


RESEARCH

Open Access



# Non-linear transformation of enzyme-linked immunosorbent assay (ELISA) measurements allows usage of linear models for data analysis

Thomas M. Lange<sup>1</sup>, Maria Rotärmel<sup>1</sup>, Dominik Müller<sup>2</sup>, Gregory S. Mahone<sup>2</sup>, Friedrich Kopisch-Obuch<sup>2</sup>, Harald Keunecke<sup>2</sup> and Armin O. Schmitt<sup>1,3\*</sup> 

## Abstract

**Background:** In research questions such as in resistance breeding against the *Beet necrotic yellow vein virus* it is of interest to compare the virus concentrations of samples from different groups. The enzyme-linked immunosorbent assay (ELISA) counts as the standard tool to measure virus concentrations. Simple methods for data analysis such as analysis of variance (ANOVA), however, are impaired due to non-normality of the resulting optical density (OD) values as well as unequal variances in different groups.

**Methods:** To understand the relationship between the OD values from an ELISA test and the virus concentration per sample, we used a large serial dilution and modelled its non-linear form using a five parameter logistic regression model. Furthermore, we examined if the quality of the model can be increased if one or several of the model parameters are defined beforehand. Subsequently, we used the inverse of the best model to estimate the virus concentration for every measured OD value.

**Results:** We show that the transformed data are essentially normally distributed but provide unequal variances per group. Thus, we propose a generalised least squares model which allows for unequal variances of the groups to analyse the transformed data.

**Conclusions:** ANOVA requires normally distributed data as well as equal variances. Both requirements are not met with raw OD values from an ELISA test. A transformation with an inverse logistic function, however, gives the possibility to use linear models for data analysis of virus concentrations. We conclude that this method can be applied in every trial where virus concentrations of samples from different groups are to be compared via OD values from an ELISA test. To encourage researchers to use this method in their studies, we provide an R script for data transformation as well as the data from our trial.

**Keywords:** Data analysis, Virus concentration, Serial dilution, Logistic regression, Generalised least squares model, Beet necrotic yellow vein virus, BNYVV

## Introduction

Since its invention in 1972, the enzyme-linked immunosorbent assay (ELISA) [1] has been used until today as a reliable tool to detect and quantify virus concentrations in humans [2–8], animals [9–11], and plants [12–18]. To quantify virus concentrations in a sample, double antibody sandwich ELISA (DAS-ELISA) is a common tool

\*Correspondence: armin.schmitt@uni-goettingen.de

<sup>1</sup> Breeding Informatics Group, University of Göttingen, Margarethe von Wrangell-Weg 7, 37075 Göttingen, Germany  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

where an enzyme-linked antibody binds specifically to the coat protein of the virus, which again is bound to the surface of a microtiter plate by another specific antibody. Subsequently, a colourless substrate is given to each sample which is decomposed over time, leading to a lower transmittance of light. This reaction is catalysed by the enzyme [13, 19]. To measure the transmittance level of the sample at a certain time point, light with a certain wavelength can be sent through the sample. It can then be measured via a sensor how much of the light was absorbed by the sample which results in an optical density (OD) value for each sample [20]. Due to the measurement of transmittance levels, the resulting ODs can only take on values within a certain interval. In case of the ELISA machine in this trial, this interval was from zero to four.

The numeric form of ODs makes it possible to compare the ODs of different groups of samples to find significant differences in protein concentrations of these groups. Such studies have been performed in the field of medicine [21–26], veterinary medicine [27], neuroscience [28, 29], pharmacology [30], and agriculture [31–34]. To find significant differences of ODs of different groups, analysis of variance (ANOVA) is the established standard. ANOVA can only be performed when normal distribution of response variables in each group can be assumed [35]. This requirement is rarely tested in data analysis of ODs, but if tested, it was rejected for at least one group [27, 29, 34].

Here, we present data from a trial where doubled haploid lines of sugar beet (*Beta vulgaris* L.) were grown in soil infested with *Beet necrotic yellow vein virus* (BNYVV) in different environments and harvested at different time points. Due to relatively large sample sizes per group and usage of genetically identical individuals, we show that the data of ODs per group are not normally distributed. This coincides with the observations from a similar trial [37] and explains the result of distribution tests in [27, 29, 34].

While performing the ELISA test, we were using a serial dilution of a sample with a high virus concentration on each of the ELISA plates. We visualise the relationship between virus concentration and OD and test different statistical models for mathematical description of the relationship. We show that a non-linear regression model with three parameters (inflection point, slope at inflection point, and asymmetry) as well as predefined values for bottom and top asymptotes could model the relationship best.

Moreover, we used the inverse of this model for data transformation of all measured OD values of the corresponding plate. We show that the transformation not only transferred the measurements on a linear scale but

also led to a normal distribution of the transformed values per group. The transformed data, however, also show unequal variances of the groups. Due to these results, we discourage researchers to analyse data from an ELISA test using ANOVA. However, we show that the transformed data can be analysed using a generalised least squares (gls) model which allows for different variances of the groups.

Since the serial dilutions in the presented study were developed without knowing the absolute protein concentration of the parent solution, we conclude that this method can be applied in any experiment where protein concentrations of samples from different groups are to be compared via OD values from an ELISA test. Moreover, we provide an R script as well as the data from this trial as supplementary material to give researchers the possibility to transform raw data from an ELISA test according to our findings.

## Materials and methods

### Design of experiment

Seven doubled haploid lines of sugar beet were used in this trial, five of which were resistant towards rhizomania (KWS A, KWS B, KWS C, KWS D, KWS E) and two susceptible (KWS F, KWS G) according to [36]. One week after seeds were germinated in sterile soil, seedlings were transplanted into a soil sand mixture (40:60) with soil from Pithivier (France) to ensure the presence of BNYVV pathotype P. Plants were grown in two environments: Environment 1 contained 560 plants which were grown in a greenhouse at around 25° at day and 16° at night. Environment 2 contained 560 plants which were grown in a climate chamber at around 25° at day and 12° at night with a period of 18 hours of light per day.

One half of the plants from each environment were harvested after seven weeks (harvest time point 1) and the other half after ten weeks (harvest time point 2), respectively. During harvesting, lateral roots were separated from the root body and plant sap was extracted from the plants by pressing the harvested lateral roots. 50 µg of plant sap was given into 500 µg of buffer solution containing 1.59g Na<sub>2</sub>CO<sub>3</sub>, 2.93g NaHCO<sub>3</sub>, and 0.2g NaN<sub>3</sub> solved in 1l distilled water. For every combination of genotype, harvest time point and environment, 40 plants were grown, thus, in total, 1120 plants were used in this trial.

