Presentation of the fuzzy expert systems, possible use in the CN-MIP project

Gianni BELLOCCHI

French National Institute for Agricultural Research, Clermont-Ferrand, France

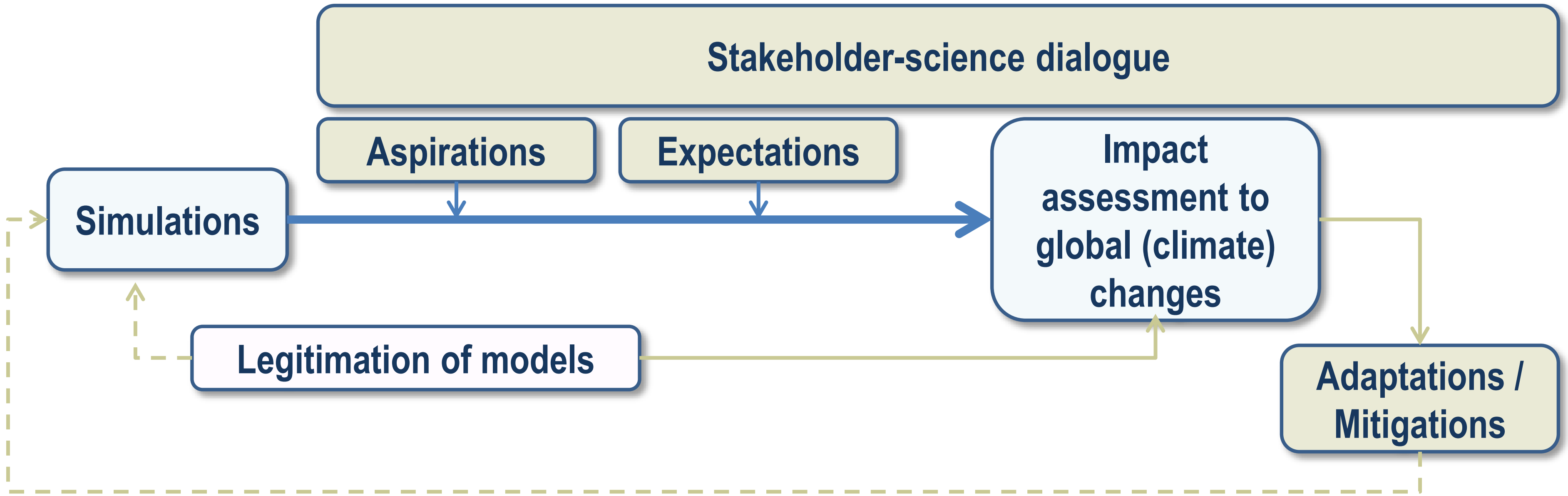
2nd CN-MIP Project Meeting

Fort Collins CO, USA

04-06 March 2015

"*We conserve many things that we don't evaluate and little of those we value*" (Geoffrey M. Heal)

Deliberative process in model-based climate change studies



Bellocchi et al., 2006, Ital. J. Agrometeorol.

Rivington et al., 2007, Environ. Modell. Softw.

Bellocchi et al., 2014, Agron. Sustain. Dev.

Institutionalising deliberative practices for context-specific model evaluations

Model evaluations are (sometimes) an (important) orientating landmark in the skyline of decisions, without replacing them

To evaluate (crop and grassland) simulation models is far more urgent as many of the (tactical and strategic) decisions (in agriculture) are based on model outcomes

Dealing with (existing) and designing (new) agricultural systems is a priority that deliberations about model evaluation contribute to accomplish in a more efficient (maybe more appropriate) manner, in any case with more awareness if (genuine) collective deliberations are possible

The central issue is to think and conceive model evaluation in a (clear) decisional perspective about type of model, operability, transparency, etc.

As several models are at hand, “mod-diversity” imposes the analysis of case-by-case issues, while also integrating the specific context in a larger-scale perspective (in space and time)

Ital. J. Agron., 3, 1, 19-39

A Review of Methodologies to Evaluate Agroecosystem Simulation Models

F. MARTORANA and G. BELLOCCHI

Agron. Sustain. Dev. (2009)
© INRA, EDP Sciences, 2009
DOI: 10.1051/agro/2009001

Available online at:
www.agronomy-journal.org

Agronomy for Sustainable Development

Review article

Validation of biophysical models: issues and methodologies. A review

Gianni BELLOCCHI[1*,**], Mike RIVINGTON[2], Marcello DONATELLI[1***], Kei...

Environmental Modelling & Software 26 (2011) 328–338

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

ELSEVIER

Technical assessment and evaluation of environmental models and software: ...etter to the Editor

...Alexandrov[a,*], D. Ames[b], G. Bellocchi[c], M. Bruen[d], N. Crout[e], M. Erechtchoukova[f], A. Hildebrandt[g], ...ffma... Matsunaga[a], S.T. Purucker[k], M. Rivington[l], ...

ELSEVIER

Contents lists available at Science...

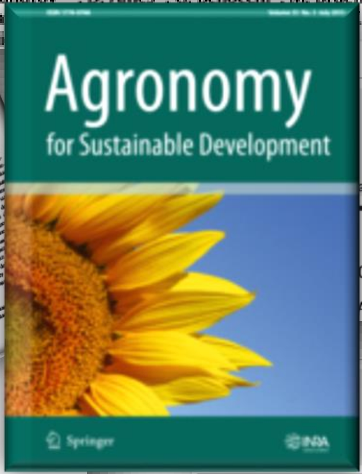
Agricultural Systems

journal homepage: www.elsevier.com/loc...

An evaluation of the statistical methods for testing the pe... of crop models with observed data

J.M. Yang[a], J.Y. Yang[b,*], S. Liu[b,c], G. Hoogenboom[d]

Deliberative processes for comprehensive evaluation of agroecological models. A review



Ecological Modelling 220 (2009) 1395–1410

Contents lists available at ScienceDirect

Ecological Modelling

...l homepage: www.elsevier.com/locate/ecolmodel

...of the models WARM, CropSyst, and WOFOST for rice

...Acutis[b], Gianni Bellocchi[c], Marcello Donatelli[d,1]

Ecological Modelling 221 (2010) 960–964

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Short communication

A proposal of an indicator for quantifying model robustness based on the relationship between variability of errors and of explored conditions

R. Confalonieri[a,*], S. Bregaglio[a,b], M. Acutis[a]

Assessment ...equacy of mathema...cal models

Luis Orlindo Tedeschi *

Journal of Applied Remote Sensing

Derivation of biophysical variables from Earth observation data: validation and statistical measures

