Use of JMP V.5 for Teaching and Research of Applied Stat/Math Methods in Life Sciences

Mohammed Al-Qinna, University of Arkansas, Fayetteville

Applied Mathematics Methods in the Life Sciences

Authors:
D. Scott
A. Mauromoustakos
M. Al-Qinna

Aim

- Need to characterize, visualize, analyze, and interpret measured data,
- Studying an arbitrary real-world phenomenon requires the use of different mathematical and statistical techniques, models and assumptions to translate the problem from its natural environment to a mathematical concept then to investigate and simplify such entity into a descriptive reality or result,
- Familiarize researchers with developments in modeling and data mining techniques using JMP 5.0 as a tool to visualize, understand and properly interpret data.

Content

- Data presentation,
- Data exploration,
- Data analyses,
- Spatial statistics,
- Modeling concepts,
- Models of growth and decay,
- Development of conservation and rate equations,
- Compartmental analysis and chemical kinetics,
- Numerical approximations of functions and equations.

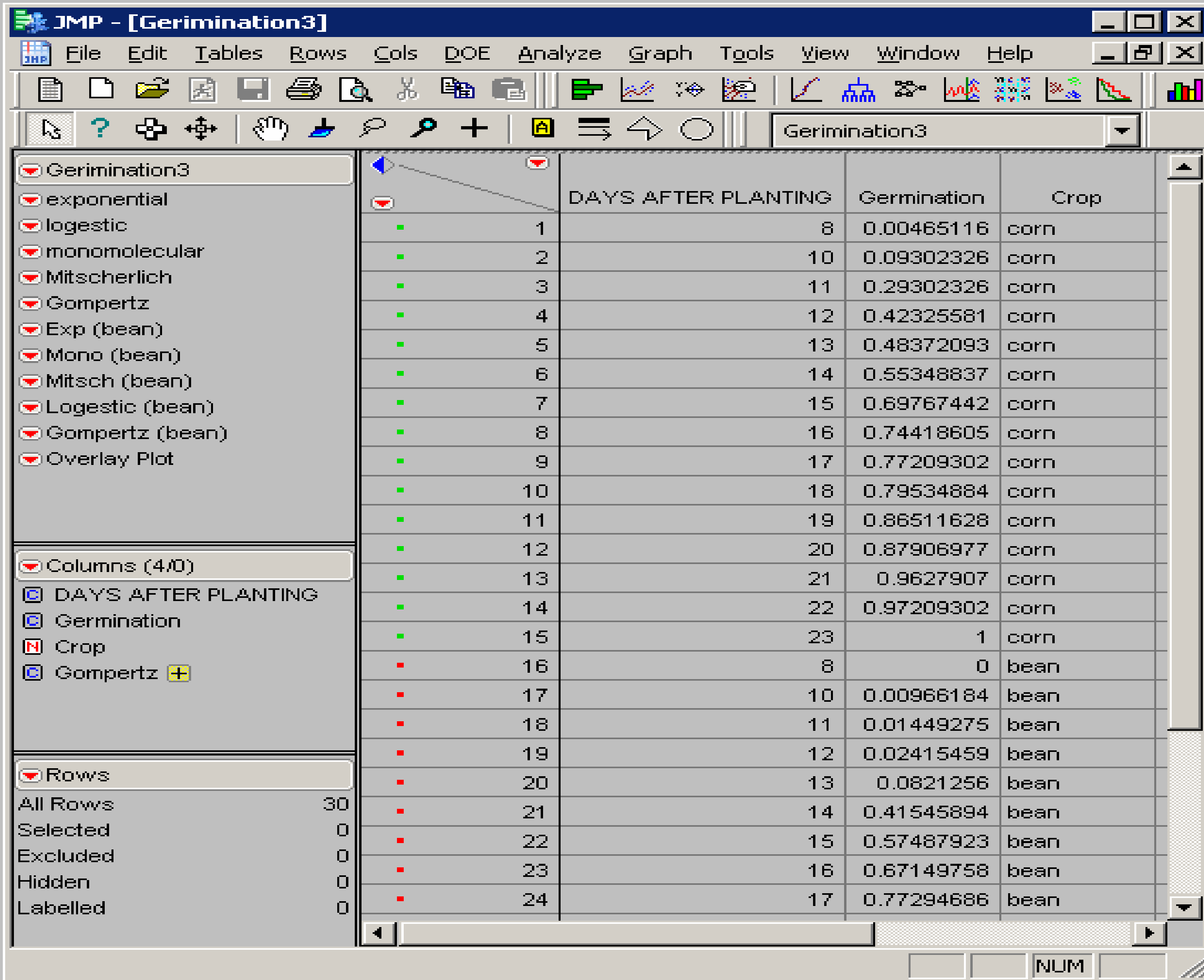
JMP is a business unit of **SAS,** the world leader in business intelligence software and services powered by market-leading analytics.

Why JMP ?

