a lack-of-fit test or estimates for $G_i$ and $Y_k$ reveal nonlinearity of trend, then the linear regression models can be replaced by nonlinear models, e.g. quadratic polynomials (Lopes et al. 2012; also see Electronic Appendix A) or splines (Verbyla et al. 1999). Subsequently, in the paper we will frequently use the term agronomic trend in place of non-genetic trend, though we are aware that long-term non-genetic trends may also be due to other causes such as climate change.

*Effect of resistance breakdown*

The trend model developed so far has two linear components. We here repeat the full fixed-effects regression part of our model for clarity:

$$\eta_{ik} = \mu + \beta r_i + \gamma t_k, \tag{4}$$

where $\eta_{ik}$ is the expected response of the $i$th genotype in the $k$th year. A common feature of new varieties is that they may loose part of their genetic potential over time due to disease-resistance breakdown (Silvey 1978; Mackay et al. 2011). This suggests that to fully describe the trend over time, we should add the time elapsed since the first harvest year as a third covariate to capture any effects due to resistance breakdown. This age at testing can be computed for the $i$th genotype in the $k$th year as

$$a_{ik} = t_k - r_i. \tag{5}$$

Adding a linear regression term for this age covariate to our trend model we have

$$\eta_{ik} = \mu + \beta r_i + \gamma t_k + \delta a_{ik}, \tag{6}$$

where $\delta$ is the fixed regression coefficient for the age covariate $a_{ik}$. As it turns out, however, because of the linear dependence of $a_{ik}$ on $r_i$ and $t_k$, and the resulting singularity in the design matrix for (5), the three regression parameters are not independently estimable. Nevertheless, the extension of the model by the age covariate $a_{ik}$ does provide some useful insights. Using the definition (5) in (6), we find

$$\eta_{ik} = \mu + \beta r_i + \gamma t_k + \delta(t_k - r_i), \tag{7}$$

which can be rearranged to give

$$\eta_{ik} = \mu + \tilde{\beta} r_i + \tilde{\gamma} t_k, \tag{8}$$

where

$$\tilde{\beta} = \beta - \delta \tag{9}$$

and

$$\tilde{\gamma} = \gamma + \delta. \tag{10}$$

The Eq. (8) shows that if there is a genetic decay of new varieties, e.g. due to the loss of resistance properties, then a regression on first year of test $r_i$ and calendar year $t_k$ does not give estimates of trend slopes $\beta$ and $\gamma$, but we are then estimating apparent trends $\tilde{\beta} = \beta - \delta$ and $\tilde{\gamma} = \gamma + \delta$, respectively.

If our trait of interest is yield, meaning that high values of the response are desirable, and if a decay of resistance and other desirable properties leads to a decreasing yield trend, then we have $\delta < 0$. Thus, we would then be overestimating the genetic trend ($\tilde{\beta} > \beta$) and we would be underestimating the agronomic trend ($\tilde{\gamma} < \gamma$). In particular, it may happen that the apparent agronomic (non-genetic) trend is estimated to be negative ($\tilde{\gamma} < 0$), although the true trend is positive ($\gamma > 0$). Conversely, if for the trait of interest smaller values for the response are desirable, the same reasoning applies with opposite signs. For example, scores for mildew resistance range from 1 (zero to lowly susceptible, desirable) to 9 (highly susceptible, undesirable). In this case, it may happen, that the apparent agronomic trend is estimated to be positive ($\tilde{\gamma} > 0$), although the true agronomic trend is negative ($\gamma < 0$).

*Estimation of an effect of resistance breakdown ($\boldsymbol{\delta}$)*

The German registration trials are conducted with two intensities. At intensity 1, no fungicides are used, while at

intensity 2, fungicides are applied. If we assume that no ageing effect occurs at intensity 2, because risk breakdown can be fully compensated by plant protection measures, then $\delta = 0$ for intensity 2, so the difference of responses for the two intensities provides information on the size of $\delta$. Alternatively, $\delta$ may also be non-zero for intensity 2, but reduced in absolute value as compared to intensity 1, meaning that compensation by plant protection is not complete. Thus, to be general, we will just assume that $\delta$ is intensity specific. Afterwards we consider the special case that $\delta = 0$ for intensity 2. Thus, from (8), (9) and (10), the regressions for the two intensities are

$$\eta_{ik1} = \mu_1 + \tilde{\beta}_1 r_i + \tilde{\gamma}_1 t_k \tag{11}$$

for intensity 1 and

$$\eta_{ik2} = \mu_2 + \tilde{\beta}_2 r_i + \tilde{\gamma}_2 t_k \tag{12}$$

for intensity 2 with

$$\tilde{\beta}_m = \beta - \delta_m \quad (m = 1, 2) \tag{13}$$

and

$$\tilde{\gamma}_m = \gamma + \delta_m \quad (m = 1, 2), \tag{14}$$

where $m$ indexes the two intensities (I1 and I2). It is assumed in (13) and (14) that the slope parameters $\beta$ and $\gamma$ are the same for both intensities. This strong assumption can be challenged by a lack-of-fit test as will be explained below. Under these assumptions, we can postulate that a regression on $r_i$ and $t_k$ of the difference of the response between intensity 2 and intensity 1 estimates the regression

