

Kansas State University Libraries

New Prairie Press

---

Conference on Applied Statistics in Agriculture

2004 - 16th Annual Conference Proceedings

---

## NONLINEAR MODELS WITH REPEATED MEASURES FOR ANALYZING DISEASE PROGRESS IN PLANT EPIDEMIOLOGY

R. Macchiavelli

W. Robles

E. Abreu

A. Pantoja

*See next page for additional authors*

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

---

### Recommended Citation

Macchiavelli, R.; Robles, W.; Abreu, E.; and Pantoja, A. (2004). "NONLINEAR MODELS WITH REPEATED MEASURES FOR ANALYZING DISEASE PROGRESS IN PLANT EPIDEMIOLOGY," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1164>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact [cads@k-state.edu](mailto:cads@k-state.edu).

---

**Author Information**

R. Macchiavelli, W. Robles, E. Abreu, and A. Pantoja

## Nonlinear Models with Repeated Measures for Analyzing Disease Progress in Plant Epidemiology

R. Macchiavelli, W. Robles, E. Abreu and A. Pantoja  
 College of Agricultural Sciences, University of Puerto Rico – Mayagüez, PO Box 9030,  
 Mayagüez, Puerto Rico, 00681-9030

### Abstract

Nonlinear models are commonly used in plant disease epidemiology to model temporal changes in the proportion of diseased plants (disease index). Most of the times they are fit using linearizing transformations or nonlinear least squares. These approaches assume that the disease index has a normal distribution, that they are independent and that they have constant variance. None of these assumptions can be justified in disease indices. In this paper we apply different strategies to model the progress of papaya ring spot virus in papaya. Using the logistic model we compare different strategies using the SAS® System. Marginal (population average) and subject-specific interpretations of the models are discussed.

### 1. Introduction

Plant diseases are normally monitored over time, assessing the amount of disease present in a population of plants: the “Disease Progress Curve” represents an interpretation of all host, pathogen and environmental effects occurring during an epidemic (Campbell and Madden, 1990). This curve is the disease index (proportion of plants showing symptoms in a plot) as a function of time. It provides a tool for analyzing plant disease epidemics, comparing different conditions (treatments) and predicting disease dynamics.

Four different models are commonly used for disease progress curves: monomolecular, logistic, Gompertz, and exponential. Table 1 presents the form of these models, as a differential equation (indicating rate of disease increase or decrease) or in the integrated form.

Table 1. Models used in disease progress curves.

Model	Differential equation	Integrated Equation
Exponential	$\frac{dY}{dt} = r_e Y$	$Y = Y_0 \exp(r_e t)$
Monomolecular	$\frac{dY}{dt} = r_m (1 - Y)$	$Y = 1 - B \exp(-r_m t)$
Logistic	$\frac{dY}{dt} = r_l Y(1 - Y)$	$Y = \frac{1}{1 + \exp(-B + r_l t)}$
Gompertz	$\frac{dY}{dt} = r_g Y [-\log Y]$	$Y = \exp[-B \exp(-r_g t)]$

Papaya, *Carica papaya* L., is a fruit commonly cultivated in the tropics and subtropics. In Puerto Rico, papaya is severely affected by the papaya ringspot virus (PRV) and the papaya mosaic virus (PMV). Both viruses can be transmitted by aphids in non-persistent manner. This means that the virus can be acquired during brief probes, retained for several hours and transmitted to a new host within minutes. So, the chemical control may not be effective preventing transmission. This makes it important to find more effective control methods against aphids and viruses on papaya. Plastic mulch has been an effective control method to delay viral diseases on vegetables (All, 1999). In the present work we evaluate the effect of plastic mulch and green cover crop (weed) on the delay of viral symptoms on papaya fields. We compare the results of fitting models under different assumptions to see which treatment is best for controlling aphids, and thus indirectly control virus spread in papaya plants.

## 2. Materials and Methods

A study was conducted in 2000-01 with the PR 6-65 variety in the Agricultural Experiment Station at Isabela, Puerto Rico. A completely randomized design with five replications was used. The plots consisted of twenty plants at a distance of 3.0 meters between rows and 1.2 meters between plants. The following treatments were evaluated: reflective plastic mulch; black plastic mulch; green cover crop (weeds) and bare ground.

In order to study the progress of the viruses, the disease index (proportion of plants with virus symptoms) was computed for each plot every two weeks. Once a plant showed symptoms, it was classified as diseased for the rest of the experiment. The response was the disease index (proportion of plants with symptoms among the 20 plants in each plot) at each period. The following disease progress models were used: exponential, monomolecular, logistic and Gompertz (Campbell and Madden, 1990).

