Research Institute of Organic Agriculture FiBL
info.suisse@fibl.org, www.fibl.org

Weighting in multi-criteria sustainability assessments: quantifying the uncertainty of expert judgements in the SMART Farm Tool

Michael Curran, Socioeconomics Dept., FiBL
michael.curran@fibl.org

LCA Discussion Forum, ETH Zürich, 09 Sept 2019

Content

1. Sustainability framework: SAFA
2. SMART MCA structure
3. Indicator weights consensus method
4. Uncertainty assessment
5. Effects on farm comparisons
6. Conclusions

smart
sustainability monitoring
and assessment routine

Sustainability framework

Food and Agriculture Organization of the United Nations

SAFA
SUSTAINABILITY ASSESSMENT OF FOOD AND AGRICULTURE SYSTEMS

FAO SAFA guidelines

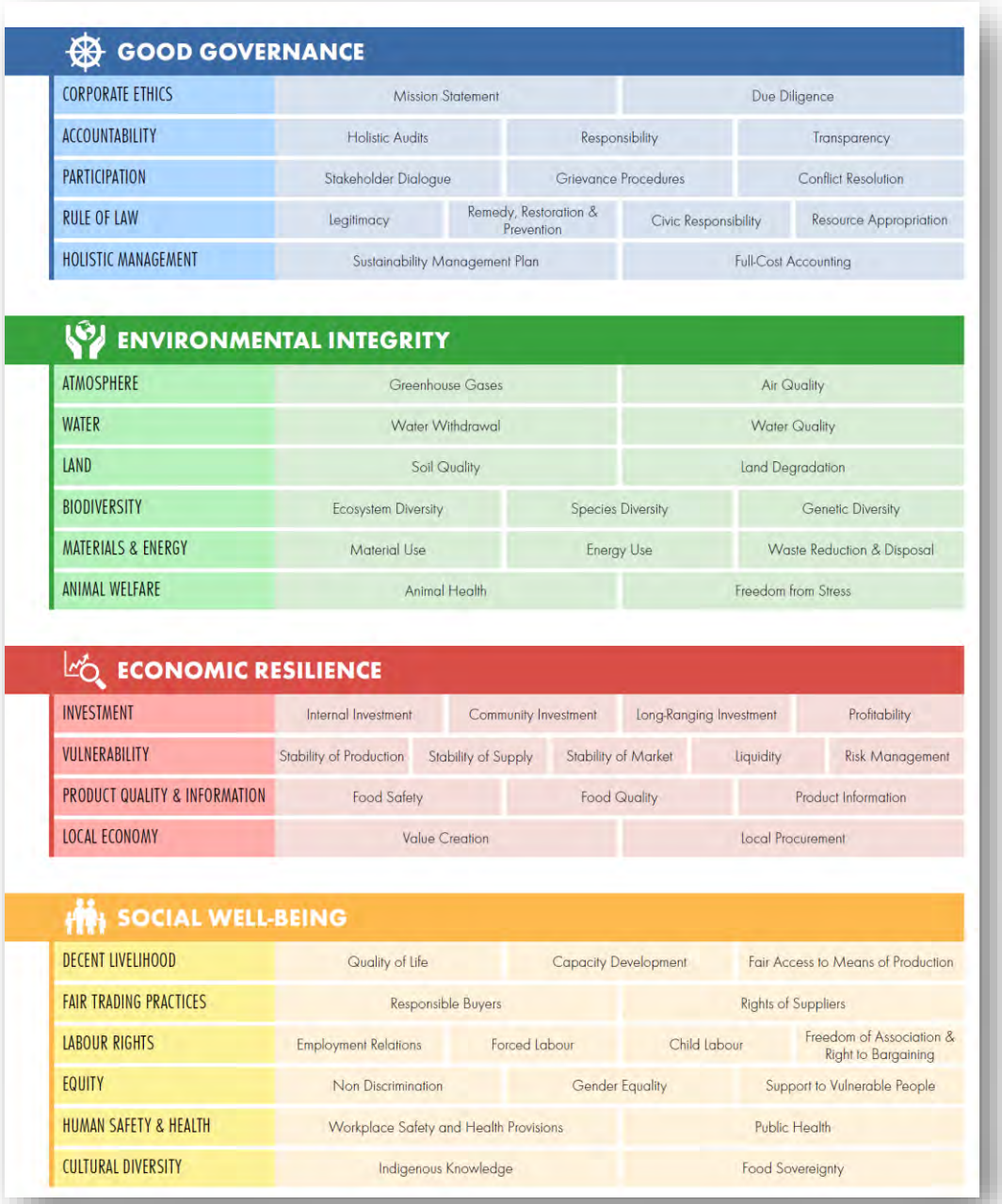
4 Dimensions, 21 Themes, 58 Subthemes

For each theme and subtheme:

- Objective
- Descriptions
- Suggested indicators (quantitative, qualitative; policy, practice, performance)

Aspiration to be:

- Universal (globally applicable)
- Holistic (combat partial truths/reporting)
- Legitimate (developed by major multilateral organization)



smart

sustainability monitoring
and assessment routine

smart

sustainability monitoring and assessment routine

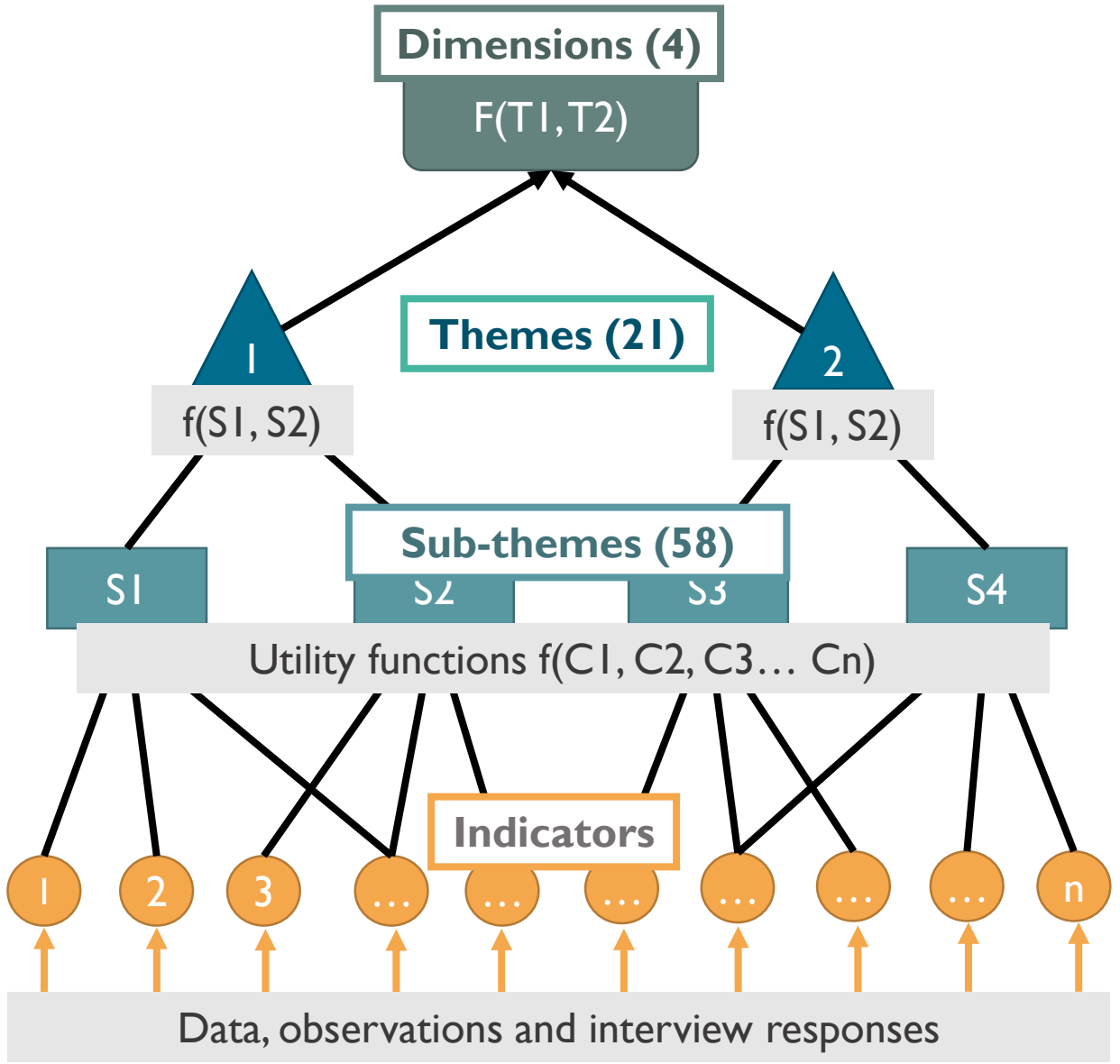
SAFA: Theory

SMART: Praxis

Multi-Criteria Assessment structure

- 327 simple indicators, multiple interactions
- Simple aggregation method: **weighted average** (compensation, commesurability)
- Indicators standardized (0-100% range)
- **Distance-to-target** MCDA method
- No aggregation beyond **SAFA theme**