The plants were grown in boxes, each containing 35 plants. In each box, 7 genotypes were planted with 5 repetitions each. The plants of one genotype were planted in rows and the rows were randomised. Each box was assigned to a harvest time point in a way that boxes were harvested in an alternating order. The samples were put onto the ELISA plates in the same order as they were

standing in the greenhouse or climate chamber, thus, on each ELISA plate were samples of each harvest time point, but plants grown in the greenhouse and in the climate chamber were analysed on different ELISA plates. The determination of OD values was performed using the protocol in [37]. This protocol stipulates 90 min between adding the substrate to the samples and measuring the OD values using the Infinite F50® (Tecan Group AG, Männedorf, Switzerland) at a wavelength of 405 nm. This machine produces OD values between zero (theoretically minimal absorbance) and four (maximum absorbance). Moreover, this protocol includes four buffer controls on each 96 well plate to estimate the background noise of the machine.

**The serial dilution**

To perform DAS-ELISA with all samples collected in the trial, eleven 96 well plates were used. On each 96 well plate, a serial dilution of twelve samples was analysed. To produce the serial dilution, a sample with a high virus concentration was produced (the “parent solution”) and its virus concentration was halved to produce the next sample of the serial dilution which has then been halved in its virus concentration to produce the next sample of the serial dilution, and so on. Though the virus concentration of the parent solution ( $C_0$ ) is not known, the virus concentration of each sample  $i$  of the serial dilution can be put into a relationship with the virus concentration of the parent solution:

$$C_i = \frac{C_0}{2^i} \tag{1}$$

where  $C_i$  is the virus concentration of the  $i$ -th sample in the serial dilution.

**Statistical modelling of the serial dilution**

The ODs of the samples in the serial dilution were compared to their relative virus concentration where the virus concentration of the last sample of the serial dilution was arbitrarily defined as  $C_{11} = 1$ . Thus,  $C_{10} = 2$ ,  $C_9 = 4$ , and so on until  $C_0 = 2048$ . If instead the dual logarithm of the relative virus concentrations is used, the difference between two successive samples in the serial dilution is 1. Thus, the ODs of each sample in the serial dilution were compared to the dual logarithm of the relative virus concentration ( $\text{ld}(C_i)$ ).

To describe the relationship between ODs and the logarithmised relative virus concentration, we have used a five parameter logistic regression (5PL) model. The general structure of the 5PL model is given in Commo and Bot [38]. Commo and Bot [38], however, used the decadic logarithm in their calculations. Here, we use the dual logarithm since each sample in the

serial dilution was created by halving the virus concentration of the previous sample in the serial dilution. Thus, the 5PL model in [38] was adapted as follows:

$$\widehat{OD}_i = B + \frac{T - B}{[1 + 2^{S \cdot (\text{ld}(I) - \text{ld}(C_i))}]^A} \tag{2}$$

where  $\widehat{OD}_i$  is the estimated OD value for the relative virus concentration  $C_i$ . This model is based on five parameters  $\Theta = \{B, T, I, S, A\}$  where  $B$  is the background noise of the test or the OD that we would expect for a sample with a virus concentration of zero.  $T$  is the maximum OD that can be measured.  $A$  describes the asymmetry of the curve.  $I$  is the relative virus concentration ( $C_i$ ) at the inflection point if  $A = 1$ .  $S$  is the slope at the inflection point.

Afterwards, we were interested if the complexity of the model can be optimised. Therefore, we have created one model where the asymmetry parameter was set to be one ( $A = 1$ ), one model where the slope was set to be one ( $S = 1$ ), and one model where the bottom asymptote was set to be zero ( $B = 0$ ). Furthermore, we have created one model where the top asymptote was set to be the technical limit of the machine ( $T = tl$ ) and one model where the bottom asymptote of the machine was set to be the median of buffer controls for each 96 well plate ( $B = \tilde{bc}$ ).

Moreover, we were interested if multiple predefined parameters can further increase the quality of the model. Therefore, we have created a model where the top asymptote was set to be the technical limit of the machine and the bottom asymptote was set to be the median of the buffer controls for each 96 well plate ( $B = \tilde{bc}$  and  $T = tl$ ). This will change Eq. 2 to

$$\widehat{OD}_i = \tilde{bc} + \frac{tl - \tilde{bc}}{[1 + 2^{S \cdot (\text{ld}(I) - \text{ld}(C_i))}]^A} \tag{3}$$

Equation 3 can be inverted such that the dual logarithm of the relative virus concentration ( $\text{ld}(\widehat{C}_i)$ ) can be calculated from the  $i$ -th OD value ( $OD_i$ ). We present the equation of the inverse logistic regression in its logarithmised form for the sake of clarity.

$$\begin{aligned} \text{ld}(\widehat{C}_i) &= \text{ld}(I) - \frac{\text{ld}\left(\left(\frac{tl - \tilde{bc}}{OD_i - \tilde{bc}}\right)^{\frac{1}{A}} - 1\right)}{S} \\ &= \text{ld}\left(\frac{I}{\left[\left(\frac{tl - \tilde{bc}}{OD_i - \tilde{bc}}\right)^{\frac{1}{A}} - 1\right]^{\frac{1}{S}}}\right) \end{aligned} \tag{4}$$

*Parameter optimisation* For each plate each model was fitted to the data points of the serial dilution by optimising the parameters in the model such that the sum of squared errors (SSE) is minimised [39]:

$$SSE = \sum_{i=1}^n w_i (OD_i - \widehat{OD}_i)^2 \tag{5}$$

with  $w_i$  being a general weights parameter defined as

$$w_i = \frac{1}{\text{ld}(C_i)} \tag{6}$$

to ensure that the curve is not over fitting data points near the upper asymptote while poorly fitting data points near the bottom asymptote [38].

The models were fitted for each 96 well plate separately using the R function `nlsLM` from the `minpack.lm` package [40, 41]. The Levenberg-Marquardt algorithm was used to to minimise SSE [42, 43]. We have used `nlsLM` instead of `nls` or `nls2` [44] which are based on the Gauss-Newton algorithm because of the robustness of the Levenberg-Marquardt algorithm and the function `nlsLM` even for poorly chosen start values [45]. Moreover, the function `nlsLM` also led to lower SSE values for the models than `nls` or `nls2` (data not shown). Since only a positive correlation between virus concentration and OD is assumed, the following assumptions can be made about the model parameters and can be used as bounds of the parameters:

$$\begin{aligned} 0 &\leq B < T \leq tl \\ 0 &< S \\ 0 &< A \end{aligned} \tag{7}$$

**Quality evaluation** For every serial dilution, twelve data points were supposed to be modelled with three to five parameters. Thus, there are a relatively large number of parameters in the models compared to the number of data points per serial dilution. Subsequently, the corrected Akaike information criterion (AICc) was used to evaluate the quality of the models since it is assumed to perform better than other quality parameters such as the Akaike Information Criterion in small samples where the quotient of observations ( $n$ ) and parameters in the model ( $m_j$ ) is smaller than 40 [46, 47]. The AICc of model  $j$  is defined as [48]

$$AICc_j = AIC_j + \frac{2(m_j + 1)(m_j + 2)}{n - m_j - 2}. \tag{8}$$

$AIC_j$  is defined as  $AIC_j = -2\log(L_j) + 2V_j$  where  $L_j$  is the maximum likelihood for the candidate model  $j$  and  $V_j$  is the number of parameters in model  $j$  [49, 50]. The calculation of AICc has been done with the R package `MuMIn` [47].