The models were fitted using the following strategy:

1. First we assumed approximate normality and fitted the nonlinear models using time, treatment and interaction effects.
2. Lack of fit tests were performed and based on the results of these tests we decided to use the logistic model (Gompertz was also a good alternative)
3. We fitted the logistic model using the SAS<sup>®</sup> System with five different methods:
  - Traditional: separate linearized regression for each plot (SAS<sup>®</sup> Proc Reg)
  - Nonlinear regression with normal distribution and treatment effects on intercepts and slopes (SAS<sup>®</sup> Proc Genmod or Proc Nlin)
  - Nonlinear regression with overdispersed binomial distribution and treatment effects (SAS<sup>®</sup> Proc Genmod / dscale option)
  - Nonlinear regression with binomial distribution and random plot effects (SAS<sup>®</sup> Proc Nlmixed)
  - Nonlinear regression with binomial distribution and correlation among repeated measures (SAS<sup>®</sup> Proc Genmod / repeated)
4. The results and interpretation of each model are discussed.

Plant pathologists (for example, Campbell and Madden, 1990) recommend to fit the linearized version of the models. Once the linear regression is done, a model is selected based on the best fit, good behavior of residuals, etc. Inference is based on the linear regression analysis, separately for each treatment. There are several problems with this approach: the errors may not be additive in the transformed scale, and hence the linearized version is not correct, the distribution may not be normal, the variances may not be constant, the treatment comparisons are not done using the full power of a single model, observations of the same plant in different weeks are dependent and observations on the same plot may not be independent (for example, because of contagion). In table 2 we show the results from fitting linearized and nonlinear models. The slopes (parameters of interest related to the spread of the disease) are very different, and the residual plots from the linearized fits show better behavior for the nonlinear fit (figure 1).

Table 2. Slopes estimated using the linearized and nonlinear fit for each of the disease progress curves models (LR=linearized regression, NLR=non linear regression).

	<b>Control (Bare soil)</b>		<b>Weeds</b>		<b>Silver Plastic</b>		<b>Black Plastic</b>	
<b>Model</b>	<b>LR</b>	<b>NLR</b>	<b>LR</b>	<b>NLR</b>	<b>LR</b>	<b>NLR</b>	<b>LR</b>	<b>NLR</b>
Monomolecular	.034	.016	.043	.018	.100	.029	.094	.027
Exponential	.079	.020	.080	.019	.072	.014	.079	.015
Logistic	.114	.059	.122	.062	.173	.113	.173	.102
Gompertz	.091	.040	.093	.041	.097	.084	.102	.074

Figure 1a. Residual plot from linearized fit of the Gompertz model in the weed treatment.

