Data Harmonization for Crop Simulation: the ICASA/AgMIP Approach

Jeffrey W. White,
USDA ARS, ALARC
Cheryl Porter,
University of Florida





EXPERIMENTAL PROCEDURE

To determine the influence of leaf area on grain yield and to study the contribution of individual leaves to seed yield, various leaf removal treatments were inflicted in a planting of Midland grain sorghum at the late boot stage in 1958 and 1959. The treatments were: check (no defoliation), complete defoliation, and alternate (every other leaf removed) defoliation. In addition, certain individual leaves were left on otherwise completely defoliated plants. Leaves numbered 1 (flag leaf), 3, 5, 7, and 9 were left on each plant in these treatments. The resulting 8 treatments were established in a randomized complete block experiment with 3 replications in 1958 and 4 in 1959. Each plot consisted of three 18-foot rows with the inside 15 feet of the center row harvested for yield. A uniform spacing of 6 inches between plants in 40-inch rows was used. Heads were cut, air dried, threshed, and subsampled for weight per 200 seeds. The number of seeds per head was inferred from seed weight and mean head weight data for each plot.

A second experiment involved removing approximately half the leaf area from grain sorghum plants at boot and anthesis stages

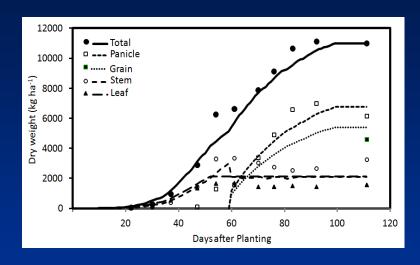


Crop simulation models

- "Ecophysiological models"
- Process-based
 - Light interception
 - Photosynthesis vs. respiration
 - Partitioning
 - Water & nutrient balances

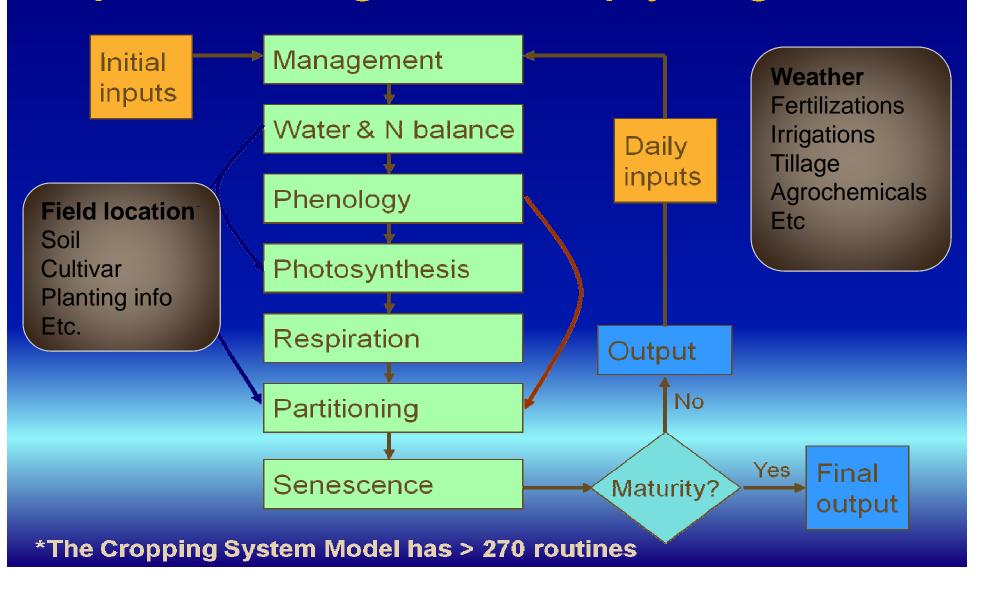


- 50 years old: C.T. de Wit, R.S. Loomis
- Used for: crop management, plant breeding, ex ante impact assessment, futures trading, etc.
- Examples: CERES, EPIC, APSIM, ecosyst, CSM, etc.