Katja Richter
Clement Atzberger
Tobias B. Hank
Wolfram Mauser

IRENE_DLL: A CLASS LIBRARY FOR EVALUATING NUMERICAL ESTIMATES

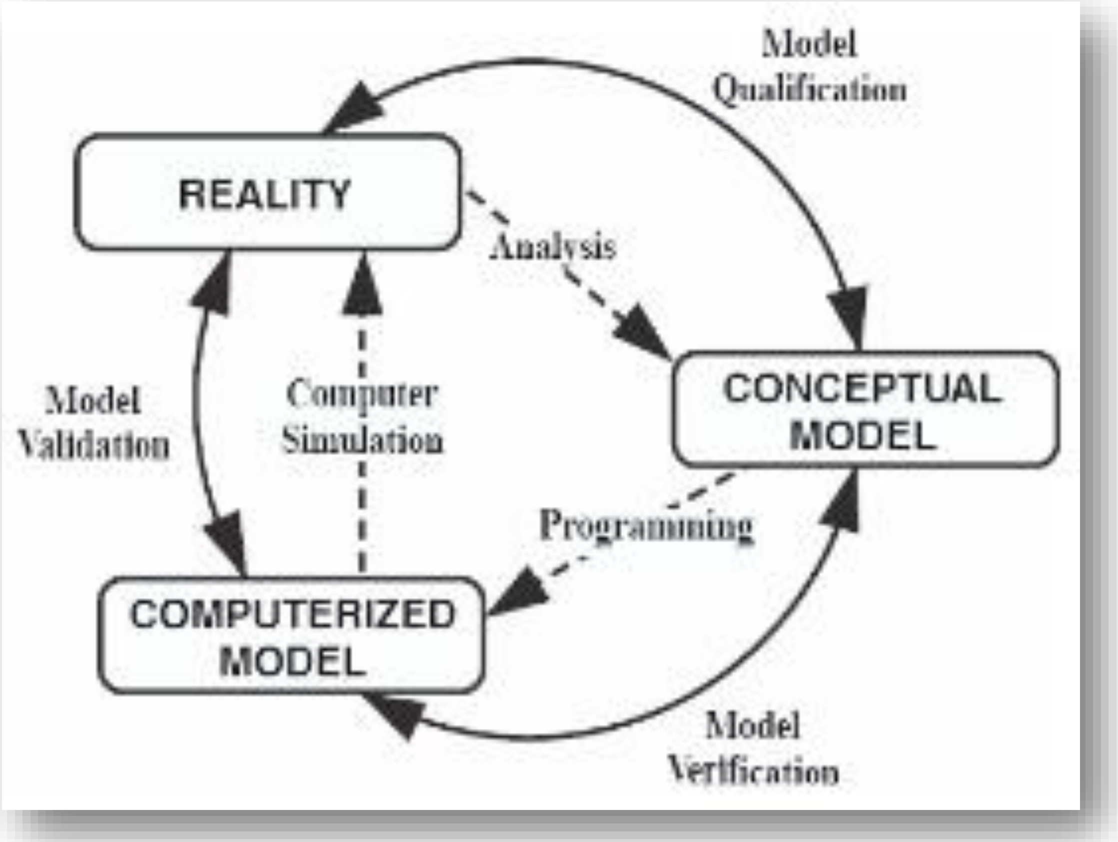
GIANNI FILA,* GIANNI BELLOCCHI, MARCELLO DONATELLI, AND MARCO ACUTIS

An Indicator of Solar Radiation Model Performance based on a Fuzzy Expert System

Gianni Bellocchi,* Marco Acutis, Gianni Fila, and Marcello Donatelli

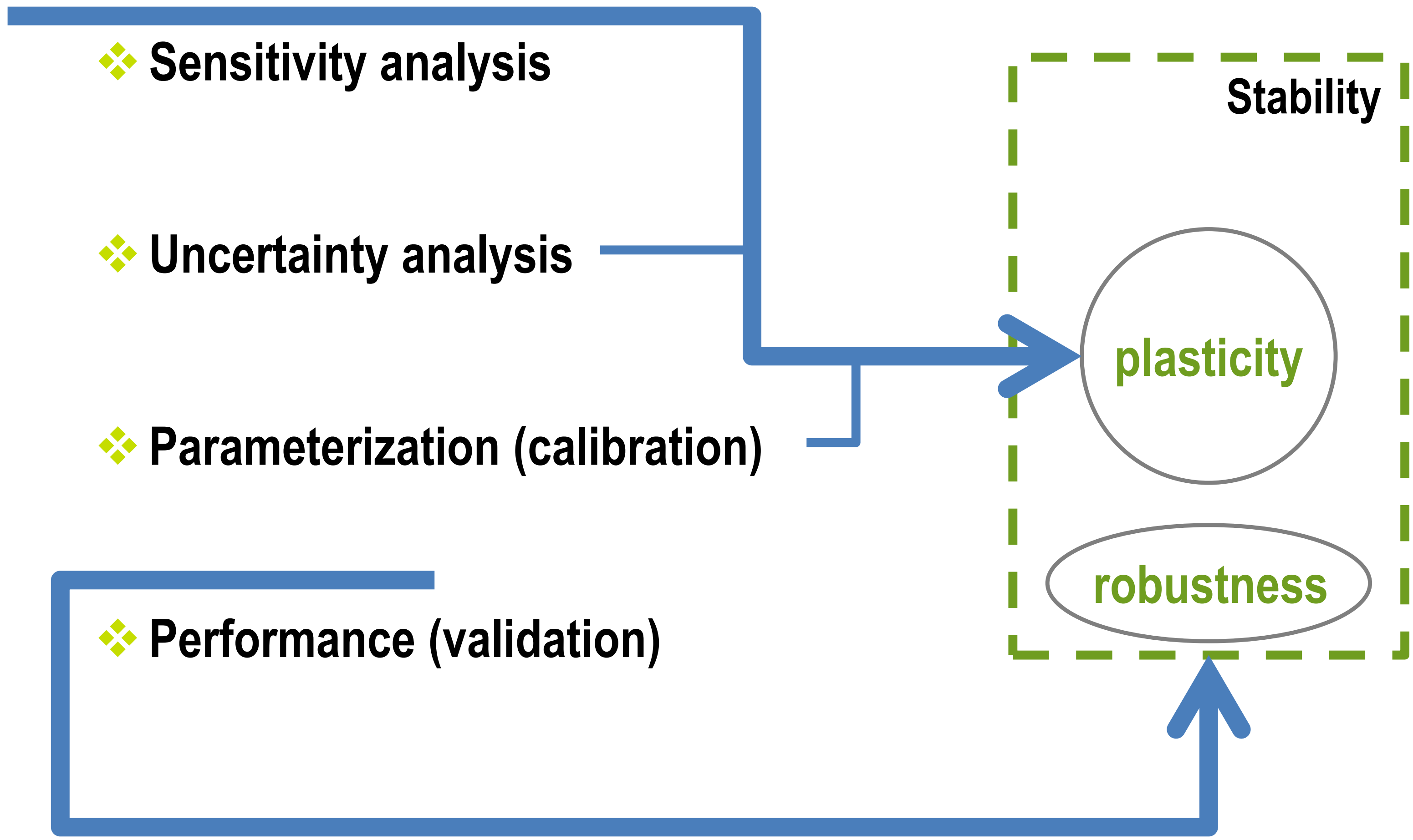
Model evaluation

Model evaluation (validation + verification):
action in which the quality of a mathematical model for specific objectives is established



Schlesinger, 1979, Simulation

Model evaluation components



Key issues

Key validation issues	Major factors to investigate				
	Modelling objective	Model inputs	Model outputs	Model structure	Modelling conditions
Validation purpose	X		X		X
Robustness of results			X		X
Interpretation of phenomena		X	X	X	
Model comparison				X	
Model predictions	X		X		X
Model complexity		X	X	X	
Data accuracy		X	X		
Time histories			X		

Bellocchi et al., 2010, Agron. Sustain. Dev.

