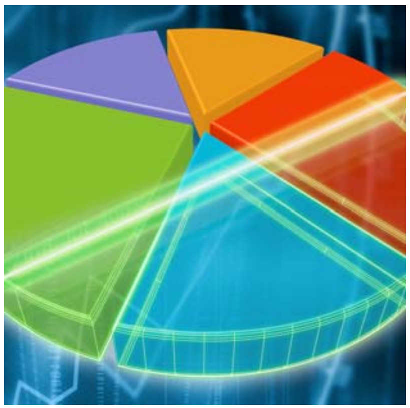
A Multi –Criteria-Based Index for the Knowledge Economy in the EU25



Outline
- General aggregation issues
- Additive aggregation
- Geometric aggregation
- MCA
- The KBE Index for the EU15 + JP + US
- When to use what

Michaela Saisana
European Commission
Joint Research Centre of Ispra, Italy
michaela.saisana@jrc.it

Fourth Workshop on Indicators in the Knowledge Economy (KEI),
Leuven,
September 5-6, 2006

Prepared with Giuseppe Munda

Based on:

•**Handbook on Constructing Composite Indicators: Methodology and User Guide (2005).** Nardo, M. M. Saisana, A. Saltelli and S. Tarantola (EC/JRC), A. Hoffman and E. Giovannini (OECD), OECD Statistics Working Paper JT00188147, STD/DOC(2005)3. http://www.olis.oecd.org/olis/2005doc.nsf/LinkTo/std-doc(2005)3

•Munda M. and Nardo M. (2005) Constructing Consistent Composite Indicators: the Issue of Weights, manuscript submitted to Economics Letters.

• Munda G. and Nardo M. (2005) Non-Compensatory Composite Indicators for Ranking Countries: A Defensible Setting, manuscript submitted to Economica.

• Munda G. (2005) Social Multi-Criteria Evaluation (SMCE): Methodological Foundations and Operational Consequences, forthcoming, J. of Operational Research.

Step 6 in the Handbook:
(Weighting and) aggregation

Aggregation rules:

- **Linear aggregation**
- **Geometric mean**
- **Multi-criteria analysis**

Additive aggregation

summation of ranks

$$CI_c = \sum_{q=1}^{Q} Rank_{qc}$$

- the simplest method
- based on ordinal information & independent to outliers BUT loses the absolute value information.

number of indicators that are above and below some benchmark

$$CI_c = \sum_{q=1}^{Q} \cdot sgn\left[\frac{I_{qc}}{I_{EUq}} - (1+p)\right]$$

- uses nominal scores
- threshold value *p* arbitrarily chosen
- Simple & unaffected by outliers BUT loses interval level information.

summation of weighted and normalized indicators

$$CI_c = \sum_{q=1}^{Q} w_q I_{qc}$$

$$\sum_q w_q = 1 \text{ and } \quad 0 \le w_q \le 1$$

- By far the most widespread method
- entails restrictions on the nature of indicators & weights
- implies full (and constant) compensability
- rewards indicators proportionally to the weights
- requires normalisation
- weights are trade offs not importance coefficients

Additive aggregation

Example: Human Poverty Index 2001

$HPI = [1/3\,(P_1^{\alpha} + P_2^{\alpha} + P_3^{\alpha})]^{1/\alpha}$; $\alpha = 3$

- P_1 = Probability at birth of not surviving to age 40
- P_2 = Adult illiteracy rate
- P_3= Unweighted average of population without sustainable access to an improved water source and children under weight for age

The 'cubing' i.e. a=3 ensures greater weight for the component with acute deprivation

Additive aggregation

Example: Gender Development Index 2001

3 dimension indices calculated for males and females and combined, *penalizing* differences in achievement

Equally distributed index =

$$\{[\text{female popn. share (female index}^{1-\varepsilon})] + [\text{male popn. share (male index}^{1-\varepsilon})]\}^{1/1-\varepsilon}$$

where $\varepsilon = 2$ (moderate penalty for gender inequality)

Additive aggregation - Linear

Restrictions and assumptions

- Indicators need to be ***mutually preferentially independent*** (i.e. every subset of these indicators is preferentially independent of its complementary set of indicators) → very strong condition from both the ***operational*** and ***epistemological*** points of view.