Crop modeling requires diverse, accurate & abundant data

Simplified* flow diagram of an ecophysiological model



ICASA Standards

ICASA: Int. Consortium for Agricultural Systems Applications

Goal: Standards provide a reliable and flexible structure for ...

- Documenting field experiments
 - Environment: soils & weather
 - Crop management
 - Phenotypic responses
- Specifying realistic conditions for dyna

other applications.

Origin: minimum dataset "agrotechnology transfe

- Henry Nix
- 1983 ICRISAT workshor

PN-A+12-332 29

Proceedings of the

International Symposium on Minimum Data Sets for Agrotechnology Transfer

> ICRISAT Center Patancheru, India 21-26 March 1983

Minimum Data Sets for Agrotechnology Transfer

Henry A. Nix*

Abstract

Prescription of appropriate technologies at farm level, at the transfer site, demands a systems understanding of the land, labor, capital, and management resources of the individual farmer. If it were possible to predict the performance of any crop production system at any location given a specified minimum set of crop-site-weather-management data, then it would be possible to prescribe appropriate technologies at the level of the farmer and his farm. It is argued that this is an ultimate and attainable objective of agricultural research, but only if there is a shift in emphasis away from reductionist and analytical research to holistic and systems-based research.

A systems-based research strategy centers on balanced development of two interactive



- Incorporate state-of-the-art climate products as well as crop and agricultural trade model improvements in coordinated regional and global assessments of future climate impacts
- Include multiple models, scenarios, locations, crops and participants to explore uncertainty and impact of data and methodological choices
- Collaborate with regional experts in agronomy, economics, and climate to build strong basis for applied simulations addressing key climate-related questions
- Improve scientific and adaptive capacity for major agricultural regions in the developing and developed world
- Develop framework to identify and prioritize adaptation strategies
- Link to key on-going efforts with numerous partners globally,

AgMIP data interoperability

- AgMIP ensemble modeling activities required a flexible, extensible means of providing data to multiple models.
- Facilitate research collaboration, model intercomparison and model improvement
- Make data available under the Creative Commons Attribution license





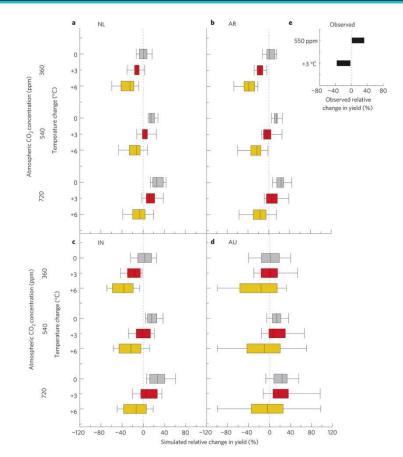
Wheat Pilot: 27 Simulation Models

Asseng et al., 2013
Uncertainty in simulating wheat yields
under climate change
Nature Climate Change

