ORIGINAL PAPER

Modelling Survival Time to Symptom Expression and Death Among Sweet Potato Crosses Inoculated with *Fusarium oxysporum* f.sp. *batatas*

Mcharo, M.*[1]; Clark, C.[2] Mutai, N.[3] and La Bonte, D.[4]

ABSTRACT

   Two hundred and six sweet potato genotypes descended from Fusarium wilt resistant Beauregard variety as one of the parents and those without Beauregard as a parent were inoculated with the Fusarium wilt pathogen, *Fusarium oxysporum* f.sp. *batatas.* The objectives of this study were to: (i) compare the efficiencies of different models in describing time to symptom expression and time to death among infected sweet potato genotypes and, (ii) develop predictive models. The Kaplan-Meier survival function and actuarial life tables were used for nonparametric modelling while probability plots using the Generalized gamma, Exponential, Weibull, Lognormal and Log-logistic models were used for parametric modelling. The three-parameter generalized gamma model was also used to determine the performance of the simpler nested two-parameter models. The effect of parental resistance level was significant (P = 0.05) for time to death. Genotypes with Beauregard as a parent took longer to die or show symptoms. From the nonparametric analysis, genotypes descended from Beauregard as a parent took on average 43.14 days to die compared to non-Beauregard genotypes that took 33.31 days. Mean time to symptom expression for all models ranged from 11.67 to 12.28 days. In conclusion, it is possible to model survival time using either the parametric Generalized gamma, Lognormal and Weibull models or the nonparametric model.

Keywords: *Ipomoea batatas*; *Fusarium oxysporum*; Kaplan-Meier; Model; Symptoms.

*Correspondence: **Mwamburi Mcharo**
 E-mail: mwamburim@ttu.ac.ke
Mcharo, M.*[1];
 https://orcid.org/0000-0002-1432-7157
Clark, C.[2]
Mutai, N.[3]
 https://orcid.org/0000-0001-9677-223X
 La Bonte, D.[4]

[1]School of Agriculture, Earth and Environmental Sciences, Taita Taveta University, P. O. Box 635-80300, Voi, KENYA
[2]Dept of Plant Pathology and Crop Physiology, Louisiana State Univ. Agricultural Center, Louisiana Agricultural Experiment Station, Baton Rouge, LA 70803, USA
[3]Berlin School of Business, and Innovation, GmbH, Alte Post Karl-Marx-Straße 97-99, 12043 Berlin, Germany
[4]School of Plant, Environmental and Soil Sciences, Louisiana State Univ. Agricultural Center, Louisiana Agricultural Experiment Station, Baton Rouge, LA 70803, USA

Introduction

   Sweet potato is the eighth most important source of starch globally, and the sixth in Africa (FAO, 2020). The commodity is used for food and for animal feed and supports thousands of households involved in its production, processing and sale. Varieties have been bred for desirable qualities like flesh and skin colour, root shape, taste and resistance to pests and diseases. Diseases that greatly affect sweet potato production include viral, bacterial and fungal infections. One of the major fungal diseases in East Africa is Fusarium wilt, caused by *Fusarium oxysporum* f.sp. *batatas* (Chalwe *et al*., 2017).

Fusarium wilt symptoms include yellowing of older leaves, brown lesions on leaves, wilting, leaf drop and ultimate death of the plant. Internal discolouration of the vascular tissue is also a diagnostic symptom (Moussa *et al.*, 2018). Symptoms are more severe under conditions of low soil moisture. Additionally, the fungus can persist in soil for many years thus perpetuating infections for long. The most effective and economical means of controlling this disease is the use of resistant cultivars like Beauregard and 'Jewel'. In addition, a combination of cultural practices like field sanitation, use of certified disease-free cuttings and crop rotation may be used for disease control.

It is not possible to know whether a field is infested with the Fusarium wilt pathogen unless symptoms are observed on growing plants or the soil tested. Also, plant pathogens have great diversity with different races and pathotypes adapted to different environments (Sseruwu *et al.*, 2020). As a result, cultivars that were previously resistant may progressively become susceptible to infection. Development of field-testing and early disease detection methods using test plants will therefore aid in disease control. The ability to model the time to appearance of disease symptoms and the time to death of the plant is important in instituting such disease control measures (Giroux *et al.*, 2016)). In sweet potatoes, the models are especially important since the crop is widely grown globally and is economically important to the livelihoods of many resource-poor farmers. Prediction of the level of crop damage and mortality at various times due to the disease will also assist in modelling possible levels of crop loss. Predictive and descriptive survival analysis models like simple probability parametric models, accelerated failure time (AFT) models and proportional hazards models have been used to explain mortality (Bogaerts *et al.*, 2018). This paper focuses on the following simple probability parametric models Weibull, Lognormal, Exponential, Log-logistic and the Generalized gamma models.

In an attempt to describe analysis of mortality data different workers have chosen to approach the subject from a 'survival analysis' point of view. Singh and Dlamini (2021) define survival analysis as "a group of statistical procedures for data analysis, for which the outcome variable of interest is time until an event occurs." Events of interest may include any observable change like death, onset of disease, seed germination or flowering in plants. Survival analysis was therefore developed for application to longitudinal data (Bogaerts *et al.*, 2018).

Survival analysis commonly uses terms like (i) *survival time* which indicates the period from onset of an event to its endpoint, (ii) *survival rate* as the proportion of individuals that survive a process for a given time, and (iii) *time-to-event* which represents the period until an event of interest occurs (Rathod *et al.*, 2020). In defining time we must differentiate between 'calendar time' and 'event time' in collecting survival data. 'Calendar time' refers to specific dates in a particular time period when the study commences and terminates (Moore, 2016 ; Bogaerts *et al.*, 2018). 'Calendar time' is important since not every case is entered into a study at the same time. The second concept of time is the 'event time'. 'Event time' is the total time period from when the study commences until the event of interest occurs. Two subjects may, therefore, be observed in two different calendar years in the same study but have identical event times.

Other commonly used terms are *hazard*, *life-table*, and *censoring*. When a study is completed before all individuals under observation have experienced the event of interest, those individuals are censored or right-censored. If individuals experience the event of interest before the study begins and the individuals are

included in the study, those are left-censored. If individuals in a study experience the event of interest while the study is ongoing but the time the event occurs is unknown, those are interval censored (Bogaerts *et al*., 2018; Singh and Dlamini, 2021). However, censoring is a feature in survival data that is difficult to incorporate into common analytical methods like regression (Rathod *et al*., 2020). Survival analysis, therefore, uses maximum likelihood and partial likelihood procedures to facilitate handling of such data with censored and uncensored cases (Moore, 2016; Bogaerts *et al*., 2018). Using survival analysis an investigator is, therefore, able to model survival distribution for a population of organisms, compare distributions from different populations and analyze the effect of any covariates on the survival times.