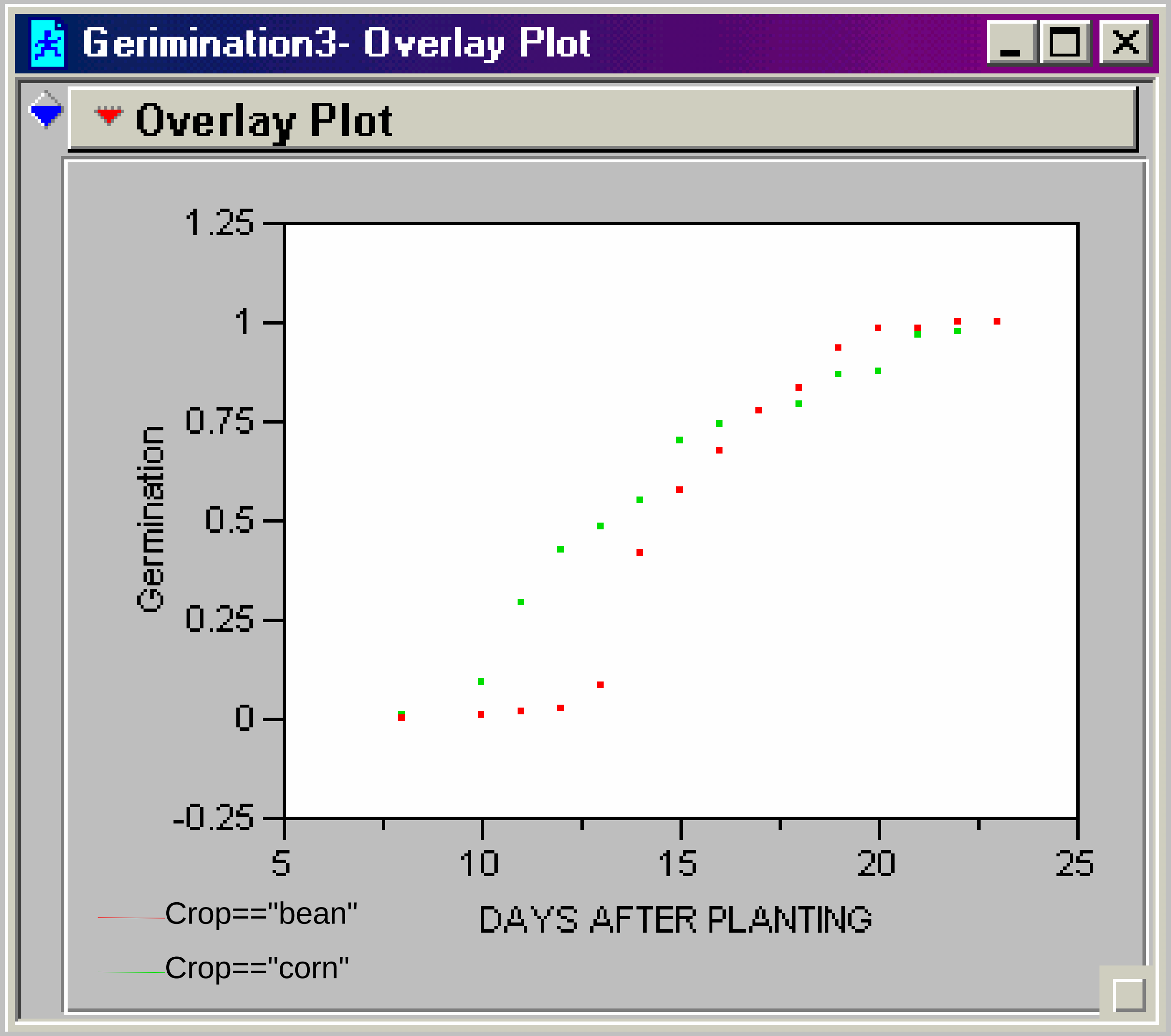
- JMP provides a comprehensive set of statistical tools as well as Design of Experiments and Statistical Quality Control in a single package.
- Empower researchers and students at all levels to explore and conduct data analysis closer to the actual process.
- Reduce time and risk of errors with fewer steps.
- Visual investigation and modeling of data.
- Designed for anyone who wants to discover relationships and outliers in their data.

Example

An experiment conducted at UARK to compare germination models for two plant species (corn and bean) over time.



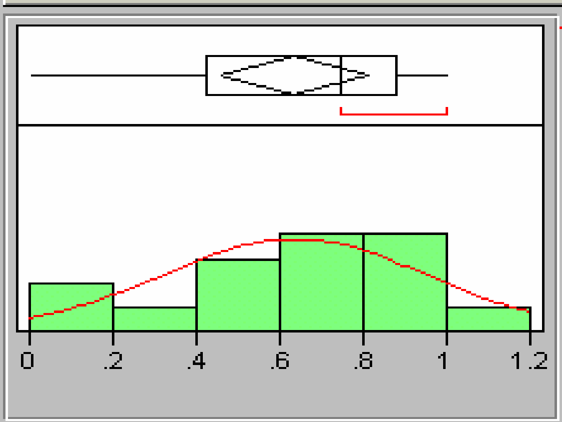
	DAYS AFTER PLANTING	Germination	Crop
1	8	0.00465116	corn
2	10	0.09302326	corn
3	11	0.29302326	corn
4	12	0.42325581	corn
5	13	0.48372093	corn
6	14	0.55348837	corn
7	15	0.69767442	corn
8	16	0.74418605	corn
9	17	0.77209302	corn
10	18	0.79534884	corn
11	19	0.86511628	corn
12	20	0.87906977	corn
13	21	0.9627907	corn
14	22	0.97209302	corn
15	23	1	corn
16	8	0	bean
17	10	0.00966184	bean
18	11	0.01449275	bean
19	12	0.02415459	bean
20	13	0.0821256	bean
21	14	0.41545894	bean
22	15	0.57487923	bean
23	16	0.67149758	bean
24	17	0.77294686	bean


Gerimination3- Overlay Plot
Overlay Plot
1.25
1
0.75
0.5
0.25
0
-0.25
Germination
5
10
15
20
25
DAYS AFTER PLANTING
Crop=="bean"
Crop=="corn"

Crop=corn

Distributions

Germination



Normal(0.63597,0.31731)

Quantiles

100.0%	maximum	1.0000
99.5%		1.0000
97.5%		1.0000
90.0%		0.9833
75.0%	quartile	0.8791
50.0%	median	0.7442
25.0%	quartile	0.4233
10.0%		0.0577
2.5%		0.0047
0.5%		0.0047
0.0%	minimum	0.0047

Moments

Mean	0.635969
Std Dev	0.3173057
Std Err Mean	0.081928
upper 95% Mean	0.811687
lower 95% Mean	0.460251
N	15
Sum Wgts	15
Sum	9.5395349
Variance	0.1006829
Skewness	-0.800066
Kurtosis	-0.394157
CV	49.893262