Governance	Environmental	Social	Economic
Work permits, audits, sourcing policies	Extensive/semi-natural areas	Wage level, work overload, equal pay	Adequate liquidity, access to credit
Cases of environmental or social impacts	Pesticide use (active substances, toxicity)	Collective bargaining, unionisation	Diversification of income, collective marketing
Conflict prevention mechanisms	Fertilization and soil management	Inputs from social hotspot countries	Secure land tenure, succession



Indicator weights consensus method

Delphi and Nominal Group Technique (NGT)

Consensus Methods: Characteristics and Guidelines for Use

Arlene Fink, PhD, Jacqueline Kosecoff, PhD, Mark Chassin, MD, MPP, MPH, and Robert H. Brook, MD, ScD

Is There a Consensus on Consensus Methodology? Descriptions and Recommendations for Future Consensus Research

Jane Waggoner, MS, Jan D. Carline, PhD, and Steven J. Durning, MD, PhD

Abstract

Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2015, **6**, 1097–1109 doi: 10.1111/2041-210X.12387

The Delphi technique in ecology and biological conservation: applications and guidelines

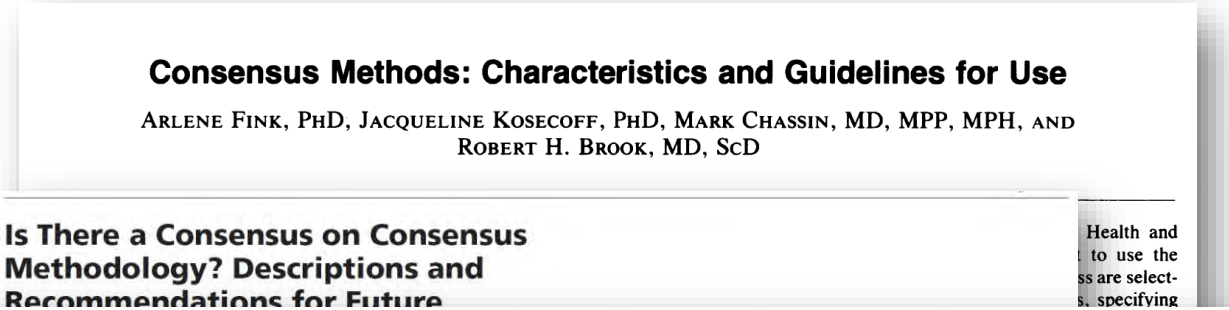
Nibedita Mukherjee[1,2,3]*, Jean Hugé[3,4], William J. Sutherland[1], Jeffrey McNeill[5], Maarten Van Opstal[3,6,7], Farid Dahdouh-Guebas[2,3,†] and Nico Koedam[2,†]

[1]Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK; [2]Plant Biology & Nature Management, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium; [3]Laboratory of Systems Ecology and Resource Management, Université Libre de Bruxelles, CP 264/1, Avenue F.D. Roosevelt 50, 1050 Brussels, Belgium; [4]Centre for Sustainable Development, Ghent University, Poel 16, 9000 Gent, Belgium; [5]School of People, Environment and Planning, Massey University, Private Bag 11222, Palmerston North, New Zealand; [6]Center of Cultural Anthropology, Université Libre de Bruxelles, CP 124, Avenue F.D. Roosevelt 50, 1050 Brussels, Belgium; and [7]Public Health Department, Vrije Universiteit Brussel, Laarbeeklaan 103, 1090 Brussels, Belgium

Indicator weights consensus method

Delphi and Nominal Group Technique (NGT)

- Crabbe et al. (2009)



Consensus Methods: Characteristics and Guidelines for Use

ARLENE FINK, PHD, JACQUELINE KOSECOFF, PHD, MARK CHASSIN, MD, MPP, MPH, AND ROBERT H. BROOK, MD, SCD

Is There a Consensus on Consensus Methodology? Descriptions and Recommendations for Future

Health and
to use the
ss are select-
s, specifying

Table 2. Comparison of the Delphi technique to other commonly used techniques in group decision-making. Anon, Anonymous. + means that the method is effective for achieving the set objective, – means that the method is not suited for achieving the set objectives. The Delphi technique combines the benefits of group discussion, iterations and anonymity without the added constraint of geographic proximity.

	Method	Possibility of iterations	Impacted by social pressure	Suitable for conflict issues	Possibility of anon. discussion	Possibility of voting	Possibility of anon. voting	Requirement of geographic proximity	Requirement of skilled facilitator	Requirement of expert judgment	References
No discussion possible	Questionnaire	–	–	–	–	+	–	+	–	–	White *et al.* (2005)
	Statistical Aggregation	–	–	–	–	–	–	–	–	–	Kerr & Tindale (2011)
	Confidential voting	–	–	+	–	+	+	–	–	–	Redpath *et al.* (2004)
	Public Voting	–	+	–	–	+	–	–	–	–	Burgman *et al.* (2014)
	Prediction markets	+	–	+	–	–	–	–	–	–	Kerr & Tindale (2011)
Discussion possible	Focus group discussion	–	+	–	–	+	–	+	+	+	Fischer & Young (2007)
	Nominal Group technique	+	+	–	–	+	–	+	+	+	Sutherland (2006)
	Delphi technique	+	–	+	+	+	+	–	+	+	Hasson, Keeney & McKenna (2000)

Indicator weights consensus method

Delphi

- Anonymous rating of indicators
- Summary results of ratings
- Exchange of justifications