Survival models have been formulated using nonparametric and parametric survival and hazard analysis techniques (Nesi *et al*., 2015; Minzat *et al*., 2018; Rathod *et al*., 2020; Mills *et al*., 2021). Survival and hazard functions have also previously been described by Benali *et al*., (2015) and Nsobinenyui *et al*., (2022) as

1. MATERIALS AND METHODS
1.1.Plant Material :

The project was conducted under greenhouse conditions at the Horticultural Hill Farm teaching facility, Louisiana State University. Crosses from the parental varieties Beauregard, Wagabolige, Tanzania, Kyukei No. 64, Jonathan-W218, Jonathan-W154 and CN1732-4 were used in this study. Beauregard was selected as a parent because it is resistant to Fusarium wilt (Clark *et al*., 1998). The other parents were selected because they were susceptible to Fusarium wilt and were previously tested for adaptability to East African conditions (Mcharo *et al*., 2001). Crosses from these parents provided an opportunity to investigate how resistance

an effort to analyse survival times using assumed distributions. According to Shrestha *et al*., (2019) other analytical methods do not make any assumptions that the survival function has a particular distribution. These procedures are called semiparametric methods, and they include Cox's proportional hazards method (Cox, 1972; Harris *et al*., 2015).

Survival analysis, has so far, been rarely used to examine time-to-event data generated by plant pathologists (Nesi *et al*., 2015) and in agriculture generally (Singh and Dlamini, 2021). This scenario is more pronounced in food crops as evidenced by the dearth of literature involving survival analysis of events in crop plants. Non-parametric analysis for this study was done using the Kaplan-Meier (KM) estimate (Kaplan and Meier, 1958) while parametric analyses used single failure-time probability distribution models (Meeker and Escobar, 1998). The objectives of this study were to: (i) compare the efficiencies of different models in describing time to symptom expression and time to death among infected sweet potato genotypes and, (ii) develop predictive models.

is inherited. First-generation seeds from two different categories of crosses (Beauregard and non-Beauregard) were germinated and the seedlings left to grow for about two months. Beauregard crosses were: Beauregard x Wagabolige and Beauregard x Tanzania. Non-Beauregard crosses consisted of offspring from the crosses Kyukei No. 64 x Jonathan-W218, Wagabolige x Jonathan-W154 and CN1732-4 x Jonathan-W218. Beauregard crosses, therefore, had Beauregard as one of the parents while non-Beauregard crosses did not have Beauregard as one of the parents. Each seedling was a unique genotype, but all seedlings from a particular cross belonged to one family.

1.2.Culturing the Pathogen:

Isolate WJM-7 of *F. oxysporum* f.sp *batatas* from infected sweet potato plants was cultured in Czapek's broth with constant agitation for 5 days. The inoculum was prepared by filtering the culture through four layers of cheesecloth and adjusting with the aid of a hemacytometer to 10^6 spores per ml.

1.3.Inoculation of the Plants:

Two hundred and ten distinct progeny from the specific crosses of sweet potatoes were inoculated with the fungus by dipping the wounded parts of the plants into the culture as described by Clark *et al*., (1998). Vine cuttings 15 cm long were cut from the seedlings and all the leaves except the top two to three were removed by breaking the petiole away from the stem. Removing the older leaves in this manner provided wounds for entry of the pathogen and reduced the likelihood that leaf senescence would occur during symptom development.

1.4.Assessment of the Plants:

Development of the inoculated cuttings was observed for 47 days when no more symptoms developed or death occurred for at least two weeks. Data on days to symptom appearance and days to plant death were recorded throughout the observation period. A seedling was considered infected when at least one necrotic fungal lesion was observed on a leaf or stem (Figure 1). Severity of the symptoms was not important for this study hence no severity scale was used. Death was recorded when the seedling was fully necrotic, wilted and toppled over (Figure 2).

1.5.Data Collected:

i. Days to symptom appearance (right-censored variable)
ii. Days to death (right-censored variable)
iii. Parentage or parental cross from which the genotype was derived. This is the categorical covariate included in the analysis which was made up of either the Beauregard and non-Beauregard categories of crosses.

Figure 1: Plant showing symptoms setting in 9 days after inoculation.

50

Egypt. J. Phytopathol., Vol. 52, No. 1, pp46 –66

Figure 2: Dead plant from a non-Beauregard cross 9 days after inoculation.

1.6. Data Analysis:

The data obtained from this study were right-censored data. There were plants that either did not show symptoms at the end of the study or were not dead at the end of the study. These plants survived the events of interest and hence were right-censored. Consequently, survival analysis models were used to analyze the data. Survival time was the dependent variable while the category of plants was the independent variable. The parent 'Beauregard' is highly resistant to Fusarium wilt while the other genotypes are susceptible. Therefore, the two categories of the covariate were expected to have different survival distributions from each other because of the parents used to derive the individuals.

Both nonparametric and parametric modelling was done. The methods used for nonparametric estimation of survival and hazard functions were the Kaplan-Meier (KM) method and actuarial life tables (Cox, 1972). The Kaplan-Meier method, also called the product-limit estimator, was developed by Kaplan and Meier (Kaplan and Meier, 1958) as a nonparametric maximum likelihood estimator and it estimates the survival function $\hat{S}(t)$. Parametric modelling used five models, namely the Exponential, Weibull, Lognormal, Log-logistic and Generalized gamma models. However, because the Generalized gamma model has more parameters than the other four, its hazard function is complicated thus making it difficult to interpret the shape of the hazard function. Also development of the model is much more difficult for the same reason. Therefore, only minimal results for

the Generalized gamma are presented for purposes of comparing it with other models.

The survival and hazard functions used in this study were of the following general form:

The survival function, $S(t)$, for time t was defined as :

$$S(t) = 1 - F(t) = P(T > t), \quad t \geq 0 \qquad (1)$$

Where:

$F(t)$ was the probability that survival time will be less than or equal to time t

T was a continuous random variable measured as any time beyond the defined survival time t

The survival function $S(t)$ was also described by the closely related cumulative distribution function $F(t)$ thus,

$$F(t) = P(T \leq t), \qquad t \geq 0 \qquad (2)$$

. The hazard function, $h(t)$, was the instantaneous rate of change of the probability of death at time t, given that the individual survived to time t.

$$h(t) = \lim_{\Delta t \to \infty} \frac{P(t < T \leq t + \Delta t \mid T > t)}{\Delta t}, \qquad t \geq 0$$

Where:

Δt was the change of time or increment in time.

The hazard function was simplified to:

$$\square(t) = \frac{f(t)}{1 - F(t)}, \quad t \geq 0 \qquad (3)$$

Where:

$f(t)$ was the probability density function for T and also the derivative of $F(t)$

2.7.1. Nonparametric Estimation of Survival:

The KM estimator that was used is defined as:

$$\hat{S}(t) = \prod_{j=1}^{i} \left[\frac{n_i - d_i}{n_i} \right], \qquad (4)$$

Where:

n_i is the number of individuals surviving up to the beginning of time t_i and d_i is the number of individuals who die at time t_i. However, the Kaplan-Meier estimator has lower estimating efficiency compared to parametric methods because it will give a larger variance for $\hat{S}(t)$ than the variance obtained by parametric analysis for a particular data set (Zee and Xie, 2018).

