

APM 630 Regression Analysis

Project #9 – Nonlinear Regression Models

Data: Tree.xls

The tree data contains 900 pairs of tree diameter at breast height (DBH, cm) and total height (HT, m) for western white pine. The **purpose** of the study is to develop the "best" model for predicting the tree HT using observed DBH in a long-term forest growth and yield simulator. Five nonlinear asymptotic growth functions are considered as candidate models:

[1] Gompertz Function:

$$HT = a * e^{-b * e^{-c * DBH}}$$

[2] Lundqvist Function:

$$HT = a * e^{(-b * DBH^{-c})}$$

[3] Modified Logistic Function:

$$HT = \frac{a}{1 + \left(\frac{1}{b}\right) * DBH^{-c}}$$

[4] Richards Function:

$$HT = a * \left[1 - e^{-b * DBH}\right]^c$$

[5] Weibull Function:

$$HT = a * \left[1 - e^{-b * DBH^c}\right]$$

where a, b, and c are regression coefficients to be estimated.

1. Draw a scatter plot of tree total height (HT) against diameter (DBH), and compute descriptive statistics.
2. Select **any one** of the five growth functions. Fit the function to the data using three iteration methods available in PROC NLIN (GAUSS, MARQUARDT, and NEWTON). Compare the parameter estimates, asymptotic standard errors for the parameters, and asymptotic correlations among the parameters for the three methods.
3. Fit the **five** candidate functions to the data using **any** iteration method. Which model(s) has the "best" statistical properties? Why?
4. Conduct simulations using the five models. The tree DBH should be ranged from 0 to 220 cm (by 10 cm). Compare your simulated "asymptote" with observed big tree record (**DBH = 208 cm and HT = 65 m**, according to the 1995 Idaho Champion Big Tree List). Which model(s) has the "best" prediction performance for large-sized trees?