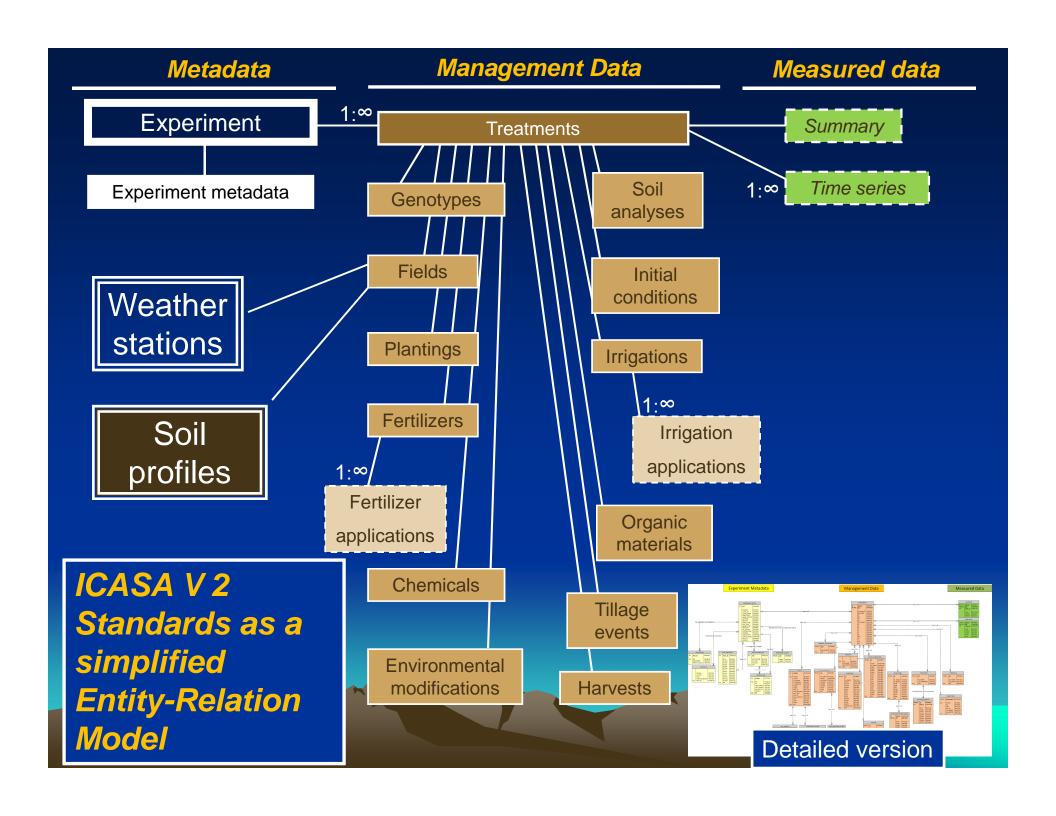
"This study . . . breaks new ground in discriminating between the uncertainties ascribed to crop models and those due to ensemble climate model projections.

Timothy R. Carter
Nature Climate Change News
and Views August 28, 2013

Responses to high temperature and high CO2 still remain significant uncertainties



a–d, Simulated relative mean (30-year average, 1981–2010) grain yield change for increased temperatures (no change, grey; $+3\,^\circ$ C, red; $+6\,^\circ$ C, yellow) and elevated atmospheric CO2 concentrations for the Netherlands (NL; **a**), Argentina (AR; **b**), India (IN; **c**) and Australia (AU; **d**). For each box plot, vertical lines represent, from left to right, the 10th percentile, 25th percentile, median, 75th percentile and 90th percentile of simulations based on multi-models. **e**, Observed range of yield impacts with elevated CO2 (refs23, 24). Observed range of yield impacts with increased temperature10, 24 (extrapolated, based on separate experiments with 40–345 ppm elevated CO2 and 1.4–4.0 $^\circ$ C temperature increase, Supplementary Information).



"Treatments" section is the key! FACTOR LEVELS section links management & field conditions

```
$MANAGEMENT
*TREATMENTS
                                                           --FACTOR LEVELS----
@TRTNO RP SQ OP CO TRT NAME.....
              1 1 0N, dryland+weeds Repl
          1 1 1 0N, dryland+weeds Rep2
                  2 ON, dryland+weeds
          1 1 1 180N, dryland
        0 1 1 1 0N,irrigated
       0 1 1 1 60N, irrigated
       0 1 1 1 180N, irrigated
       0 1 1 1 180N, irrig(d2p)
       0 1 1 1 180N, irrig(automatic)
                  1 RedClover Post-harvest
            2 2 1 RedClover2 Postharvest
*GENOTYPES
@ GE CRID CUL ID CUL NAME.... CUL NOTES
   1 WHB IB0488 Newton Hard red winter, Central Plains
2 GRW IB0001 Garytown Central Gt. Plains quackgrass
3 RCL IB0001 Bigyield General double cut
4 RCL IB0002 Earlywonder Single cut
*FIELDS
        ID FIELD.. FL NAME
@ FL
        KSTASH0001 Research Park F105
       FL LOC 1
                      FL LOC 2 FL LOC 3
@ FL
        USA
                                 Research Park, Ashland
@ FL
        FL LAT
                      FL LONG
                                  FLELE FLSL FLSLL FLSLA FAREA FLOB FLLWR
           <del>-</del>37 1167
                      -90 4521
                                                 200
                                                               226
```