**Model selection and data transformation** The OD values were transformed using Eq. 4 for each 96 well plate. When transforming data using non-linear regression

with top and bottom asymptote, only data in the interval of monotonic growth of the regression model can be transformed. If a sample did not contain any virus, its OD should result from the background noise of the machine. In this way, it could be lower or equal to the bottom asymptote set in the model. Due to this, no distinct virus concentration can be estimated for an OD value that is equal or smaller than the bottom asymptote. The same holds for ODs that equal the upper asymptote of the model. Thus, only ODs between the bottom and upper asymptote of model  $j$  could be transformed, other data had to be removed.

### Analysing data distributions

For each combination of genotype (7 levels), environment (2 levels), and harvest time point (2 levels), the data distribution was analysed by testing for skewness using the D’Agostino  $K^2$  test [51, 52] and testing for non-normality using Shapiro-Wilk test [53]. We performed the D’Agostino  $K^2$  test using the R function `agostino.test` from the R package `moments` with two-sided alternative hypothesis that the data are skewed [54]. If the  $p$  value from the D’Agostino test was smaller than 0.05, it was assumed that the data distribution was skewed. We performed the Shapiro-Wilk test with the base R function `shapiro.test` with default settings. We assumed non-normality of the data if  $p < 0.05$  [55]. Moreover, we tested if the variances of the 28 groups differed significantly. Therefore, we used Levene’s test of equality of variances [56]. To perform Levene’s test, the R function `levene.test` from the package `lawstat` was used [57]. With default settings, this function uses the modified Brown-Forsythe Levene-type procedure [58]. We assumed unequal variances if  $p < 0.05$ .

### Data analysis

To estimate the effect of the environment and simultaneously the effect of the harvest time point on ODs and transformed data, a `gls` model was used that allowed for different variances per group, following suggestions in [59, 60]. Therefore, a `gls` model was created using the R function `gls` where the variance structure was integrated using the R function `varIdent`. Both functions are from the package `nlme` [61]. Afterwards, differences between the models were examined using the base R function `anova` with the `gls` model. The `gls` models included all possible combinations of interaction terms of the three variables. The model formulation was set to be

$$\begin{aligned} Y_{ijkl} = &\mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} \\ &+ (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{ijkl} \end{aligned} \tag{9}$$

where  $\mu$  is the overall mean response,  $\alpha_i$  is the effect due to the  $i$ -th level of the genotype,  $\beta_j$  is the effect due to the  $j$ -th level of the environment, and  $\gamma_k$  is the effect due to the  $k$ -th level of the harvest time point. Moreover,  $(\alpha\beta)_{ij}$  is the effect due to any interaction between the  $i$ -th level of the genotype and the  $j$ -th level of the environment,  $(\alpha\gamma)_{ik}$  is the effect due to any interaction between the  $i$ -th level of genotype and the  $k$ -th level of the harvest time point,  $(\beta\gamma)_{jk}$  is the effect due to any interaction between the  $j$ -th level of the environment and the  $k$ -th level of the harvest time point, and  $(\alpha\beta\gamma)_{ijk}$  is the effect due to any interaction between the  $i$ -th level of genotype, the  $j$ -th level of environment, and the  $k$ -th level of the harvest time point. The response variable ( $Y_{ijkl}$ ) represents either the OD value or the transformed data.  $\varepsilon_{ijkl}$  is an error term for the  $l$ -th subject with genotype  $i$ , environment  $j$ , and harvest time point  $k$  and was assumed to be

$$\varepsilon_{ijkl} \sim N(0, \sigma_{ijkl}^2). \quad (10)$$

### Robustness of the method

We were moreover interested in the robustness of this method. Therefore, we have used a cross validation where we have estimated the parameters of the model in ten of the eleven plates. Subsequently, we have calculated the median of the ten estimations for the parameters and have transformed the data on the eleventh plate using the median from the parameter estimations from the other ten plates. Afterwards, the results from both transformations were compared regarding the normal distribution of the resulting transformed data and regarding the number of data points that were outside the monotonic growth region of the model and could, thus, not be transformed.

## Results

### Statistical modelling and model selection

As described above, a serial dilution with 12 samples was used on each of the eleven 96 well plates. Moreover, four buffer controls were used on every 96 well plate. Each serial dilution was modelled using the logistic regression models as described above. The serial dilutions as well as the corresponding regression models can be seen in Fig. 1.

One can see in Fig. 1 that the relationship between OD and  $\ln(C)$  is non-linear. Moreover, one can see in Fig. 1 that the 5PL model as well as the 5PL model with predefined values for bottom and top asymptote were able to model the data points of the serial dilution very well. Both lines in Fig. 1 are displaying a non-linear relationship. Based purely on visual examination, it is difficult

to see a difference between the two models in Fig. 1. It is reasonable to assume that if the models are very similar, the lower complexity of the model with predefined values will lead to a smaller AICc.

This assumption is confirmed by the AICc values for each combination of serial dilution and regression model, shown in Table 1. The mean of the AICc values from the 5PL model where no parameters have been predefined is  $-16.4$ . One can see that the AICc increased dramatically when the parameters for asymmetry or slope have been predefined as 1. This shows that these parameters are increasing the quality of the model if they can be optimised freely. This coincides with results from [39].