NGT

- Physical discussion and voting (non-anonymous)
- Summary results of voting

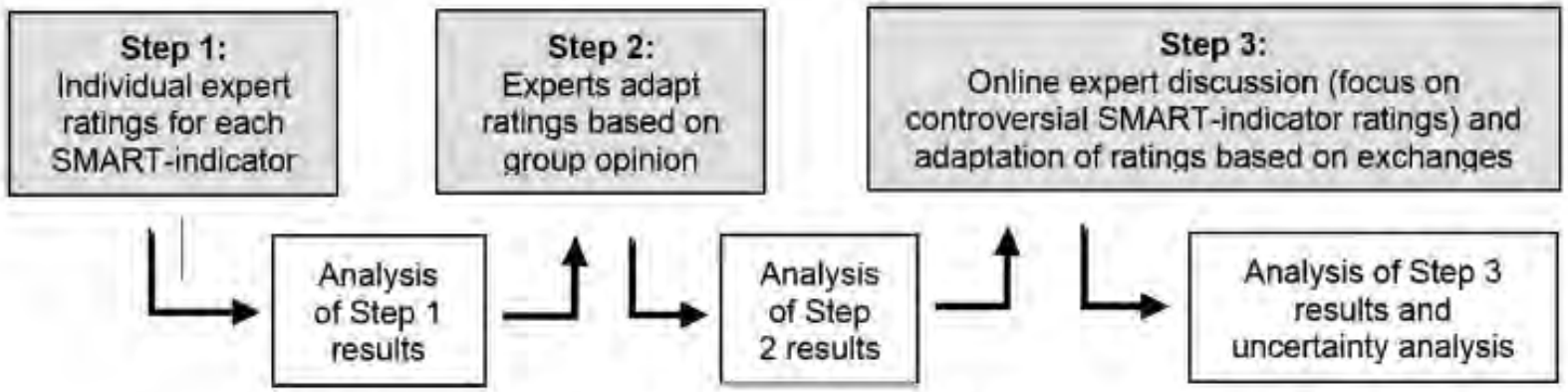
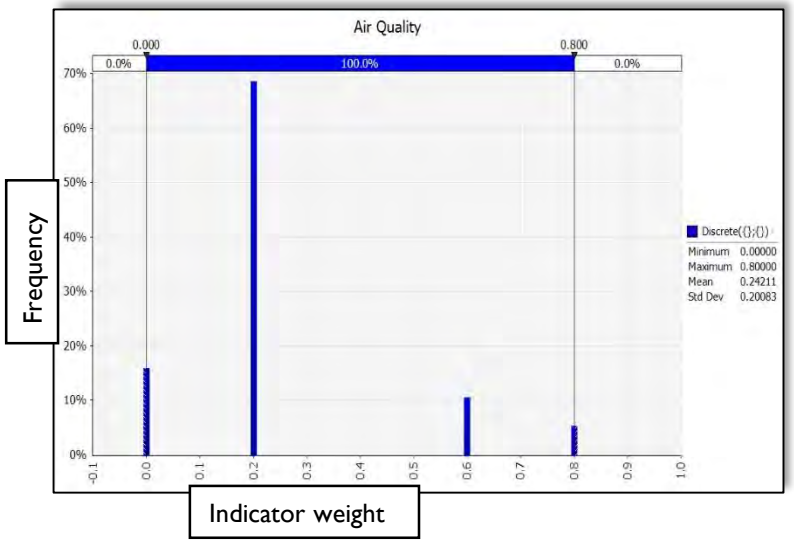
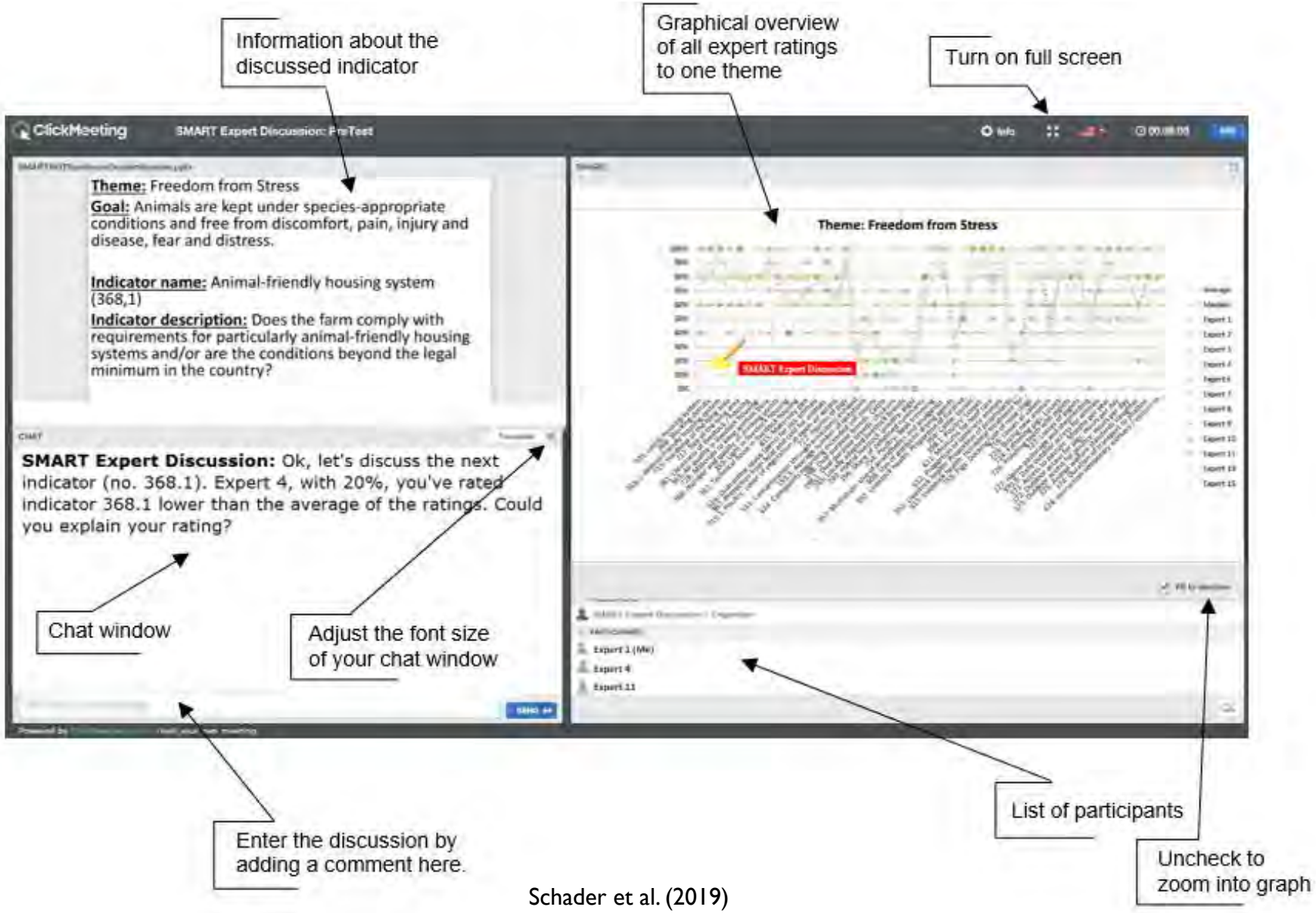


Fig. 1. Overview of steps of the hybrid Delphi/NGT process implemented in this study.



Indicator weights consensus method

"ClickMeetings" as NGT element

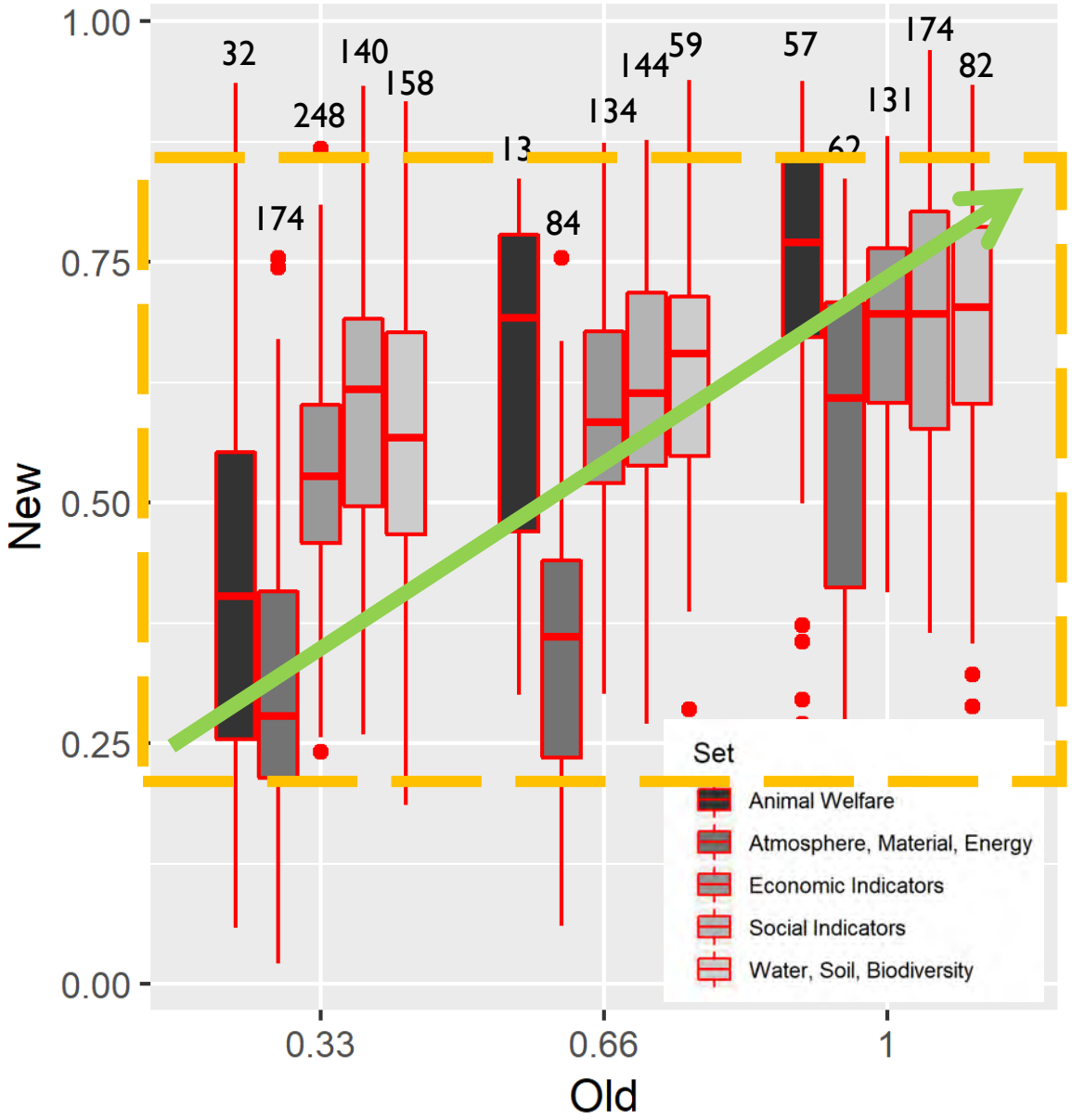


Schader et al. (2019)