Approaches to model performance evaluation

Metrics

- Indices
  - difference-based indices (RMSE, CRM, MAE, ...)
  - association-based indices (slope, intercept, r, r^2, EF, ...)
  - pattern indices (range-based, F-based)
- Test statistics
  - Student-t, chi square, …
- Probability distributions
  - cumulative, density functions

Procedures

- Time mismatch analysis
- Aggregation of statistics
  - hierarchical, fuzzy logic based indicators

$$CRM = \frac{\sum_{i=1}^{n} O_i - \sum_{i=1}^{n} P_i}{\sum_{i=1}^{n} O_i}$$

$$MAE = \frac{\sum_{i=1}^{n} |P_i - O_i|}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (P_i - O_i)^2}{n}}$$

simple

absolute

squared

difference-based metrics

non-parametric

$$MdAE = \mathrm{median}_{i=1,\dots,n} |P_i - O_i|$$

Elaboration of new metrics

Setting of thresholds

Meaning and limitations

Intercorrelation

Disaggregation

Aggregation

association-based metrics

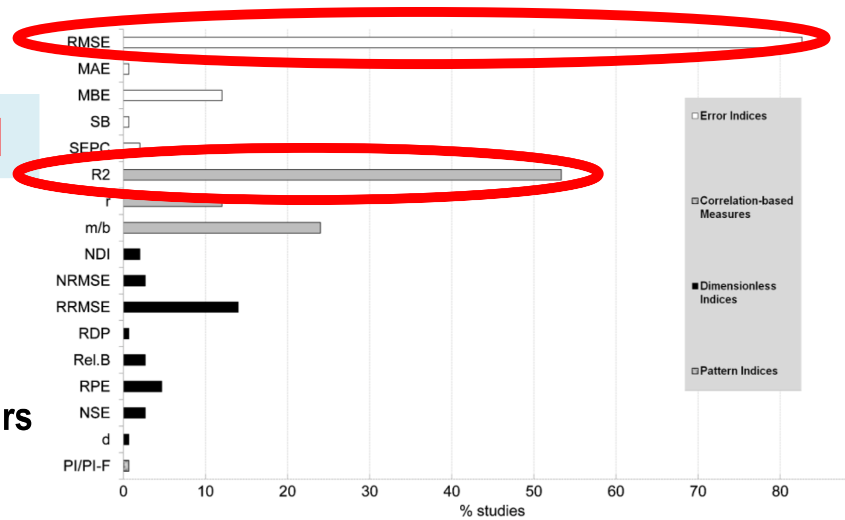
$$r = \frac{\sum_{i=1}^{n} (P_i - \bar{P}) \cdot (O - \bar{O})}{\sqrt{\sum_{i=1}^{n} (P_i - \bar{P})^2 \cdot \sum_{i=1}^{n} (O_i - \bar{O})^2}} \qquad r^2, slope, intercept$$

combined

$$EF = 1 - \frac{\sum_{i=1}^{n} (P_i - O_i)^2}{\sum_{i=1}^{n} (O_i - \bar{O})^2}$$

$$d = 1 - \frac{\sum_{i=1}^{n} (P_i - O_i)^2}{\sum_{i=1}^{n} \left(|P_i - \bar{O}| + |O_i - \bar{O}|\right)^2}$$

P = predicted; O = observed; i = i^{th} O/P pair; n = number of O/P pairs



Richter et al., 2012, J. Appl. Remote Sens.

Setting of thresholds

Performance measure	Unit	Value range and purpose	Reliability criteria
Coefficient of determination (**R^2**) of the linear regression estimates versus measurements	dimensionless	0 (absence of fit) to 1 (perfect fit): the closer values are to 1, the better the model	> 0.8
Willmott (1982) index of agreement (**d**)	dimensionless	0 (absence of agreement) to 1 (perfect agreement): the closer values are to 1, the better the model	> 0.8
Mean absolute error over the mean of the measured values (**MAE(%)**)	%	0 (optimum) to positive infinity: the smaller MAE(%), the better the model performance	< 20

De Jager, 1994, Water SA

Properties of evaluation metrics

- Non-dimensionality
  - to avoid influences from the magnitude of the values
- Lower and upper bounding
  - for easy comprehension (e.g. between 0-no agreement and 1-full agreement)
- Symmetry
  - datasets interchangeability (e.g. Reduced Major Axis instead of Least Squares regression)
- Same unit of data (difference-based metrics)
  - to understand the magnitude of the error
- Prediction capability
  - comparison to measurement statistics (e.g. efficiency with respect to observed mean value)

Latest developments

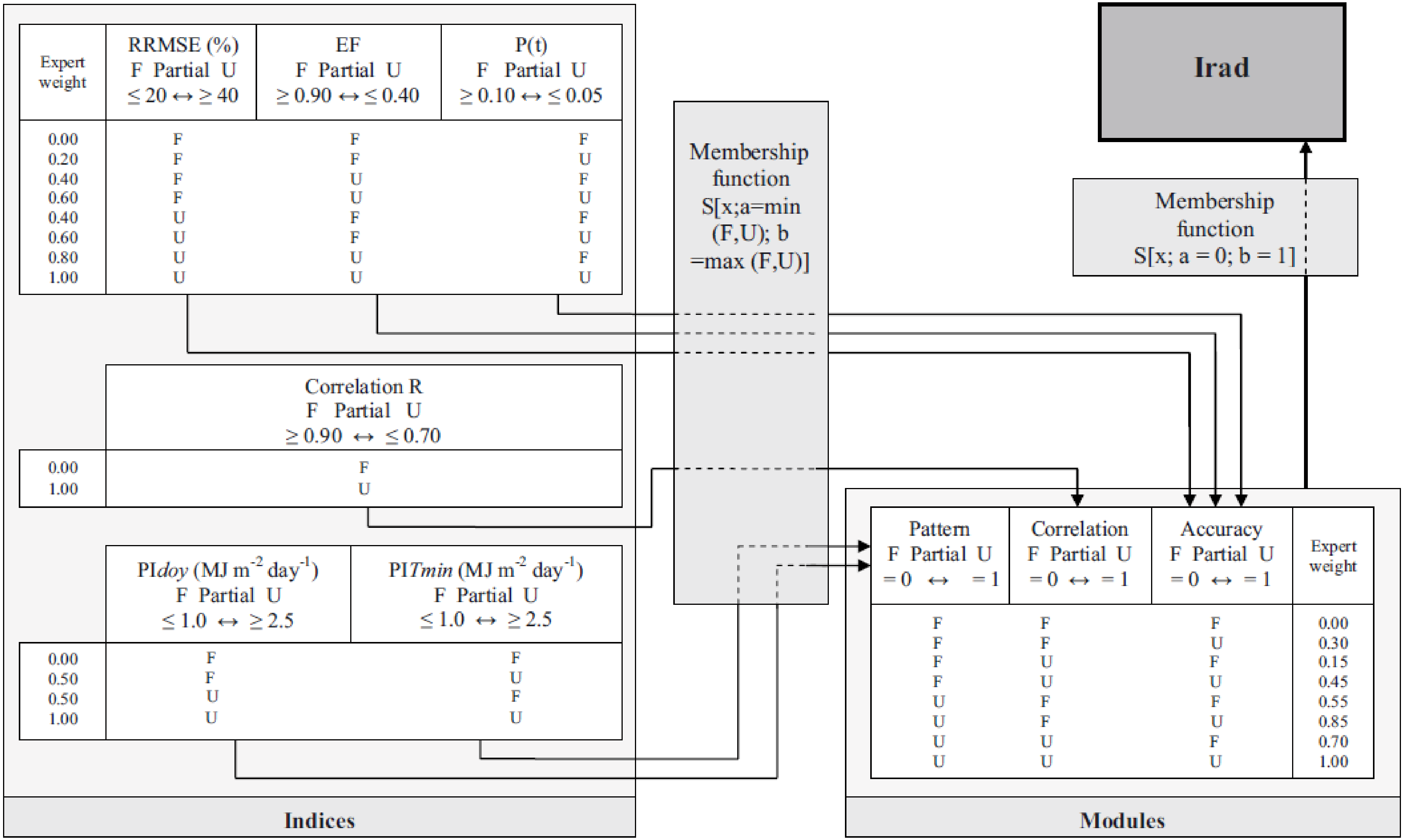
- Aggregation of multiple evaluation metrics into integrated indicators (fuzzy logic principle, after Bellocchi et al., 2002, Agron. J.)

- Inclusion, in the evaluation of models, of other measures than performance metrics, such as sensitivity analysis measures and information criteria for model selection (Confalonieri et al., 2009, Ecol. Modell.)