Three criteria were used to compare survival functions between the two categories of crosses. The criteria, which all follow a chi-square distribution, are the likelihood ratio test (Özen *et al.*, 2021).), the log-rank test and the Wilcoxon test. The likelihood ratio test assumes a constant hazard function in each group thus suggesting an Exponential distribution for time to event. This assumption is however not practical for biological organisms. The log-rank test is suitable for the proportional hazards model while the Wilcoxon test is superior to the log-rank test where time to event has a Lognormal distribution. The log-rank test is more sensitive in detecting differences between strata at later times while the Wilcoxon is better at early times (Bogaerts *et al.*, 2018).

2.7.2. **Parametric Estimation of Survival**:

Parametric modelling of survival data in this study was done using simple probability parametric models because there were no measureable explanatory variables to make use of regression functions. These models are extensively discussed by Meeker and Escobar (1998). This study did not assume proportional hazards and thus excluded proportional hazards models. Since the categories that made up the covariate were nominal, AFT regression methods were not used either. The study, therefore, used single failure-time distribution parametric models for

analysis. These distributions have the following advantages: (i) They can be described concisely with just a few parameters instead of reporting an entire curve, (ii) They are amenable to extrapolating to lesser or greater times, and (iii) They provide smooth estimates of survival distributions. Chiang *et al*., (2024) further discussed the use of survival analysis methods for interval-censored data with a quantitative covariate. Five parametric models, namely the Generalized gamma, the Exponential, Weibull, Lognormal and Log-logistic were tested to determine the model(s) that best describe the distribution of time to symptoms or death (Meeker and Escobar, 1998). The Generalized gamma model, which has three parameters, was the unrestricted (larger) model while the other four were restricted (smaller) models. The Exponential, Weibull and Lognormal models are derivatives of the Generalized gamma model but the Log-logistic model is not derived from the Generalized gamma model. The model comparisons were done using the following restrictions on the scale parameter and the shape parameter for the Generalized gamma model: Scale = 1 for Weibull; scale = 1, shape = 1 for Exponential; shape = 0 for Lognormal. The performances of different survival models were compared using twice the difference of log-likelihood for any two models being compared. Results for the survival, cumulative and hazard functions were obtained for all the five tested models. The statistical software R version 4.1.0 was used for data analysis. The mathematical foundations for these models are described in Meeker and Escobar (1998) and Bogaerts *et al*., (2018). Therefore, this paper will restrict itself to the general descriptions of the models.

Generalized Gamma model:

$$T \sim GENG(\eta, \beta, \kappa),$$

$$F(t; \eta, \beta, \kappa) = \Gamma_I \left[\left(\frac{t}{\eta} \right)^{\beta} ; \kappa \right], \quad t > 0$$

Where:

$\eta > 0$ is a scale parameter, $\beta > 0$ and $\kappa > 0$ are shape parameters, and $\Gamma_I(\nu, \kappa)$ is the gamma function.

Exponential model. $T \sim EXP(\eta)$:

$$F(t) = 1 - exp\left(\frac{t}{\eta} \right), \quad \eta \geq 0, \ t \geq 0$$

Where:

η is the scale parameter and the constant hazard. This model is a one-parameter distribution and the scale parameter η is constrained to $\eta = 1$. The log of T has an extreme-value distribution and consequently T has an Exponential distribution. The Exponential distribution of T has a constant hazard function.

Weibull Model. $T \sim WEIB(\eta, \beta)$,

$$F(t) = 1 - exp\left[-\left(\frac{t}{\eta} \right)^{\beta} \right], \quad \beta \geq 0, \ \eta \geq 0, \ t \geq 0,$$

Where:

$\beta > 0$ is the shape parameter and $\eta > 0$ is the scale parameter. Consequently, the Exponential survival function is a derivative of the Weibull with the restriction $\beta = 1$. In this model the value of η varies (Escobar, 1998 and Moore, 2016 Meeker). The hazard function decreases with time when $\eta > 1$ but it increases at a decreasing rate when $0.5 < \eta < 1$. The hazard increases at an increasing rate when $0 < \eta < 0.5$ and the hazard function is an increasing straight line with origin 0 when $\eta = 0.5$.

Log-normal model. $T \sim LOGNOR(\mu, \sigma)$,

$$F(t) = \Phi nor \left[\frac{log(t) - \mu}{\sigma} \right], \quad t > 0$$

Where:

exp (μ) is a scale parameter at the median (= $t_{0.5}$) and $\sigma > 0$ is a shape parameter. Φ_{nor} is the cumulative distribution function for a $NOR(0,1)$. When $t = 0$ the hazard is 0, rising to a peak with time and decreasing to 0 as t approaches infinity.

Log-logistic model:

$Y \sim LOGIS(\mu, \sigma),$

therefore:

$T = exp(Y) \sim LOGLOGIS(\mu, \sigma),$

$F(t) = \Phi logis \left[\dfrac{log(t) - \mu}{\sigma} \right], \quad t > 0$

Where:

exp(μ) is a scale parameter and $\sigma > 0$ is a shape parameter. Φ_{logis} is the cumulative distribution function for a $LOGIS(0,1)$. Bogaerts *et al.*, (2018) further stated that this model assumes that the error term has a logistic distribution with mean = 0. It then follows that log T has a logistic distribution and T has a Log-logistic distribution with the hazard function described above. As opposed to the Exponential, Weibull and Lognormal models, the Log-logistic model, is not nested within the Generalized gamma model and so a likelihood ratio test to compare the performance of the Log-logistic against the Generalized gamma could not be computed. The efficacy of Lognormal and Log-logistic models in modelling time to germination has been documented by Romano and Stevanato (2020).

3. RESULTS AND DISCUSSION
3.1. Censored and Un-censored Data:
Table (1) is the summary of censored and un-censored data from the experiment. Censored individuals were plants that did not show any symptoms or die at the end of the study. For symptom expression, 6% of Beauregard progeny were censored while 3% of non-Beauregard progeny were censored at 32 days. These did not show symptoms until the close of the study on the 47th day. Results in Table (2) suggest that symptoms set in faster in the crosses not involving Beauregard as a parent. The earliest symptoms were necrotic spots on leaves, and they were observed two days after inoculation while the earliest death was recorded eight days after inoculation. Symptoms set in the Beauregard crosses on average 2 days after the onset of symptoms in the non-Beauregard crosses. These values were underestimates because the Kaplan-Meier nonparametric estimation method was restricted to the largest time to symptom expression while in reality there were censored individuals. Nesi *et al.*, (2015) also reported having encountered limitations of the Kaplan-Meier test when analysing post-harvest diseases of peaches.

Table 1: Summary of the number of censored and uncensored plants for symptom expression

Cross	Total plants	Plants with symptoms	Number censored	Percent censored
Beauregard	56	53	3	5.36
Non-Beauregard	150	146	4	2.67
Total	206	199	7	3.40

3.2. Nonparametric tests:
Tests of equality of time to symptom expression and plant death were conducted involving all crosses. The probabilities of a greater Chi-square for all the three (Log-Rank, Wilcoxon, and Likelihood Ratio) tests of time to symptom expression were greater than the critical value of α = 0.05 as presented in Table (3). We, therefore, did not reject the null hypothesis that cross

54

Egypt. J. Phytopathol., Vol. 52, No. 1, pp46 –66

type had no effect on time to symptom expression.