Indicator weights consensus method

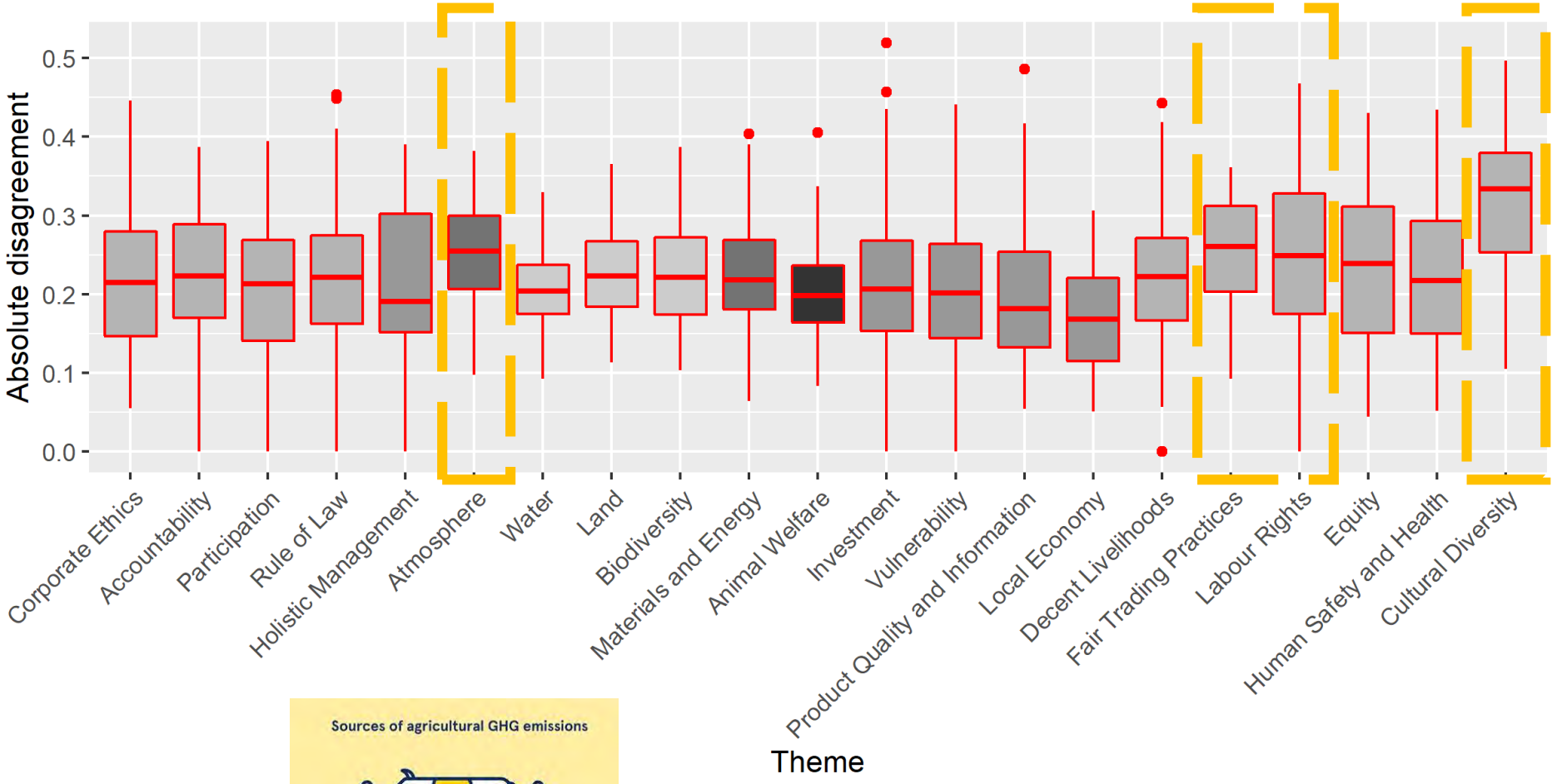
Validation

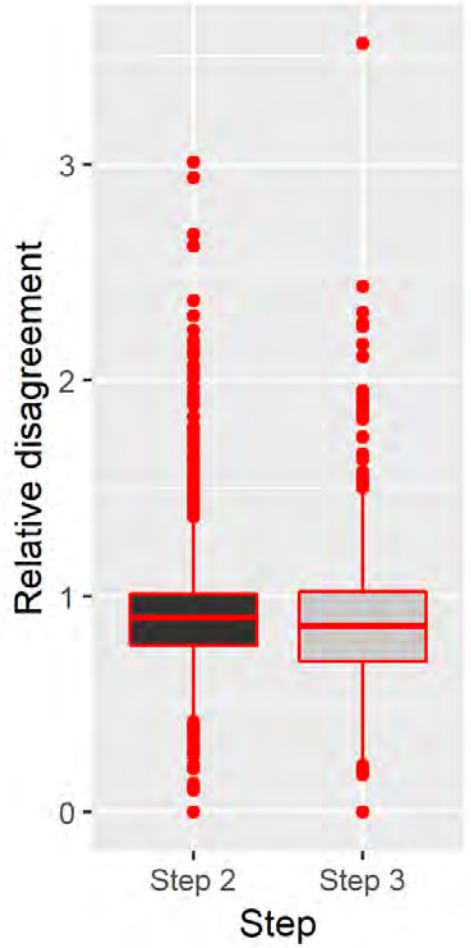
- Dataset of preliminary indicator weights ("Old") taken from an independent set of FiBL experts (ranging -3 to +3)
- Correlation with the final values ("New") provided by the Delphi/NGT process (ranging -10 to +10)
- "Flattening" of values when provided with a larger (10-point) scale - sign of fatigue? (Saisana 2015)



Uncertainty assessment

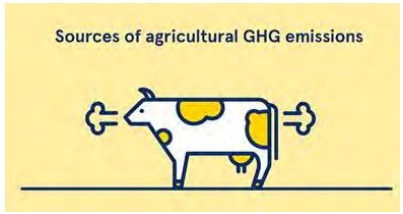
Patterns in uncertainty (s.d. of expert weights)