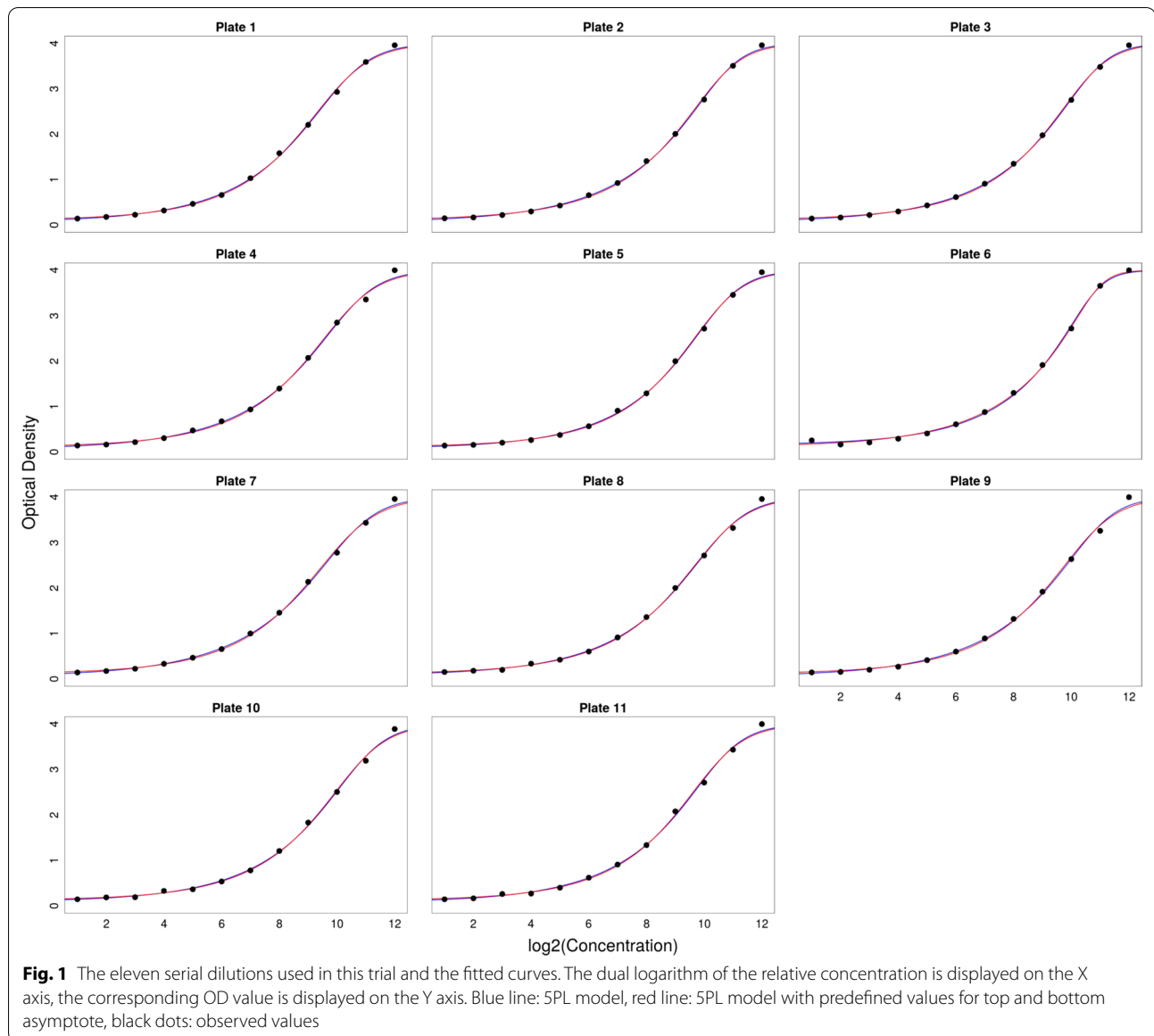
If the upper or bottom asymptote have been predefined on the other hand, the mean AICc was decreased. One can see that in many cases, the AICc was decreased if the bottom asymptote has been predefined as zero or as the median of the buffer controls. The same holds if the upper asymptote has been predefined as the technical limit of the machine. Basing on these results, we have created one more model in which the upper and bottom asymptote have been predefined simultaneously, the upper asymptote as the technical limit and the bottom asymptote as the median of buffer controls.

Based on the results shown in Table 1, we have selected the 5PL model with predefined values for the upper and bottom asymptote as the best statistical model since it has shown the smallest AICc for nine of the eleven serial dilutions and provided the smallest AICc on average. Thus, data were transformed using Eq. 4. Parameters  $\Theta = \{I, S, A\}$  were optimised freely to minimise the SSE (Eq. 5) for the serial dilution on each plate.

From the 1120 plants grown in this trial, 816 ODs could be measured in the ELISA test. From these ODs, four were at the technical limit of the machine ( $OD = 4$ ) and five were smaller than the median of the buffer controls of the plate. Thus, these 9 ODs could not be transformed and were removed from the analysis. Subsequently, 807 measurements could be transformed using the inverted logistic regression model and analysed further. Afterwards, sample sizes of the 28 groups reached from  $n = 3$  to  $n = 40$ . Since we considered a sample size of  $n = 3$  as too small, this group (KWS D, environment 2, harvest time point 2) had to be removed from the data as well. The other groups showed a sample size of  $n \geq 9$ .

### Analysing data distributions

Each combination of genotype, environment and harvest time point was considered a distinct group and data distribution of OD values and transformed data for each group were analysed using the D'Agostino  $K^2$  test [51, 52] and the Shapiro-Wilk test [53]. The resulting  $p$  values are given as an Additional file 3. The null hypothesis



of symmetrical data distribution was rejected if the  $p$  value from the D'Agostino  $K^2$  test was smaller than 0.05 and the null hypothesis of normally distributed data was rejected if the  $p$  value from the Shapiro-Wilk test was smaller than 0.05, respectively. Moreover, Levene's test was used to check for equal variances. If the  $p$  value from Levene's test was smaller than 0.05, unequal variances were assumed [56].

The D'Agostino test with the OD values as response variable led to  $p$  values smaller than 0.05 for 21 of the 27 groups under investigation. Four of the six groups that showed a  $p$  value greater than 0.05 belonged to the susceptible genotype KWS G. The Shapiro-Wilk test with the OD values as response variable led to  $p$  values

smaller than 0.05 for 24 of the 27 groups under investigation. From the six groups that showed a  $p$  value greater than 0.05 with the D'Agostino test, only two groups also showed a  $p$  value greater than 0.05 with the Shapiro-Wilk test and only one of them belongs to the genotype KWS G. It can be assumed that the data distribution of OD values is symmetrical for susceptible genotypes, thus, for samples with high virus concentrations but nevertheless being non-normally distributed.