Table 2: Mean and median time to symptom expression

Cross	Mean (days)	Standard error for mean	Median (days)	95% Confidence intervals for median	
				Lower	Upper
Beauregard	12.875	1.408	7.500	5.000	15.000
Non-Beauregard	10.787	0.773	6.000	5.000	7.000
Both combined	11.354	0.682	6.000	5.000	7.000

Table 3: Test of equality of time to symptom expression over crosses

Test	Chi-Square	Df	Pr > Chi-Square
Log-Rank	1.168	1	0.280
Wilcoxon	0.807	1	0.369
-2Log(LR)	1.685	1	0.194

This result suggests that the two categories of crosses had the same survival function, and consequently, modelling of distribution of time to symptom expression ignored the cross covariate. Yellareddygari *et al*., (2017) also reported results of the Log-Rank and Wilcoxon tests to model survival time of the potato mop virus.

Actuarial life table results in Table (4) indicate that the hazard is stable for the first 10 days as indicated by the median residual lifetime. At commencement of the study the mean residual lifetime for the whole population of plants is 7.5 days. It then increased to a peak of 16.41 days for the plants without symptoms at 10 days after inoculation. At 25 days after inoculation, the plants still without symptoms had a median residual lifetime of 3.08 days. These results are further discuss in conjunction with those in Figure (1)

Figure (3) is a non-parametric mortality and hazard plot for time to symptom expression, with point wise confidence bands for the two categories of sweet potato crosses. Hazard plots have been used effectively in the past to detect symptomatic and asymptomatic plants by Benali *et al*., (2015).

Figure (1) indicates high mortality rates within the first 7 days. The rate thereafter declined gradually until the 13th day before it flattened up to the 22nd day for Beauregard crosses. Thereafter the proportion with disease symptoms increased up to the 27th day for Beauregard crosses before flattening. For non-Beauregard crosses, the trend flattened from the 13th to the 26th day before suddenly increasing and thereafter flattening on the 27th day. The trend of the hazard plot (Figure 3) corresponds well with that of the mortality plot. In the hazard plot, there is a general increase in the cumulative hazard up to the 9th day for Beauregard crosses before flattening up to the 22nd day. A gentle increase was observed up to the 26th day and thereafter a rapid increase in the cumulative hazard. The cumulative hazard for the non-Beauregard crosses flattened from the 13th up to the 26th day before rapidly increasing. This was also considered to be the day when the hazard function peaked thus indicating a higher propensity for

Table 4: Life table for symptom expression showing hazard rates for 5 day time intervals

Days interval		Plants with symptoms	Number censored	Median Residual Lifetime	PDF[1]	Evaluated at the Midpoint of the Interval		
Lower	Upper					PDF Standard Error	Hazard	Hazard Standard error
0	5	74	0	7.500	0.072	0.007	0.088	0.010
5	10	58	0	7.500	0.056	0.006	0.113	0.014
10	15	16	0	16.410	0.016	0.004	0.048	0.012
15	20	1	0	12.436	0.001	0.001	0.003	0.003
20	25	9	0	7.500	0.009	0.003	0.034	0.011
25	30	39	0	3.077	0.038	0.005	0.274	0.032
30	33	2	7	.	0.005	0.003	0.148	0.102

[1]Probability density function
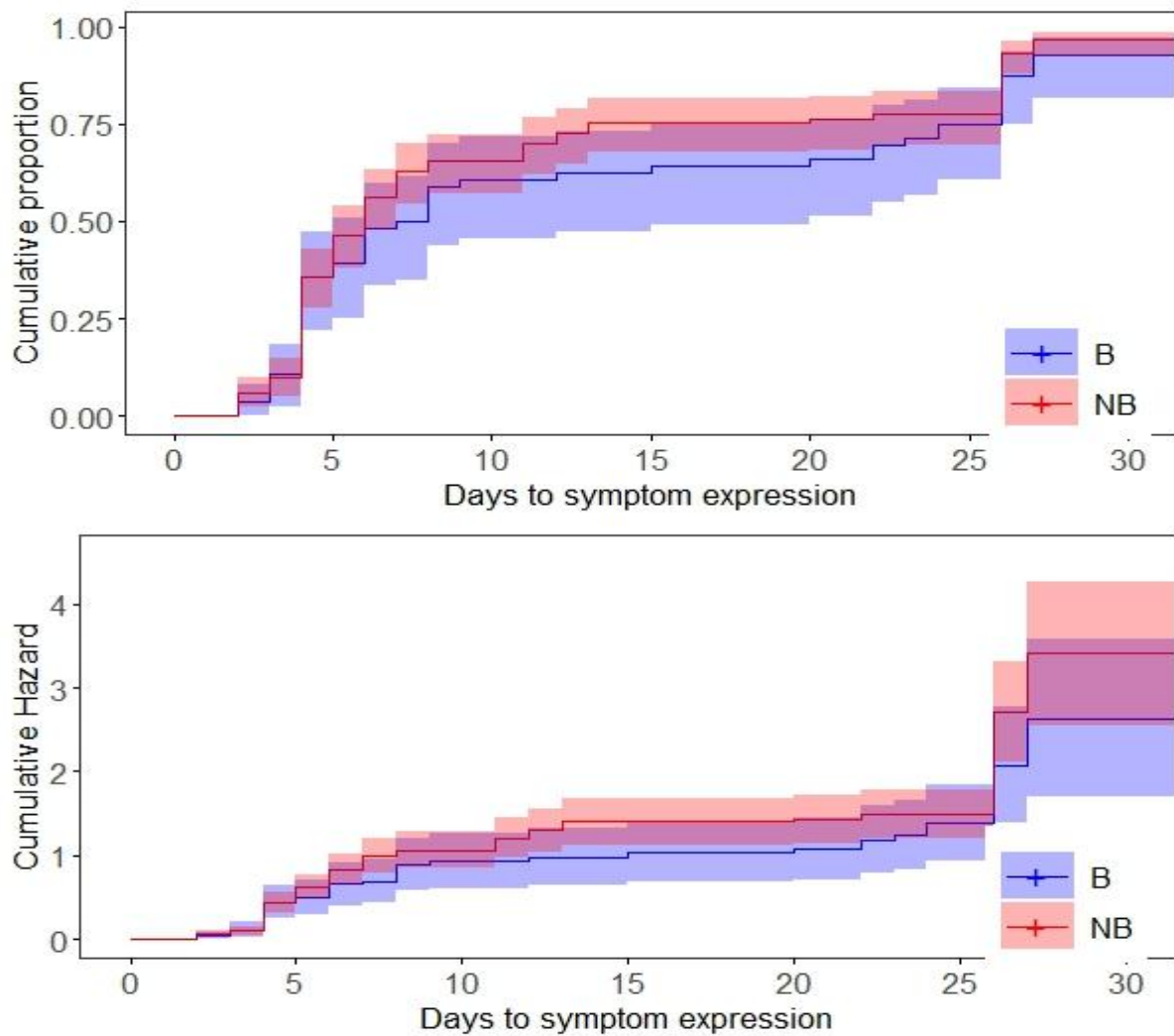

Figure 3. Nonparametric mortality and hazard trends of days to Fusarium wilt symptom expression (B – Beauregard, NB – Non-Beauregard).