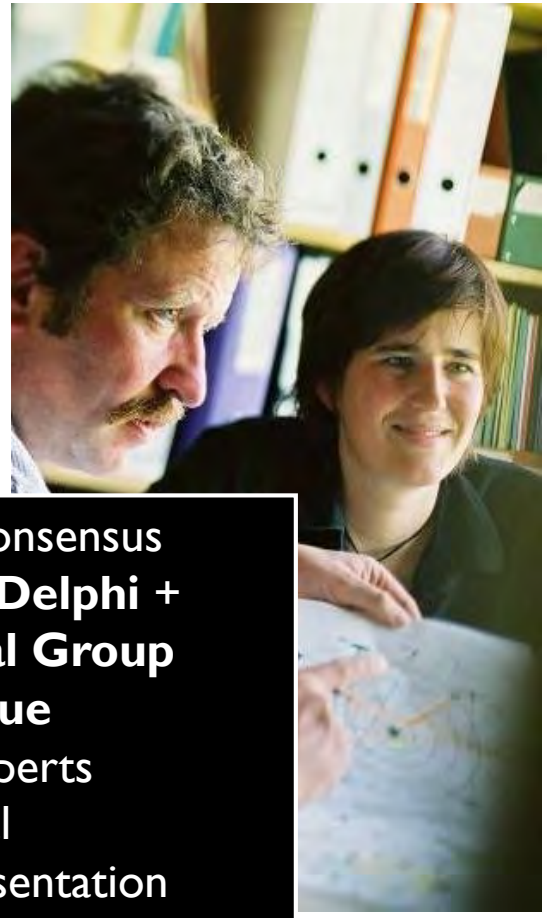
-11%* -14%

Consensus building



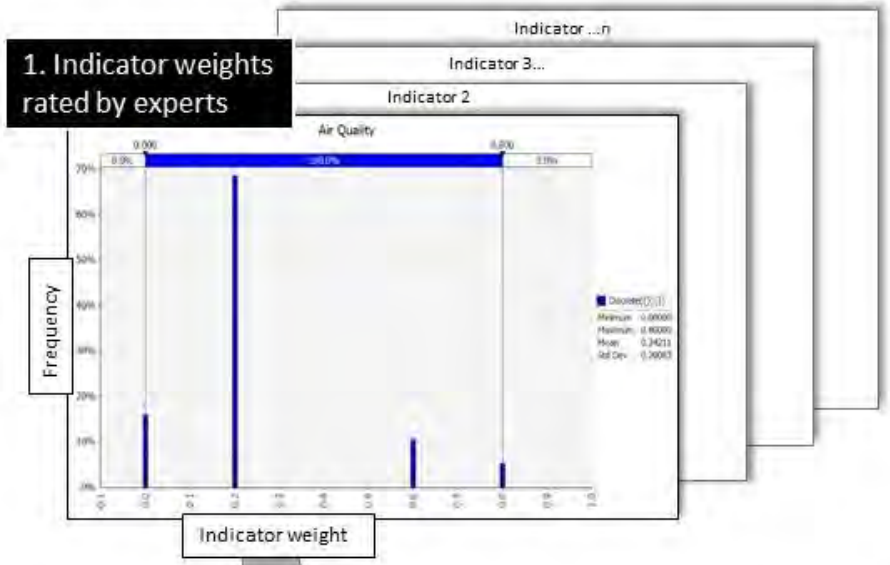
FiBL

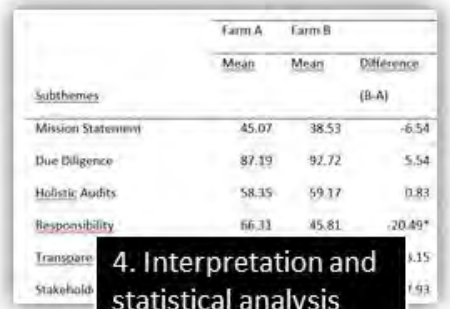
Uncertainty assessment



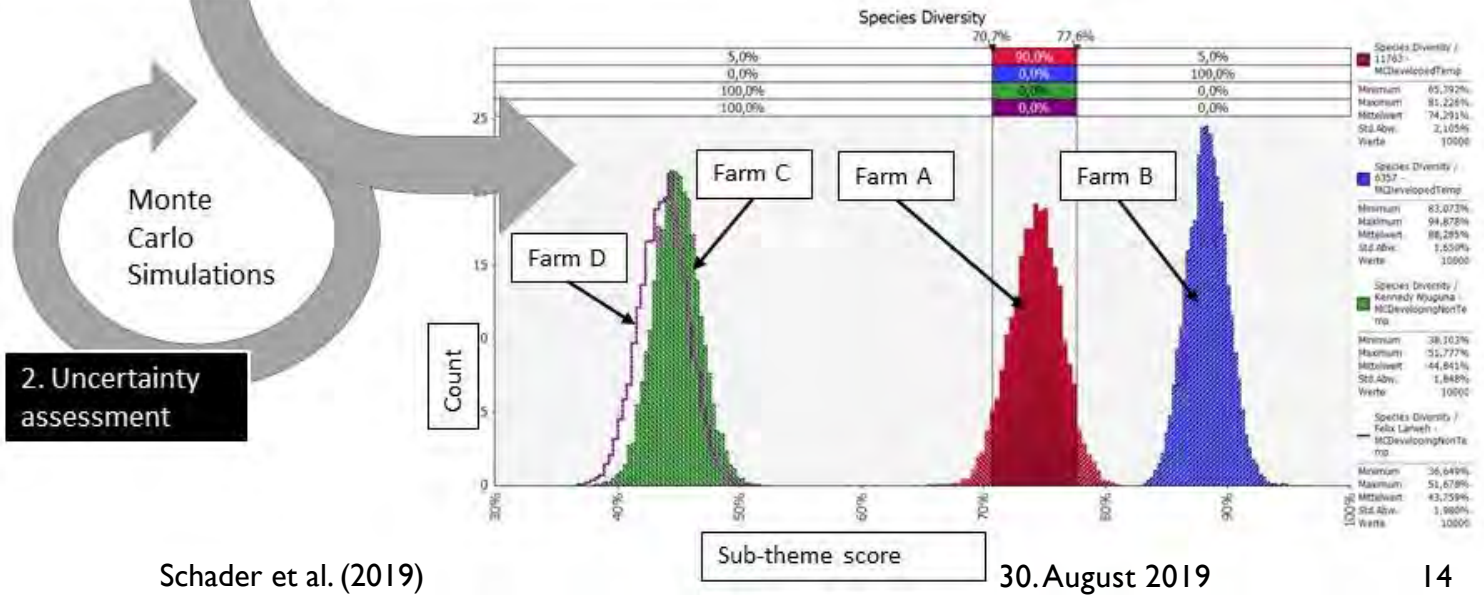
Group consensus method: **Delphi + Nominal Group Technique**

- 64 experts
- Global representation
- Separate assessment for each SAFA subtheme