$$\begin{aligned}\eta_{ik2} - \eta_{ik1} &= (\mu_2 - \mu_1) + \tilde{\beta}_2 r_i + \tilde{\gamma}_2 t_k - \left(\tilde{\beta}_1 r_i + \tilde{\gamma}_1 t_k\right) \\ &= (\mu_2 - \mu_1) + \bar{\beta} r_i + \bar{\gamma} t_k,\end{aligned} \tag{15}$$

where

$$\bar{\beta} = \tilde{\beta}_2 - \tilde{\beta}_1 \tag{16}$$

and

$$\bar{\gamma} = \tilde{\gamma}_2 - \tilde{\gamma}_1. \tag{17}$$

If the relations (13) and (14) hold, we find that $\bar{\beta} = \delta_1 - \delta_2$ and $\bar{\gamma} = \delta_2 - \delta_1$, so that the regression model for the intensity differences becomes

$$\begin{aligned}\eta_{ik2} - \eta_{ik1} &= (\mu_2 - \mu_1) + (\delta_1 - \delta_2) r_i - (\delta_1 - \delta_2) t_k \\ &= (\mu_2 - \mu_1) - (\delta_1 - \delta_2) a_{ik}\end{aligned} \tag{18}$$

Model (18) also follows directly from (7) if we assume that only the regression coefficient $\delta$ differs between intensities. To test the fit of the reduced model (18) compared to the full model (15), we can regress the response difference of intensities 2 and 1 on both $r_i$ and $t_k$ using the model (14) and then test the null hypothesis $H_0 : \bar{\beta} = -\bar{\gamma}$. If no evidence of lack-of-fit is found in this test, the best estimate of $(\delta_1 - \delta_2)$ is obtained by a regression of the difference on $-a_{ik}$ according to model (17). Furthermore, if we assume that $\delta_2 = 0$, we can replace $(\delta_1 - \delta_2)$ with $\delta_1$ in the above derivations. Note, however, that $H_0 : \delta_2 = 0$ is an untestable assumption. Even so, if we find that $(\delta_1 - \delta_2) \neq 0$ in our analyses, we may conclude that there is indeed an ageing effect, although we cannot quantify that effect separately for the two intensities unless we are willing to make the strong, and generally untestable assumption that $\delta_2 = 0$. Thus, the value of the estimate of the coefficient for the regression on $a_{ik}$ should probably not be over-interpreted. Some progress can be made if independent evidence is available on $\beta$, e.g. from yield trials concurrently evaluating old and new varieties (see "Discussion").

*Use of time of first testing ($r_i$) as categorical variable*

In (2) we modelled genetic trend by a linear regression on the time variable $r_i$. Alternatively, we can group genotypes according to levels of the time variable. Thus, we may define a fixed categorical effect $C_p$ for groups $p = 1, \ldots, P$, where $P$ is the number of levels of $r_i$, such that each group is represented by at least one genotype and at least some groups comprise more than one genotype. Thus, the genetic effect can be modelled by

$$G_i = C_p + H_i, \tag{19}$$

where $H_i$ is the random deviation from the trend, as in (2). For visually inspecting trend, it is useful to compute adjusted means for $C_p$ and plot these against first year of testing ($r_i$). We can also use the categorical group effect to test the lack-of-fit of the linear trend model (2). Thus, we may extend (2) as

$$G_i = \beta r_i + C_p + H_i. \tag{20}$$

If in this model the fixed effect $C_p$ is significant, it may be concluded that there is a significant lack-of-fit. In that event, the model (2) may be extended, e.g. by a quadratic term (see Electronic Appendix A), and re-tested for lack-of-fit.

*Use of age at testing ($a_{ik}$) as categorical variable*

To study the effect of the age of a variety, age may be fitted as a categorical effect, i.e. we use the baseline model (1) and replace the genotype-year interaction effect by

$$(GY)_{ik} = D_q + (ZH)_{ik}, \tag{21}$$

where $D_q$ ($q = 1, \ldots, Q$) is the fixed effect for the $q$th age class, $Q$ is the maximum age of a variety in the dataset, and $(ZH)_{ik}$ is a random residual genotype-year interaction with zero mean and variance $\sigma^2_{ZH}$. To visualize trend, we may plot adjusted means for $D_q$ against age ($a_{ik}$).

*Graphical displays for each intensity*

All effects in the proposed models are random, except the intercept $\mu$, the regression coefficients ($\beta$, $\gamma$ and $\delta$) and the group effects $C_p$ and $D_q$. The fixed regression part of the model, i.e. $\mu + \beta r_i + \gamma t_k$ (Eq. 4), is not straightforward to depict in two dimensions because there are two covariates and a response, which involves three dimensions. The following plots can be considered based on the proposed models:

(1) Apparent genetic trend can be depicted by plotting adjusted genotype-group means for $C_p$ based on (19), inserted in the baseline model (1), against time ($r_i$).

(2) Apparent agronomic trend can be depicted by plotting adjusted year means for $Y_k$ using the baseline model (1), inserting (19) to model $G_i$.

(3) Age effects can be assessed by plotting adjusted means for $D_q$ using (21), inserted in the baseline model (1), against age ($a_{ik}$).