- Elaboration of the model robustness concept (Confalonieri et al., 2010, Ecol. Modell.)

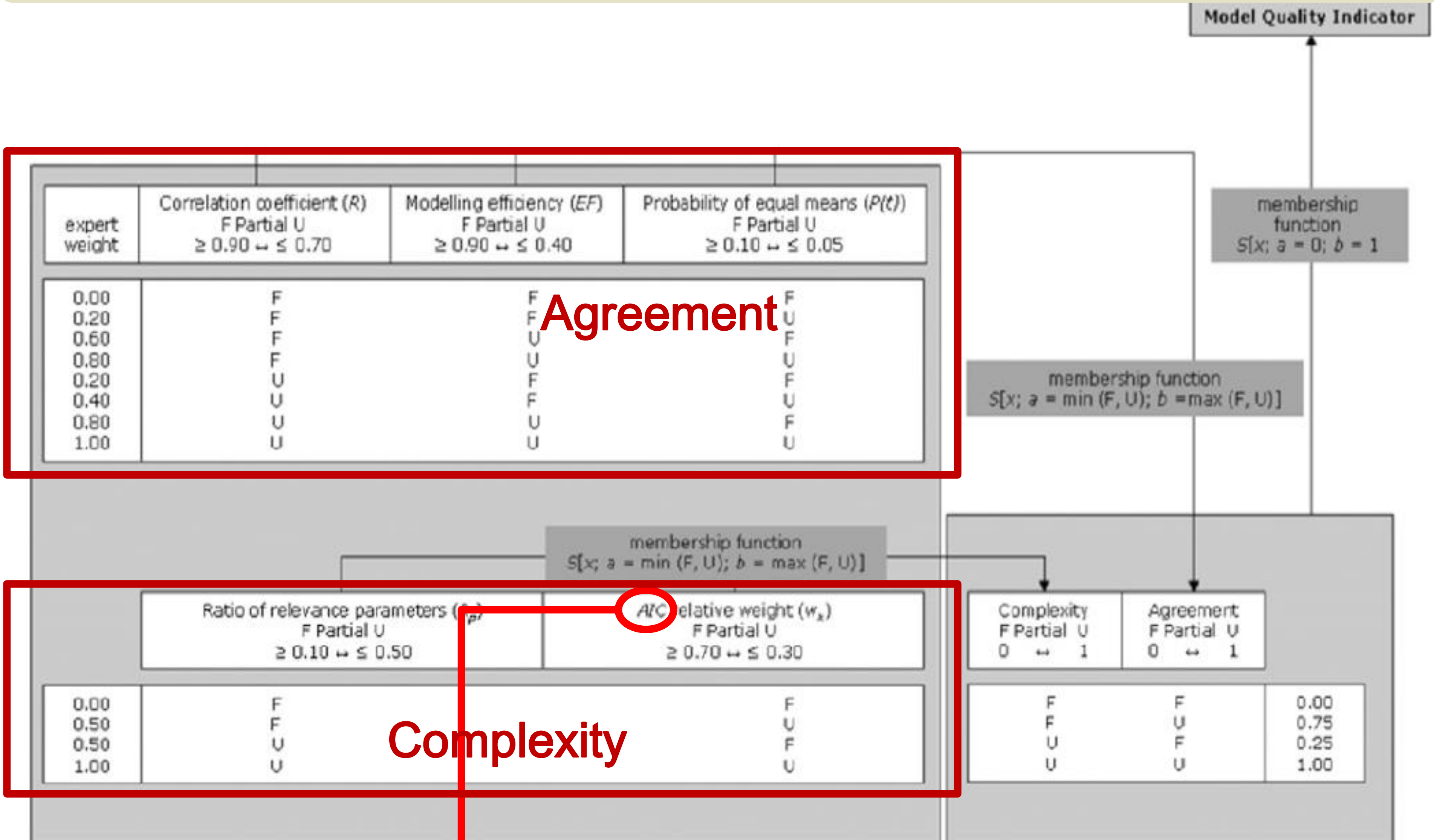
- Elaboration of the model plasticity concept (Confalonieri et al., 2012, Ecol. Modell.)

Fuzzy logic-based integrated indicators / 1



Bellocchi et al., 2002, Agron. J.
Rivington et al., 2007, Agr. Forest Meteorol.

Fuzzy logic-based integrated indicators / 2



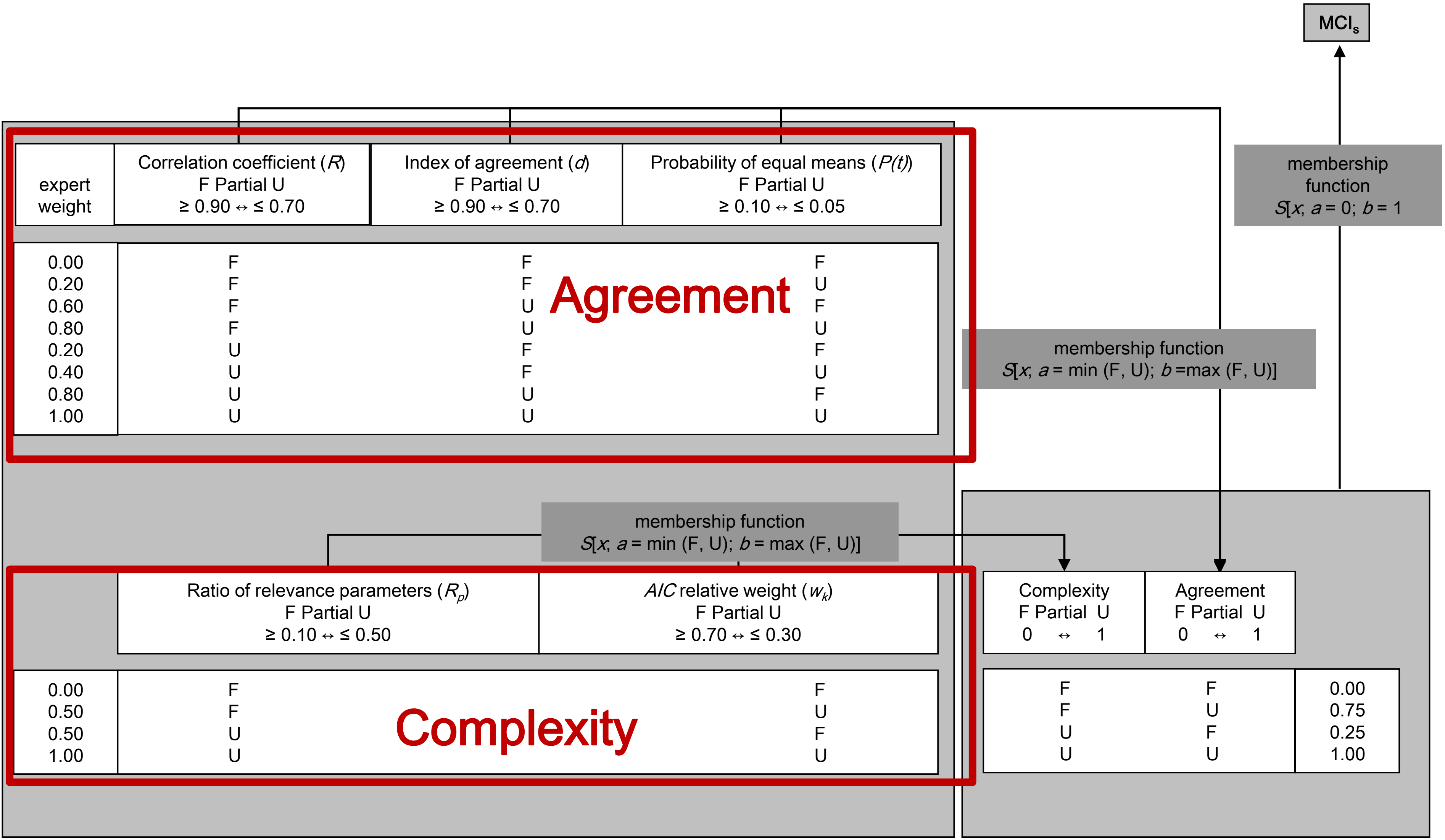
expert weight	Correlation coefficient (R)	Modelling efficiency (EF)	Probability of equal means (P(t))
0.00	F	F	F
0.20	F	F	U
0.60	F	U	F
0.80	F	U	U
0.20	U	F	F
0.40	U	F	U
0.80	U	U	F
1.00	U	U	U

expert weight	Ratio of relevance parameters	AIC relative weight
0.00	F	F
0.50	F	U
0.50	U	F
1.00	U	U