56

Egypt. J. Phytopathol., Vol. 52, No. 1, pp46 –66

these plants to die later as time progressed. Trends for both categories of crosses flattened on the 27th day. The initial rapid increases were due to highly susceptible plants while the flattening and gentle increases reflected the fact that the remaining plants had higher levels of resistance. The second rapid increase in mortality correlates well with the pathogen overcoming the first line of resistance of the remaining symptom-less plants. This occurs when the pathogen in the plant system multiplies to a threshold level where symptom expression due to tissue breakdown is inevitable. The final flattening of the trend indicates the fact that most of the remaining plants had very high levels of resistance and were unlikely to succumb to the fungus. The results, therefore, suggest that there were: three classes of plants which were susceptible, intermediate, and resistant. The susceptible plants account for the first increase in mortality, the intermediate the second

increase after the first flattening, and the resistant the long-term survivors. Because the greenhouse environment was generally homogeneous, resistance was most likely genetic. However, in another study, Lanza *et al*. (2019) suggested that the expression of citrus canker lesions and premature fruit drop were influenced by the season of the year thus underscoring the importance of the environment and not just genetics in modelling disease management.

By the end of the study, only about 29% of Beauregard plants and 41% of non-Beauregard plants had died (Table 5). Therefore, no median time to death could be computed (Table 6). The earliest death was recorded 8 days after inoculation in a cross involving non-Beauregard parents. Also, an earlier mean time to death of 33.31 days was recorded for non-Beauregard crosses while crosses involving Beauregard as one of the parents took on average 10 days longer to die (Table 11).

Table 5: Summary of the number of censored and uncensored plants for plant death

Cross	Total	Plants Dead	Percent Dead	Plants Censored	Percent Censored
Beauregard	56	16	28.6	40	71.43
Non-Beauregard	150	61	40.7	89	59.33
Total	206	77	37.4	129	62.6

Table 6: Mean and median time to plant death

Cross	Mean (days)	Standard error of mean	Median (days)
Beauregard	43.14	1.18	-
Non-Beauregard	33.31	1.16	-

There was evidence of a significant effect of the category of cross as a covariate on time to plant death as indicated by the significant probability of the Chi-square tests presented in Table 7.

Consequently, modelling of survival and hazard functions considered crosses involving Beauregard and non-Beauregard crosses as coming from different independent populations.

Table 7: Test of equality of time to plant death over crosses

Test	Chi-square	df	Pr > Chi-square
Log-Rank	4.230	1	0.0397
Wilcoxon	6.456	1	0.0111
-2Log(LR)	4.398	1	0.0360

There was a higher rate of plants dying in the non-Beauregard category than in the other category of plants as illustrated in Figure 4. The rate of dying is lower in Beauregard crosses than in non-Beauregard crosses, and this may be attributed to the higher level of resistance that is found in the Beauregard parent and which is imparted onto the progeny. The fraction of dead individuals from the non-Beauregard crosses increased rapidly from about 0.005 on the eighth day to almost 0.26 on the fifteenth day as modelled by the KM estimate in Figure 4. These early deaths were primarily due to the low resistance levels in the plants. Conversely, the Beauregard group does not seem to exhibit the distinctive feature of early or statistical infant mortality. This was possibly because the Beauregard crosses were derived from three different parents with the more resistant Beauregard being the common parent while the non-Beauregard crosses were derived from five genetically heterogeneous parents with lower levels of resistance. In another study by Pereira *et al*., (2019), the effect of genetics was also observed on the survival of Passiflora spp when they analysed the incidence of *Fusarium oxysporum* f. sp. *Passiflorae*. At the end of our study on the 47th day, the rate of survival in the Beauregard plants was about 0.74 while in the non-Beauregard plants it was about 0.60 (Figure 4). We conclude that Beauregard plants are resistant during the early days of the disease but become weaker with time while plants in the non-Beauregard group succumb much earlier but the ones still surviving are relatively more resistant.
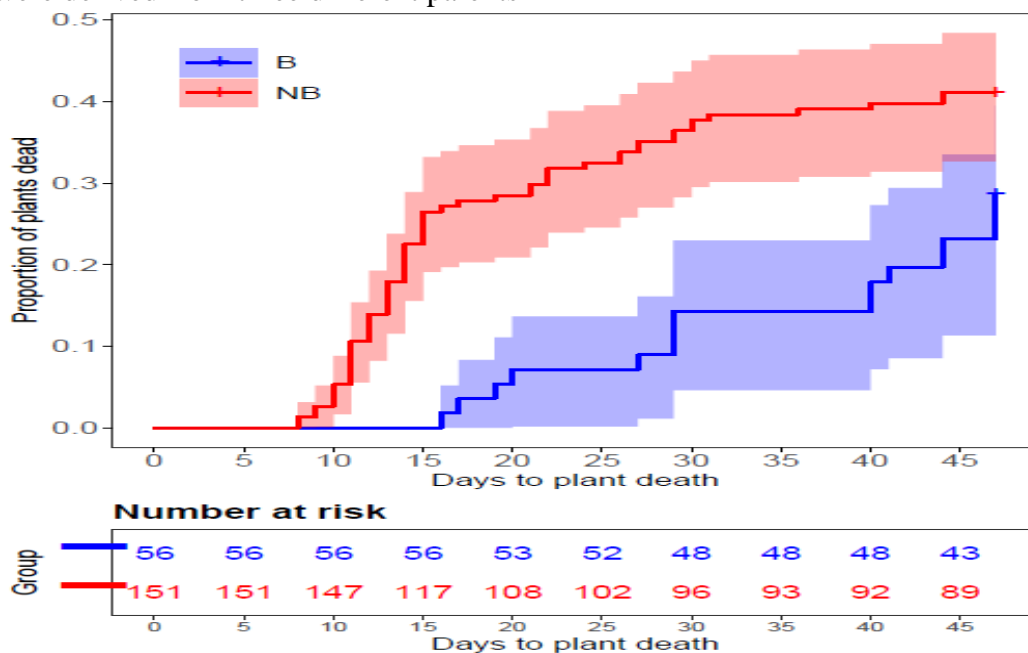

Figure 4. Nonparametric mortality curve for plant death with 95% point wise confidence bands (B – Beauregard, NB – Non-Beauregard).

58

Egypt. J. Phytopathol., Vol. 52, No. 1, pp46 –66

3.3. Parametric Tests:

The Wald Chi-square test statistic (P = 0.05) results obtained from modelling using the five parametric models found no significant effect of the category of cross on time to symptom expression. It was concluded that the Beauregard and non-Beauregard genotypes do not differ significantly in the time it takes for them to express symptoms. The significant Wald Chi-square test statistics from modelling for time to plant death suggested that there was a significant effect of the category of a cross on time to plant death. The Generalized gamma model elucidated the greatest differences between the two categories of crosses (P < 0.0001) followed by the Lognormal (P = 0.003). It was, therefore, concluded that the Beauregard and non-Beauregard crosses differed significantly in the time it took for them to die. This was not surprising because Beauregard is likely to have imparted some of its resistance to its progeny. These results confirm the non-parametric results that the two types of crosses belong to different distributions. The likelihood ratio tests (Giroux *et al.*, 2016) compared parametric fits of the simpler, or restricted, models with the more complicated, or unrestricted Generalized gamma model (Nesi *et al.*, 2015). The hypothesis tested was equality of restricted and unrestricted models. Fitting a Generalized gamma model for the Beauregard crosses for time to plant death was not possible because the model could not converge. The Exponential model was also tested against the Weibull model. The Chi-Square tests for plant death were all significant except when the Exponential model was tested against the Weibull model for plant death. These results suggest that none of the restrictions imposed on the three-parameter Generalized gamma model to derive the

two-parameter models resulted in a better fit. We conclude that the Generalized gamma model has a significantly better fit than its subsets.