For the  $p$  values from the D'Agostino test with the transformed data as response variable, only four showed a  $p$  value smaller than 0.05. Similar results can be noticed for the  $p$  values from the Shapiro-Wilk test with the transformed data as response variable. In

**Table 1** AICc of the 5PL model with different predefined values for the eleven serial dilutions as well as the mean AICc for all eleven serial dilutions

Plate	Predefined parameters						
	none	$A = 1$	$S = 1$	$B = 0$	$B = \tilde{b}c$	$T = tl$	$B = \tilde{b}c$ $T = tl$
	AICc of the model						
1	-22.714	-8.724	-6.421	-24.688	-26.267	-31.514	-32.552
2	-25.508	-4.557	-2.636	-26.232	-26.789	-34.308	-33.074
3	-30.735	-4.473	-2.052	-26.85	-29.595	-39.535	-35.881
4	-11.73	-4.226	-3.198	-17.608	-17.548	-20.53	-23.833
5	-19.368	-6.963	-3.423	-18.951	-25.138	-28.168	-31.423
6	-5.105	0.005	4.449	-1.516	-12.002	-13.905	-18.288
7	-17.987	-6.82	-6.319	-23.749	-20.692	-26.787	-26.978
8	-13.647	-6.684	-5.212	-16.764	-20.742	-22.447	-27.028
9	-9.264	-3.486	-2.567	-16.195	-14.567	-18.064	-20.852
10	-12.951	-6.583	-4.548	-14.438	-19.944	-21.751	-26.23
11	-10.923	-4.824	-2.335	-15.069	-17.42	-19.723	-23.705
mean	-16.4	-5.2	-3.1	-18.4	-21	-25.2	-27.3

this case, only two groups resulted in a  $p$  value smaller than 0.05. The transformation has reduced the number of significant results for distribution tests indicating non-normal data distribution from 21 to four groups regarding the D’Agostino  $K^2$  test and from 24 to two groups, regarding the Shapiro-Wilk test.

Levene’s test led to a  $p$  value of  $4 \cdot 10^{-31}$  with the OD values as response variable and to a  $p$  value of  $3 \cdot 10^{-13}$  with the transformed values as response variable. One can see that the  $p$  value was increased due to the transformation but with a significance threshold of  $\alpha = 0.05$ , the assumption of equal variances must be rejected for both response variables.

**Data analysis**

Regarding the results from the distribution tests with the transformed data, it can be assumed that the transformed data are normally distributed but show unequal variances. Thus, a gls model was used which incorporated unequal variances. Afterwards, an ANOVA output can be produced using the base R function `anova`. The results from analysis suggest that the genotype had a significant effect on the result as well as the environment (Table 2). The harvest time point had no significant effect. Nevertheless, the interaction between genotype and harvest time point was significant as well as the interaction between genotype and environment. The interaction between harvest time point and environment as well

**Table 2** Results from the gls model with transformed data as response variable

	numDF	F-value	p-value
(Intercept)	1	9481.088	<.0001
Genotype	6	255.132	<.0001
HarvestTimePoint	1	0.011	0.9163
Environment	1	4.181	0.0412
Genotype:HarvestTimePoint	6	4.195	0.0004
Genotype:Environment	6	5.714	<.0001
HarvestTimePoint:Environment	1	0.022	0.8824
Genotype:HarvestTimePoint:Environment	6	0.885	0.5051

as the interaction between all three variables were not significant.

**Robustness of the method**

Moreover, a cross validation was used to analyse the robustness of the method. Therefore, the transformation was done with ten of the eleven serial dilutions and the median of each parameter was calculated. Subsequently, the data on the eleventh plate were transformed with these medians and with the serial dilution on the eleventh plate itself. Afterwards, the results were compared regarding applicability and data distribution per group.

As described above, the data transformation with the 5PL model and predefined values for the upper and bottom asymptote resulted in four data points that were at

the technical limit of the machine and could, thus, not be transformed as well as five data points that were smaller than the median of the buffer controls and could not be transformed for that reason. The transformation resulted in data that were mostly normally distributed with five groups where the D'Agostino  $K^2$  test resulted in a  $p$  value smaller than 0.05 and with two groups where the Shapiro-Wilk test resulted in a  $p$  value smaller than 0.05. The transformation via this cross-validation led to similar results. The D'Agostino  $K^2$  test resulted again five times in a  $p$  value smaller than 0.05 and the Shapiro-Wilk test resulted two times in a  $p$  value smaller than 0.05. Nevertheless, the cross-validation resulted in 28 data points that were smaller than the median of the buffer controls and could, thus, not be transformed.

## Discussion

In research questions where the effect of multiple variables on the response variable is of interest, ANOVA is the established standard for data analysis even though it requires assumptions to be made about the distribution of the response variable. We show that in the present study, these requirements were not met for the ODs measured in an ELISA test. Similar findings were made in experiments where protein concentrations were measured with ELISA and the normality assumption was tested. None of these tests suggested a normal distribution [27, 29, 34].

Furthermore, we show that the relationship between ODs and virus concentration in this trial was non-linear. Assuming that the relationship between protein concentration and OD always follows a non-linear relationship, it could be concluded that although the virus concentration might follow a normal distribution, the data distribution of ODs might not. If normal distribution of the response variable is not given, simple methods for data analysis such as ANOVA cannot be used to reliably analyse the data [62]. Thus, ANOVA should not be used to analyse the raw ODs from an ELISA test. Due to our experience, we presume that a simple transformation of the data (i.e. decadic logarithm of the OD values) is not sufficient to transform the data into normally distributed data. Following this assumption, some researchers use non-parametric hypothesis tests (such as  $U$  test or Kruskal-Wallis test) to analyse ODs from an ELISA test [22, 27, 29, 63]. Nevertheless, non-parametric hypothesis tests have less test power than their parametric equivalents which might explain why the majority of researchers use parametric hypothesis tests such as ANOVA or  $t$  test to analyse ODs from an ELISA test [24, 26, 28, 34, 64–66].

Here, we present an alternative to these methods by transforming the ODs with an inverse logistic function

where parameters were fitted based upon serial dilutions. The OD values were on a non-linear scale, i.e. an OD value of four cannot be assumed to be twice an OD value of two. The transformed data, however, are on a linear scale, thus, a transformed value of four can be assumed to be twice a value of two. We show that the transformed data are normally distributed per group but still do not provide equal variances. This means that linear models can be used if an appropriate variance structure is integrated into the model. Thus, if only one variable is analysed, Welch's ANOVA can be applied, and if multiple variables are analysed, a gls model can be used which incorporates a variance structure with a different variance per group. In this way, linear models can be used to reliably analyse data from an ELISA test.

We assume that the presented transformation can be applied in any experiment where ELISA is used to differentiate between protein concentrations since we have performed the data transformation with a serial dilution for which we have not known the distinct virus concentrations in the sample but only referred each sample's virus concentration to the parent solution ( $C_0$ ) in the serial dilution. Thus, this method can be applied in any ELISA trial even if a serial dilution with known protein concentrations cannot be produced.