Fitted Normal

Parameter Estimates

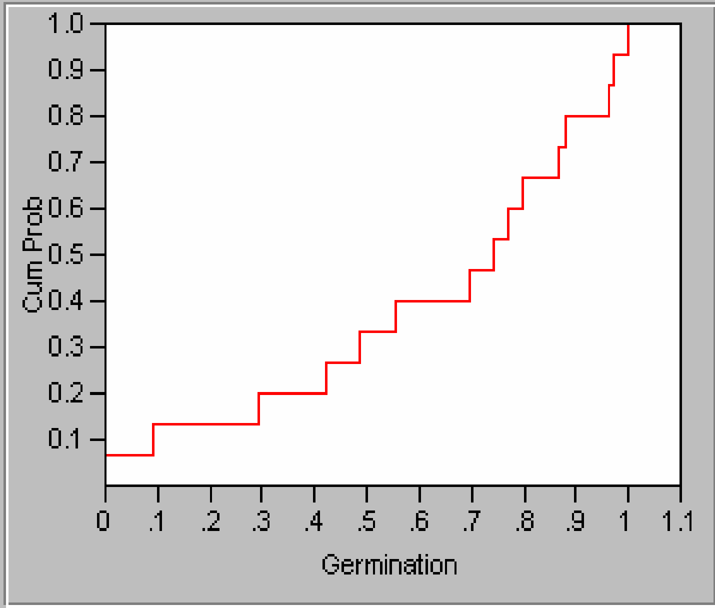
Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	Mu	0.6359690	0.4602510	0.8116870
Dispersion	Sigma	0.3173057	0.2323079	0.5004225

Goodness-of-Fit Test

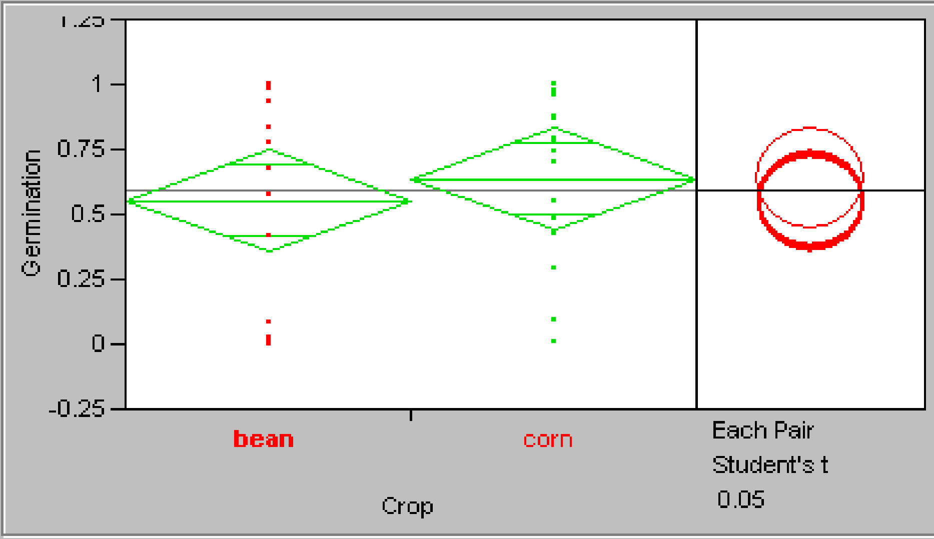
Shapiro-Wilk W Test

W	Prob<W
0.910688	0.1394

CDF Plot



Oneway Analysis of Germination By Crop



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Crop	1	0.0512548	0.051255	0.3700	0.5479
Error	28	3.8786901	0.138525		
C. Total	29	3.9299449			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
bean	15	0.553301	0.09610	0.35645	0.75015
corn	15	0.635969	0.09610	0.43912	0.83282

Std Error uses a pooled estimate of error variance

Means Comparisons

Dif=Mean[i]-Mean[j]

	corn	bean
corn	0.00000	0.08267
bean	-0.08267	0.00000

Alpha= 0.05

Comparisons for each pair using Student's t

t	Alpha
2.04841	0.05

Abs(Dif)-LSD

	corn	bean
corn	-0.27839	-0.19572
bean	-0.19572	-0.27839

Positive values show pairs of means that are significantly different.

Level		Mean
corn	A	0.63596899
bean	A	0.55330113

Levels not connected by same letter are significantly different

Oneway Anova

Summary of Fit

Rsquare	0.013042
Adj Rsquare	-0.02221
Root Mean Square Error	0.372189
Mean of Response	0.594635
Observations (or Sum Wgts)	30

