Metadata for Describing Field Experiments

Jeffrey W. White, USDA ARS, ALARC

I. Introduction

A. Challenge:

1. Field phenomics research

a) Costly

b) Difficult to reproduce genetic materials

c) Need to maximize value

d) Often involve multiple research partners, who will only be familiar with certain aspects of a given experiment or nursery

B. Solution:

1. Ensure that metadata allow other researchers to fully interpret experiment

a) "Metadata" are "data about data"

(1) Management practices

(2) Weather conditions

(3) Soil conditions

(4) Problems

(5) Measurement protocols

2. Metadata are long been recognized as difficult

a) Crop simulation modeling

b) Meta-analysis

c) GIS

d) "Minimum dataset" of Henry Nix

3. ICASA standards created by modeling community

a) IBSNAT and SARP projects

b) DSSAT models

c) AgMIP model intercomparison project (www.agmip.org)

C. Outline

1. Review ICASA standards in relation to field phenomics

2. Detail look at metadata

a) Management

b) Weather

c) Soils

II. Crop simulation models as gold standard for describing an experiment

A. "Ecophysiological models"

B. Process-based

1. Light interception
2. Photosynthesis vs. respiration
3. Partitioning
4. Water & nutrient balances

C. Complete life cycle of annual crops

D. In use for:

1. crop management,
2. plant breeding,
3. *ex ante* impact assessment,
4. futures trading, etc.

E. Example 5200 maize NAM lines for QTL analysis of physiological traits

III. ICASA Standards

A. Goal: to provide a reliable and flexible structure for describing field situations

1. Field experiments
2. Any production situation: real or hypothetical
3. Not just for modeling

B. Basic data domains

1. Experimental background information
   a) Who? Where? Why?
2. Crop management
3. Environment
   a) Weather
   b) Soils
4. Phenotypic data
   a) Summary
      (1) Yield and related traits
      (2) Phenology
      (3) Growth summary
   b) Time series
      (1) Growth

(2) Pest damage

(3) [Phenomics]

IV. Specifying realistic conditions for dynamic simulations or other applications.

A. Origin: minimum dataset concept for

1. "agrotechnology transfer"
2. Henry Nix at 1983 ICRISAT workshop
3. ICASA: International Consortium for Agricultural Systems Analysis

a) Grew out of two major crop modeling groups

(1) Systems Analysis and Simulation for Rice Production (SARP)

(a) IRRI

(b) Wageningen, NL modeling community

(2) International Benchmark Sites Network for Agrotechnology Transfer

(a) USAID

(b) Univ. Florida, Michigan State Univ. & others

b) 1993 to 2012

c) No funding

d) Similar/related activities: SEAMLESS, DSSAT, AgMIP, GRACEnet

V. Goal of the ICASA standards

A. Allow as complete reproducibility of an experiment as possible through conventional means

B. Perfect reproducibility is impossible in field research

1. Weather
2. Soil conditions
3. Biotic factors

C. "Materials & Methods" descriptions are seldom adequate

1. In Journals
2. Multi-Location Breeding Trials

D. Minimum data standards: H Nix

VI. ICASA Standards: Architecture

A. Datasets divided into subsets

1. Experiments
2. Metadata

3. Management

B. Measurements ("phenotypes" sensu latu)

C. Weather data

1. Individual weather stations
2. Emphasis on daily weather
    a) Variables: Solar radiation, max and min air temperature, precipitation, etc.
3. Discussion ongoing about hourly or more frequent data
    a) Essential for interpreting IRT data (A French lecture on Monday)
    b) Valuable for reflectance data (K Thorp lecture on Monday)

D. Soil profile data

1. Individual soil profiles
    a) In papers or reports:
        (1) Phrases like "deep, sandy loam"
        (2) Taxonomy
    b) Difficult to relate to
2. Physical properties (quantitative data)
    a) Surface traits: albedo, runoff
    b) Profile (by depth in horizontal layers)
        (1) Water retention
        (2) Bulk density
        (3) Soil organic matter
        (4) Nutrient concentrations
    c)

VII. ICASA Standards: Implementations

A. Digital formats: agnostic about format

1. Plain text
2. Spreadsheet
3. Relational database
4. noSQL database

B. Architecture and vocabulary are what determine compliance!

C. Emphasis is on describing single experiments accurately.

D. Example: ICASA Standards as formatted plain text

1. Treatments are key!

FACTOR LEVELS link management sensu latu

E. AgMIP.org implementation for crop model intercomparisons

VIII. Detailed look at sections

A. Experiment metadata

B. Treatments linking to environment, management, phenotypes

C. Management

D. Weather

E. Soils

F. Phenotypic data

IX. ICASA V2.0 Standards: Strong points

A. 30+ years of experience

B. Balance necessary detail vs. approximations

C. Completeness in specifying environment

1. Initial soil conditions

2. Daily weather

X. Completeness in specifying management

A. Irrigation types

B. Fertilizers

C. Tillage – but likely won't capture problems like wheel pass compaction

D. Chemical applications

E. Organic matter, residue, mulches, etc.

F. 600+ variables for metadata and management

XI. Where work is needed

A. Software tools to make easier for non-specialists to use

B. Complex experimental designs

C. Perennial crops

D. Measurement errors – especially in metadata

E. Time series < 1 day

F. Very large experiments

1. Challenge for phenomics

G. Germplasm descriptions

H. Biotic interactions

1. Pests: disease & insect damage

I. Measurement protocols

1. Protocols for phenomics seem especially important

XII. Conclusions

A. Need to fully characterize experiments

1. Management
2. Weather
3. Soils
4. Protocols

B. ICASA standards are a promising start

C. Challenges

1. Lack of software
2. Sub-daily weather
3. Protocols

References

Hunt L.A., White J.W., Hoogenboom G. 2001. Agronomic data: advances in documentation and protocols for exchange and use. Agricultural Systems 70:477-492.

White J.W., Hunt L., Boote K.J., Jones J.W., Koo J., Kim S., Porter C.H., Wilkens P.W., Hoogenboom G. 2013. Integrated description of agricultural field experiments and production: The ICASA Version 2.0 data standards. Computers and Electronics in Agriculture 96:1-12.

Links

ICASA Master Variable List:
https://docs.google.com/spreadsheet/pub?key=0AiUhvGN3ZVgXdGJud3lDNThFSnlGWGlkTjgxWXRwbWc&output=html

AgMIP climate change project: www.agmip.org

Figure 1. Schematic of how data are grouped and linked in the ICASA version 2 standards for documenting experiments (White et al., 2013).