Moreover, we analysed if the complexity of the 5PL model can be optimised by predefining some of the parameters. Since the number of samples that can be analysed on an ELISA plate is limited, the number of samples in a serial dilution is limited, too. Thus, one has to assume that relatively few data points can only be available for modelling the serial dilution and, thus, a reduction of the number of parameters in the logistic regression model will be beneficial. We found that predefining the asymmetry parameter or the slope parameter decreased the quality of the model dramatically. Similar findings regarding the asymmetry parameter were published in [39]. We found that the upper and bottom asymptote of the model can be predefined as the median of the buffer controls for the bottom asymptote and as the technical limit of the machine for the upper asymptote. Nevertheless, it should be pointed out that this method can only be applied if the protocol of the ELISA test indicates that the upper asymptote of the serial dilution can approximate the technical limit of the machine. If there is too little time between adding the substrate and measuring the ODs, the upper asymptote will form its plateau below the technical limit of the machine. In this case, the technical limit of the machine should not be set as upper asymptote of the logistic model.

Finally, we have used a cross-validation to analyse the robustness of this method. We found that data that were transformed in this way can also be assumed to be



normally distributed with five of the 27 groups under investigation leading to a  $p$  value smaller than 0.05 with the D'Agostino  $K^2$  test and four of the 27 groups leading to a  $p$  value smaller than 0.05 with the Shapiro-Wilk test. Nevertheless, 28 data points could not be transformed in the cross validation, showing the necessity of measuring the background noise on each ELISA plate. In this case it can be argued that once a model for data transformation is derived, it can be used for transformation of data from an ELISA test without running a serial dilution on each ELISA plate. In this way, resources could be saved. Nevertheless, it must be kept in mind that this assumption only holds if serial dilutions are very similar to each other as it was the case for the eleven serial dilutions in this trial. Therefore, high amount of precision is necessary which has to be kept over the course of multiple trials. It can be argued that such a precision might be reached through automatization of work processes in the performance of the ELISA test but further research would be necessary to answer this question.

## Conclusion

In research questions where the effect of multiple variables on the response variable is of interest, ANOVA is the established standard for data analysis. One requirement of ANOVA is the normal distribution of the response variable per group. This makes non normality a common problem in many situations. We show that in this trial, the ODs from the ELISA test were not normally distributed which might be due to the non-linear relationship between virus concentration and OD. To model a non-linear relationship with two plateaus, logistic regression models are an appropriate choice.

We show that a logistic regression model with predefined values for the bottom and top asymptote as well as three free parameters which can be estimated using a serial dilution can model the relationship between virus concentration and OD values accurately. Using the inverse of the logistic regression model, data can be transformed to estimate the virus concentration for every OD value that has been measured in the ELISA test. Furthermore, we show that though the measured OD values are not normally distributed, the estimated virus concentrations are normally distributed and can be analysed using linear models.

Since the serial dilutions were prepared without knowing the absolute protein concentrations in these samples, we anticipate that this method can be applied in every experiment where protein concentrations of samples from different groups are to be compared via ODs from an ELISA test.

## Abbreviations

5PL: Five parameter logistic regression; AICc: Corrected Akaike information criterion; ANOVA: Analysis of variance; BNYVV: *Beet necrotic yellow vein virus*; ELISA: Enzyme-linked immunosorbent assay; gls: Generalised least squares; OD: Optical density; SSE: Sum of squared errors.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12985-022-01804-3>.

**Additional file 1.** R script for transformation of optical density values from an ELISA test. The script expects a column sample that contains the string bc for each buffer control, the string sample for each sample in the trial and the string serialdilution\_ and the corresponding number of the sample of the serial dilution for each sample of the serial dilution. The R script is also available at <https://github.com/tmlange/TransformELISA>.

**Additional file 2.** Data from the presented trial with six columns. One column describes the sample, one column describes the ELISA plate, and one column contains the corresponding optical density value from the ELISA test. Three more columns describe the factors to be analysed in the trial: The genotype, the environment, and the harvest time point.

**Additional file 3.** Description of the 28 groups via sample size ( $n$ ),  $p$  value from the D'Agostino  $K^2$  test ( $pA$ ) for each response variable, and  $p$  value from the Shapiro-Wilk test ( $pS$ ) for each response variable. For all analyses, OD values ( $OD$ ) and transformed data ( $C$ ) were used as response variables. For the group with genotype KWS D grown in environment 2 and harvested at time point 2, only 3 data points were available. Thus, for this group neither Shapiro-Wilk test nor D'Agostino  $K^2$  test were performed.

## Acknowledgements

We would like to thank the Phytopathology group of KWS SAAT SE & Co. KGaA, most of all A. Ebeling, J. Gersch, and M. Kohlmorgen for performing the greenhouse trial and laboratory work. Moreover, we would like to thank F. Heinrich for valuable communication regarding the comparison of statistical models.

## Author contributions

Conceptualization, AOS and TML; methodology, TML, MR, DM, and AOS; software, TML, MR, and DM; validation, GSM, FK-O, and HK; resources, DM, FK-O, and HK; data curation, FK-O and HK; writing—original draft preparation, TML and AOS; writing—review and editing, GSM, FK-O, HK, and AOS; visualization, TML; supervision, AOS; project administration, AOS; funding acquisition, AOS. All authors have read and agreed to the published version of the manuscript.

## Funding

No funding to declare.

## Availability of data and materials

The data set supporting the conclusions of this article is included within the article (and its Additional file 1, 2, 3).

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Breeding Informatics Group, University of Göttingen, Margarethe von Wrangell-Weg 7, 37075 Göttingen, Germany. <sup>2</sup>KWS SAAT SE & Co. KGaA,

37574 Einbeck, Germany. <sup>3</sup>Center of Integrated Breeding Research (CiBreed), University of Göttingen, Carl Sprengel-Weg 1, 37075 Göttingen, Germany.