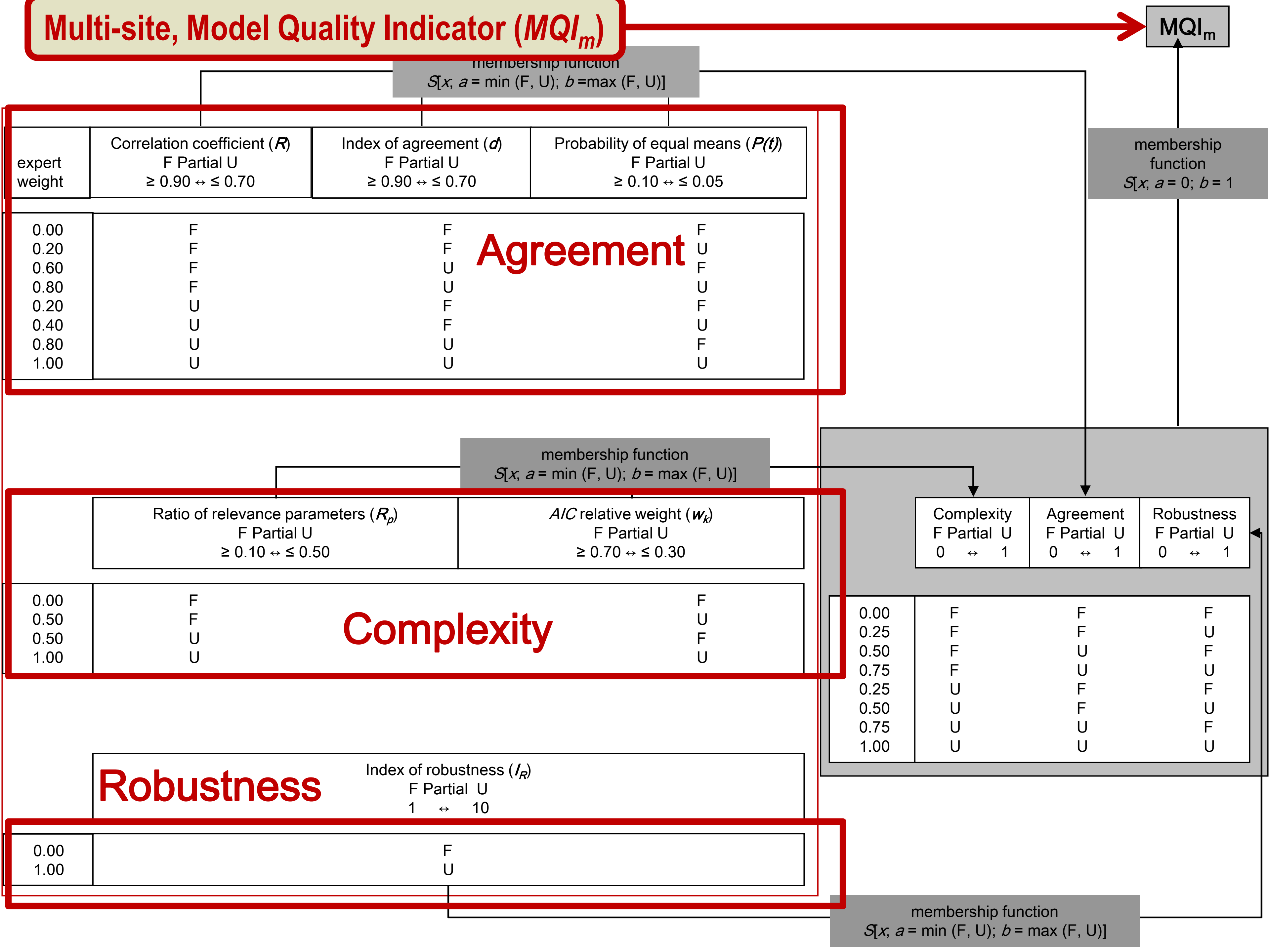
Complexity	Agreement	expert weight
F	F	0.00
F	U	0.75
U	F	0.25
U	U	1.00

$$AIC = n \cdot log(MSE) + 2 \cdot T$$

Confalonieri et al., 2009, Ecol. Modell.