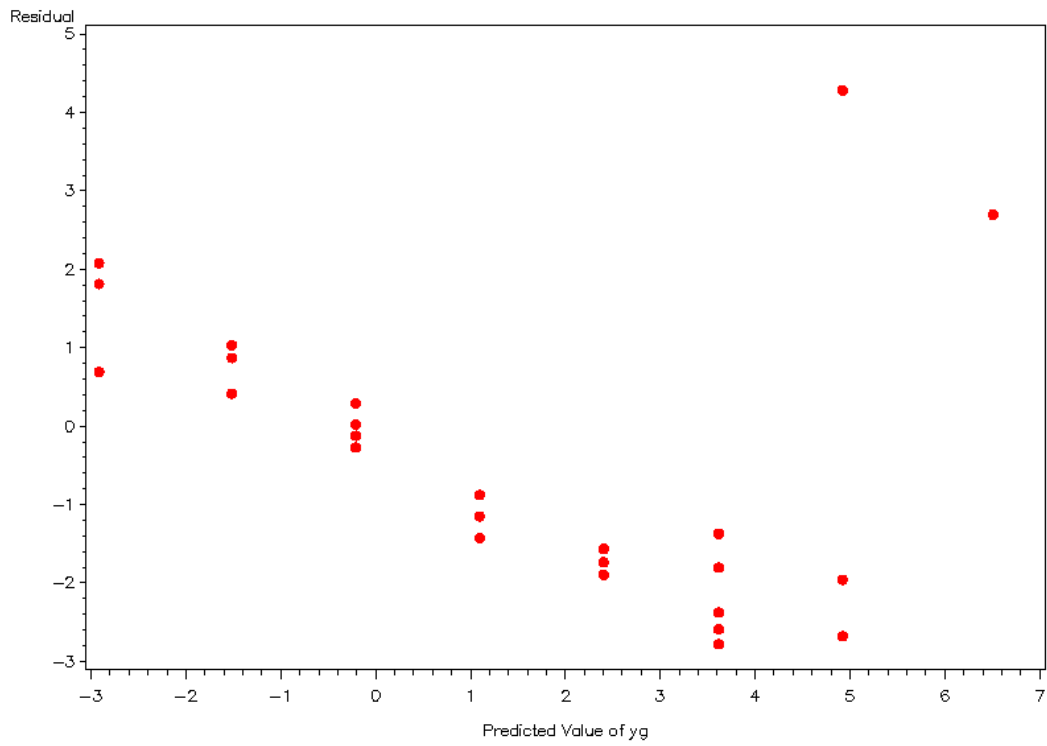
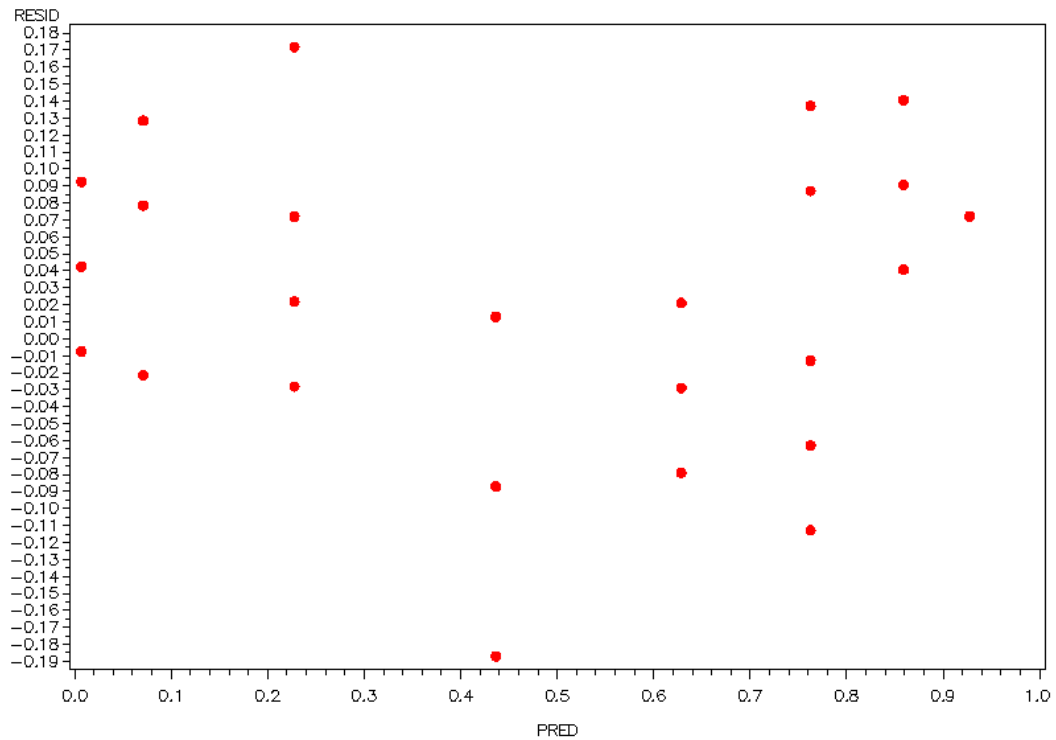


Figure 1b. Residual plot from nonlinear fit of the Gompertz model in the weed treatment.



Some solutions to these problems are to use nonlinear models incorporating treatment effects on the intercepts and/or slopes, to consider distributions other than normal (binomial, beta-binomial) and to consider longitudinal data methodology. In order to use nonlinear models incorporating treatment effects, SAS<sup>®</sup> PROC NLIN can be used, but one needs to set up dummy variables for treatments and needs initial values of parameters. (As initial values, one could use results from linear regression.) But, since every linearizable model can be written as a generalized linear model, SAS<sup>®</sup> PROC GENMOD can be used. This procedure does not require initial values or the specification of dummy variables for treatments (it has a CLASS statement). If none of the available link functions is appropriate, a specific one can be written with the FWDLINK command.

Since there are replicates, tests for lack of fit can be used. In generalized linear models with normal distribution and identity link, the deviance equals the error sum of squares. To obtain the “pure error” we need a model with dummy variables for each time (treatment x time factorial structure). In this case,

$$F = \frac{\frac{\text{Error SS (reg)} - \text{Error SS (ANOVA)}}{\text{Error df (reg)} - \text{Error df (ANOVA)}}}{\frac{\text{Error SS (ANOVA)}}{\text{Error df (ANOVA)}}}$$

For other distributions, the same idea yields likelihood ratio tests for lack of fit:

$$\chi^2 = \text{Deviance (reg)} - \text{Deviance (ANOVA)}$$

As shown in Table 3, the F lack-of fit tests for all models under the assumptions of normality and independence indicate that the logistic and Gompertz models seem appropriate for these data.