t Test

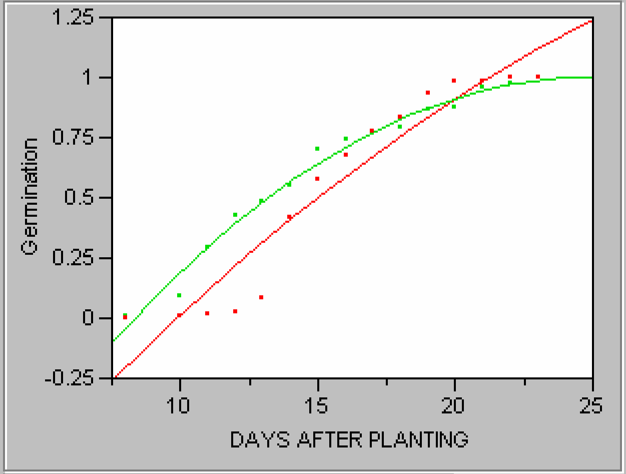
Assuming equal variances

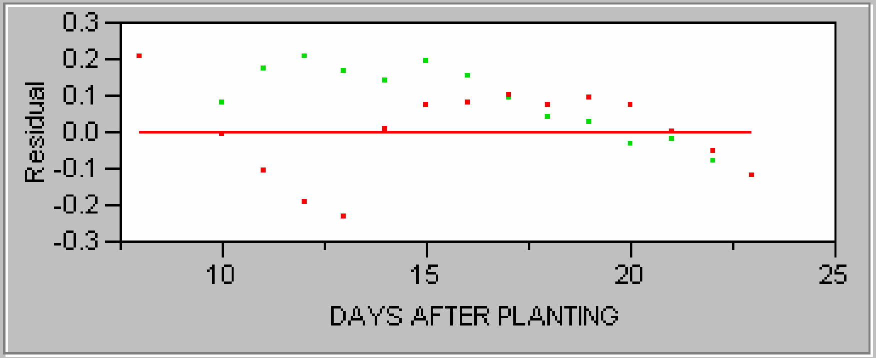
	Difference	t Test	DF	Prob > |t|
Estimate	-0.08267	-0.608	28	0.5479
Std Error	0.13590			
Lower 95%	-0.36106			
Upper 95%	0.19572			

UnEqual Variances

	Difference	t Test	DF	Prob > |t|
Estimate	-0.08267	-0.608	26.0556	0.5483
Std Error	0.13590			
Lower 95%	-0.36199			
Upper 95%	0.19666			

Bivariate Fit of Germination By DAYS AFTER PLANT





Polynomial Fit Degree=2 Crop=="bean"
Polynomial Fit Degree=2 Crop=="corn"

Polynomial Fit Degree=2 Crop=="bean"

Germination = -0.798536 + 0.0868023 DAYS AFTER PLANTING - 0.0015875 (DAYS AFTER PLANTING-15.9333)^2

Summary of Fit

RSquare	0.917596
RSquare Adj	0.903862
Root Mean Square Error	0.130213
Mean of Response	0.553301
Observations (or Sum Wgts)	15

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	2.2656645	1.13283	66.8124
Error	12	0.2034651	0.01696	Prob > F
C. Total	14	2.4691296		<.0001

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.798536	0.133699	-5.97	<.0001
DAYS AFTER PLANTING	0.0868023	0.007612	11.40	<.0001
(DAYS AFTER PLANTING-15.9333)^2	-0.001587	0.001811	-0.88	0.3979

Polynomial Fit Degree=2 Crop=="corn"

Germination = -0.338082 + 0.0656046 DAYS AFTER PLANTING - 0.0036236 (DAYS AFTER PLANTING-15.9333)^2

Summary of Fit

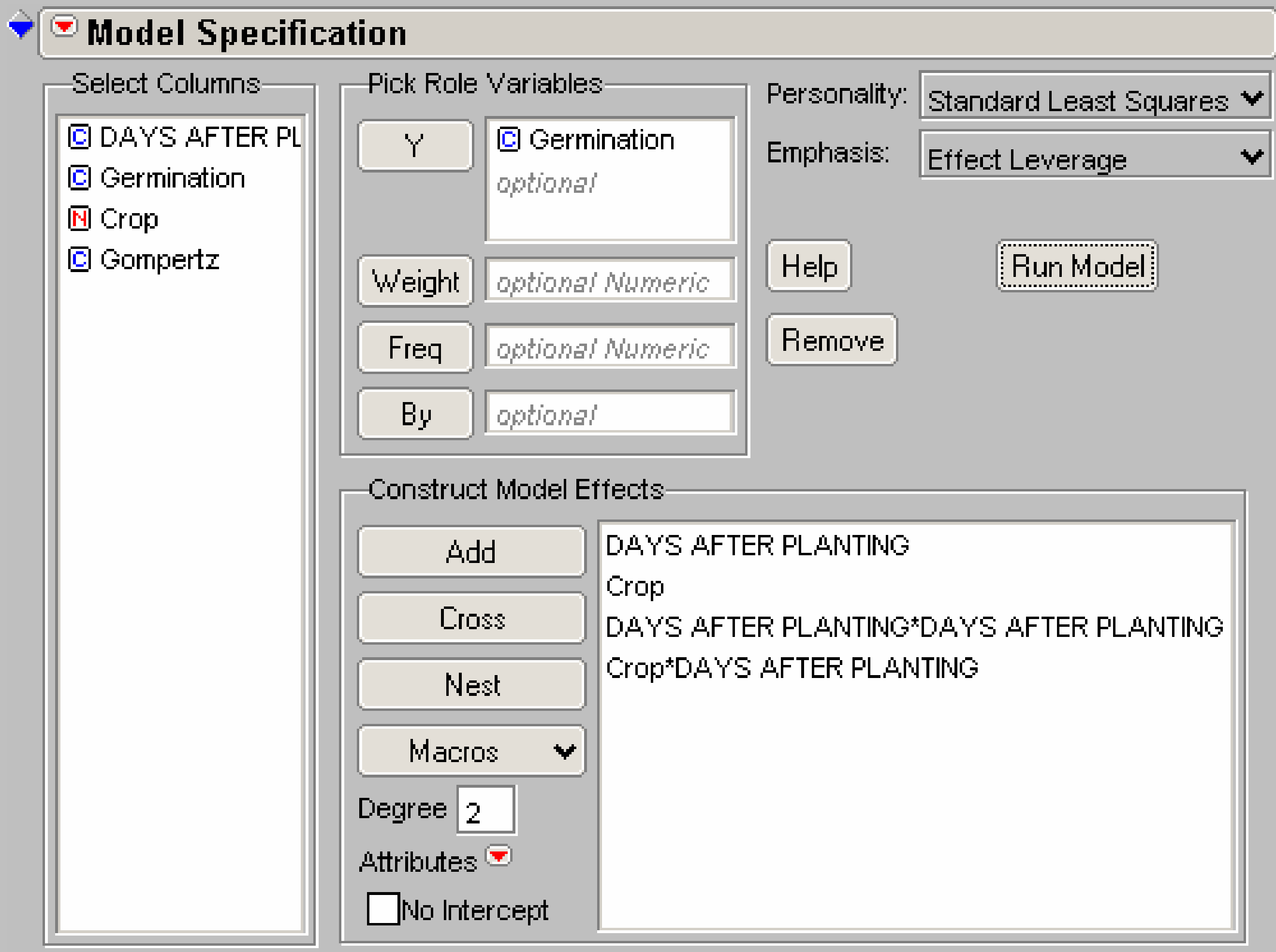
RSquare	0.986197
RSquare Adj	0.983896
Root Mean Square Error	0.040267
Mean of Response	0.635969
Observations (or Sum Wgts)	15