The likelihood ratio tests in Table (8) assume that the unrestricted model fits the data well when compared against the Generalized gamma. The Lognormal model was the best fitting for both symptom expression and plant death. From the results, one may conclude that there was no significant difference between the Exponential and the Weibull models for time to plant death. However, when the Weibull model was tested against the Generalized gamma model we rejected the Weibull model. The Weibull model does not fit as well as the Generalized gamma. This result suggests that the restriction imposed on the Generalized gamma model to derive the Weibull model was not useful. Consequently, we would also reject the Exponential model regardless of the Exponential versus Weibull test results. On the other hand, Zaluski *et al.* (2018) found the Exponential model to provide a reliable fit for willow (*Salix L.*) survival times and lifespan. The Akaike information criterion (AIC) model fit statistics (Table 9) were interpreted as discussed by (Chang *et al.*, 2018). The results indicate that the Generalized gamma model had the best fit followed by the Lognormal for symptom expression. The Weibull model provided the best fit for time to death among Beauregard crosses while among non-Beauregard crosses the best restricted model was the Lognormal. Overall, the Generalized gamma fitted the data the best. The time to death results follow the trends of the data observed for time to symptom expression. All the tested models suggest that Beauregard crosses take longer to die. Parameter estimates for model formulation are presented in Table (10). The estimates generated by all the models are very

similar with mean time to symptom expression ranging from 11.67 days for the

Lognormal model to 12.28 days for the Log-logistic model (Table 10).

.

Table 8. Likelihood ratio tests for model fit

Event	Test	df	χ^2	$Pr > \chi^2$
Symptom expression	Lognormal vs. Generalized Gamma	1	827.00	<0.005
	Weibull vs. Generalized Gamma	1	867.20	<0.005
	Exponential vs. Weibull	1	8.80	<0.005
	Exponential vs. Generalized Gamma	2	876.00	<0.005
Plant death (non-'Beauregardard')	Lognormal vs. Generalized Gamma	1	382.60	<0.005
	Weibull vs. Generalized Gamma	1	397.00	<0.005
	Exponential vs. Weibull	1	0.60	0.439
	Exponential vs. Generalized Gamma	2	397.60	<0.005

The quantile estimates presented in Table (8) indicate the number of days it would take for a certain probability of the plant population to express symptoms. This is a useful value in predicting crop damage levels at given time periods. Distributions with a larger log-likelihood or a low AIC value have a better fit for the data. Consequently, the best restricted model would be the Lognormal. The unrestricted Generalized gamma model used has one scale parameter and two shape parameters. These parameters give it greater flexibility to fit better models for these data as compared to the other two-parameter models. As previously noted, several models are available to analyse survival data, and these models can allow multiple covariates and multiple parameters for increased robustness (Onofri *et al*., 2019 and Gianinetti, 2020). Madden *et al*. (2018) compared multiple Generalized linear models to fit grape downy mildew data. They found that the Generalized gamma distribution could be used with Generalized linear modelling methods.

Table 9. Akaike information criterion (AIC) model performance statistics

Event	Distribution	AIC
Symptom expression	Exponential	1380.80
	Log-logistic	1345.80
	Lognormal	1331.80
	Weibull	1372.00
	Generalized Gamma	505.80
Plant death (Beauregard)	Exponential	194.56
	Log-logistic	182.76
	Lognormal	182.60
	Weibull	182.50
Plant death (non-Beauregard)	Exponential	667.80
	Log-logistic	660.60
	Lognormal	652.80
	Weibull	667.20
	Generalized Gamma	271.20

3.4. Utility of the models:

There is evidence that the Kaplan-Meier nonparametric method and the Lognormal parametric model are adequate to describe survival functions for symptom development in sweet potato plants inoculated with the Fusarium wilt pathogen (Benali *et al*., 2015).

The significant difference in mortality curves between Beauregard and non-Beauregard crosses suggests that Beauregard as a parent imparted a significant amount of resistance to its progeny. Survival functions for plant death were well described by both nonparametric and parametric methods for Beauregard crosses.The lack of significant covariate effects in time to symptom expression was not unexpected because symptoms are mainly indicators of the presence of a disease. In contrast, the severity of symptoms reflects the inherent capacity of the plant to resist disease. Using a Kaplan-Meier approach, the severity of symptom expression was previously reported to differ among citrus varieties as the disease progressed (Frare *et al*., 2019). Nonetheless, disease severity analysis was not an objective of this study. Although the survival functions between the categories of crosses were not significantly different for symptom expression, the hazard function (Rathod *et al*., 2020), seems to have been a result of more factors than just the virulence of the pathogen. Such factors may have included the genetic make-up of the cross. It is

Table 10. Parameter estimates and model fit criteria for the best fitting models.

Event	Distribution	Parameter	Maximum likelihood estimate	Standard error	95% confidence intervals Lower	Upper
Symptom Expression	Lognormal	μ	2.080	0.061	1.962	2.199
		Σ	0.868	0.044	0.786	0.959
Plant death (Beauregard)	Weibull*	μ	4.265	0.142	3.987	4.543
		σ^1	0.377	0.091	0.236	0.604
		η^2	71.158	10.085	53.900	93.942
		β	2.652	0.637	1.656	4.246
	Lognormal	μ	4.243	0.163	3.923	4.563
		σ	0.679	0.139	0.455	1.013
Plant death (non-Beauregard)	Weibull	μ	4.391	0.144	4.110	4.673
		σ	0.915	0.108	0.727	1.152
		η	80.757	11.587	60.960	106.983
		β	1.093	0.129	0.868	1.376
	Lognormal	μ	4.049	0.143	3.769	4.328
		σ	1.251	0.132	1.017	1.539
	Generalized Gamma	μ^3	2.417	0.149	2.124	2.710
		σ^4	0.519	0.135	0.254	0.784
		κ	-5.966	1.774	-9.443	-2.490

* Weibull η {= exp(μ)}, Weibull β (= 1/σ), Gamma η {= exp(μ)} and Gamma β (= 1/σ).

worth noting that the hazard function is also called the instantaneous probability of a plant to express symptoms at time *t* given that it had not expressed the symptoms up to time *t*. Sweet potato is genetically a hexaploid, meaning that it has six sets of chromosomes as opposed to the two sets that most food crops have. This multiple chromosomal setup complicates the genetics of host-plant resistance of the sweet potato to diseases to the extent that it is very difficult to predict the behaviour of a genotype derived from cross-pollination. Consequently, it is possible to have progeny from the same parents behaving very differently. Disease resistance is primarily a product of the combined action of the genetic make-up of the plant and the environment in which the plant is growing. Since each plant is a unique genotype and the experiment was conducted in a homogeneous greenhouse environment, the genotype by environment interaction was minimal, and hence it is postulated that differences in time to death were highly influenced by the genotype of the plant. Therefore, disease modelling approaches should take into account both the environment and the genotype (Chen *et al.*, 2019). However, Onofri *et al.* (2022) suggest that interactions between the biotic and abiotic environments can make such modelling complicated.