Table 3. Lack of fit tests assuming normality and independence

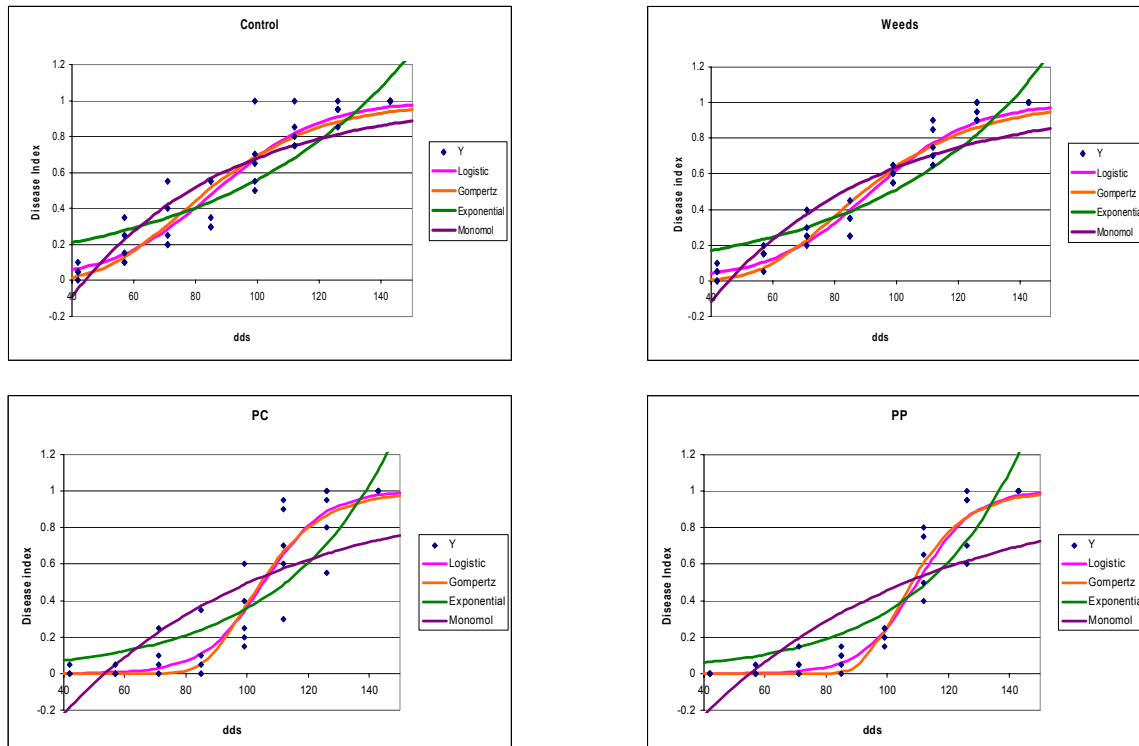
<b>Model</b>	<b>Error SS</b>	<b>Error df</b>	<b>F (lack of fit)</b>	<b>p-value</b>
ANOVA	1.5110	128		
Monomolecular	5.6349	152	14.5560	0.0000
Exponential	3.5923	152	7.3463	0.0000
Logistic	1.6916	152	0.6375	0.8999
Gompertz	1.9131	152	1.4163	0.1104

From here on, the rest of the modeling strategies use the logistic curve as representative of the disease progress curve of these data.

### 3. Results

Figure 2 shows the fit of the different models to the data under normality and independence. It is clear from this figure that both the logistic and the Gompertz model provide a good fit in all treatments. Neither the exponential nor the monomolecular models are able to model the inflexion point which the data clearly show.

Figure 2. Nonlinear fit of the different models under normality and independence.



If we do not assume normality or independence in the data, different modeling options can be used. The first possibility is to consider the disease index (proportion of plants with symptoms) at each occasion as a binomial proportion. If we define  $Y^* = 20Y$  (number of diseased trees in the plot), the logistic curve and binomial assumption imply that

$$Y_{ijk}^* \sim \text{Binomial}(20, \pi_{ij})$$

$$\text{logit } \pi_{ij} = \mu + \alpha_i + r_i \text{dap}_j$$

The main difficulty dealing with this as a regular generalized linear model is that  $Y^*$  is not generally binomial because of the possible contagion. Alternatives are using the beta binomial distribution (difficult to consider treatment effects), generalized linear models methods with overdispersion, generalized linear models using marginal expected values and particular correlation structures, and nonlinear mixed models.



