The AgMIP IT Systems: New Data Interoperability Tools

The AgMIP IT Team
presented by Cheryl H. Porter
November 4, 2013
Why do we need interoperability?

- Multiple inconsistent data sources
- Incomplete information to parameterize multiple crop models
- Multiple input and output formats for crop models
1. Define a common vocabulary for crop modeling variables
2. Design a harmonized data format
3. Develop data translation libraries
   a) ACE input
   b) ACE output
   c) ACMO
4. Uniformly apply model assumptions
5. Serve harmonized data in an online, searchable database
6. Maintain data provenance
7. Develop the harmonized data system collectively with input from the global modeling community.
• ICASA Master Variable List
  – Flexible, extensible ontology
  – Comprehensive documentation of field experiments
  –Modified for use by AgMIP
    • Added metadata
    • New variables required for participating models
    • Clarifications of definitions

\(^1\) White et al., 2013; Hunt et al., 2001
ACE Harmonized data format

- Key-value pairs, (where key = ICASA variable) stored in JSON objects
- Non-relational schema conforms to highly irregular data from diverse sources
- Efficient, flexible, open ended storage
- “buckets” for experiment / field, soils, and weather data
- Sub-structures for initial conditions, management events, soil layer data, etc.
- Variables conform to ICASA master variable list
- Metadata describe simulation, scenarios, institutional information (22 variables)
- Experimental data summary (23 variables)
- Harmonized outputs from crop models (10 variables)

(see research.agmip.org for full list of ACMO variables)
• ACE input, ACE output and ACMO translators
• 5 multi-model development sprints so far
• 13 models now have translators
  • APSIM
  • CropGrow-NAU
  • DSSAT
  • STICS
  • WOFOST
• All AgMIP software projects are open source

Data translation libraries

• APSIM
• CropGrow-NAU
• DSSAT
• STICS
• WOFOST
• AquaCrop
• CropSyst
• EPIC
• InfoCrop
• ORYZA2000
• RZWQM2
• SALUS
• SarraH
Uniformly apply model assumptions

The DOME: Data Overlay for Multi-model Export

- Fill missing data prior to data translation so that all models use the same assumptions
  - Start of simulation dates
  - Planting densities
  - Initial soil conditions
  - Fertilizer & irrigation details
- Overlay a management regimen, for applying RAPs to future climate scenarios
  - Fertilizer rates & schedules
  - Cultivars
  - Automatic planting rules
Online, searchable database
• Desktop utilities – tools.agmip.org
  – Used by AgMIP Regional Research Teams for Regional Integrated Assessments
  – QuadUI - conversion of raw input data from multiple formats and translation to model-ready formats for multiple models
  – ACMOUI - conversion of model outputs to ACMO harmonized format.

• Gridded modeling applications for multiple models
  – pSIMS – running multiple models on global gridded modeling framework, University of Chicago
  – Plans for implementing translators in IFPRI gridded modeling platforms
  – CCAFS Regional Agricultural Forecast Tool (CRAFT) application

• CCAFS – AgTrials and AgMIP data sharing and exposure

• Next generation applications
  – Desktop application for online search, upload, download, with local translation, DOME, ACMO support
  – Web-based interactive user interface for creating, sharing and using complex workflows
  – Interfaces to Linked Open Data
Data flow diagram for Regional Integrated Assessments