3.5. Developing predictive models:

The quintile estimates using Lognormal and Weibull models as presented in Table (11) indicate the number of days it would take for a certain probability of the plant population to express symptoms or to die. According to the Lognormal prediction model, 25% of the plants will express symptoms by the fourth day and 75% by the 15th day. By this time most of the

susceptible plants are predicted to have expressed symptoms and an additional 45 days required for 99% of the plants to show symptoms. The standard errors are low thus suggesting an accurate model. This is a useful value in predicting crop damage levels at given periods. Our results in Table (11) suggest that suitability of a model to predict plant death depends on the genotype. Also, accuracy of prediction is seemingly restricted to 50% death because beyond that the standard errors for both Weibull and Lognormal models become large. Other models like the Logistic and Gompertz have also been used for plant disease prediction (Kebede and Golla, 2020; Singh and Deepankar, 2020) but their suitability vary with diseases. Singh *et al.*, (2019) reported the efficacy of Exponential and Gompertz models to fit epidemic data for Ber powdery mildew. These reports suggested that the accelerated failure time (AFT) models, were adequate in delineating the effects of covariates on survival times. Benali *et al.*, (2015), on the other hand, used the PH model to adequately model incubation period of Ascochyta blight in pea correctly simulate fruit setting times. Survival analysis of wheat infected with wheat blast has also been conducted using parametric models (Mills *et al.*, 2021). These workers used Exponential and Weibull models to fit survival and hazard functions on life data of adult leafhoppers that had been exposed to maize mycoplasma under greenhouse conditions. In that controlled environment, the Weibull distribution gave them a better fit.

4. CONCLUSION

This study has shown that it is possible to develop models that approximate the time it takes for symptoms and death to be observed on inoculated plants. The resistance genes in Beauregard-related plants conferred to

these plants better survival chances. The utility of non-parametric Kaplan-Meier method was limited because it ignored censored data. The three-parameter Generalized gamma model provided the best fit for the data for symptom expression among all the crosses and time to plant death among the non-Beauregard crosses. However, the gamma model needs sufficient data to converge and provide a well-fitting model. The two-parameter Lognormal model was the best restricted predictor for symptom expression and plant death among non-Beauregard plants. The Weibull model was the best restricted predictor among Beauregard-related plants. Therefore, when data are inadequate, the Weibull model is the best predictive model.

Table 11. Parametric maximum likelihood quintile estimates for the best restricted models

Event	Distribution	p value[a]	Quintile (days)	Standard error	Confidence intervals (95%)	
					Lower	Upper
Symptom expression	Lognormal	0.25	4.46	0.30	3.91	5.09
		0.50	8.01	0.49	7.11	9.02
		0.75	14.38	0.98	12.59	16.43
		0.99	60.33	7.22	47.72	76.27
Plant death (Beauregard)	Weibull	0.25	44.48	4.21	36.96	53.54
		0.50	61.97	7.38	49.08	78.25
		0.75	80.49	13.28	58.25	111.21
		0.99	126.57	33.12	75.78	211.38
Plant death (non-Beauregard)	Lognormal	0.25	24.66	2.89	19.60	31.02
		0.50	57.32	8.18	43.34	75.82
		0.75	133.26	27.59	88.81	199.97
		0.99	1051.82	427.53	474.20	2333.07

[a]p value is the proportion of plants expressing the event of interest

6.   CONFLICT OF INTEREST

The author(s) declare no conflict of interest.

6.   REFERENCES

Benali, S.; Mohammed, B. and Eddine, H.J. 2015. Survival analysis for estimating risk factors in incubation period of *Didymella pinodes* of pea (*Pisum sativum L.*). African Journal of Microbiology Research., 9(9):604-610. https://doi.org/10.5897/AJMR2014.7051.

Bogaerts, K.; Komarek, A. and Lesaffre, E. (2018). Survival analysis with

interval-censored data: a practical approach with examples in R, SAS, and BUGS. Chapman and Hall/CRC. Boca Raton, FL.Chang, A.L.S.; Phin, C.K. and Mun, H.C. 2018. Comparing nonlinear models in describing disease progress curve of cocoa black pod. AIP Conference Proceedings., 1974(1):040008. https://doi.org/10.1063/1.5041682.

Chalwe, J.; Adebola, P. and Pillay, M. 2017. Assessing the Genetic Diversity of *Alternaria bataticola* in South Africa using Molecular Markers. Asian Journal of Applied Sciences., 5:280-290.

Chang, A.L.S., Phin, C.K., and Mun, H.C. 2018. Comparing nonlinear models in describing disease progress curve of cocoa black pod. AIP Conference Proceedings., 1974(1):040008. https://doi.org/10.1063/1.5041682.

Chen, M.; Brun, F.; Raynal, M. and Makowski, D. 2019. Timing of grape downy mildew onset in bordeaux vineyards. Phytopathology.,109:787-795. https://doi.org/10.1094/PHYTO-12-17-0412-R. .

Chiang, K.S.; Chang, Y.M.; Liu, H.I.; Lee, J.Y.; Jarroudi, M.E. and Bock, C.H. 2024. Survival Analysis as a Basis for Testing Hypotheses when Using Quantitative Ordinal Scale Disease Severity Data. Phytopathology., 114(2). PHYTO-02. https://doi.org/10.1094/phyto-02-23-0055-r.

Clark, C.A; Hyun, J.W. and Hoy, M.W. 1998. Relationships among wilt-inducing isolates of *Fusarium oxysporum* from sweet potato and tobacco. Plant Disease., 82:530-536.

Cox, D.R. 1972. Regression models and life tables. Journal of the Royal Statistical Society B., 34:187-220. FAO. 2020. FAOSTAT: Crops and Livestock Products. https://www.fao.org/faostat/en/#data/QCL

Frare, G.F.; Silva-Junior, G.J.; Lanza, F.E.; Bassanezi, R.B.; Ramires, T.G. and Amorim, L. 2019. Sweet orange fruit age and inoculum concentration affect the expression of citrus black spot symptoms. Plant Disease, 103:913-921 https://doi.org/10.1094/pdis-03-18-0492-re.

Gianinetti, A. 2020. Basic features of the analysis of germination data with Generalized linear mixed models. Data., 5:6. https://www.mdpi.com/2306-5729/5/1/6.

Giroux, M.E.; Bourgeois, G.; Dion, Y.; Rioux, S.; Pageau, D.; Zoghlami, S.; Parent, C.; Vachon E. and Vanasse, A. 2016. Evaluation of forecasting models for Fusarium head blight of wheat under growing conditions of Quebec, Canada. Plant Disease, 100:1192-1201. https://apsjournals.apsnet.org/doi/pdf/10.1094/PDIS-04-15-0404-RE.