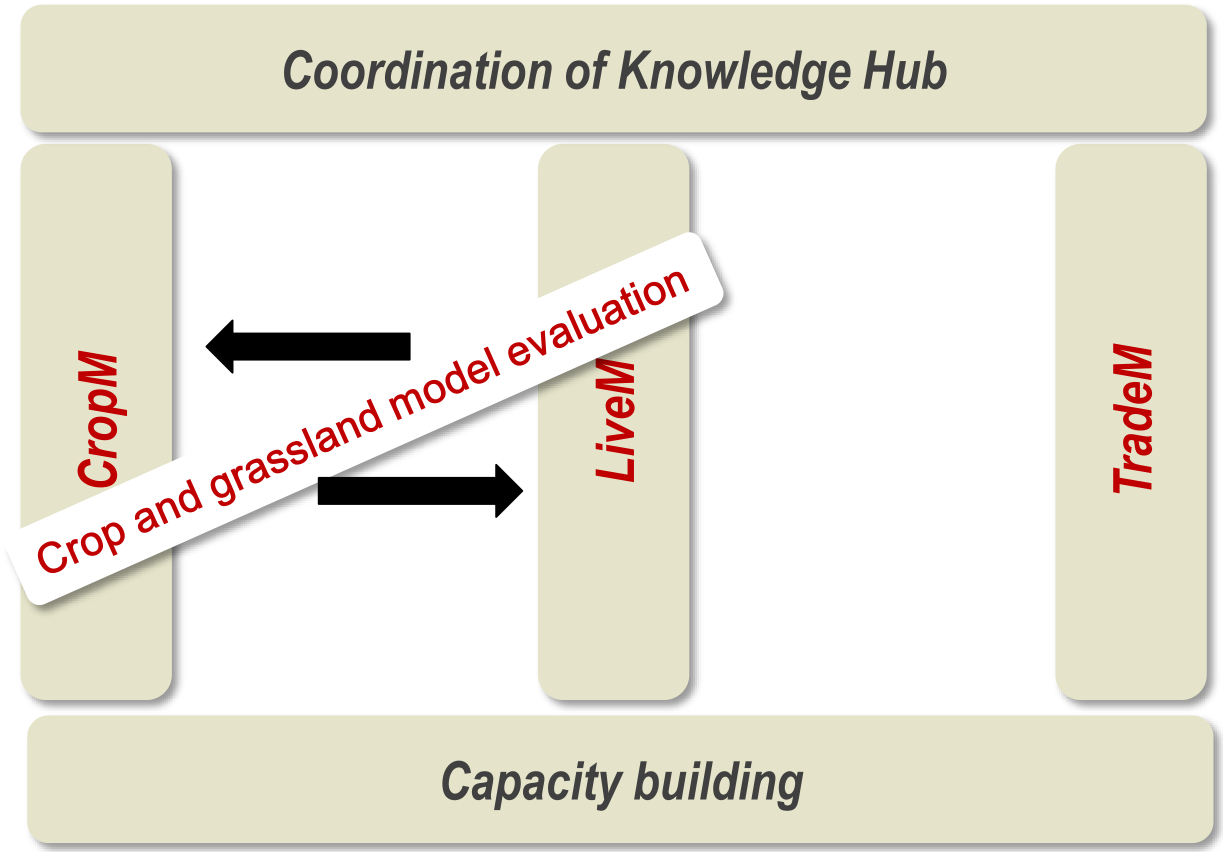


MCI$_s$
membership function
$S[x; a = 0; b = 1$
expert weight
Correlation coefficient (R)
F Partial U
≥ 0.90 ↔ ≤ 0.70
Index of agreement (d)
F Partial U
≥ 0.90 ↔ ≤ 0.70
Probability of equal means ($P(t)$)
F Partial U
≥ 0.10 ↔ ≤ 0.05
0.00 F F F
0.20 F F U
0.60 F U F
0.80 F U U
0.20 U F F
0.40 U F U
0.80 U U F
1.00 U U U
Agreement
membership function
$S[x; a = \min (F, U); b = \max (F, U)]$
membership function
$S[x; a = \min (F, U); b = \max (F, U)]$
Ratio of relevance parameters (R_p)
F Partial U
≥ 0.10 ↔ ≤ 0.50
AIC relative weight (w_k)
F Partial U
≥ 0.70 ↔ ≤ 0.30
0.00 F F
0.50 F U
0.50 U F
1.00 U U
Complexity
Complexity
F Partial U
0 ↔ 1
Agreement
F Partial U
0 ↔ 1
F F 0.00
F U 0.75
U F 0.25
U U 1.00



Multi-site, Model Quality Indicator (MQI_m)
MQI_m
membership function
$S[x; a = \min(F, U); b = \max(F, U)]$
expert weight
Correlation coefficient (R) F Partial U ≥ 0.90 ↔ ≤ 0.70
Index of agreement (d) F Partial U ≥ 0.90 ↔ ≤ 0.70
Probability of equal means ($P(t)$) F Partial U ≥ 0.10 ↔ ≤ 0.05
0.00 F F F
0.20 F F U
0.60 F U F
0.80 F U U
0.20 U F F
0.40 U F U
0.80 U U F
1.00 U U U
Agreement
membership function
$S[x; a = 0; b = 1]$
membership function
$S[x; a = \min(F, U); b = \max(F, U)]$
Ratio of relevance parameters (R_p) F Partial U ≥ 0.10 ↔ ≤ 0.50
AIC relative weight (w_k) F Partial U ≥ 0.70 ↔ ≤ 0.30
0.00 F F
0.50 F U
0.50 U F
1.00 U U
Complexity
Complexity F Partial U 0 ↔ 1
Agreement F Partial U 0 ↔ 1
Robustness F Partial U 0 ↔ 1
0.00 F F F
0.25 F F U
0.50 F U F
0.75 F U U
0.25 U F F
0.50 U F U
0.75 U U F
1.00 U U U
Robustness
Index of robustness (I_R) F Partial U 1 ↔ 10
0.00 F
1.00 U
membership function
$S[x; a = \min(F, U); b = \max(F, U)]$

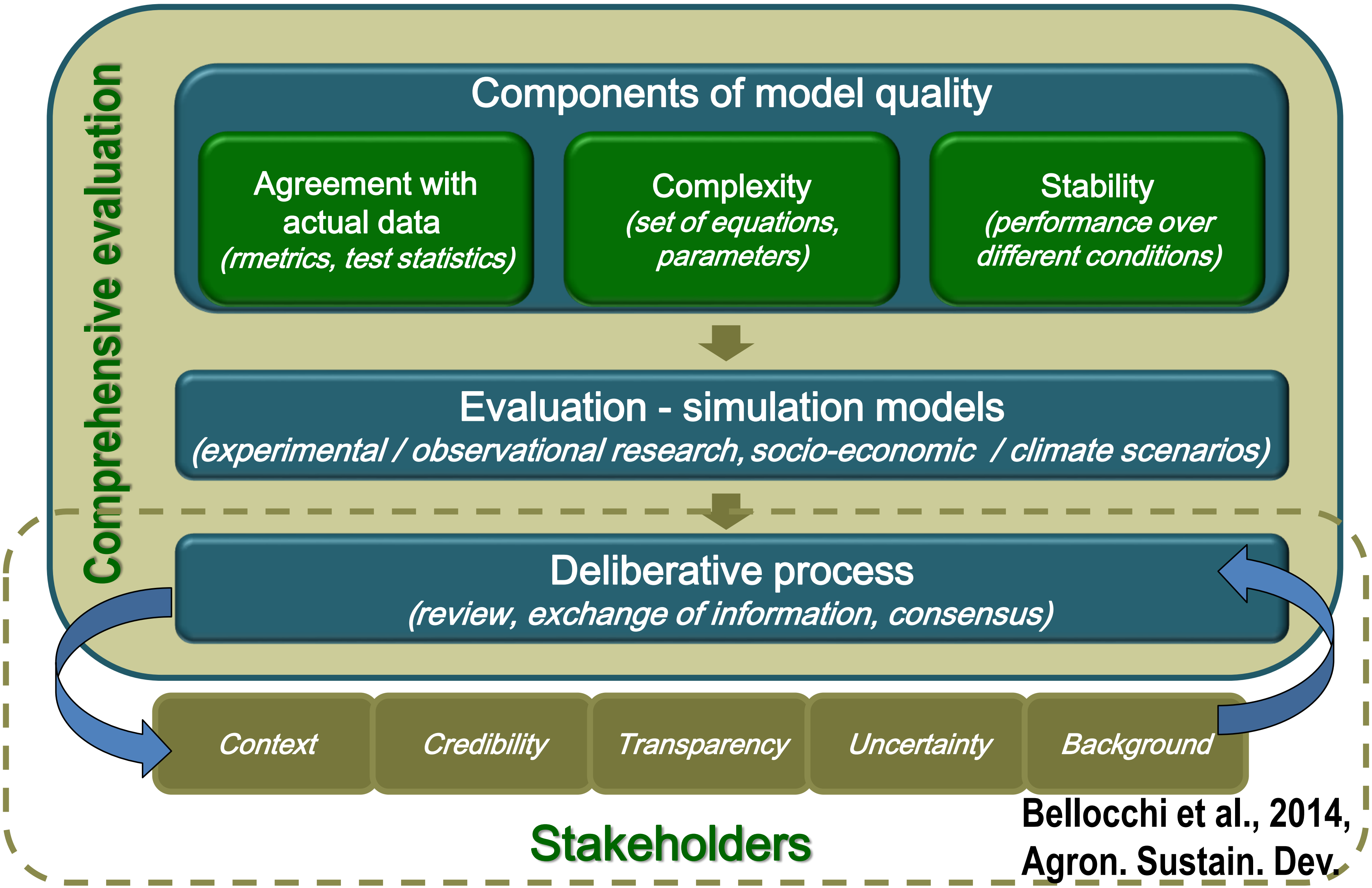
MACSUR cross-cutting activities



CropM-LiveM

- Definition of model performance indicators
- Elaboration of model evaluation protocols

Model evaluation / deliberative process



Bellocchi et al., 2014, Agron. Sustain. Dev.