In the binomial distribution the variance is determined by the mean. If the binomial assumptions fail (i.e., contagion) the variance may be larger than the one predicted by the binomial mean (Stokes et al., 2001). In this case an overdispersion parameter ( $\phi$ ) must be estimated and standard errors and tests can be adjusted to account for this extra variation. Using an overdispersion makes a “quasi-likelihood” function instead of a true likelihood, but the asymptotic properties on which the inference is based do not change (Stokes et al., 2001). Using SAS<sup>®</sup> PROC GENMOD or other software, the overdispersion parameter  $\phi$  can be estimated as the deviance divided by its degrees of freedom, and the tests are adjusted accordingly.

In order to use marginal models, we specify the marginal means and the marginal correlation between observations taken repeatedly on the same unit:

$$E(Y_{ij}) = \mu_{ij}, \quad h(\mu_{ij}) = X_{ij}' \beta$$

$$Var(Y_{ij}) = v(\mu_{ij})\phi$$

$$Corr(Y_{ij}, Y_{ik}) = \rho(\mu_{ij}, \mu_{ik}; \alpha)$$

Regression coefficients are interpreted as average effects through the subpopulation having same values of  $x$  (“population average inference”). Except in normal linear models (and a few other cases), these regression coefficients are different from the ones estimated in a “subject specific” model, which we will discuss later. SAS<sup>®</sup> PROC GENMOD permits fitting the marginal models by using the REPEATED command. In general, there is no likelihood available, and hence the inference is based on a “quasi-likelihood”. PROC GENMOD reports “quasi score” tests when fitting these models.

For nonlinear mixed models, the repeated observations from the same plot are correlated because they share the same value of one or more random plot effects. In its simplest form for generalized linear models, the random effect adds a value to the intercept of the linear predictor. In the logistic model that we are using for the papaya data,

$$Y_{ijk}^* | u_k \sim \text{Binomial}(20, \pi_{ijk})$$

$$\text{logit } \pi_{ijk} = \mu + \alpha_i + u_k + r_i \text{ dap}_j$$

$$u_k \sim \text{Normal}(0, \sigma^2)$$

The interpretation of regression coefficients corresponds to a “typical” subject with a value of the random effect  $u=0$  (“subject specific interpretation”). The fixed effects (for example difference in treatment effects) are controlled by the plot (random effect). This model can be fitted in SAS<sup>®</sup> using PROC NLMIXED.

A more complicated model can be formulated if individual trees responses (0,1) are used. This formulation models the correlation between repeated observations on the same tree (longitudinal) and between trees on the same plot (contagion):

$$Y_{ijkl} | u_k, v_{l(k)} \sim \text{Bernouilli}(\pi_{ijkl})$$

$$\text{logit } \pi_{ijkl} = \mu + \alpha_i + u_k + v_{l(k)} + r_i \text{ dap}_j$$

$$u_k \sim \text{Normal}(0, \sigma_u^2)$$

$$v_{l(k)} \sim \text{Normal}(0, \sigma_v^2)$$

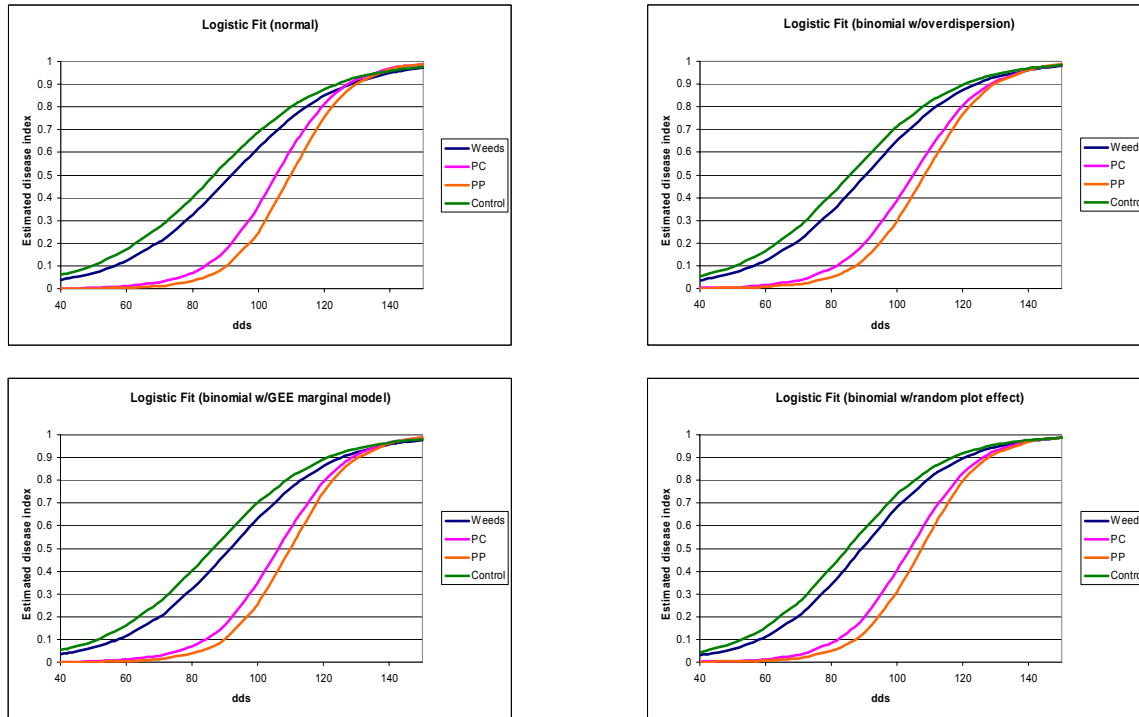
This formulation would require two RANDOM commands in SAS<sup>®</sup> PROC NLMIXED, which is not a current option (SAS v.9). Alternatively, a more complex formulation could be done using only one RANDOM command by specifying 20 random effects, and indicating the covariance matrix of these random effects as a function of  $\sigma_u^2, \sigma_v^2$ . The fit of this alternative using SAS<sup>®</sup> was not possible because of memory problems.