ICASA Standards: Structured plain text

```
$EXPERIMENT: KANKSTS200401WHB
   $EXPERIM
     EXPERI
               $MANAGEMENT
   @ EXNAME
               *TREATMENTS
                                                           -----FACTOR LEVELS-----
     Fictit
               @TRTNO R
   @ LOCAL
     WHEX20
                   1
                           *IRRIGATIONS
   @ OBJECT
                           @ IR
     To det
                             1
                                        *CHEMICALS
                    3
                             2
   @ MAIN E
                                        @ CH
                                               CH NAME
                             3
     FE(N)
                                               Pre-plant weed control
                           @ IR
*TIME COURSE
                       YEAR
                                                                               CWAD
                                                                                      PWAD
                                                                                             SHAD
                                                                                                    SH%D
@TRTNO RP
           SQ OP
                               DOY
                                     L#SD
                                            LAID
                                                   P#AD
                                                          SWAD
                                                                 GWAD
                                                                        LWAD
                                                                                                           SLAD
                  CO
                       1986
                               289
                                            1.12
                                                    0.0 104.0
                                                                 0.00 250.0 354.0
                                                                                      0.00
                                                                                             0.00
                                                                                                          446.8
                       1986
                               289
                                      5.1
                                            0.83
                                                    0.0 98.00
                                                                 0.00 242.0 340.0
                                                                                      0.00
                                                                                             0.00
                                                                                                          342.4
                       1986
                               289
                                            0.58
                                                                                                          379.2
                                      5.0
                                                    0.0 64.00
                                                                 0.00 152.0 216.0
                                                                                      0.00
                                                                                             0.00
                               289
                       1986
                                            0.69
                                                        78.00
                                                                      204.0 282.0
                                                                                      0.00
                                                                                                          337.2
                                                                                             0.00
                       1986
                               296
                                            2.21
                                                                               1192
                                                                                                          279.3
                                                    0.0 402.0
                                                                 0.00 790.0
                                                                                      0.00
                                                                                             0.00
                       1986
                               296
                                            1.62
                                                                                                          292.0
                                                    0.0 304.0
                                                                 0.00 556.0 860.0
                                                                                      0.00
                                                                                             0.00
                       1986
                               296
                                            1.43
                                                    0.0 202.0
                                                                 0.00 444.0 646.0
                                                                                      0.00
                                                                                             0.00
                                                                                                        . 321.0
                               296
                                            1.35
         0
                       1986
                                                    0.0 154.0
                                                                 0.00 438.0 592.0
                                                                                      0.00
                                                                                             0.00
                                                                                                          308.2
         0
                       1986
                               305
                                            5.38
                                                          1086
                                                                 0.00
                                                                        1214
                                                                               2300
                                                                                      0.00
                                                                                             0.00
                                                                                                          442.9
                                                    0.0
                       1986
                               305
                                            4.46
                                                          1004
                                                                 0.00
                                                                        1128
                                                                               2132
                                                                                      0.00
                                                                                             0.00
                                                                                                         395.2
                       1986
                               305
                                            4.52
                                                                        1062
                                                                               1888
                                                                                                         425.3
                                                    0.0 826.0
                                                                 0.00
                                                                                      0.00
                                                                                             0.00
                                            4.91
                       1986
                               305
                                                    0.0 988.0
                                                                 0.00
                                                                        1182
                                                                               2170
                                                                                      0.00
                                                                                             0.00
                                                                                                          415.5
                       1986
                               311
                                                   68.0
                                                          1312
                                                                        1296
                                                                               2612
                                                                                             4.00
         0
                                            5.67
                                                                 0.00
                                                                                      4.00
                                                                                                    0.00 437.8
                                                                               3318
                       1986
                               311
                                                  122.0
                                                          1580
                                                                        1728
         0
                                            4.55
                                                                 0.00
                                                                                    10.00
                                                                                           10.00
                                                                                                    0.00 263.3
                                                   36.0
                       1986
                               311
                                            1.42
                                                          1016
                                                                 0.00
                                                                        1008
                                                                               2026
                                                                                      2.00
                                                                                             2.00
                                                                                                    0.00 140.4
                                                                               2856
                       1986
                               311
                                            3.93
                                                   60.0
                                                          1286
                                                                 0.00
                                                                        1566
                                                                                      4.00
                                                                                             4.00
                                                                                                    0.00 250.9
                       1986
                               316
                                                 488.0
                                                          1888
                                                                 0.00
                                                                        1658
                                                                               3792
                                                                                     246.0
                                                                                            246.0
                                                                                                    0.00 437.3
                       1986
                               316
                                                 492.0
                                                          1962
                                                                 0.00
                                                                        1678
                                                                               3894 254.0 254.0
                                                                                                    0.00 378.7
```