Synthetic indicators

Aggregation rules: fuzzy-logic based weighing system

Non-dimensionality

Lower and upper bounding

I. Agreement

- Correlation coefficient
- Index of agreement
- Probability of equal means

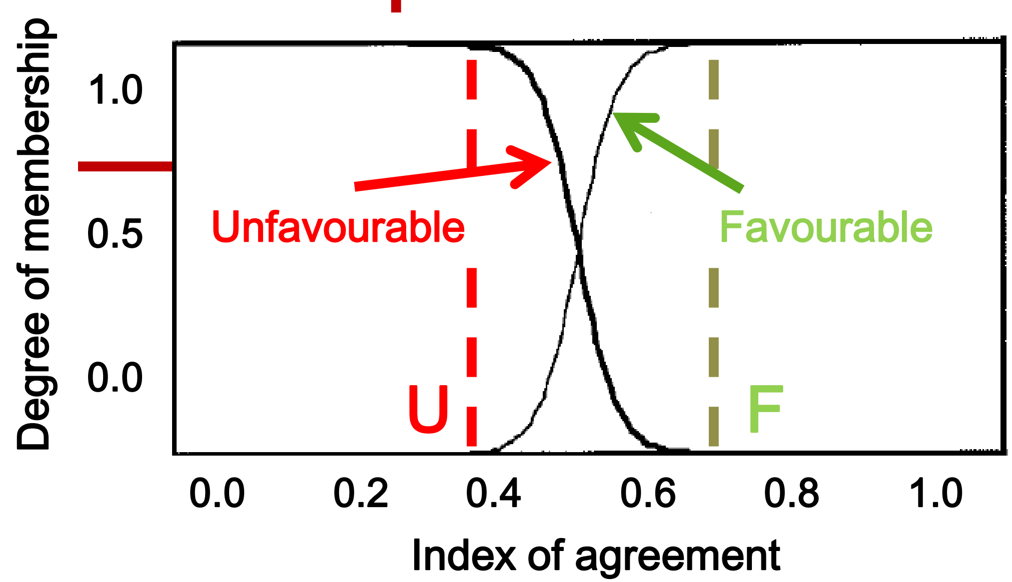
II. Complexity

- Ratio of relevant parameters
- Parameters-agreement criterion

III. Stability (robustness)

- Index of robustness

Model Quality Indicator



Hindrances to overcome: thresholds and weights

MQI_m – Questionnaire

Questionnaires answered / commented: 16 (13 online + 3 offline) + 1 comment

7. Over the range 0 (best) to 1 (worst) of MQI_m, may crisp threshold values be set to interpret results (e.g. >0.66: poor model performance?

6. Do the expert weights assigned to Modules reflect the importance of each of them?

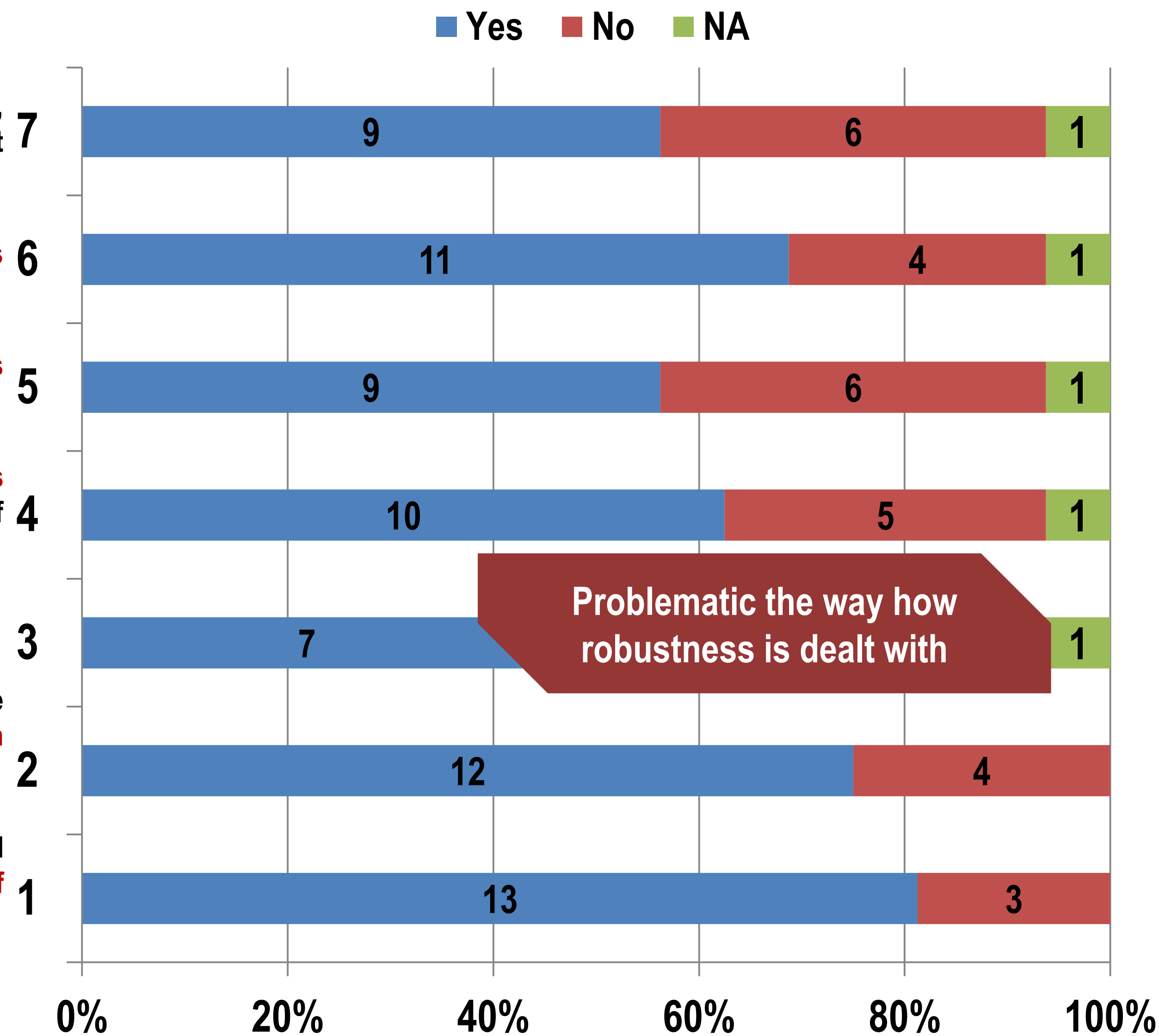
5. Do the expert weights assigned to metrics within a Module reflect their relative importance?

4. Do the favourable / unfavourable thresholds assigned to each metric reflect the perception of the quality of model performance?

3. Do the equations of the metrics need changes?

2. Do the metrics of MQI_m represent a good choice to cover aspects of model evaluation (quantification of error, bias, efficiency, etc.)?

1. Do the fuzzy-logic based assessment method (MQI_m) account for all the relevant aspects of model inter-comparison?



Robustness of a model

A robustness measure would account for model performance stability over a wide range of conditions (single site versus multiple sites)

How the variability of model performance can be quantified with the variability of conditions?