Table 4 shows the estimates obtained by the four fitting methods used. The parametrization used for the treatment effects on the intercept and the slope is the standard one used by SAS<sup>®</sup> PROC GLM and GENMOD. We can see that the coefficients are similar under the different modeling strategies, and the standard errors differ markedly. Figure 3 shows the fitted curves for the four different treatments.

Table 4. Fitted models.

Parameter	Normal NLReg	Std. Error	Binom overdis	Std. Error	Binom margin	Std. Error	Binom mixed	Std. Error
<b>Intercept</b>	-5.1339	0.5073	-5.4031	0.5109	-5.4164	0.3678	-5.7562	0.4927
<b>trat</b>   <b>W</b>	-0.5297	0.7628	-0.4827	0.7463	-0.5185	0.6277	-0.2265	0.6995
<b>trat</b>   <b>BP</b>	-5.6118	1.5042	-4.4469	1.0723	-5.0679	1.4240	-5.4332	0.9936
<b>trat</b>   <b>SP</b>	-7.1632	1.7807	-5.8215	1.2254	-6.5040	1.3186	-5.9576	1.0310
<b>trat</b>   <b>C</b>	0	0	0	0	0	0	0	0
<b>dap</b>	0.0593	0.0056	0.0631	0.0057	0.0627	0.0027	0.0673	0.0045
<b>dap*tr</b>   <b>W</b>	0.0023	0.0082	0.0020	0.0081	0.0021	0.0058	-0.0010	0.0062
<b>dap*tr</b>   <b>BP</b>	0.0425	0.0145	0.0308	0.0106	0.0359	0.0130	0.0395	0.0089
<b>dap*tr</b>   <b>SP</b>	0.0535	0.0167	0.0404	0.0117	0.0457	0.0136	0.0408	0.0091
<b>dap*tr</b>   <b>C</b>	0	0	0	0	0	0	0	0

Figure 3. Fitted models



Contrasts of interest to the researchers are to compare the plastic covers with the control and weed treatments. This can be done in all cases using F or chi-square tests with 4 degrees of freedom (here the contrasts test simultaneously the equality of 2 intercepts and 2 slopes). The results from these tests are presented in Table 5 (for comparison purposes, F and chi-square tests are indicated as generated by SAS<sup>®</sup>, and in parentheses we present the equivalent test when not presented by SAS<sup>®</sup>).

Table 5. Contrasting equality of intercepts and slopes for control and weeds, and for silver and black plastic.

Model	Num df	Den df	F	Pr>F	Chisq	Pr>Chi	Type
Normal NLReg	4	152	2.12	0.0814	8.47	0.0759	LR
Binom overdis	4	152	1.43	0.2269	5.72	0.2212	QLR
Binom margin	4		(0.49)		1.95	0.7447	Score
Binom mixed	4	19	0.34	0.8486	(1.32)		LR

Another inferential procedure in which there was interest is in predicting the disease index at 90 days after planting, which is an important phenological stage in papaya: at 90 days flowers start to bloom, and if there are fewer plants with virus symptoms at this moment, the damage caused by the disease will be small. Table 6 shows the estimated disease indices for 90 dap for each treatment and each fitting method. This table also presents the width of the confidence interval for this estimate.

Table 6. Predicting disease index at 90 days after planting.