ICASA Standards: Implementations

Digital formats: AGNOSTIC

- Plain text (specific syntax)
- Spreadsheet
- Relational database
- noSQL database using JSON objects

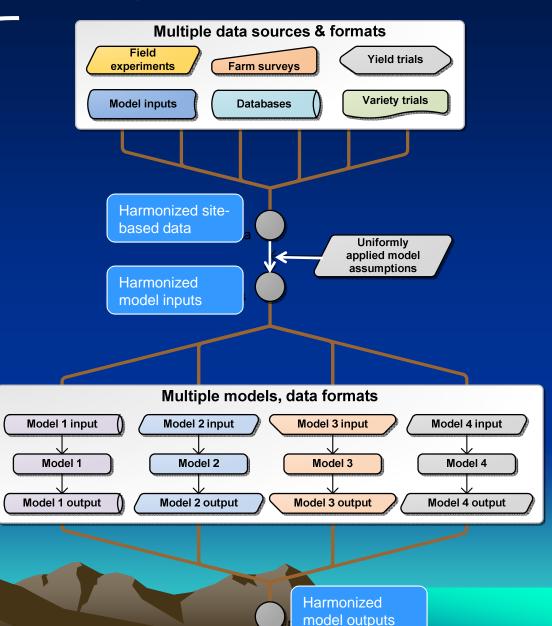
Architecture and vocabulary are what determine compliance!

Emphasis is on describing single experiments accurately.

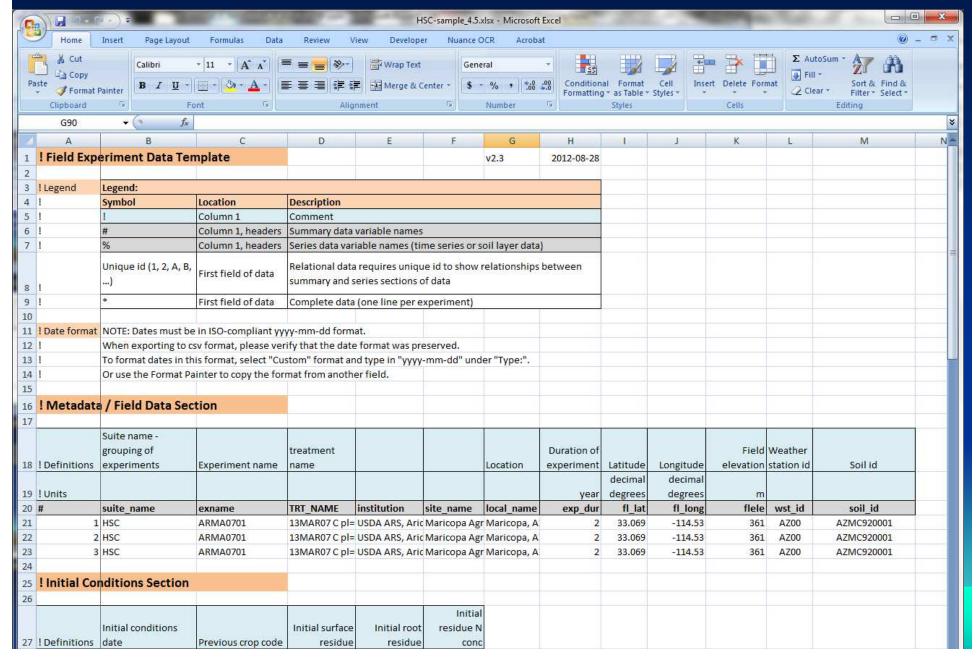


The ICASA Standards provide the foundation for harmonizing data from different sources