Harris, C.M.; Sadykova, D.; DeRuiter, S.L.; Tyack, P.L.; Miller, P.J.O.; Kvadsheim, P.H.; Lam, F.P.A. and Thomas, L. 2015. Dose response severity functions for acoustic disturbance in cetaceans using recurrent event survival analysis. Ecosphere., 6(11):236. http://dx.doi.org/10.1890/ES15-00242.1.

Kaplan, E.L. and Meier, P. 1958. Nonparametric estimation from incomplete observations. Journal of

American Statistical Association, 53:457-481.

Kebede, A.A. and Golla, W.N. 2020. Model selection in describing disease progress curve of Fusarium wilt (*Fusarium oxysporum* f.sp. *Sesami*) disease in sesame varieties. International Journal of Pathogen Research, 5(2):30-38.

Lanza, F.E.; Marti, W.; Silva Jr., G.J. and Behlau F. 2019. Characteristics of citrus canker lesions associated with premature drop of sweet orange fruit. Phytopathology, 109:44-51. https://doi.org/10.1094/phyto-04-18-0114-r.

Madden, L.V.; Hughes, G.; Moraes, W.B.; Xu, X.M. and Turechek, W.W. 2018. Twenty-five years of the binary power law for characterizing heterogeneity of disease incidence. Phytopathology, 108(6):656-680. https://doi.org/10.1094/phyto-07-17-0234-rvw.

Mcharo, M.; Carey, E.E. and Gichuki, S.T. 2001. Performance of selected sweet potato varieties in Kenya. African Crop Science Journal., 9(1):49-57. http://dx.doi.org/10.4314/acsj.v9i1.2 7624.

Meeker, W.Q. and Escobar, L.A. 1998. Statistical Methods for Reliability Data. John Wiley and Sons, Inc. USA. pp 679.

Mills, K.B.; Salgado, J.D.; Cruz, C.D; Valent, B.; Madden, L.V. and Paul, P.A. 2021. Comparing the temporal development of wheat spike blast epidemics in a region of Bolivia where the disease is endemic. Plant Disease, 105(1):96-107.

Minzat, D.; Breaban, M. and Luchian, H. 2018. Modelling real estate dynamics using survival analysis. 2018. 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC). https://www.researchgate.net/publica tion/334167194_Modeling_real_estat e_dynamics_using_survival_analysis .

Moore, D.F. 2016. Applied survival analysis using R. Springer. Piscataway, NJ. USA. pp 226.

Mousa, T.A. ; Farag, F.M..; Armanious, H.A.; Salem A.A. and Galal. A.A. 2018. Fusarium wilt of sweet potato caused by *Fusarium oxysporum* f. sp. *batatas* in Egypt. Egyptian Journal of Phytopathology, 46(1):21-35. https://ejp.journals.ekb.eg/article_87 431_53c6d0ec9ee638e0fd84a4971f7 e453f.pdf.

Nesi, C.N.; Shimakura, S.E.; Junior, P.J.R. and De Mio, L.L.M. 2015. Survival analysis: a tool in the study of post-harvest diseases in peaches. Review of Ceres and Viçosa, 62(2):052-061. http://dx.doi.org/10.1590/0034-737X201562010007.

Nsobinenyui, D.; Mbafor, L.F.; Mensiy, W.S.; Abdulai, A.N.; Mapiemfu-Lamare, D.; Sama, V.; Achiri, D.T. and Njualem, D.K. 2022. Survival analysis of maize weevil *Sitophilus zeamais* Motschulsky 1855 (Coleoptera: Curculionidae) exposed to some medicinal plant powders. World News of Natural Sciences, 43:126-137.

Onofri, A.; Mesgaran, M.B. and Ritz, C. 2022. A unified framework for the analysis of germination, emergence, and other time-to-event data in weed

science. Weed Science., 70:259–271. https://doi.org/10.1017/wsc.2022.8.

Onofri, A.; Piepho, H-P. and Kozak, M. 2019. Analysing censored data in agricultural research: a review with examples and software tips. Annals of Applied Biology, 174:3–13. https://doi.org/10.1111/aab.12477..

Özen, H.; Bal, C. and Çolak, E. 2021. Comparison of the Statistical Tests for Two Crossing Survival Functions. Türkiye Klinikleri Biyoistatistik., 13(1):13-27.

Pereira, P.P.A.; Lima, L.K.S.; Soares, T.L.; Laranjeira, F.F. and de Jesus, O.N. 2019. Initial vegetative growth and survival analysis for the assessment of Fusarium wilt resistance in *Passiflora* spp. Crop Protection, 121:195-203. https://doi.org/10.1016/j.cropro.2019 .03.018.

Rathod, S.; Saha, A. and Sinha, K. 2020. Introduction to survival analysis and application in agricultural research. Food and Scientific Reports, 1(4):28-30.

Romano, A. and Stevanato, P. 2020. Germination data analysis by time-to-event approaches., Plants 9:617. https://www.mdpi.com/2223-7747/9/5/617.

Shrestha, S.; Elderd, B.D. and Dukic, V. 2019. Bayesian-based survival analysis: inferring time to death in host-pathogen interactions. Environmental and Ecological Statistics, 26:17-45. https://doi.org/10.1007/s10651-019-00418-3.

Sing, A.S. and Dlamini. S.G. 2021. Analytical models of survival analysis: Concepts and their applications. International Journal of Science and Research Methodology, 19:2150-169.

Singh, R. and Deepankar. 2020. Modelling for anthracnose development in mango in relation to weather parameters. Australasian Plant Pathology, 49:285–294. https://doi.org/10.1007/s13313-020-00704-w.

Singh, R.; Kumar, M.; Mamta, D. and Baloda, S. 2019. Development of growth model for Ber powdery mildew in relation to weather parameters. Indian Phytopathology, 72:235–241. https://doi.org/10.1007/s42360-019-00124-x.

Sseruwu, G.; Nanyanzi, M.; Alajo, A.; Kituuka, G. and Benywanira, I. 2020. Evaluation of Sweetpotato (*Ipomoea batatas* (L.) Lam.) genotypes for resistance to Alternaria leaf petiole and stem blight (*Alternaria* spp.) in Uganda. Journal of Agricultural Science, 12(10):263-277. https://ccsenet.org/journal/index.php/ jas/article/download/0/0/43679/4588 6.

Yellareddygari, S.; Domfeh, O.; Bittara, F.G. and Gudmestad, N. 2017. Analysis of potato mop-top virus survival probability in post-harvest storage. American Journal of Potato Research, 94:632–637. https://link.springer.com/article/10.1 007/s12230-017-9601-5 .

Zaluski, D.; Mielniczuk, J.; Bronowicka-Mielniczuk, U.; Stolarski, M.J.; Krzyżaniak, M.; Szczukowski, S. and Tworkowski, J. 2018. Survival analysis of plants grown in long-term field

experiments. Agronomy Journal, 110:1791– 1798. https://acsess.onlinelibrary.wiley.com/doi/full/10.2134/agronj2018.01.0062.

Zee, J. and Xie, S.X. 2018. The Kaplan–Meier method for estimating and comparing proportions in a randomized controlled trial with dropouts. Biostatistics and Epidemiology, 2(1):23-33.