(4) Overall trend can be depicted by plotting the regression line

$$\eta_k = \mu + (\beta + \gamma)t_k \tag{22}$$

against time ($t_k$). This is the frontier line representing the varieties released in year $t_k$ as well as the agronomic advance in year $t_k$. Moreover, we can draw genotype-specific regression lines starting at the frontier line, given by Eq. (4), although the graph can become too crowded when there are many genotypes. We can also draw the corresponding line with a genotype-specific starting point depending on the residual effect $H_i$, i.e.

$$\eta_{ik} = \mu + \beta r_i + H_i + \gamma t_k. \tag{23}$$

This kind of plot is illustrated in Fig. 1 for yield trend of three winter wheat genotypes at intensity 1. When the trend is found to be quadratic, plots can be defined analogously.

*Graphical displays for differences of intensities 2 and 1*

For a plot of differences, the frontier line can be defined as

$$\eta_{ik2} - \eta_{ik1} = (\mu_2 - \mu_1) + \left(\bar{\beta} + \bar{\gamma}\right)t_k \tag{24}$$

In addition, all plots used for individual intensities can also be produced for the yield difference.

## Results

We considered using the Kenward and Roger (1997) method for adjusting the denominator degrees of freedom for the Wald-type $F$ and $t$ tests of fixed effects. Looking at a few traits, it turned out that because of the large size of the datasets this had virtually no effect on the outcome of the tests. Thus, subsequently we did not use the Kenward–Roger method in our final analyses for all traits because this saved considerable computing time and allowed us to use the computationally efficient procedure HPMIXED in SAS, which exploits sparse matrix methods. Some of the SAS code we used is given in Electronic Appendix E.

We here report results for a linear trend. For mildew susceptibility trends have a slightly nonlinear component (Fig. 2b, d). Despite some apparent nonlinearities and indication of lack-of-fit (Table 3), we report linear regression results for mildew for ease of interpretation (Tables 2, 4, 5). In addition, we also fitted quadratic models for mildew, which yielded similar conclusions. The quadratic models are described in Electronic Appendices A and B and the numerical results are shown in Electronic Appendix C.

For both crops and traits, the genotype-location and genotype-year interaction variances are small compared to that for the three-way interaction (Table 2). Also, variances for year and location main effects, as well as for year-location interaction are relatively large compared to variances for effects involving genotype. For grain yield the genotype variance is larger than the year variance whereas the reverse relation is true for mildew (Table 2). These results are in good agreement with estimates typically found in these trials (Laidig et al. 2008).

Interestingly, the apparent non-genetic yield trend for wheat under intensity 1 is negative. This result is most likely due to ageing effects. If we assume that intensity 2 captures trends free of age effects ($\delta_2 = 0$), then it may be concluded for winter wheat that most of the trend is due to genetic causes ($\hat{\beta} = 0.53$ dt ha$^{-1}$ year$^{-1}$), amounting to about three times the contribution of non-genetic causes ($\hat{\gamma} = 0.19$ dt ha$^{-1}$ year$^{-1}$) (Table 4). The situation is similar in spring barley ($\hat{\beta} = 0.39$ dt ha$^{-1}$ year$^{-1}$, $\hat{\gamma} = 0.11$ dt ha$^{-1}$ year$^{-1}$) (Table 5). By contrast, for mildew resistance, almost all trends are genetic for intensity 1 (Tables 4, 5). It needs to be re-iterated, however, that this interpretation of results assumes that $\delta_2 = 0$, which is an untestable assumption. In the discussion some evidence is presented suggesting that for winter wheat yield $\delta_2 < 0$. Thus, the regression of yield differences on the age covariate $-a_{ik}$ probably estimates the difference $(\delta_1 - \delta_2)$ rather than $\delta_1$.

The frontier lines for wheat yield have similar slopes for both intensities (Fig. 2a), whereas slopes for apparent genetic and non-genetic yield trends are quite different for both intensities (Table 4). This is most likely due to ageing effects operating mainly under intensity 1. For yield in spring barley, differences in slopes for the two intensities are somewhat more pronounced (Fig. 2c; Table 5). For mildew susceptibility we find that genetic trends are decreasing under intensity 1 for both winter wheat and spring barley (Fig. 2b, d; Tables 4, 5), which likely reflects an ageing

**Table 2** Variance component estimates for winter wheat and spring barley and both traits (yield, mildew score) at intensities 1 and 2 (I1 and I2) based on model (1) using linear trend models (2) and (3)

| Effect | Winter wheat | | | | Spring barley | | | |
|---|---|---|---|---|---|---|---|---|
| | Yield (dt/ha) | | Mildew susceptibility | | Yield (dt/ha) | | Mildew susceptibility | |
| | I1 | I2 | I1 | I2 | I1 | I2 | I1 | I2 |
| $H_i$ (genotype) | 16.5 | 17.2 | 0.426 | 0.089 | 10.5 | 11.8 | 0.506 | 0.158 |
| $Z_k$ (year) | 20.6 | 24.5 | 0.083 | 0.009 | 14.7 | 15.9 | 0.036 | 0.010 |
| $L_j$ (location) | 37.0 | 57.2 | 0.482 | 0.139 | 30.3 | 27.0 | 0.216 | 0.089 |
| $(GL)_{ij}$ | 3.2 | 2.2 | 0.098 | 0.027 | 1.2 | 0.9 | 0.191 | 0.073 |
| $(GY)_{ik}$ | 4.3 | 3.5 | 0.049 | 0.013 | 1.3 | 1.0 | 0.101 | 0.031 |
| $(LY)_{jk}$ | 74.1 | 81.6 | 0.784 | 0.421 | 71.9 | 77.7 | 0.409 | 0.318 |
| $(GLY)_{ijk}$ | 22.4 | 19.6 | 0.615 | 0.237 | 13.2 | 12.9 | 0.673 | 0.407 |