AgMIP and ICASA



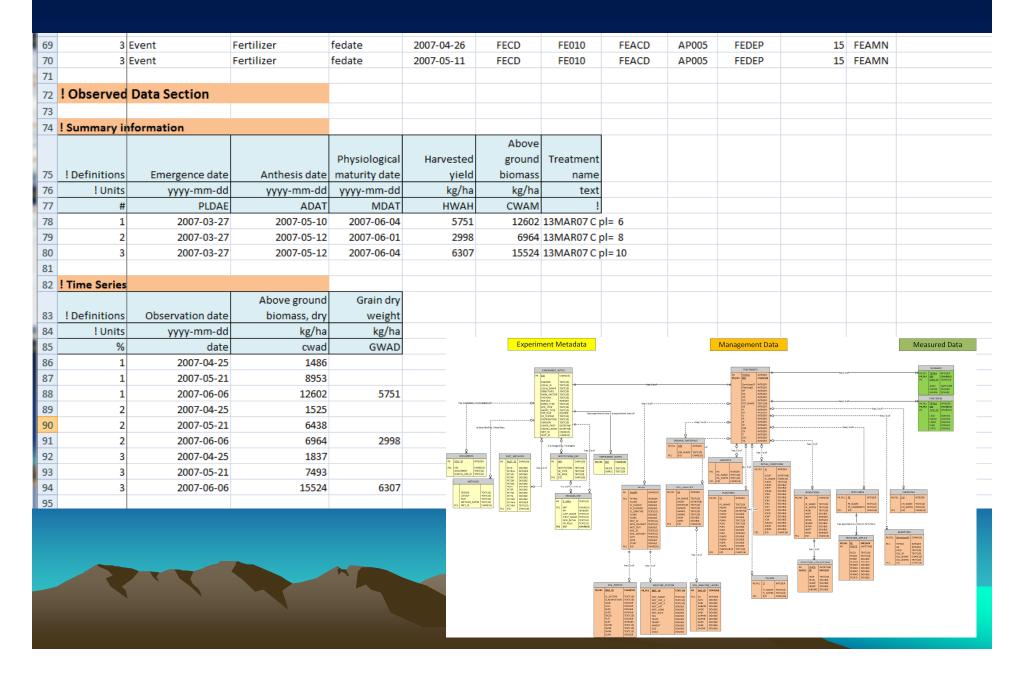
ICASA/AgMIP template

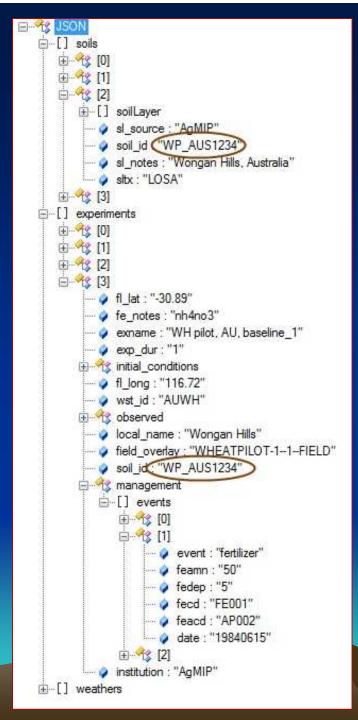


ICASA/AgMIP template

4	Α	В	С	D	Е	F	G	Н	1	J	K	L	M	
37	! Managen	nent Events Sec	ction											
38	_													
										Additional		Plant		
										data about		populati		
39	! Definitions	Type of event		Date of event	yyyy-mm-dd	Crop ID	Wheat	Cultivar ID	text	cultivar	text	on	#/n	n2
10	1	Event	Planting	pdate	2007-03-13	CRID	WHB	CUL_ID	Yecora Ro	CUL_NOTES	Spring wheat	PLPOP	2	65 PLR
1	2	Event	Planting	pdate	2007-03-13	CRID	WHB	CUL_ID	Yecora Ro	CUL_NOTES	Spring wheat	PLPOP	3:	20 PLR
12	3	Event	Planting	pdate	2007-03-13	CRID	WHB	CUL_ID	Yecora Ro	CUL_NOTES	Spring wheat	PLPOP	2	76 PLR
13														
						Irrigation		Irrigation						
4	! Definitions		Type of event	Date of event	yyyy-mm-dd	type code	Sprinkler	amount	mm					
5	1	Event	Irrigation	idate	2007-03-19	IROP	IR005	IRVAL	94.1					
6	1	Event	Irrigation	idate	2007-03-27	IROP	IR005	IRVAL	9.7					
7	1	Event	Irrigation	idate	2007-03-28	IROP	IR005	IRVAL	10					
8	1	Event	Irrigation	idate	2007-04-04	IROP	IR005	IRVAL	32.3					
9	2	Event	Irrigation	idate	2007-03-21	IROP	IR005	IRVAL	94.1					
0	2	Event	Irrigation	idate	2007-03-27	IROP	IR005	IRVAL	8.6					
1	2	Event	Irrigation	idate	2007-03-28	IROP	IR005	IRVAL	8.5					
2	2	Event	Irrigation	idate	2007-04-04	IROP	IR005	IRVAL	31.1					
3	3	Event	Irrigation	idate	2007-03-21	IROP	IR005	IRVAL	94.1					
4	3	Event	Irrigation	idate	2007-03-27	IROP	IR005	IRVAL	7.7					
5	3	Event	Irrigation	idate	2007-03-28	IROP	IR005	IRVAL	7.9					
6	3	Event	Irrigation	idate	2007-04-04	IROP	IR005	IRVAL	32.7					
7														
								Fertilizer	Fertilizer					