Treatment	Normal NLReg		Binom overdis		Binom margin		Binom mixed	
	Predicti on	Width of C.I.	Predicti on	Width of C.I.	Predicti on	Width of C.I.	Predicti on	Width of C.I.
Control	0.5516	0.1025	0.5693	0.1267	0.5561	0.2608	0.5738	0.3104
Weed	0.4715	0.1070	0.4947	0.1294	0.4734	0.1043	0.4946	0.3234
Black plastic	0.1712	0.1378	0.1989	0.1244	0.1664	0.2493	0.1709	0.1952
Silver plastic	0.1055	0.1228	0.1298	0.1080	0.1033	0.0533	0.1208	0.1513

Another question of interest is the inverse prediction problem: when would, for example, 50% of the plants show symptoms? This can be expressed as a nonlinear function of the parameters of the model, and thus can be estimated directly using the nonlinear mixed model approach. The other modeling approaches do not directly lead to estimating nonlinear functions of the parameters. In SAS<sup>®</sup> PROC NLMIXED, the command ESTIMATE generates point estimates and Wald's confidence intervals. Table 7 shows these estimates.

Table 7. Estimating number of days after planting until 50% of the plants show symptoms

<b>Treatment</b>	<b>95% lower limit</b>	<b>Point estimate</b>	<b>95% upper limit</b>
<b>Control</b>	76.2	85.6	95.0
<b>Weed</b>	80.8	90.3	99.9
<b>Black plastic</b>	98.6	104.8	110.9
<b>Silver plastic</b>	102.3	108.4	114.4

#### 4. Conclusions and Summary

In order to model disease progress curves, it is important to use modeling strategies which recognize the distribution of the dependent variable being modeled and the longitudinal nature of the data. Linearized models with linear regression (PROC REG) yield very different estimates, while non linear models with normality (PROC NLIN or GENMOD) need to take into account actual distribution or longitudinal nature. This can be obtained with binomial distribution and overdispersion (PROC GENMOD / dscale), marginal binomial models (PROC GENMOD / repeated), or non linear models with binomial distribution and “plot” random effects (PROC NLMIXED).

The marginal model permit “population average” inference, while the parameters in the mixed model have a “plot specific” interpretations (i.e., compare effects in the same plot). The nonlinear mixed model permits the estimation of any function of parameters, while there are no direct ways of doing this with the other approaches.

Currently there is no simple way of fitting in SAS a nonlinear mixed model including both a tree effect and a plot effect (nested random effects).

The models fitted for the papaya data set indicate clearly that there is a treatment effect: both plastic cover treatments delay the presence of symptoms with respect to the weed or bare ground treatments. Among both plastic treatment there are no significant differences, and hence the black plastic cover appears as the recommended treatment (because it is cheaper than silver plastic).

## References

- Agresti, A. (2002) *Categorical Data Analysis*. 2nd edition. New York: Wiley.
- Campbell, C. Lee and L.V. Madden (1990) *Introduction to Plant Disease Epidemiology*. Wiley: N. York
- Davidian, M. and D. Giltinan (1995) *Nonlinear Models for Repeated Measurement Data*. Boca Raton, FL: Chapman & Hall/CRC.
- Diggle, P., P. Heagerty, K. Liang and S. Zeger (2002). *Analysis of Longitudinal Data*. 2nd edition. London: Oxford University Press.
- Littell, R., G. Milliken, W. Stroup and R. Wolfinger (1996). *SAS System for Mixed Models*. SAS Institute Inc.: Cary, NC
- Macchiavelli, R., W. Robles, E. Abreu and A. Pantoja (2003). Nonlinear models with repeated measures for analyzing disease progress. Abstracts of the II Regional Meeting, International Biometric Society, Mayagüez, Puerto Rico.
- McCulloch, C. and S. Searle (2001) *Generalized, Linear, and Mixed Models*. New York: Wiley.
- SAS Institute (1999) *SAS OnlineDoc, Version 8*.
- Stokes, M., C. Davis and G. Koch (2000). *Categorical Data Analysis using SAS*. 2nd edition. Cary, NC: SAS Institute, Inc.