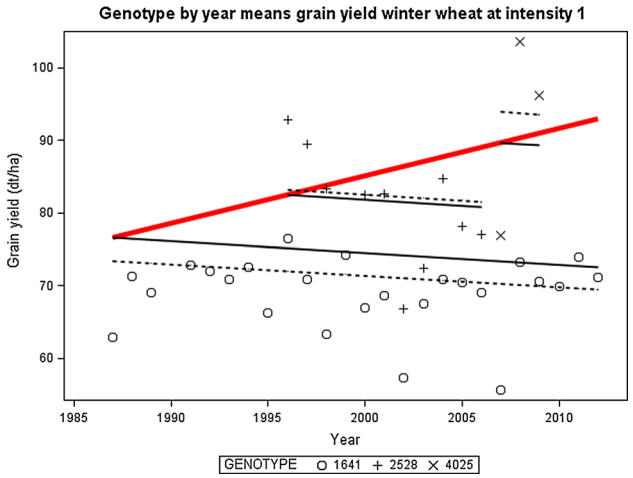


**Fig. 1** Plot for winter wheat grain yield (dt/ha). *Solid red line* Eq. (22), frontier line. *Solid black line* (shown for three genotypes) Eq. (4). *Dashed black line* (shown for three genotypes) Eq. (23). Scatter plot (*three different symbols* for three different genotypes): adjusted genotype-by-year means based on Eq. (1), taking year, genotype and the corresponding interaction as fixed (color figure online)

effect. Under intensity 2, there appears to be little genetic trend. For both traits and crops, there is clear evidence of an ageing effect (Fig. 3; Tables 4, 5). For Septoria leaf blotch in winter wheat, the ageing effect seems to be more pronounced than that for net blotch in barley (Table C5 in Electronic Appendix C and Figures S1 and S2 in Electronic Appendix D).

## Discussion

The linear regression approach did not give a perfect fit in all four cases studied, as evidenced by some significant lack-of-fit tests (Table 3). The problem persisted when quadratic trends were fitted (Electronic Appendix C), so there may be some irregularities in the data preventing a simple trend model from giving a perfect fit. Nonetheless, we think that the linear trend models capture the main features, and this is supported by visual inspection of Figs. 2 and 3. It should be kept in mind that the datasets are rather large, so the power to detect even minor departures from model assumptions such as linearity or independence of regression coefficients from intensities is expected to be high.

In winter wheat, the mildew susceptibility dropped notably for both intensities (Fig. 2b). These changes correlate with intensive resistance breeding efforts up to about the year 1992. Figure 2b, d indicates that there is a gradual increase of mildew susceptibility with increasing age of a genotype due to loss of resistance.

Our results confirm the general experience and the observation by other authors (Silvey 1978, 1981, 1986; Mackay et al. 2011) of a gradual breakdown of disease resistance within a few years after variety release. The resulting decline in yield potential of varieties is probably one of the major driving forces behind the adoption of new varieties with higher yield potentials by farmers. Most of the genotypes in our wheat and barley datasets were in trial for only 3 years (Table 1). The declining effect due to disease-resistance breakdown was probably still fully operative for most varieties for most of the time, including the standard varieties which were in trials for a longer period than the other varieties. Out of the 286 winter wheat genotypes, only 16 % were standard varieties, however, 48 % of the total number of observations are from these standards. For spring barley 14 % were standards comprising 41 % of the total number of observations. It is important to realize that the ageing effect due to disease breakdown confounds estimates of long-term genetic and non-genetic trend, as can be seen from Eq. (8). This problem cannot be fully resolved, unless one is prepared to make some strong assumptions. As was demonstrated in this paper, some progress can be made if two agronomic intensities are tested, one of which (I2) is believed to mask the ageing effect. This assumption is underlying our