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	1.3901037	0.695052	428.6744
Error	12	0.0194568	0.001621	Prob > F
C. Total	14	1.4095605		<.0001

Parameter Estimates

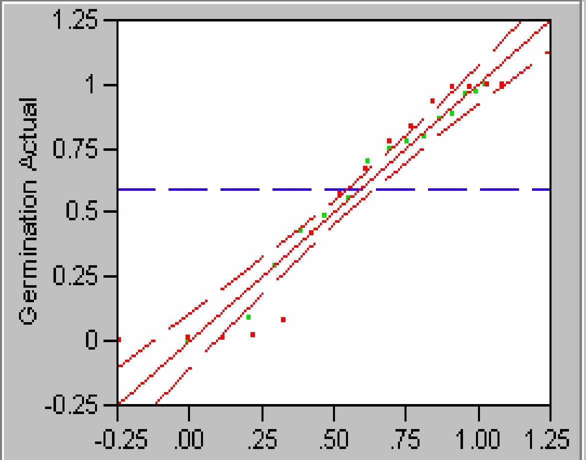
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.338082	0.041345	-8.18	<.0001
DAYS AFTER PLANTING	0.0656046	0.002354	27.87	<.0001
(DAYS AFTER PLANTING-15.9333)^2	-0.003624	0.00056	-6.47	<.0001


Model Specification
Select Columns
DAYS AFTER PL
Germination
Crop
Gompertz
Pick Role Variables
Y
Germination
optional
Weight
optional Numeric
Freq
optional Numeric
By
optional
Personality:
Standard Least Squares
Emphasis:
Effect Leverage
Help
Run Model
Remove
Construct Model Effects
Add
Cross
Nest
Macros
Degree
2
Attributes
No Intercept
DAYS AFTER PLANTING
Crop
DAYS AFTER PLANTING*DAYS AFTER PLANTING
Crop*DAYS AFTER PLANTING

Response Germination

Whole Model

Actual by Predicted Plot



Germination Predicted P<.0001 RSq=0.94
RMSE=0.0967

Summary of Fit

RSquare	0.940549
RSquare Adj	0.931037
Root Mean Square Error	0.096672
Mean of Response	0.594635
Observations (or Sum Wgts)	30

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	3.6963066	0.924077	98.8790
Error	25	0.2336383	0.009346	Prob > F
C. Total	29	3.9299449		<.0001

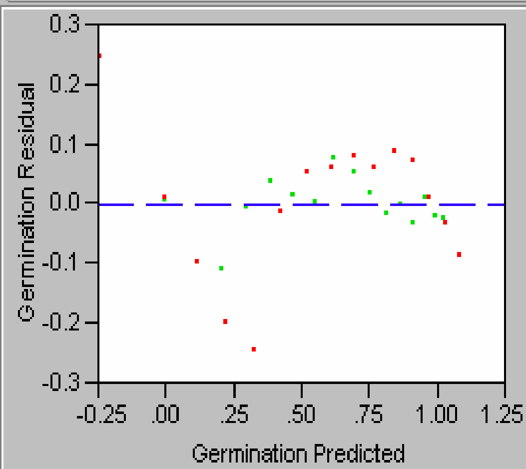
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.568309	0.070188	-8.10	<.0001
DAYS AFTER PLANTING	0.0762034	0.003996	19.07	<.0001
Crop[bean]	-0.041334	0.01765	-2.34	0.0275
(DAYS AFTER PLANTING-15.9333)*(DAYS AFTER PLANTING-15.9333)	-0.002606	0.000951	-2.74	0.0112
Crop[bean]*(DAYS AFTER PLANTING-15.9333)	0.0102191	0.00398	2.57	0.0166

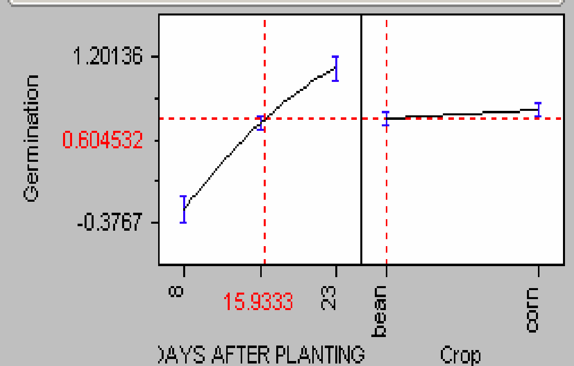
Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
DAYS AFTER PLANTING	1	1	3.3983595	363.6346	<.0001
Crop	1	1	0.0512548	5.4844	0.0275
DAYS AFTER PLANTING*DAYS AFTER PLANTING	1	1	0.0701924	7.5108	0.0112
Crop*DAYS AFTER PLANTING	1	1	0.0616002	6.5914	0.0166