Received: 31 January 2022 Accepted: 22 April 2022

Published online: 18 May 2022

## References

- Engvall E, Perlmann P. Enzyme-linked immunosorbent assay (ELISA). *J Immunol.* 1972;109:1.
- Vestergaard BF, Grauballe PC. Elisa for herpes simplex virus (HSV) type-specific antibodies in human sera using HSV type 1 and type 2 polyspecific antigens blocked with type-heterologous rabbit antibodies. *Acta Pathologica Microbiologica Scandinavica Series B Microbiol.* 1979;87B:261–3.
- Feinstein S, Akov Y, Lachmi B-E, Lehrer S, Rannon L, Katz D. Determination of human IgG and IgM class antibodies to West Nile virus by enzyme linked immunosorbent assay (ELISA). *J Med Virol.* 1985;17(1):63–72.
- Homsy J, Thomson-Honniebier GA, Cheng-Mayer C, Levy JA. Detection of human immunodeficiency virus (HIV) in serum and body fluids by sequential competition ELISA. *J Virol Methods.* 1988;19(1):43–56.
- Beavis KG, Matushek SM, Precy A, Abeleda F, Bethel C, Hunt C, Gillen S, Moran A, Tesic V. Evaluation of the EUROIMMUN anti-SARS-CoV-2 ELISA assay for detection of IgA and IgG antibodies. *J Clin Virol.* 2020;129:104468.
- Krüttgen A, Cornelissen CG, Dreher M, Hornef M, Imöhl M, Kleines M. Comparison of four new commercial serologic assays for determination of SARS-CoV-2 IgG. *J Clin Virol.* 2020;128:104394.
- Dolscheid-Pommerich R, Bartok E, Renn M, Kümmerer BM, Schulte B, Schmithausen RM, Stoffel-Wagner B, Streeck H, Saschenbrecker S, Steinhagen K, Hartmann G. Correlation between a quantitative anti-SARS-CoV-2 IgG ELISA and neutralization activity. *J Med Virol.* 2021;94(1):388–92.
- Pereckaitė L, Dambrauskienė A, Urbonienė D, Sadauskas S, Petrikonis K, Naudžiūnas A, Vitkauskienė A. SARS-CoV-2 seroprevalence in healthcare workers of Kaunas hospitals during the first wave of the COVID-19 pandemic. *Medicina.* 2021;57(2):148.
- Van Wyngaardt W, Mashau C, Wright I, Fehrens J. Serotype-and serogroup-specific detection of African horsesickness virus using phage displayed chicken scFvs for indirect double antibody sandwich ELISAs. *J Vet Sci.* 2013;14(1):95–8.
- Fan B, Sun J, Zhu L, Zhou J, Zhao Y, Yu Z, Sun B, Guo R, He K, Li B. Development of a novel double antibody sandwich quantitative enzyme-linked immunosorbent assay for detection of porcine epidemic diarrhea virus antigen. *Front Vet Sci.* 2020;7:868.
- Wang W, Li J, Fan B, Zhang X, Guo R, Zhao Y, Zhou J, Zhou J, Sun D, Li B. Development of a novel double antibody sandwich ELISA for quantitative detection of porcine deltacoronavirus antigen. *Viruses.* 2021;13:12.
- Voller A, Bartlett A, Bidwell DE, Clark MF, Adams AN. The detection of viruses by enzyme-linked immunosorbent assay (ELISA). *J Gen Virol.* 1976;33(1):165–7.
- Clark MF, Adams AN. Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *J Gen Virol.* 1977;34(3):475–83.
- Lister RM. Application of the enzyme-linked immunosorbent assay for detecting viruses in soybean seed and plants. *Phytopathology.* 1978;68:139.
- Gonsalves D. Tomato spotted wilt virus in papaya and detection of the virus by ELISA. *Plant Dis.* 1986;70:6.
- Broccanello C, McGrath JM, Panella L, Richardson K, Funk A, Chiodi C, Biscarini F, Barone V, Baglieri A, Squartini A, Concheri G, Stevanato P. A SNP mutation affects rhizomania-virus content of sugar beets grown on resistance-breaking soils. *Euphytica.* 2017;214:1.
- Capistrano-Gossmann GG, Ries D, Holtgräwe D, Minoche A, Kraft T, Freichmann SLM, Rosleff Soerensen T, Dohm JC, González I, Kopisch-Obuch FJ. Crop wild relative populations of *Beta vulgaris* allow direct mapping of agronomically important genes. *Nat Commun.* 2017;8:1–8.
- Özmen CY, Khabbazi SD, Khabbazi AD, Gürel S, Kaya R, Oğuz MÇ, Turan F, Rezaei F, Kibar U, Gürel E, Ergül A. Genome composition analysis of multipartite BNYVV reveals the occurrence of genetic re-assortment in the isolates of Asia minor and thrace. *Sci Rep.* 2020;10:1–11.
- Fecker LF, Koenig R, Obermeier C. Nicotiana benthamiana plants expressing Beet necrotic yellow vein virus (BNYVV) coat protein-specific scFv are partially protected against the establishment of the virus in the early stages of infection and its pathogenic effects in the late stages of infection. *Arch Virol.* 1997;142:9.
- Lottspeich F, Engels JW. *Bioanalytik*. 1st ed. Heidelberg: Spektrum; 2012.
- Carlier Y, Bout D, Dessaint JP, Capron A, Van Knapen F, Ruitenberg EJ, Bergquist R, Huldt G. Evaluation of the enzyme-linked immunosorbent assay (ELISA) and other serological tests for the diagnosis of toxoplasmosis. *Bull World Health Organ.* 1980;58(1):99–105.
- Heeschen C, Goldmann BU, Langenbrink L, Matschuck G, Hamm CW. Evaluation of a rapid whole blood ELISA for quantification of troponin I in patients with acute chest pain. *Clin Chem.* 1999;45(10):1789–96.
- Dutaud D, Aubry L, Henry L, Levieux D, Hendil KB, Kuehn L, Bureau JP, Ouali A. Development and evaluation of a sandwich ELISA for quantification of the 20S proteasome in human plasma. *J Immunol Methods.* 2002;260(1–2):183–93.
- Shafat I, Zcharia E, Nisman B, Nadir Y, Nakhoul F, Vlodavsky I, Ilan N. An ELISA method for the detection and quantification of human heparanase. *Biochem Biophys Res Commun.* 2006;341(4):958–63.
- Breuninger S, Erl J. Quantitative analysis of liposomal heat shock protein 70 (Hsp70) in the blood of tumor patients using a novel LipHsp70 ELISA. *J Clin Cell Immunol.* 2014;5:5.
- Lee BJ, Sukri NM, Ogden H, Vine C, Thake CD, Turner JE, Bilzon JLL. A comparison of two commercially available ELISA methods for the quantification of human plasma heat shock protein 70 during rest and exercise stress. *Cell Stress Chaperones.* 2015;20:6.
- Warnken T, Huber K, Feige K. Comparison of three different methods for the quantification of equine insulin. *BMC Vet Res.* 2016;12:1.
- Jensen M, Hartmann T, Engvall B, Wang R, Ulljon SN, Sennvik K, Näslund J, Muehlhauser F, Nordstedt C, Beyreuther K, Lannfelt L. Quantification of Alzheimer amyloid beta peptides ending at residues 40 and 42 by novel ELISA systems. *Mol Med.* 2000;6:4.
- Órpez-Zafra T, Pavia J, Hurtado-Guerrero I, Pinto-Medel MJ, Bada JLR, Urbaneja P, Suardiaz M, Villar JM, Comabella M, Montalban X, Alvarez-Cermeño JC, Leyva L, Fernández Ó, Oliver-Martos B. Decreased soluble IFN-β receptor (sIFNAR2) in multiple sclerosis patients: a potential serum diagnostic biomarker. *Multiple Sclerosis J.* 2016;23:7.
- Shkedy Z, Straetemans R, Molenberghs G, Desmidt M, Vinken P, Goeminne N, Coussement W, Van Den Poel B, Bijnsens L. Modeling anti-KLH ELISA data using two-stage and mixed effects models in support of immunotoxicological studies. *J Biopharm Stat.* 2005;15(2):205–23.
- Meyer UM, Spotts RA, Dewey FM. Detection and quantification of botrytis cinerea by ELISA in pear stems during cold storage. *Plant Dis.* 2000;84(10):1099–103.
- Elzo MA, Rae DO, Lanhart SE, Wasdin JG, Dixon WP, Jones JL. Factors associated with ELISA scores for paratuberculosis in an Angus-Brahman multibreed herd of beef cattle. *J Anim Sci.* 2006;84(1):41–8.
- Elzo MA, Rae DO, Lanhart SE, Hembry FG, Wasdin Driver JD. Association between cow reproduction and calf growth traits and ELISA scores for paratuberculosis in a multibreed herd of beef cattle. *Tropical Anim Health Prod.* 2008;41(6):851–8.
- Liebe S, Wibberg D, Maiss E, Varrelmann M. Application of a reverse genetic system for beet necrotic yellow vein virus to study Rz1 resistance response in sugar beet. *Front Plant Sci.* 2020;10:1703.
- Kitchen CMR. Nonparametric vs parametric tests of location in biomedical research. *Am J Ophthalmol.* 2009;147(4):571–2.
- Büttner G, Märlander B, Manthey R. Breeding for resistance to rhizomania in sugar-beet (*Beta vulgaris* L.). *Plant Breed.* 1995;114(2):160–4.
- Lange TM, Wutke M, Bertram L, Keunecke H, Kopisch-Obuch F, Schmitt AO. Decision strategies for absorbance readings from an enzyme-linked immunosorbent assay - a case study about testing genotypes of sugar beet (*Beta vulgaris* L.) for resistance against Beet Necrotic Yellow Vein Virus Agric. 2021;11:10.
- Commo F, Bot BM. nplr: N-parameter logistic regression. R package version 0.1-7. 2016. <https://CRAN.R-project.org/package=nplr>.