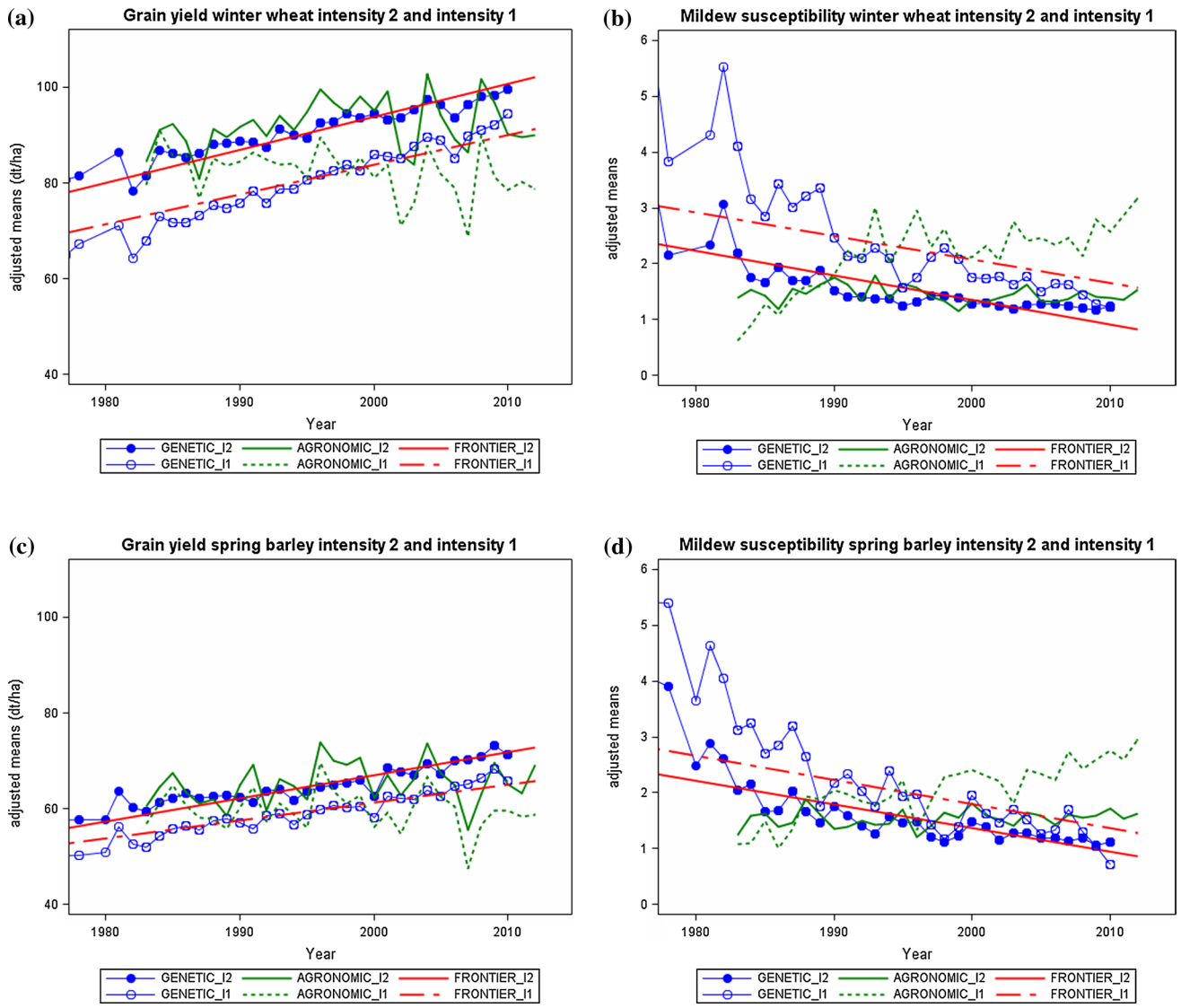


**Fig. 2** Adjusted means for grain yield of winter wheat for intensity 2 (I2) and intensity 1 (I1). **a** Winter wheat, yield. **b** Winter wheat, mildew score. **c** Spring barley, yield. **d** Spring barley, mildew score. GENETIC: variety group means [effect $C_p$ in Eq. (19)]; this assesses genetic trend. AGRONOMIC: year means [Eq. (1), using Eq. (19) to model $G_i$]; this assesses non-genetic trend. FRONTIER: frontier line [Eq. (22)]

**Table 3** Lack-of-fit tests based on three different null hypotheses [see Eqs. (14), (19), and (21)]

| Null hypothesis | Intensity | Winter wheat | | | | Spring barley | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Yield | | Mildew susceptibility | | Yield | | Mildew susceptibility | |
| | | *F* value | *p* value | *F* value | *p* value | *F* value | *p* value | *F* value | *p* value |
| $C_p = 0$ | Intensity 1 (I1) | 0.86 | 0.7029 | 3.29 | <0.0001 | 0.99 | 0.4803 | 1.65 | 0.0109 |
| | Intensity 2 (I2) | 0.75 | 0.8470 | 7.80 | <0.0001 | 1.03 | 0.4121 | 4.08 | <0.0001 |
| | I2–I1 | 1.72 | 0.0062 | 1.70 | 0.0078 | 1.95 | 0.0007 | 3.07 | <0.0001 |
| $D_q = 0$ | I2–I1 | 1.43 | 0.0773 | 5.03 | <0.0001 | 1.01 | 0.4484 | 1.69 | 0.0282 |
| $\bar{\beta} = -\bar{\gamma}$ | I2–I1 | 1.30 | 0.2551 | 0.02 | 0.8757 | 8.19 | 0.0042 | 0.13 | 0.7228 |

$H_0 : C_p = 0$ implies linearity of genetic trend. $H_0 : D_q = 0$ corresponds to linearity of the regression on age

$H_0 : \bar{\beta} = -\bar{\gamma}$ holds if genetic and non-genetic trend can be assumed to be the same for both intensities