						Fertilizer	Fertilizer	application	applicati	Incorporatio		Fertilizer		
8	%		Type of event	Date of event	yyyy-mm-dd	type code	type value	code	on value	n depth	cm	N	kg[N]/ha	
9		Event	Fertilizer	fedate	2007-03-13	FECD	FE007	FEACD	AP002	FEDEP		FEAMN		51
0		Event	Fertilizer	fedate	2007-04-13	FECD	FE010	FEACD	AP005	FEDEP	15			57
1		Event	Fertilizer	fedate	2007-04-26	FECD	FE010	FEACD	AP005	FEDEP	15			56
2		Event	Fertilizer	fedate	2007-05-11	FECD	FE010	FEACD	AP005	FEDEP	15			56
3		Event	Fertilizer	fedate	2007-03-13	FECD	FE007	FEACD	AP002	FEDEP	15			51
4		Event	Fertilizer	fedate	2007-04-13	FECD	FE010	FEACD	AP005	FEDEP	15			55
5		Event	Fertilizer	fedate	2007-04-26	FECD	FE010	FEACD	AP005	FEDEP	15			56
6		Event	Fertilizer	fedate	2007-05-11	FECD	FE010	FEACD	AP005	FEDEP	15			56
7		Event	Fertilizer	fedate	2007-03-13	FECD	FE007	FEACD	AP002	FEDEP	15			51
8		Event	Fertilizer	fedate	2007-04-13	FECD	FE010	FEACD	AP005	FEDEP	15			55
9		Event	Fertilizer	fedate	2007-04-26	FECD	FE010	FEACD	AP005	FEDEP	15			56
70	3	Event	Fertilizer	fedate	2007-05-11	FECD	FE010	FEACD	AP005	FEDEP	15	FEAMN		56

ICASA/AgMIP template





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.

ICASA/AgMIP list of terms – From a knowledge management perspective, what is it?

Candidate terms

- Ontology
- Data dictionary
- Thesaurus
- Glossary
- Controlled vocabulary

Content

- Name & synonyms
- Meaning
- Hierarchy of relations
- Units
- Data type
- Validation criteria

Wikipedia: A data dictionary, as defined in the *IBM Dictionary* of *Computing*, is a "centralized repository of information about data such as meaning, relationships to other data, origin, usage, and format."