39. Gottschalk PG, Dunn JR. The five-parameter logistic: a characterization and comparison with the four-parameter logistic. *Anal Biochem.* 2005;343(1):54–65.
40. R Core Team. R: A language and environment for statistical computing, Vienna, Austria. R Version 4.0.3.2020. <https://www.R-project.org/>.
41. Elzhov TV, Mullen KM, Spiess AN, Bolker B, minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds. R package version 1.2-1;2016. <https://CRAN.R-project.org/package=minpack.lm>.
42. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *J Soc Ind Appl Math.* 1963;11(2):431–41.
43. Pujol J. The solution of nonlinear inverse problems and the Levenberg-Marquardt method. *Geophysics.* 2007;72(4):W1–16.
44. Grothendieck G. Non-linear regression with brute force. R package version 0.2; 2013.
45. Nash JC. Nonlinear parameter optimization using R tools. New York: Wiley; 2014.
46. Wong CS, Li WK. A note on the corrected Akaike information criterion for threshold autoregressive models. *J Time Ser Anal.* 1998;19(1):113–24.
47. Barton K. MuMIn: multi-model inference. R package version 1.43.17; 2020. <https://CRAN.R-project.org/package=MuumIn>.
48. Hurvich CM, Tsai C-L. Regression and time series model selection in small samples. *Biometrika.* 1989;76(2):297–307.
49. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. *Selected papers of Hirotugu Akaike.* 1st ed. New York: Springer; 1998. p. 199–213.
50. Wagenmakers E-J, Farrell S. AIC model selection using Akaike weights. *Psychonomic Bull Rev.* 2004;11(1):192–6.
51. D'Agostino RB. Transformation to normality of the null distribution of  $g_1$ . *Biometrika.* 1970;57(3):679–81.
52. D'Agostino RB, Pearson ES. A further development of test departure from normality. *Biometrika.* 1973;60(3):613.
53. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika.* 1965;52(3/4):591–611.
54. Komsta L, Novomestky F. Moments: moments, cumulants, skewness, kurtosis and related tests. R package version 0.14; 2015. <https://CRAN.R-project.org/package=moments>.
55. Royston JP. An extension of Shapiro and Wilk's W test for normality to large samples. *Appl Stat.* 1982;31(2):115–24.
56. Levene H. Robust tests for equality of variances. In: Olkin I, Hotelling H, editors. *Contributions to probability and statistics: essays in honor of Harold Hotelling.* Stanford University Press; 1961. p. 278–92.
57. Gastwirth JL, Gel YR, Hui WLW, Lyubchich V, Miao W, Noguchi K. lawstat: tools for biostatistics, public policy, and law. R package version 3.4; 2020. <https://CRAN.R-project.org/package=lawstat>.
58. Brown MB, Forsythe AB. Robust tests for the equality of variances. *J Am Stat Assoc.* 1974;69(346):364–7.
59. Pinheiro JC, Bates D. *Mixed-effects models in S and S-PLUS.* Berlin: Springer; 2009.
60. Zuur A, Ieno EN, Walker N, Saveliev AA, Smith GM. Dealing with heterogeneity. In: *Mixed effects models and extensions in ecology with R. Statistics for biology and health; 1st edition.* New York: Springer; 2009. pp. 71–100.
61. Pinheiro J, Bates D, DebRoy S, Sarkar D. R Core Team nlme: linear and nonlinear mixed effects models. R package version 3.1-153; 2021. <https://CRAN.R-project.org/package=nlme>.
62. Stroup WW. Rethinking the analysis of non-normal data in plant and soil science. *Agron J.* 2015;107(2):811–27.
63. Ibarra F, Montenegro N, Vera Y, Boulard C, Quiroz H, Flores J, Ochoa P. Comparison of three ELISA tests for seroepidemiology of bovine fasciolosis. *Vet Parasitol.* 1998;77(4):229–36.
64. Giunchedi L, De Biaggi M, Poggi PC. Correlation between tolerance and Beet necrotic yellow vein virus in Sugar-beet genotypes. *Phytopathologia Mediterranea.* 1987;26(1):23–8.
65. Mannerlöf M, Lennerfors B-L, Tenning P. Reduced titer of BNYVV in transgenic sugar beets expressing the BNYVV coat protein. *Euphytica.* 1996;90(3):293–9.
66. Zare B, Niazi A, Sattari R, Aghelpasand H, Zamani K, Sabet MS, Moshiri F, Darabie S, Daneshvar MH, Norouzi P, Kazemi-Tabar SK, Khoshnami M, Malboobi MA. Resistance against rhizomania disease via RNA silencing in sugar beet. *Plant Pathol.* 2014;64(1):35–42.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