**Table 4** Estimates of regression coefficients in mixed model linear regression analyses for winter wheat

| Trait | Intensity | Slope estimates for regression on | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Year of first trial ($r_i$) | | | Calendar year ($t_k$) | | | Years since application ($a_{ik} = t_k - r_i$) | | |
| | | Estimate | *SE* | *p* | Estimate | *SE* | *p* | Estimate | *SE* | *p* |
| Yield | 1 | 0.817 | 0.044 | <0.0001 | −0.197 | 0.111 | 0.0752 | – | – | – |
| | 2 | 0.530 | 0.042 | <0.0001 | 0.161 | 0.119 | 0.1748 | – | – | – |
| | 2–1 | −0.286§ | 0.024 | <0.0001 | 0.360§ | 0.068 | <0.0001 | 0.287 | 0.024 | <0.0001 |
| Mildew | 1 | −0.095 | 0.006 | <0.0001 | 0.052 | 0.009 | <0.0001 | – | – | – |
| | 2 | −0.042 | 0.003 | <0.0001 | −0.002 | 0.005 | 0.6507 | – | – | – |
| | 2–1 | 0.053§ | 0.005 | <0.0001 | −0.055§ | 0.008 | <0.0001 | −0.054 | 0.004 | <0.0001 |

*SE* standard error, *p* *p* value of *t* test

§ Test of $H_0 : \bar{\beta} = -\bar{\gamma}$ not significant at 5 % level (Table 3)

**Table 5** Estimates of regression coefficients in mixed model linear regression analyses for spring barley

| Trait | Intensity | Slope estimates for regression on | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Year of first trial ($r_i$) | | | Calendar year ($t_k$) | | | Years since application ($a_{ik} = t_k - r_i$) | | |
| | | Estimate | *SE* | *p* | Estimate | *SE* | *p* | Estimate | *SE* | *p* |
| Yield | 1 | 0.455 | 0.036 | <0.0001 | −0.081 | 0.095 | 0.3973 | – | – | – |
| | 2 | 0.391 | 0.036 | <0.0001 | 0.093 | 0.098 | 0.3420 | – | – | – |
| | 2–1 | −0.065* | 0.014 | <0.0001 | 0.165* | 0.037 | <0.0001 | 0.067 | 0.014 | <0.0001 |
| Mildew | 1 | −0.098 | 0.009 | <0.0001 | 0.055 | 0.009 | <0.0001 | – | – | – |
| | 2 | −0.052 | 0.005 | <0.0001 | 0.010 | 0.006 | 0.1031 | – | – | – |
| | 2–1 | 0.046§ | 0.006 | <0.0001 | −0.040§ | 0.006 | <0.0001 | −0.046 | 0.005 | <0.0001 |

*SE* standard error, *p* *p* value of *t* test

§ Test of $H_0 : \bar{\beta} = -\bar{\gamma}$ not significant at 5 % level (Table 3)

* Test of $H_0 : \bar{\beta} = -\bar{\gamma}$ significant at 5 % level (Table 3)

approach based on intensity differences. It must be stressed, however, that this approach also relies on the strong assumption that true genetic and non-genetic trends are identical between both intensities. Moreover, analysis based on two intensities only allows estimation of the contrast of ageing effects ($\delta_1 - \delta_2$) for the two intensities, but not of the ageing effects themselves ($\delta_1$, $\delta_2$). Additional experiments would need to be conducted that allow an unequivocal separation of long-term trend from ageing effects.

There are a few studies that compare old and current varieties in joint variety trials (Perry and D'Antuono 1989; Ahlemeyer and Friedt 2011; Lopes et al. 2012). It is interesting to compare results of these experiments with the trends observed in long-term data. When all varieties from different eras are grown together, then for the older varieties the decay due to resistance breakdown is probably complete, while the decay only continues for the newer releases. This could explain the relatively small observed genetic trend for winter wheat in the treated trials by Ahlemeyer and Friedt (2011) (about 0.34 d$t$ ha$^{-1}$ year$^{-1}$ during 1996–2007) under conditions comparable to intensity 2 compared to the genetic trends in long-term data in Mackay et al. (2011) (about 0.74 d$t$ ha$^{-1}$ year$^{-1}$ during 1982–2007) or the apparent trend of 0.53 d$t$ ha$^{-1}$ year$^{-1}$ for wheat yield with intensity 2 in our analyses (Table 4). It may be speculated that the trend in the trials of Ahlemeyer and Friedt (2011) reflects purely genetic trend ($\beta$), while the trends based on long-term multi-environment trial (MET) data provide estimates of apparent genetic trends ($\tilde{\beta}_2 = \beta - \delta_2$ with $\delta_2 < 0$), which overestimate true genetic trends as shown in this paper. Thus, if we assume that $\beta = 0.34$ d$t$ ha$^{-1}$ year$^{-1}$ based on results in Ahlemeyer and Friedt (2011) and that $\tilde{\beta}_2 = \beta - \delta_2 = 0.53$ from our analysis (Table 4), then we may conclude that $\delta_2 = -0.19$ d$t$ ha$^{-1}$ year$^{-1}$, implying a substantial ageing effect even under intensity 2. Moreover, since our estimate of ($\delta_1 - \delta_2$) equals $-0.29$ d$t$ ha$^{-1}$ year$^{-1}$ (Table 4), the ageing effect under intensity 1 can be concluded to equal $\delta_1 = -0.48$ d$t$ ha$^{-1}$ year$^{-1}$. The difference in ageing effects under the two intensities ($\delta_1 - \delta_2$) of $-0.29$ d$t$ ha$^{-1}$ year$^{-1}$ would be due to plant protection measures buffering the effect of disease-resistance breakdown. In summary, our modelling approach, combined with these results, suggests that long-term trial data analyses tend to overestimate genetic trend.