Index of robustness

$$I_R = \frac{\sigma_{EF}}{\sigma_{SAM}}$$ (0, best; +∞, worst)

Modelling efficiency

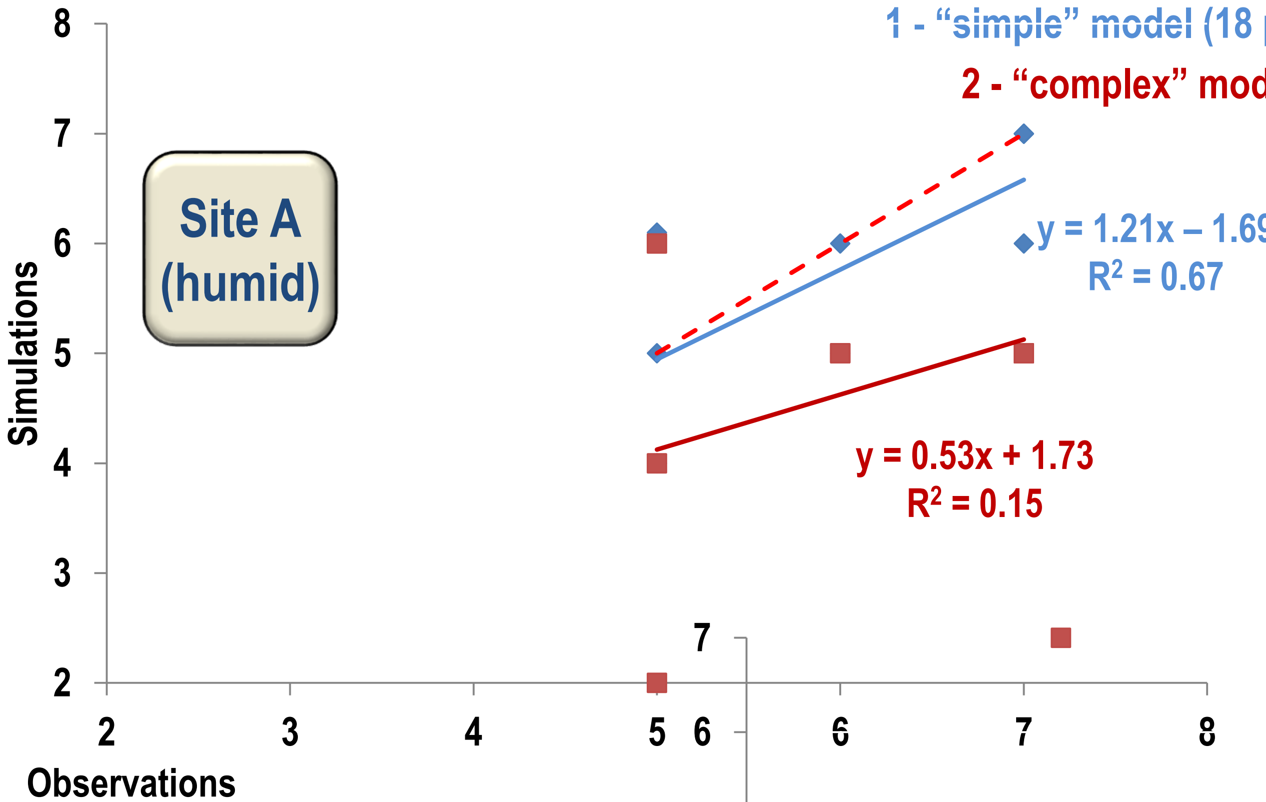
$$EF = 1 - \frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}$$ (-∞, worst; 1, best)

Synthetic Agro-Meteorological Indicator

$$SAM = \frac{Rain - ET_0}{Rain + ET_0}$$ (-1, +1)

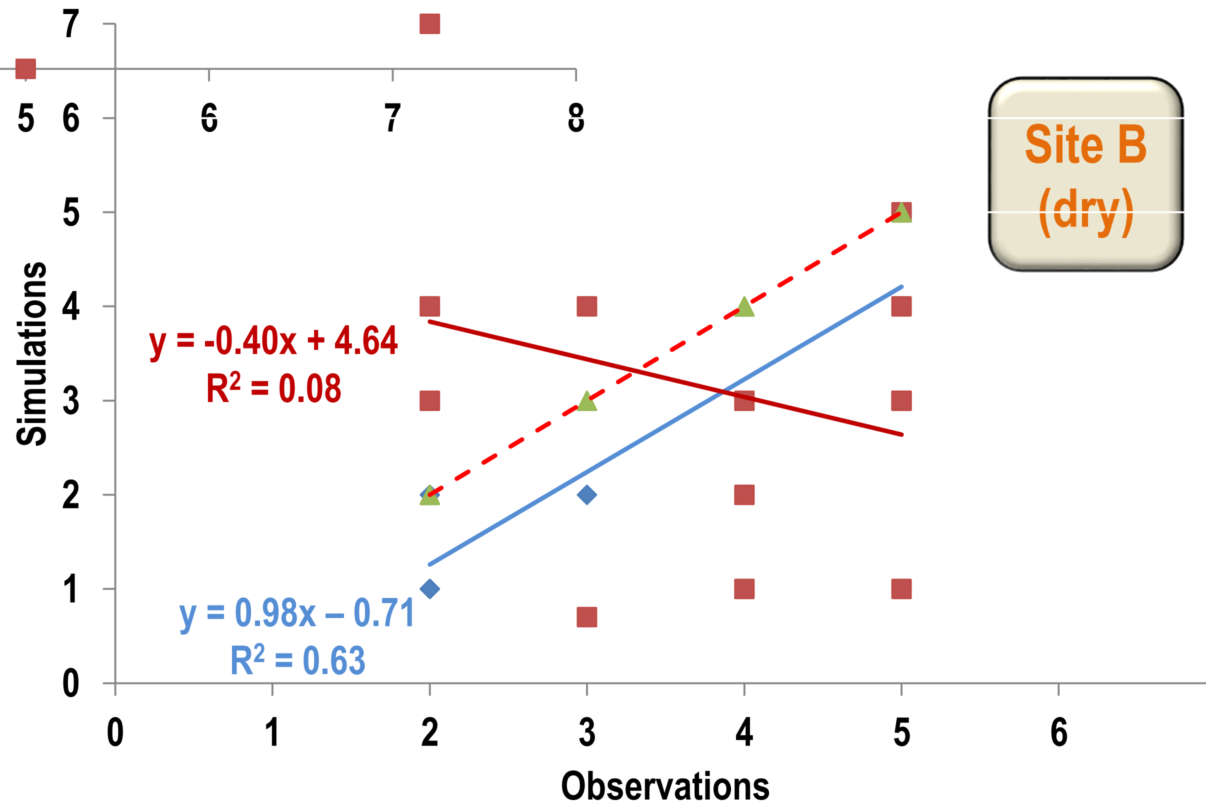
From the questionnaires:

- Need to test the index on a variety of rainfall patterns (e.g. monsoonal areas)
- Whole year versus growing season, or winter and summer?
- Accounting for soil properties if water limited simulations are performed



Evaluation	Model 1
Agreement	0.329
Complexity	0.016
Robustness	0.000
MQI_m	0.109

Evaluation	Model 2
Agreement	0.800
Complexity	0.500
Robustness	0.006
MQI_m	0.556



Exemplary results

Above-ground rice biomass (kg DM m^{-2})

Three models: A (intermediate), B (simple), C (complex)

MQI_s	A	B	C
C. d'Agogna	**0.0313**	0.1250	0.2174
Vercelli	0.1070	**0.0853**	0.1372
Mortara	0.2188	**0.0000**	0.2174
Rosate	**0.0313**	0.2284	0.2388
MQI_m	**0.0750**	0.1940	0.3356

AIC	A	B	C
C. d'Agogna	**34**	37	**79**
Vercelli	**33**	34	**73**
Mortara	**26**	28	**67**
Rosate	**20**	49	**91**

Complexity

EF	A	B	C
C. d'Agogna	0.90	0.95	0.93
Vercelli	0.92	0.97	0.96
Mortara	0.96	**0.98**	**0.98**
Rosate	0.92	**0.62**	**0.48**
I_R	**0. 16**	1.24	1.71

Robustness

Towards an internationally-agreed protocol to evaluate C-N models: what does go forth?

- Review of settings
  - Selection of metrics,
  - Attribution of thresholds and weights

- Extension to multiple outputs (three GHGs)

- **Working group to be established**
- **Protocol to be elaborated**

IMΦ

Thank you

for your attention !