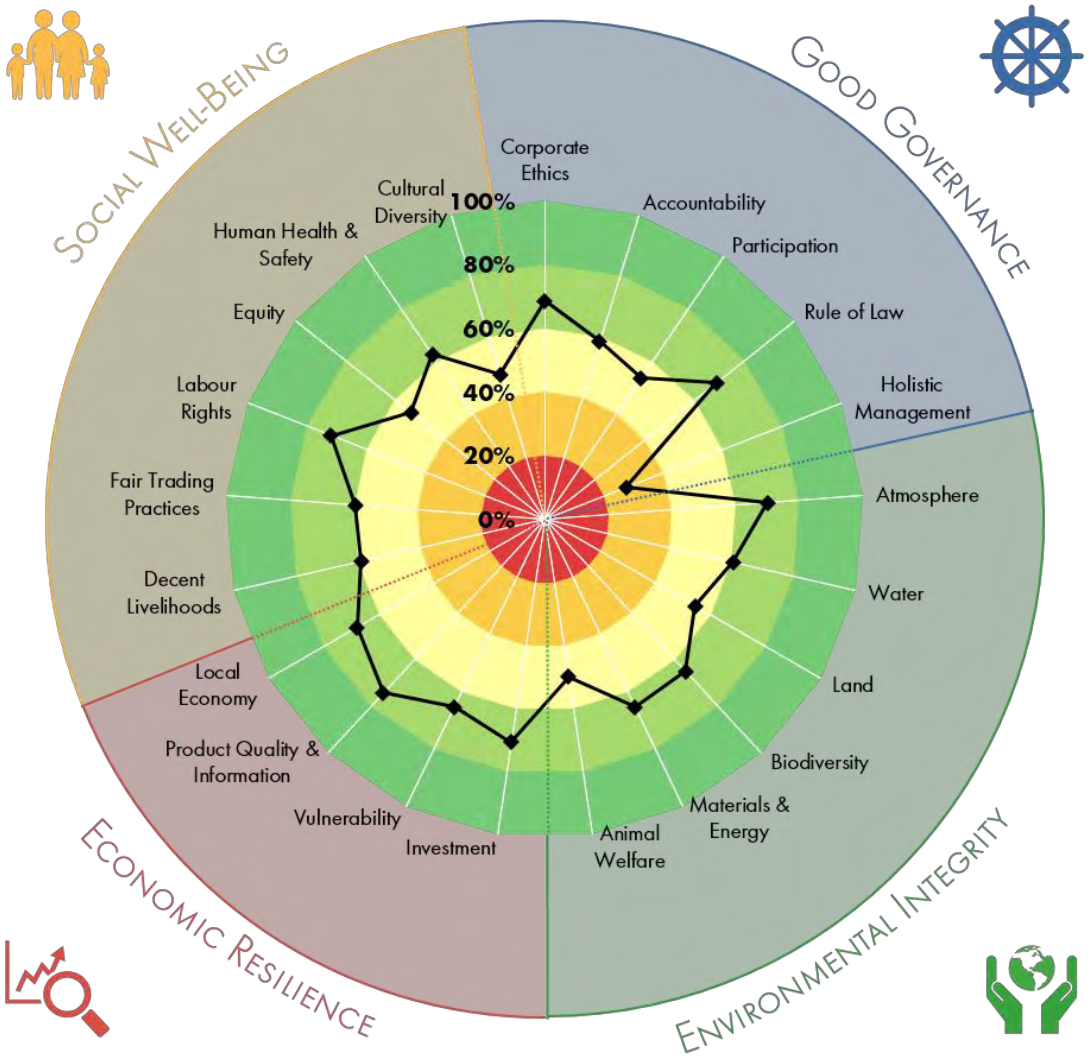
3. Farm-level comparisons per sub-theme



Uncertainty assessment

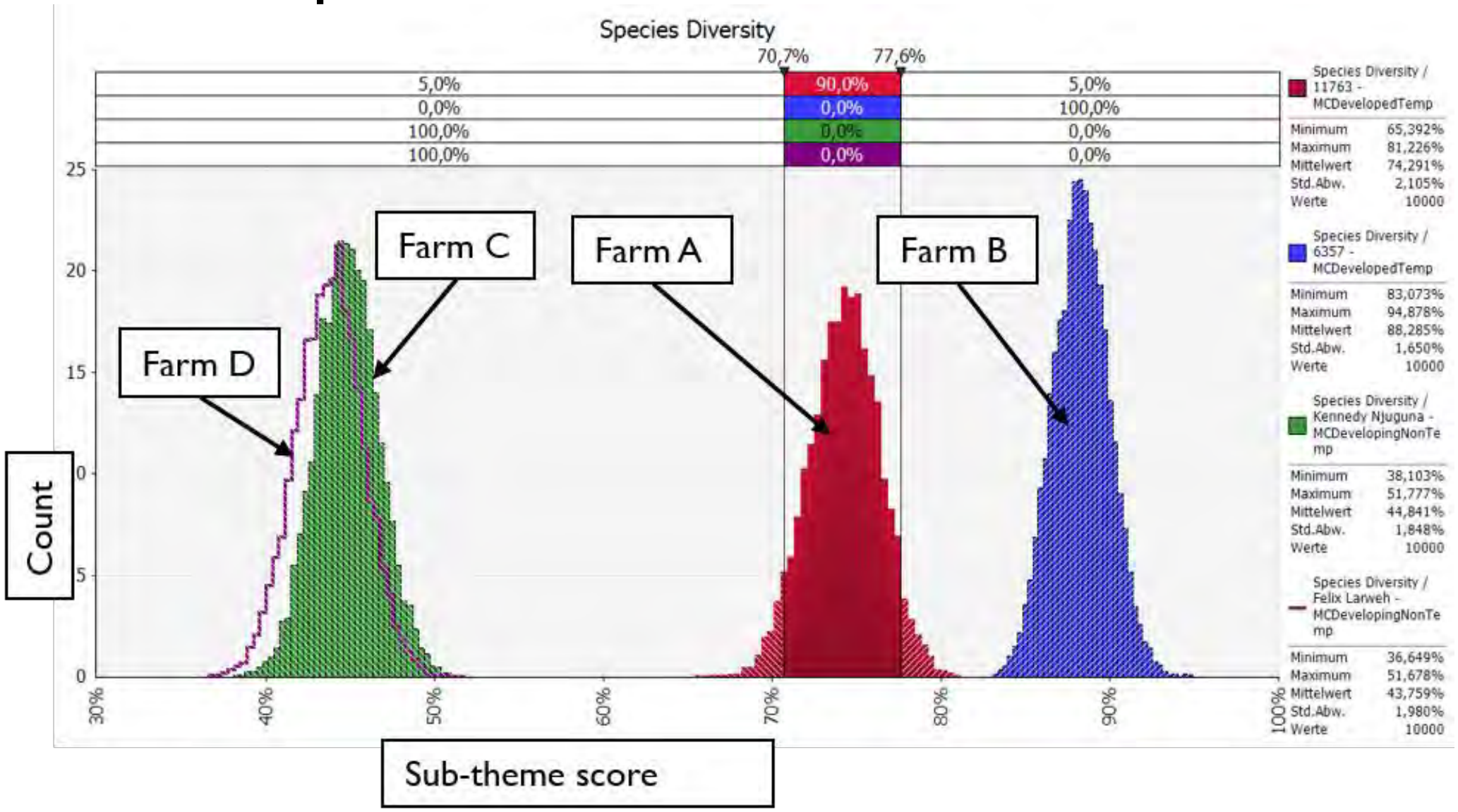
Monte Carlo Simulation

- Refsgaard et al. (2007)
- @ Risk Software
- Error propagation to simulate uncertainty distribution in model output
- Outputs of SMART are SAFA subtheme performance scores ranging 0-100%



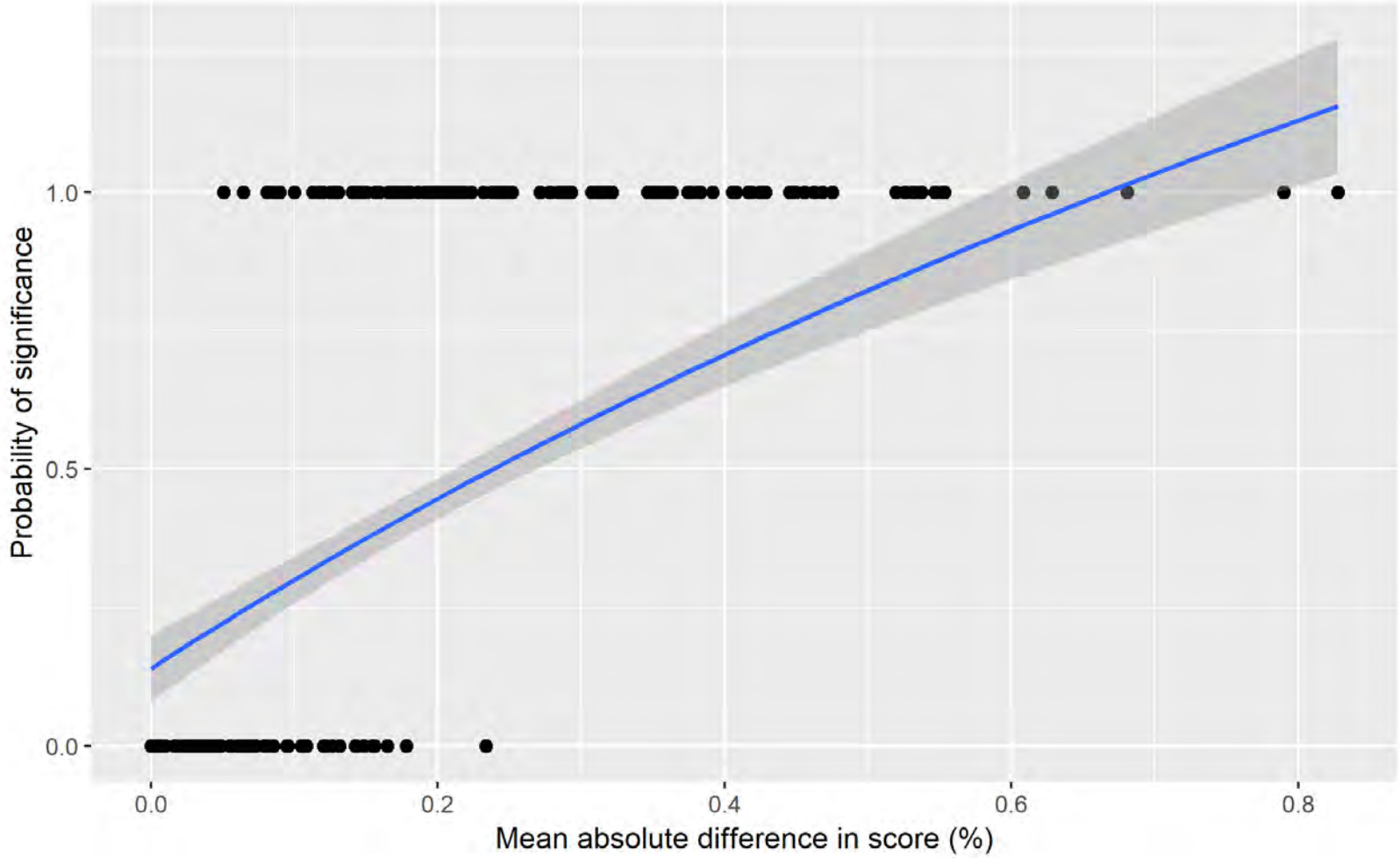
Palisade Corporation. 2009. Guide to using@ RISK.: Risk analysis and simulation add-in for Microsoft Excel. https://www.palisade.com. USA, Newfield NY.

Effects on farm comparisons



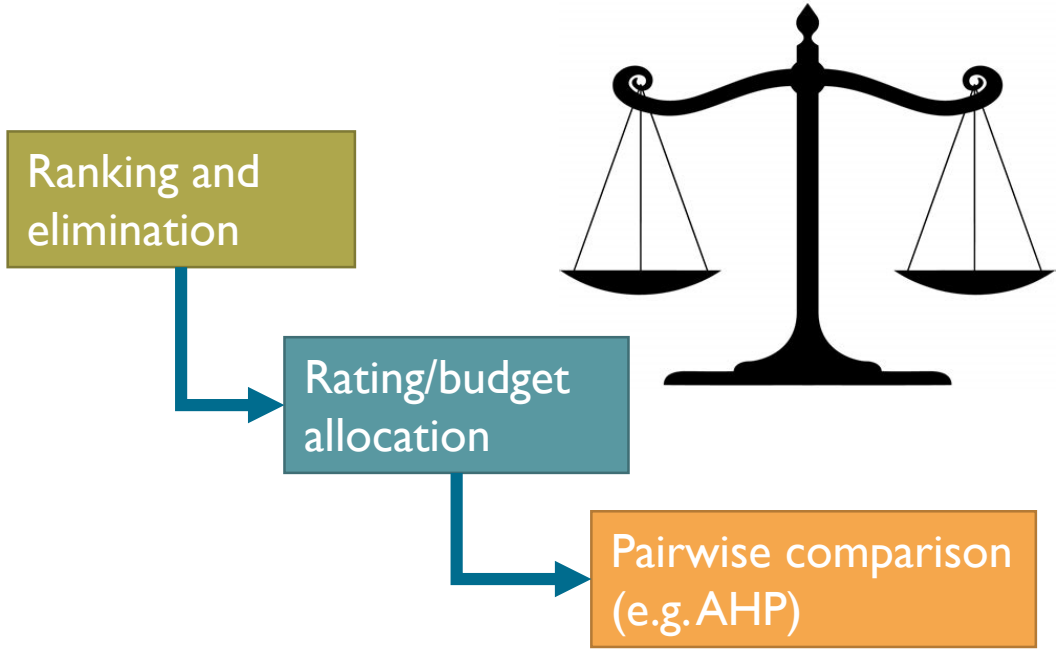
Effects on farm comparisons

- Four example farms (Schader et al. 2016)
  - 2 farms in Europe (Livestock and Arable)
  - 2 farms in Sub-Saharan Africa (Livestock and Agroforestry)
- Non-significant differences between farms ranging up to 20%
- High uncertainty in the governance dimension
- **Mean of ca. 6%**