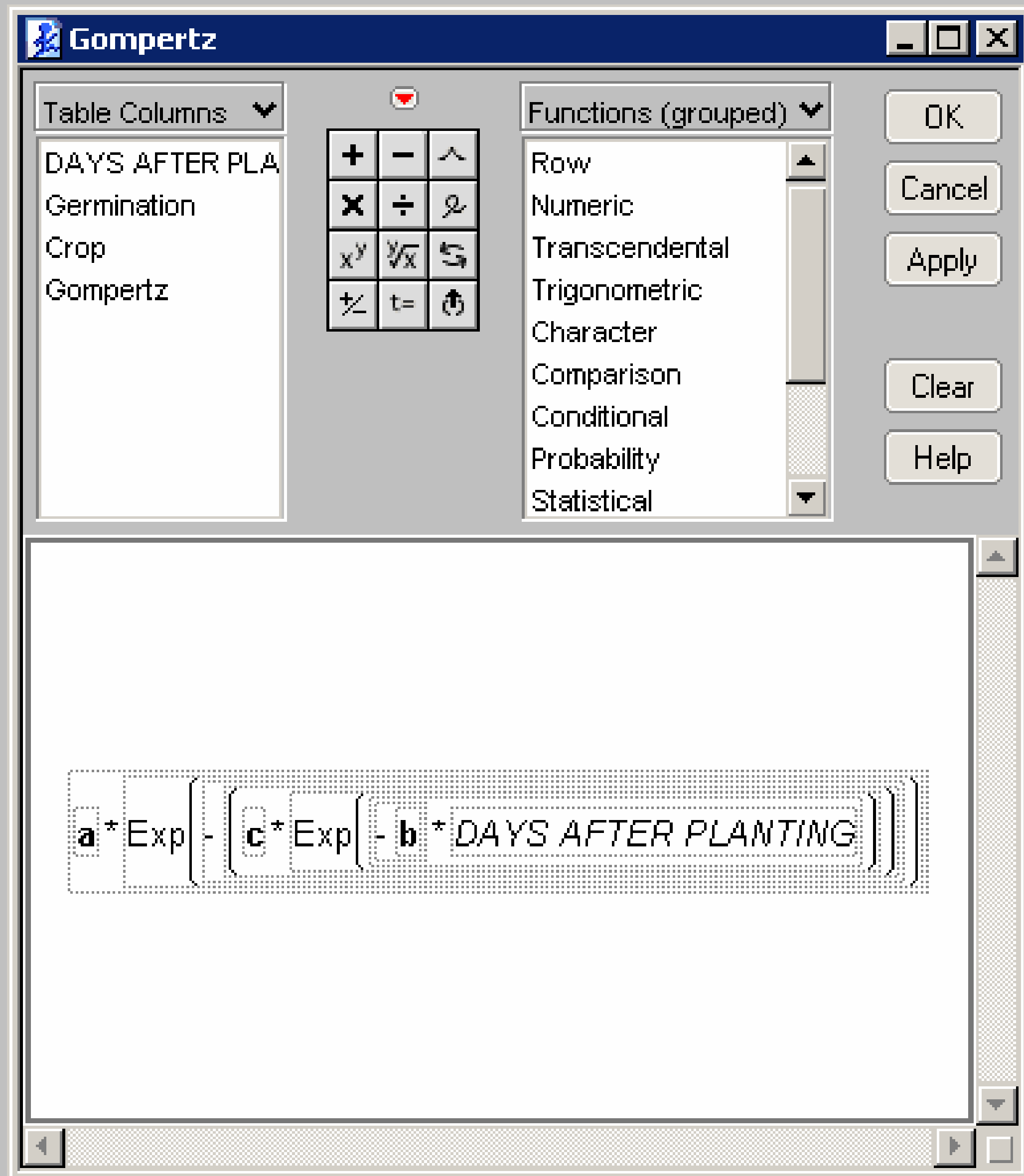
Residual by Predicted Plot

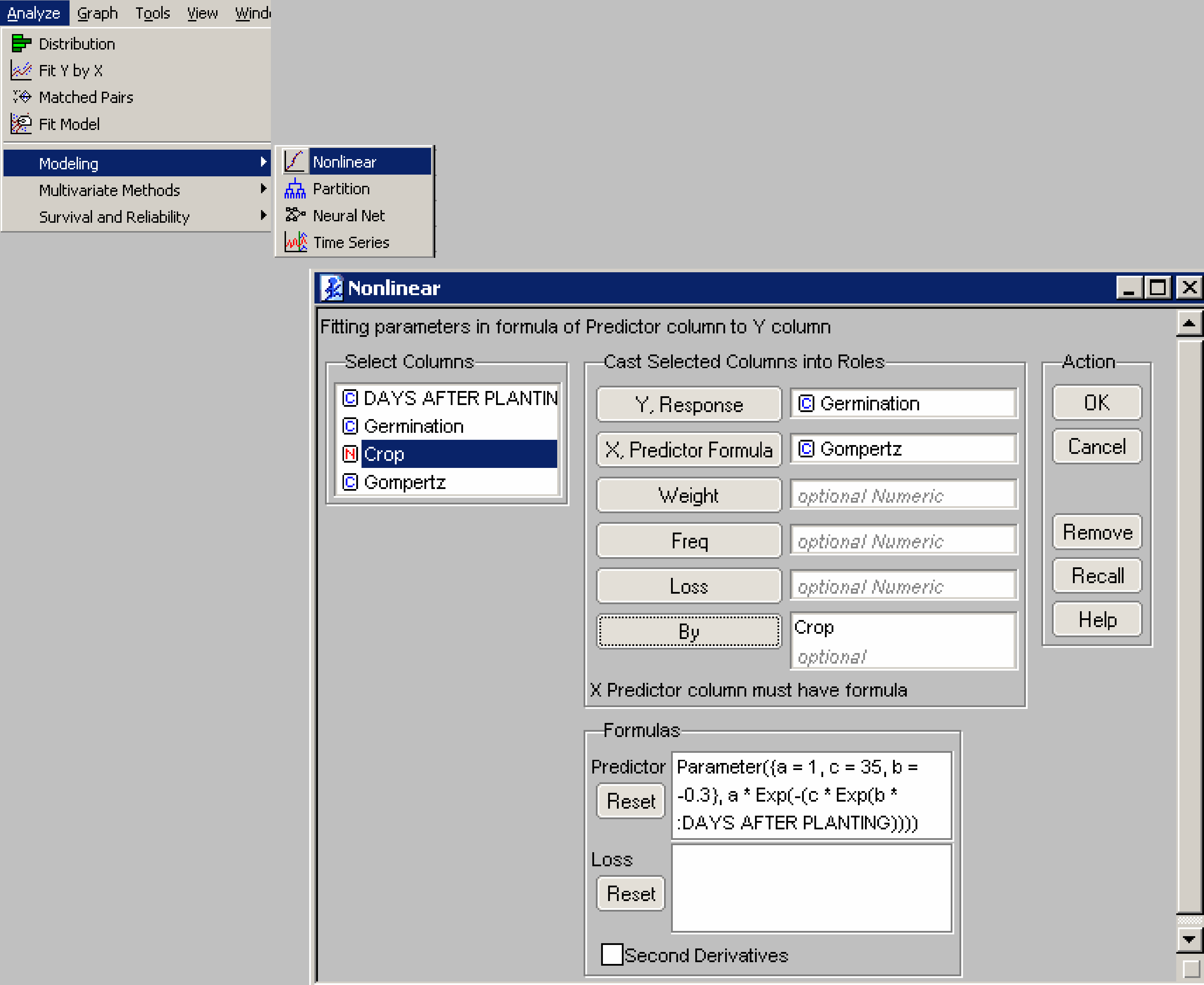


Prediction Profiler



- Exponential Growth: $Y = \alpha \times e^{\beta X}$
- Monomolecular: $Y = \alpha \times (1 - \gamma \times e^{-\beta X})$
- Mitscherlich: $Y = \alpha \times (1 - e^{-\beta (X + \delta)})$
- Logistic: $Y = \alpha / (1 + \gamma \times e^{-\beta X})$
- Gompertz: $Y = \alpha \times EXP(-\gamma \times EXP (-\beta X))$


Gompertz
Table Columns
DAYS AFTER PLA
Germination
Crop
Gompertz
Functions (grouped)
Row
Numeric
Transcendental
Trigonometric
Character
Comparison
Conditional
Probability
Statistical
OK
Cancel
Apply
Clear
Help
a * Exp(-(c * Exp(-b * DAYS AFTER PLANTING)))


Analyze
Graph
Tools
View
Windo
Distribution
Fit Y by X
Matched Pairs
Fit Model
Modeling
Multivariate Methods
Survival and Reliability
Nonlinear
Partition
Neural Net
Time Series
Nonlinear
Fitting parameters in formula of Predictor column to Y column
Select Columns