James and Segal (1982) describe a similar confounding problem as the one presented here in the context of so-called age-period-cohort analysis of cancer mortality data. This analysis involves effects of calendar year, age at death and epoch of birth (cohort effects), which are

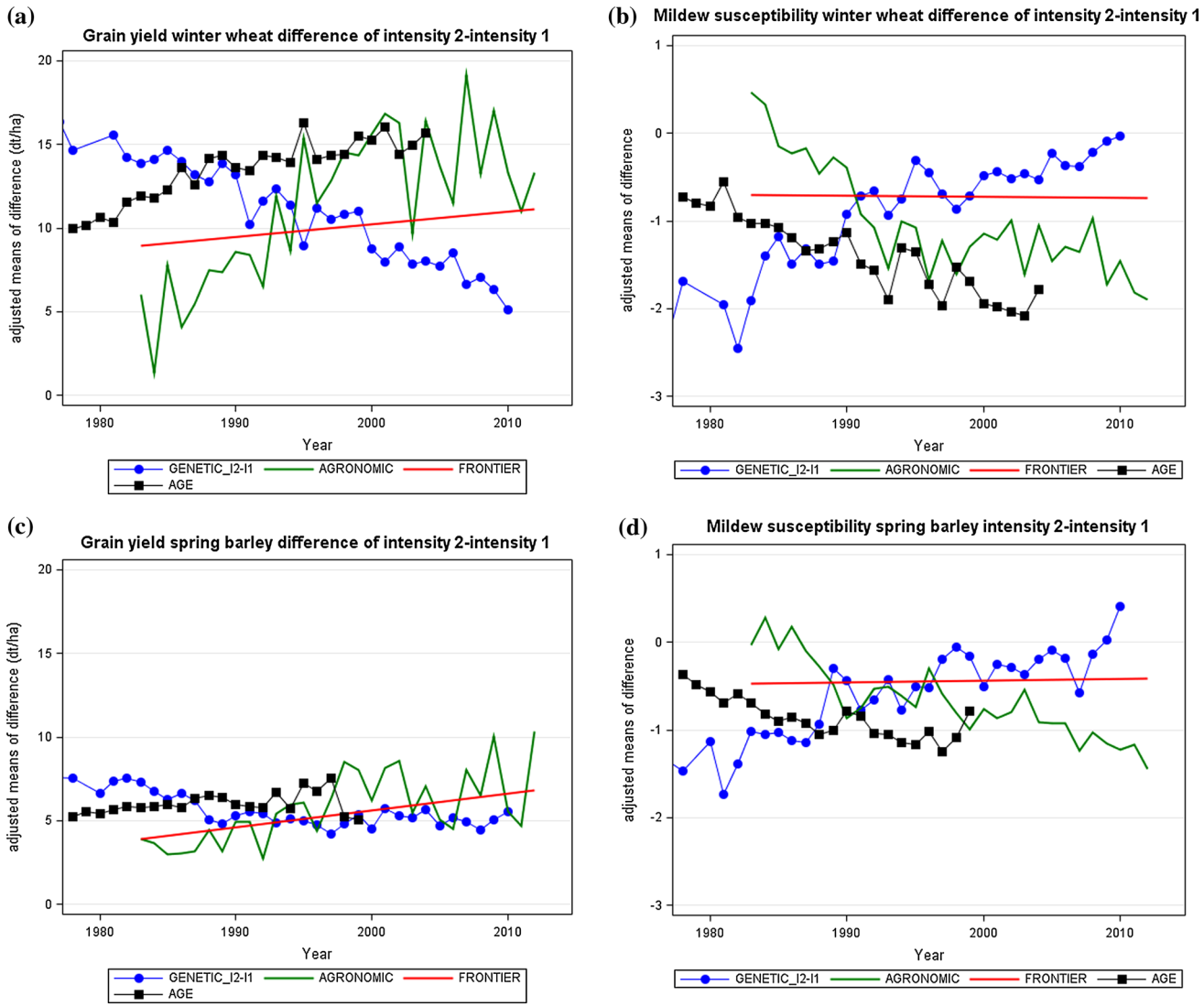


**Fig. 3** Difference between adjusted means of intensity 2 (I2) and intensity 1 (I1) for grain yield of winter wheat and effect of variety age. **a** Winter wheat, yield. **b** Winter wheat, mildew score. **c** Spring barley, yield. **d** Spring barley, mildew score. GENETIC: variety group means [effect $C_p$ in Eq. (19)]; this assesses genetic trend. AGRONOMIC: year means [Eq. (1), using Eq. (19) to model $G_i$]; this assesses non-genetic trend. FRONTIER: frontier line [Eq. (22)]. AGE: effect of variety age (1970 corresponds to age 0) on difference between intensity 2 and 1) (model (1), combined with [Eq. (21)])

similarly confounded as in our models. Recently, Bayesian approaches have been proposed to overcome the non-identifiability problem in these types of model (Fukuda 2008). The resolution comes at the price of strong prior assumptions built into the prior distribution. If such assumptions can be specified and quantified, a Bayesian approach may also be useful for analysing yield trends.