Conclusions

- Weights could be validated with separate dataset
- Adding NGT element to Delphi process improved consensus (by ca. 3%)
- Additional time & effort probably not worth it
- As expected, high levels of uncertainty in social/cultural themes, but also in air quality/GHG (!)
- Consistency and quality of consensus process could have been improved
  - Balanced and manageable indicator sets
  - Quantify consistency of expert ratings



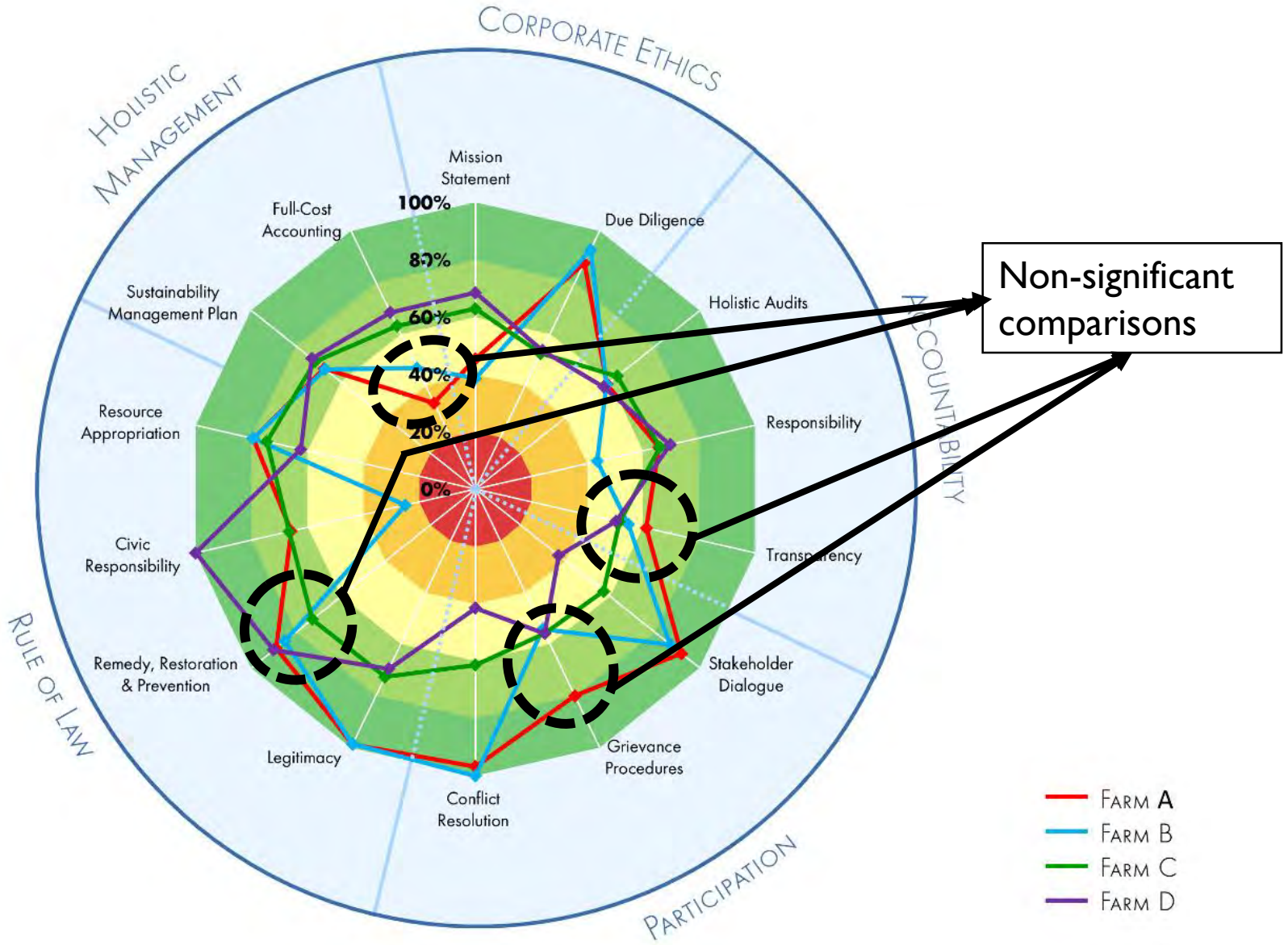
Indicator A	Criterion C2.1																	Indicator B
I2.1.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	I2.1.2
I2.1.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	I2.1.3
I2.1.1	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	I2.1.4
I2.1.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	I2.1.3
I1.1.2	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	I2.1.4
I1.1.3	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	I2.1.4

The shaded squares are the ones chosen by the expert to represent the relationship between the 2 indicators being compared.

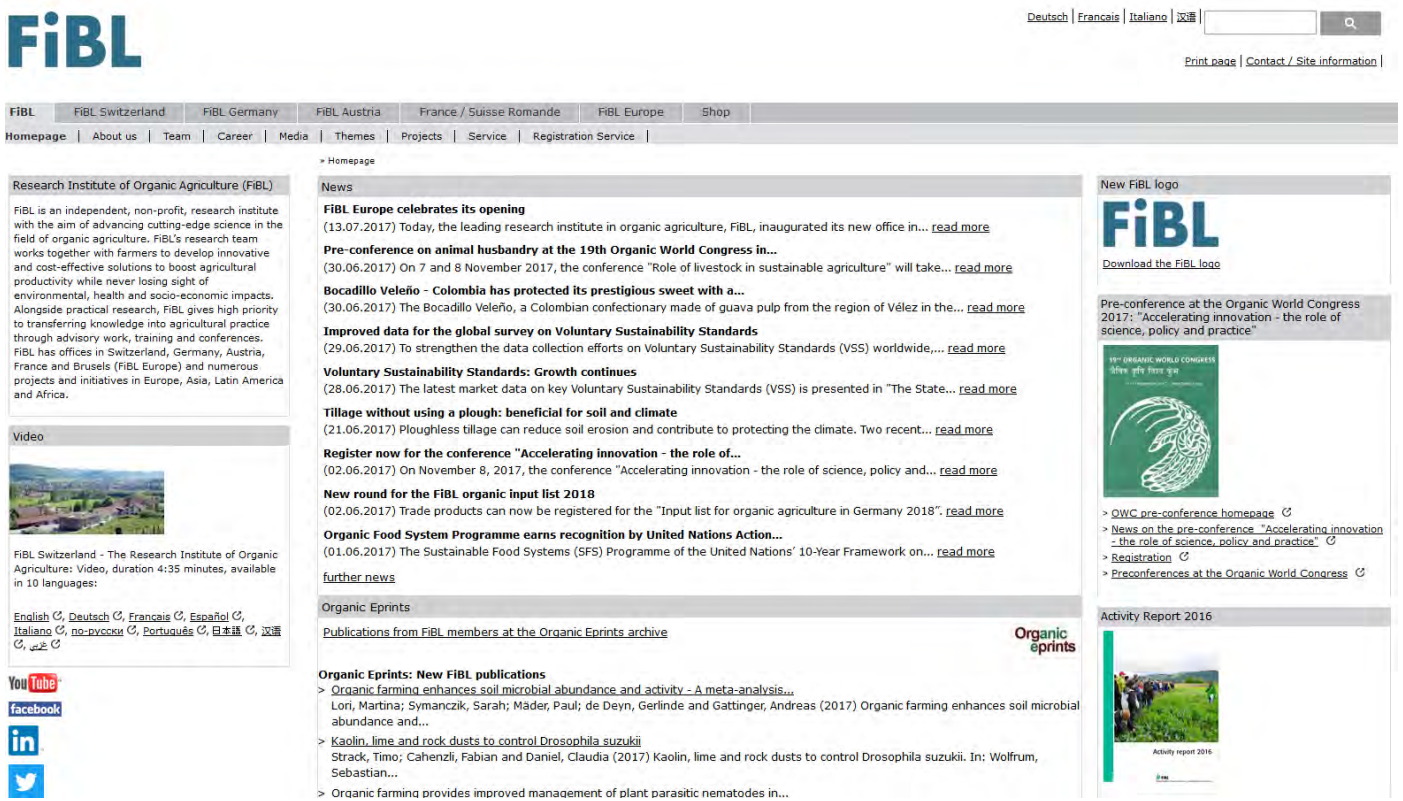
Saaty's 9 point scale

Conclusions

- Sustainability assessment tools rarely calculate uncertainty
- Clear need, although not always straightforward with expert judgement
- Present method estimates and integrates parameter uncertainty
- Users of such tools should be informed of uncertainty



