Data analysis: responses at weekly time scales
Jeff White, Wed. 13:00

I. Introduction

A. From first 2.5 days of workshop

1. Can acquire data over days using field phenomics systems

2. Can georegister to obtain plot-level data

B. How does one make the most efficient use of data?

1. Maximize the information content

2. Minimize error

C. Our approach for time series derives from work on growth analysis 1950s to 1980s

1. Consider growth curve

2. Canopy height of spring wheat over time

a) Field 105, 2013.

b) Single plot

3. What might we analyze as the phenotype?

a) Maximum height

b) Height at day 100

c) Final height

d) Maximum elongation rate

II. What is the best estimate of growth: raw vs. fitted data?

A. Fitted curves:

1. Reduce sampling error

a) Must be well chosen

2. Once fitted, you can estimate any parameter:

a) Maximum height

b) Height at any given day (e.g., 100 DAP)

c) Final height

d) Maximum elongation rate

B. Rationale for fitting growth curves

"The rationale behind the use of the fitted functions is then simple: if attempts to assess the reality of growth result in a time series of observations scattered randomly about that reality, then a suitable mathematical function fitted to those observations may be expected to regain much of the clarity with which reality is perceived by the experimenter." R. Hunt. 1979. Ann Bot 43:245

III. Curve fitting basics

  A. Linear or non-linear regression

  B. Examples of curves used to fit time series data

    1. Quadratic

    2. Cubic

    3. Various exponential functions

    4. Richard's equation

  C. "Cottage industry" of papers describing "improved curves" (see references)

IV. Accounting for time vs. environmental effects

  A. We observe crops growing and developing over *time*

  B. But crops respond to multiple environmental factors

    1. Temperature, wind, radiation, humidity

    2. Water status

    3. Nitrogen status

    4. Etc.

V. Can we improve analysis by analyzing based on environment instead of time

  A. Growth = f(time)

  B. Growth = f(weather, management, soils, etc.)

    1. "Physiological time"

    2. Full ecophysiological modeling (see lecture)

  C. "Physiological time"

  D. Cumulative sum

    1. Temperature: Heat units, growing degree days, thermal units

    2. Photothermal time: Temperature + photoperiod

    3. Water stress days

  E. What is really going on?

    1. Summing over time = integrating a rate over time

    2. Simplest model

        (1) $dG/dt = R \times f(T)$

      b) where:

        (1) G = growth trait

(2) R = potential rate

(3) f(T) = temperature effect on rate

F. Simple model: a closer look

1. $dG/dt = R \times f(T)$

2. where:

a) G = growth trait

b) R = potential rate

c) f(T) = temperature effect on rate

3. $\int dG = \int R \times f(T)\, dt$

4. $G = R\int f(T)\, dt$

5. $G = R\Sigma f(T)\, \Delta t$ ← summation over time (e.g., days)

6. Conclusion: predictions based on sums of "physiological time" are actually applying simple rate-based (process) models

G. Example for wheat height

VI. Estimating parameters from curves

A. Phenotypic values at specific times

1. Plant height at anthesis

2. Plant height at X DAP even if samples taken on other dates (useful for very large trials)

B. Fundamental phenotypic traits

1. Maximum height

2. Maximum growth rate

3. Relative growth rate (1/R dR/dt)

C. Timing of important events

1. Time maximum value is first reached

2. Time of maximum growth or extension

VII. Introduction to exercise

A. Canopy height of spring wheat grown under two irrigation regimes: Canopy_ht_exercise_V1.0.xlsx

B. Using Excel, fit different equations for three irrigation levels:

1. DAP
2. GDD

C. Estimate:
   1. Canopy height at anthesis for each
   2. Maximum rate of canopy height increase
   3. Time of maximum rate
   4. Relative growth rate

VIII. Conclusion

A. Fitting growth curves
   1. Can reduce effect of sample error
   2. Allows inclusion of simple environmental effects
      a) Temperature as "physiological time"
      b) Other stresses
   3. Allows calculation of numerous parameters besides value at *X* days after planting

B. Appropriate curve may be difficult to identify

C. Opinion: Complicated curve-fitting approaches appear less promising than explicitly developing process-based model and applying inverse modeling techniques

References

Hunt R. 1979. Plant growth analysis: The rationale behind the use of the fitted mathematical function. Annals of Botany 43:245-249.

Wu R., Ma C.-X., Yang M.C., Chang M., Littell R.C., Santra U., Wu S.S., Yin T., Huang M., Wang M. 2003. Quantitative trait loci for growth trajectories in Populus. Genetical research 81:51-64.

Yin X., Goudriaan J., Lantinga E.A., Vos J., Spiertz H.J. 2003. A flexible sigmoid function of determinate growth. Annals of Botany 91:361-371.