There are some differences of our modelling approach compared to those used in other studies. Silvey (1978, 1981, 1986) assesses the *realized* genetic trend for the UK, because she weights genotype means in a year by their growing areas or the sold amount of commercial seeds. By contrast, in the present work, we consider the *potential* genetic gain ($\beta$) that can be realized if each year the newest available genotypes are grown. In our analyses we have considered years and locations as crossed factors. Some authors either implicitly (Silvey 1978) or explicitly (Mackay et al. 2011) model locations as nested within years, which makes sense when locations change much between years which appears to be the case in many UK trials. The nested design usually provides more accurate genotype means because across years more locations are tested than in a crossed design where the same locations are seen repeatedly across years.

The susceptibility to mildew was assessed on an ordinal rating score (1–9), but analysed as metric data. This

is standard procedure in the analysis of variety trials and, in fact, decisions on whether a new genotype is registered were based on such analyses, so we followed the same type of analysis here. Strictly speaking the data cannot meet usual assumptions of normality and homogeneity of variance, although residual analysis revealed no gross departures, so analysing these data as if they were metric seems acceptable here. More specialized methods for ordinal data do exist, such as the threshold model (Thöni 1985; Hartung and Piepho 2005), but they are difficult to apply here because only a single score was available per plot, while fitting the threshold model ideally requires a sample of 10–20 plants per plot to be scored (Thöni 1985). A further option is to use non-parametric methods based on ranks, but so far such methods are available only for balanced series of trials (Bathke et al. 2010). The ordinal rating scale is based on an underlying percentage scale for the infected leaf area. Thus, if disease susceptibility could be scored as percentage, metric data would be available and such information may be more accurate than ordinal data (Hartung and Piepho 2007).

**Conflict of interest** The authors declare that they have no conflict of interest.

## References


Ahlemeyer J, Friedt W (2011) Progress in winter wheat yield in Germany—What's the share of the genetic gain? In: 61. Tagung der Vereinigung der Pflanzenzüchter und Saatgutkaufleute Österreichs. pp 19–23

Bathke AC, Harrer SW, Wang H, Zhang K, Piepho HP (2010) Series of randomized complete block experiments with nonnormal data. Comput Stat Data Anal 54:840–1857

Fukuda K (2008) Age-period-cohort decomposition of U.S. and Japanese birth rates. Popul Res Policy Rev 27:385–402

Hartung K, Piepho HP (2005) A threshold model for multi-year genebank data based on different rating scales. Crop Sci 45: 045–1051

Hartung K, Piepho HP (2007) Are ordinal rating scales better than percent ratings?—A statistical and "psychological" view. Euphytica 155:5–26

James IR, Segal MR (1982) On a method of mortality analysis incorporating age-year interaction, with application to prostate cancer mortality. Biometrics 38:433–443

Kenward MG, Roger JH (1997) Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 53:983–997

Laidig F, Drobek T, Meyer U (2008) Genotypic and environmental variability of yield for cultivars from 30 different crops in German official variety trials. Plant Breed 127:541–547

Lopes MS, Reynolds MP, Manes Y, Singh RP, Crossa J (2012) Genetic yield gains and changes in associated traits of CYMMIT spring bread wheat in a "historic" set representing 30 years of breeding. Crop Sci 52:1123–1131

Mackay IJ, Horwell A, Garner J, White J, McKee J, Philipott H (2011) Reanalyses of the historical series of UK variety trials to quantify the contributions of genetic and environmental factors to trends and variability in yield over time. Theor Appl Genet 122:225–238

Peltonen-Sainio P, Jauhiainen L, Laurila IP (2009) Cereal yield trends in northern European conditions: changes in yield potential and its realization. Field Crops Res 110:85–90

Perry MW, D'Antuono MF (1989) Yield improvement and associated characteristics of some Australian spring wheat cultivars introduced between 1860 and 1982. Aust J Agric Res 40:457–472

Schuster WH (1997) Welchen Beitrag leistet die Pflanzenzüchtung zur Leistungssteigerung von Kulturpflanzen. Pflanzenbauwissenschaften 1:9–18

Schuster W, Schreiner W, Müller GR (1977) Über die Ertragssteigerung bei einigen Kulturpflanzen in den letzten 24 Jahren in der Bundesrepublik Deutschland. Zeitschrift für Acker- und Pflanzenbau 145:119–141

Silvey V (1978) The contribution of new varieties to increasing cereal yield in England and Wales. J Natl Inst Agric Bot 14:367–384

Silvey V (1981) The contribution of new varieties to cereal yields in England between 1947 and 1978. J Natl Inst Agric Bot 15:399–412

Silvey V (1986) The contribution of new varieties to cereal yields in England between 1947 and 1983. J Natl Inst Agric Bot 17:155–168

Thöni H (1985) Auswertung von Bonituren: Ein empirischer Methodenvergleich. EDV Med Biol 16:108–114

Verbyla AP, Cullis BR, Kenward MG, Welham SJ (1999) The analysis of designed experiments and longitudinal data by using smoothing splines. Appl Stat 48:260–311