DAYS AFTER PLANTIN
Germination
Crop
Gompertz
Cast Selected Columns into Roles
Y, Response
Germination
X, Predictor Formula
Gompertz
Weight
optional Numeric
Freq
optional Numeric
Loss
optional Numeric
By
Crop
optional
X Predictor column must have formula
Action
OK
Cancel
Remove
Recall
Help
Formulas
Predictor
Reset
Parameter({a = 1, c = 35, b = -0.3}, a * Exp(-(c * Exp(b * :DAYS AFTER PLANTING))))
Loss
Reset
Second Derivatives

Crop=bean

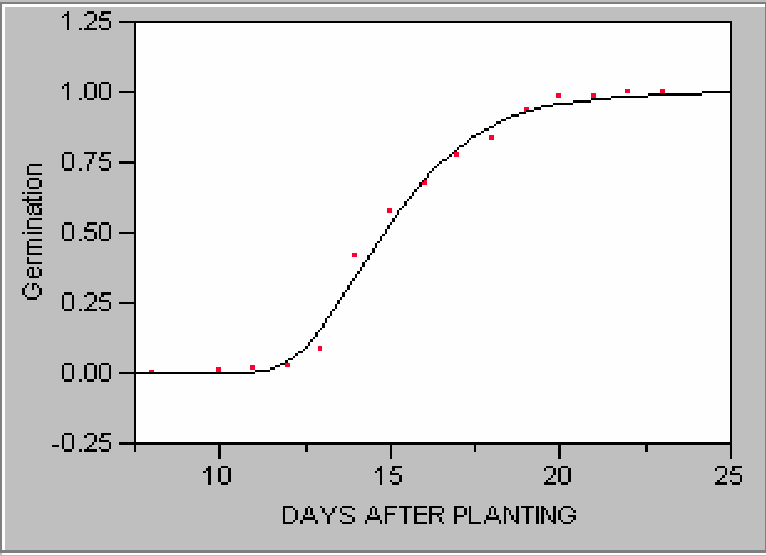
Nonlinear Fit

Control Panel

Criterion	Current	Stop Limit
Iteration	59	60
Shortening	0	15
Obj Change	0.0000010943	0.0000001
Prm Change	0.0027220276	0.0000001
Gradient	2.0482439e-8	0.000001

Edit Alpha	0.050
Convergence Criterion	0.00001
Goal SSE for CL	0.0265281135

Plot



Solution

SSE	DFE	MSE	RMSE
0.0190083644	12	0.001584	0.0397999

Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL
a	1.0021421732	0.02421239	0.95064683	1.06282905
c	1908.5174061	1446.78652	.	.
b	0.534408643	0.05368836	.	.