Thank you for your attention!



michael.curran@fibl.org

Research Institute of Organic Agriculture FiBL

Ackerstrasse 113 / Postfach 219

5070 Frick

Switzerland

Phone +41 62 8657-272

Fax +41 62 8657-273

info.suisse@fibl.org

www.fibl.org

References

Crabbe, M. J. C., E. Martinez, C. Garcia, J. Chub, L. Castro, and J. Guy. 2009. Is Capacity Building Important in Policy Development for Sustainability? A Case Study Using Action Plans for Sustainable Marine Protected Areas in Belize. Society & Natural Resources 23:181–190.

FAO. 2014. Sustainability Assessment of Food and Agriculture Systems (SAFA) Guidelines, Vers. 3. Food and Agricultural Organization (FAO), Rome.

Fink, A., J. Kosecoff, M. Chassin, and R. H. Brook. 1984. Consensus methods: characteristics and guidelines for use. American Journal of Public Health 74:979–983.

Mukherjee, N., J. Hugé, W. J. Sutherland, J. McNeill, M. Van Opstal, F. Dahdouh-Guebas, and N. Koedam. 2015. The Delphi technique in ecology and biological conservation: applications and guidelines. Methods in Ecology and Evolution 6:1097–1109.

Saaty, T. L. 1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15:234–281.

Saaty, T. L. 2008. Decision making with the analytic hierarchy process. International Journal of Services Sciences 1:83.

Saisana, M. 2015. Impact Assessment Tools: Multi-criteria Analysis (the Maximum Likelihood Approach). European Commission Joint Research Centre, Econometrics and Applied Statistics Unit.

Schader, C., L. Baumgart, J. Landert, A. Muller, B. Ssebunya, J. Blockeel, R. Weisshaidinger, R. Petrasek, D. Mészáros, S. Padel, C. Gerrard, L. Smith, T. Lindenthal, U. Niggli, and M. Stolze. 2016. Using the Sustainability Monitoring and Assessment Routine (SMART) for the Systematic Analysis of Trade-Offs and Synergies between Sustainability Dimensions and Themes at Farm Level. Sustainability 8:274.

Schader, C., M. Curran, A. Heidenreich, J. Landert, J. Blockeel, L. Baumgart, B. Ssebunya, S. Moakes, S. Marton, G. Lazzarini, U. Niggli, and M. Stolze. 2019. Accounting for uncertainty in multi-criteria sustainability assessments at the farm level: Improving the robustness of the SMART-Farm Tool. Ecological Indicators 106:105503.

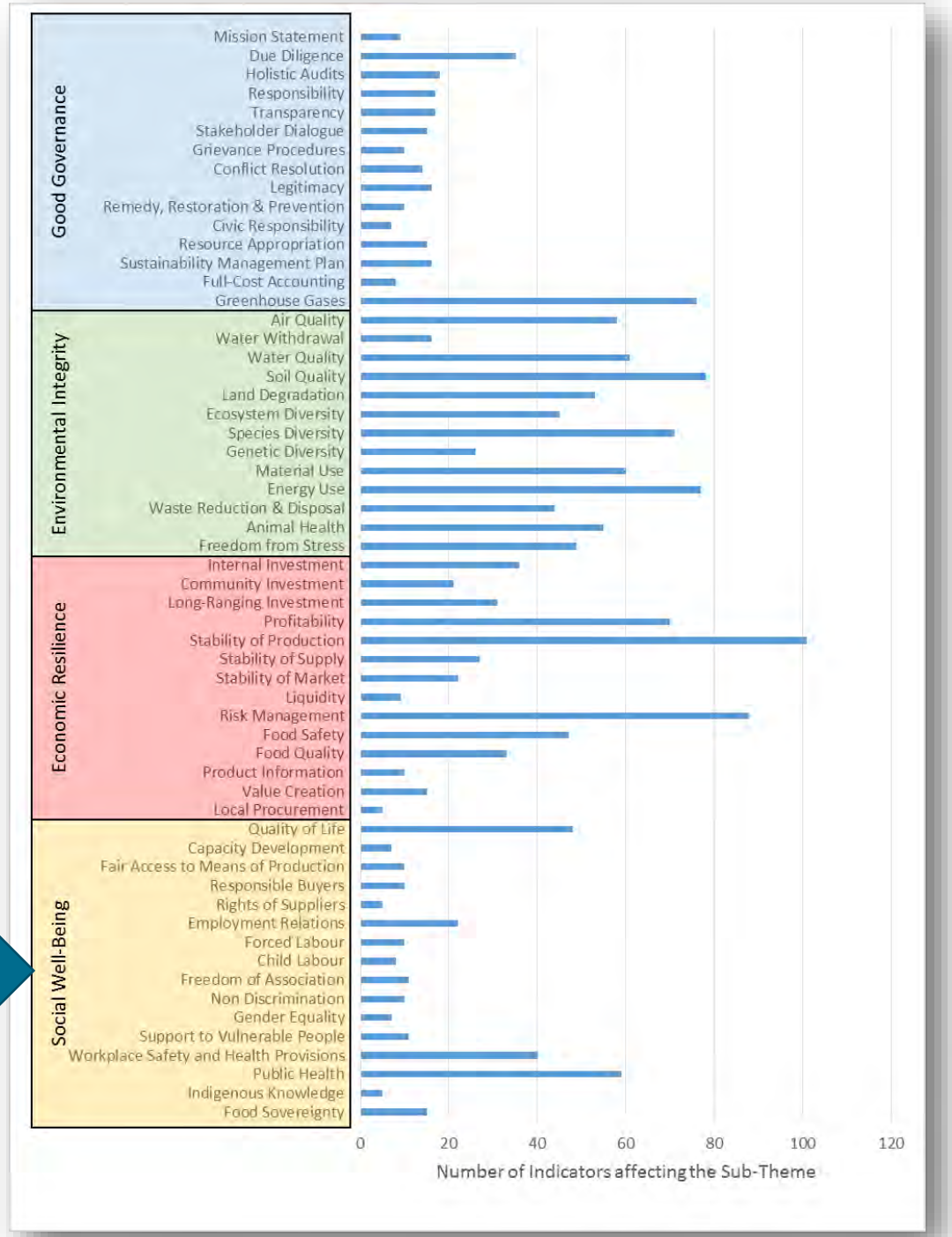
Waggoner, J., J. D. Carline, and S. J. Durning. 2016. Is There a Consensus on Consensus Methodology? Descriptions and Recommendations for Future Consensus Research: Academic Medicine 91:663–668.

Extra slides

Indicator selection

- Indicators influence multiple subthemes (+ve/-ve)
- >1700 indicator weights (influences) between indicators and SAFA sub-themes
- Ca. 30 indicators per SAFA sub-theme
- Emphasis on simple, easy to measure indicators

- Work overload
- Relationship with suppliers
- Local procurement of inputs
- Self-sufficiency, local/direct sales
- Equipment and safety protocols
- Level of mechanization
- Working hours, overtime compensation
- Formal contracts, collective bargaining
- Child labour, forced labour



Schader et al. (2016)

SMART methodology – aggregation

Sub-theme for which a SAFA objective is defined

Impacts of each indicator in the sub-theme

Capacity Development	Impact Weight	Indicator Rating	Impact weight * score	Sustainability Score at the Sub-theme level
Further training for farm staff	0.93	100%	0.93	44.36%
Apprenticeships	0.71	100%	0.71	
Training on sustainability	0.73	0%	0.00	
Employees: Access to external training	0.85	25%	0.21	
Access to advisory services	0.91	25%	0.23	
Education and training at master level (master crafts men places)	0.56	0%	0.00	
	4.69		2.08	

Relevant indicators impacting on degree of goal achievement

Sum of impacts = maximum achievable points

Sum of impacts * actual performance = achieved points

Data collection on farm

- Similar to a compliance check
- Familiarity and acceptance by farmers
- Semi-structured interview
  - 2-3 hours for a small to medium sized farm (<100 ha)
  - Up to several days for large farms (>1000 ha)