Appendix. SAS Code to fit the different models

```
*NON LINEAR FITS USING GENMOD, ASSUMING NORMALITY AND INDEPENDENCE;

Proc genmod data=medias;
class treat repet;
model y=treat dds treat*dds /dist=normal link=log type3 dscale;
title 'Exponential';
run;

Proc genmod data=medias;
class treat repet;
model y1=treat dds treat*dds /dist=normal link=log type3 dscale;
title 'Monomolecular';
run;

Proc genmod data=medias;
class treat repet;
model y=treat dds treat*dds /dist=normal link=logit type3 dscale;
title 'Logistic';
run;

Proc genmod data=medias;
class treat repet;
fwdlink link=log(-log(_mean_));
invlink ilink=exp(-exp(_xbeta_));
model y=treat dds treat*dds /dist=normal type3 dscale;
title 'Gompertz';
run;

*ANOVA MODEL (INCLUDING DDS AS A CLASS VARIABLE) TO TEST LACK OF FIT;
* remember that Deviance in this model is the same as Error SS in ANOVA;

Proc genmod data=medias;
class treat repet dds;
model y=treat dds treat*dds /dist=normal link=id type3 dscale;
title 'ANOVA';
run;
```

```

*Contrasts and estimations using Logistic best fitting curve;
Proc genmod data=medias;
class treat repet;
model y=treat dds treat*dds /dist=normal link=logit type3 dscale;
contrast 'PP=PC, C=W' treat 1 0 0 -1 , treat 0 1 -1 0,
          treat*dds 1 0 0 -1, treat*dds 0 1 -1 0;
estimate 'mean at 90 days, control'
intercept 1 treat 0 0 0 1 dds 90 treat*dds 0 0 0 90;
estimate 'mean at 90 days, weed'
intercept 1 treat 1 0 0 0 dds 90 treat*dds 90 0 0 0;
estimate 'mean at 90 days, pc'
intercept 1 treat 0 1 0 0 dds 90 treat*dds 0 90 0 0;
estimate 'mean at 90 days, pp'
intercept 1 treat 0 0 1 0 dds 90 treat*dds 0 0 90 0 ;
title 'Normal';
run;

* Logistic model, binomial distribution, overdispersion;
Proc genmod data=medias;
class treat;
model ndiseased/trees = treat dds treat*dds /dist=bin link=logit type3
dscale;
contrast 'PP=PC, C=W' treat 1 0 0 -1 , treat 0 1 -1 0,
          treat*dds 1 0 0 -1, treat*dds 0 1 -1 0;
estimate 'mean at 90 days, control'
intercept 1 treat 0 0 0 1 dds 90 treat*dds 0 0 0 90;
estimate 'mean at 90 days, weed'
intercept 1 treat 1 0 0 0 dds 90 treat*dds 90 0 0 0;
estimate 'mean at 90 days, pc'
intercept 1 treat 0 1 0 0 dds 90 treat*dds 0 90 0 0;
estimate 'mean at 90 days, pp'
intercept 1 treat 0 0 1 0 dds 90 treat*dds 0 0 90 0 ;
title 'Binomial w/overdispersion';
run;

*Logistic model, binomial data, marginal model;
Proc genmod data=medias;
class treat repet;
model ndiseased/trees = treat dds treat*dds /dist=bin link=logit type3
dscale;
contrast 'PP=PC, C=W' treat 1 0 0 -1 , treat 0 1 -1 0,
          treat*dds 1 0 0 -1, treat*dds 0 1 -1 0;
estimate 'mean at 90 days, control'
intercept 1 treat 0 0 0 1 dds 90 treat*dds 0 0 0 90;
estimate 'mean at 90 days, weed'
intercept 1 treat 1 0 0 0 dds 90 treat*dds 90 0 0 0;
estimate 'mean at 90 days, pc'
intercept 1 treat 0 1 0 0 dds 90 treat*dds 0 90 0 0;
estimate 'mean at 90 days, pp'
intercept 1 treat 0 0 1 0 dds 90 treat*dds 0 0 90 0 ;
repeated subject=repet / type=ar(1) corrw;
title 'Binomial w/repeated';
run;

```



Applied Statistics in Agriculture

```

*Logistic model, binomial data, normal random intercepts;

proc sort data=medias; by repet;

proc nlmixed data=medias;
parms b0=-5.4 bm= -0.5 bpc=-5 bpp=-6.5
      r=0.06 rm=0.002 rpc=0.036 rpp=0.046 sigma=0.3;
media =1/(1+exp(-b0-bm*tratm-bpc*tratpc-bpp*tratpp+u-r*dds-rm*tratm*dds-
rpc*tratpc*dds-rpp*tratpp*dds));

model ndiseased ~ binomial (20, media);
random u ~normal(0,sigma*sigma) subject=repet;

contrast 'PP=PC, C=W' bm , bpc-bpp , rm , rpc-rpp;

estimate 'mean at 90 days, control, typical plot'
        1/(1+exp(-b0-r*90));
estimate 'mean at 90 days, weed, typical plot'
        1/(1+exp(-b0-bm-r*90-rm*90));
estimate 'mean at 90 days, pc, typical plot'
        1/(1+exp(-b0-bpc-r*90-rpc*90));
estimate 'mean at 90 days, pp, typical plot'
        1/(1+exp(-b0-bpp-r*90-rpp*90));

estimate 'days to 50%, control' -b0/r;
estimate 'days to 50%, control' -(b0+bm)/(r+rm);
estimate 'days to 50%, control' -(b0+bpc)/(r+rpc);
estimate 'days to 50%, control' -(b0+bpp)/(r+rpp);

title 'Binomial with random effects';
run;

```