ICASA/AgMIP Data Dictionary: Next Steps

- Need to provide the Dictionary with a more intelligent interface than Google Sheets:
 - Searchable (not "Ctrl-F" among spreadsheets)
 - Allow users to submit questions and suggest terms
 - Hide information used only by developers
- Expand to meet community needs:
 - Sub-daily time scales
 - Life Cycle Analysis
 - Pest and disease evaluations Protocols/assays
 - More complete information on variables
- Harmonize with other resources
 - GRACEnet/REAP (USDA research databases)
 - CropOntology.org and AgTrials

The ICASA Data Dictionary: www.tinyurl.com\icasa-mvl

A	В	С	D	Е	F	G
1 Variable_Name	Description	Unit_or	Dataset	Subset	Group	Sub-Group
142 soil_identifier	Soil identifier	code	EXPERIMENT	MANAGEMENT	FIELDS	
143 soil_file	File containing relevant soil profile description	text	EXPERIMENT	MANAGEMENT	FIELDS	
144 water_table_depth	Water table depth	cm	EXPERIMENT	MANAGEMENT	FIELDS	
145 field_notes	Field, short notes	text	EXPERIMENT	MANAGEMENT	FIELDS	
146 field_soil_texture	Field observed soil texture	code	EXPERIMENT	MANAGEMENT	FIELDS	
147 soil_analysis_date	Soil analysis date	date	EXPERIMENT	MANAGEMENT	SOIL ANALYSES	
148 id	id		EXPERIMENT	MANAGEMENT	SOIL ANALYSES	
149 soil_analysis_level	Soil analysis, treatment level identifier	number	EXPERIMENT	MANAGEMENT	SOIL ANALYSES	
150 soil_analysis_name	Soil analysis name	text	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
151 soil_family_NRCS_system	Soil family, NRCS system	text	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
152 soil_analysis_year	Soil analysis year	year	EXPERIMENT	MANAGEMENT	SOIL ANALYSES	
153 soil_analysis_day	Soil analysis day of year or days after planting	day	EXPERIMENT	MANAGEMENT	SOIL ANALYSES	
154 soil_analys_meth_pH_buff	Soil anlysis method for pH in buffer solution	code	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
155 soil_anal_meth_P_extract	Soil analysis method for P extraction	code	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
156 soil_anal_meth_exch_K	Soil analysis method for exchangeable	code	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
157 drainage_rate_per_day	Drainage rate as fraction per day	1/d	EXPERIMENT	MANAGEMENT	SOIL ANALYSES	
158 runoff_curve_no_SCS	Runoff curve no. (Soil Conservation Service)	number	EXPERIMENT	MANAGEMENT	SOIL ANALYSES	
159 id	id		EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
160 soil_analysis_level	Soil analysis, treatment level identifier linkage	number	EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
161 soil_layer_top_depth	Soil layer depth to upper boundary of sample (top)	cm	EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
162 soil_layer_base_depth	Soil layer depth to lower boundary of sample (bottom)		EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
163 soil_bulk_dens_moist		g/cm3	EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
164 organic_carbon_conc	Organic carbon	%	EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
165 soil_N_conc_total	Nitrogen, total soil, total	%	EXPERIMENT	MANAGEMENT	SOIL ANALYSES	SOIL_ANALYSES_LAYERS
166 pH_in_water	pH of soil in water, from in-field sample	number	EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
167 pH_in_buffer	pH of soil measured in buffer solution, from in-field sai	number	EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
168 phosphorus_extractable	Phosphorus, extractable	mg/kg	EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
169 potassium_exchangeable	Potassium, exchangeable	cmol/kg	EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
170 soil_compac_fact_by_lyr	Soil compaction factor by layer (0 to 1 scale)		EXPERIMENT	MANAGEMENT		SOIL_ANALYSES_LAYERS
II I Dimate_codes / C	Other_codes / Pest_codes / Form Responses 1 / Fo	rm Respo		· ·		COIL ANALYCES LAVEDS