Crop=corn

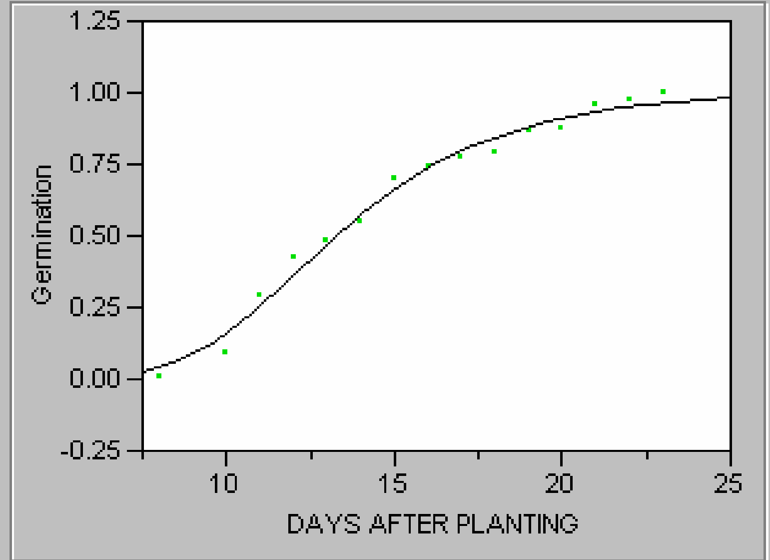
Nonlinear Fit

Control Panel

Criterion	Current	Stop Limit
Iteration	5	60
Shortening	0	15
Obj Change	0.0000187755	0.0000001
Prm Change	0.0003411888	0.0000001
Gradient	3.6294123e-7	0.000001

Edit Alpha	0.050
Convergence Criterion	0.00001
Goal SSE for CL	0.0267119687

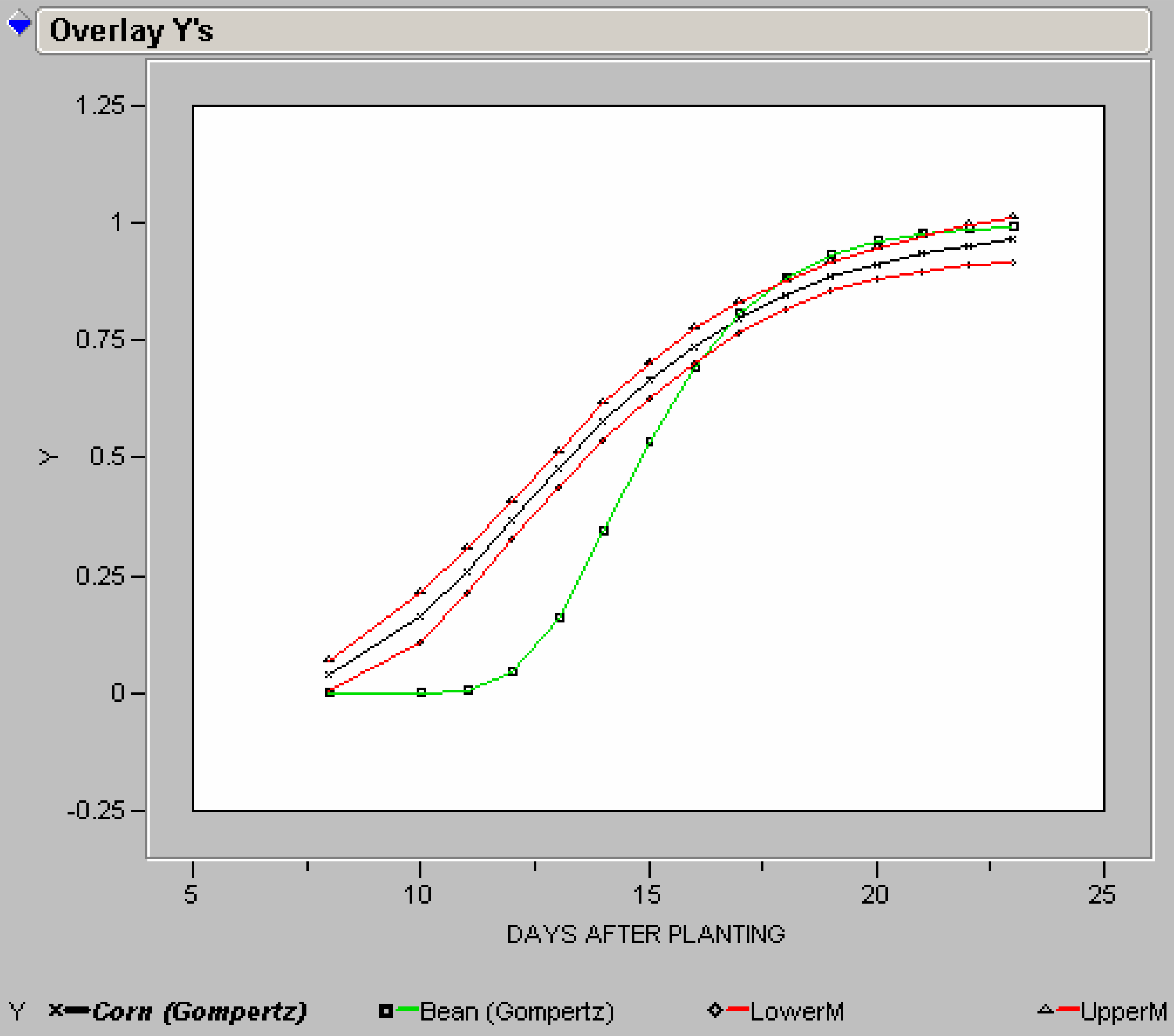
Plot



Solution

SSE	DFE	MSE	RMSE
0.0191401034	12	0.001595	0.0399376

Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL
a	1.0012386663	0.03426018	0.93305445	1.09555761
c	35.774463543	13.2104947	17.214676	88.4422536
b	0.2979902173	0.03204642	0.23224581	0.37700293



Overlay Y's
1.25
1
0.75
0.5
0.25
0
-0.25
Y
5
10
15
20
25
DAYS AFTER PLANTING
Y
Corn (Gompertz)
Bean (Gompertz)
LowerM
UpperM

Conclusions

- Germination behavior of the two crops are nonlinearly related to time after planting.

- Beans germination growth shows more delay time but faster growth rate compared to corn.

- The best nonlinear predictive model that best fit the germination is Gompertz;

  Bean germination = 1.00 × EXP (-1908.52 × EXP (-0.53 × day after planting))

  Bean germination = 1.00 × EXP (-35.77 × EXP (-0.30 × day after planting))

- JMP Just Made it easy.