- *Compensability* among the indicators is always assumed → complete substitutability among the various indicators

  E.g. in a sustainability index, economic growth can always substitute any environmental destruction or inside e.g., the environmental dimension, clean air can compensate for a loss of potable water. ***From a descriptive point of view, such a complete compensability is often not desirable***

- ***Weights*** have the meaning of trade-off ratio. Yet, in all constructions of a composite indicator, weights are used as importance coefficients, as a consequence, ***a theoretical inconsistency exists.***

- ***Synergy or conflict -*** Preferential independence implies that the trade-off ratio between two indicators is independent of the values of the n-2 other indicators

Additive aggregation - Linear

Example

A hypothetical composite: inequality, environmental degradation, GDP per capita and unemployment

Country A: 21, 1, 1, 1 → 6
Country B: 6, 6, 6, 6 → 6

Obviously the two countries would represent very different social conditions that would not be reflected in the composite.

Geometric aggregation

$$CI_c = \prod_{q=1}^{Q} x_{q,c}^{w_q}$$

Example

A hypothetical composite: inequality, environmental degradation, GDP per capita and unemployment

Country A: 21, 1, 1, 1 → 2.14
Country B: 6, 6, 6, 6 → 6

• Countries with low scores in some indicators would prefer a linear rather than a geometric aggregation (the simple example above shows why).
• Yet, the marginal utility from an increase in low absolute score would be much higher than in a high absolute score under geometric aggregation

Country A: 21, 2, 1, 1 → 2.54 → 19% increase in the score
Country B: 6, 7, 6, 6 → 6.23 → 4% increase in the score

Lesson: a country should be more interested in increasing those sectors/ activities/ alternatives with the lowest score in order to have the highest chance to improve its position in the ranking if the aggregation is geometric rather than linear (Zimmermann and Zysno, 1983).

The absence of synergy or conflict effects

among the indicators & weights express trade-offs between indicators

are necessary conditions to admit

either **linear or geometric aggregation**

Multi-criteria type of aggregation

When different goals are equally legitimate and important, then a non compensatory logic may be necessary.

Example: physical, social and economic figures must be aggregated. If the analyst decides that an increase in economic performance can not compensate a loss in social cohesion or a worsening in environmental sustainability, then neither the linear nor the geometric aggregation are suitable.

Instead, a non-compensatory ***multicriteria approach*** will assure non compensability by formalizing the idea of finding a compromise between two or more legitimate goals.

+ does not reward outliers

+ different goals are equally legitimate and important

+ no normalisation is required

BUT

- computational cost when the number of countries is high

Multi-criteria type of aggregation

(Munda 2003, Munda & Nardo 2003)

	Indic.	GDP	Unemp. Rate	Solid wastes	Income disp.	Crime rate
Country		↑	↓	↓	↓	↓
A		25,000	0.15	0.4	9.2	40
B		45,000	0.10	0.7	13.2	52
C		20,000	0.08	0.35	5.3	80
weights		0.165	0.165	0.333	0.165	0.165

	A	B	C
A	0	0.666	0.333
B	0.333	0	0.333
C	0.666	0.666	0

ABC = 0.666 + 0.333 + 0.333 = 1.333

BCA = 0.333 + 0.666 + 0.333 = 1.333

CAB = 0.666 + 0.666 + 0.666 = 2

ACB = 0.333 + 0.666 + 0.666 = 1.666

BAC = 0.333 + 0.333 + 0.333 = 1

CBA = 0.666 + 0.333 + 0.666 = 1.666

Linear aggregation: CBA

Multi-criteria type of aggregation

The Computational problem

Moulin (1988, p. 312) clearly states that the Kemeny method is ***"the correct method"*** for ranking alternatives, and that the *"only drawback of this aggregation method is the difficulty in computing it when the number of candidates grows"*.

With only 10 countries → 10! = 3,628,800 permutations

A NP-hard problem

The *complexity class* of *decision problems* that are intrinsically harder than those that can be solved by a *nondeterministic Turing machine* in *polynomial time*. When a decision version of a combinatorial *optimization problem* is proved to belong to the class of *NP-complete* problems, then the optimization version is NP-hard.

(definition given by the National Institute of Standards and Technology, http://www.nist.gov/dads/HTML/nphard.html)

Multi-criteria type of aggregation

This NP-hardness has discouraged the development of algorithms searching for exact solutions, thus the majority of the algorithms which have been proposed in the literature; are mainly

- ***heuristics based on artificial intelligence,***
- ***branch and bound approaches and***
- ***multi-stage techniques***

(see e.g., Barthelemy et al., 1989; Charon et al.,1997; Cohen et al., 1999; Davenport and Kalagnam, 2004; Dwork et al., 2001; Truchon, 1998b).

Multi-criteria type of aggregation

A ***new numerical algorithm*** aimed at solving the computational problem connected to linear median orders by finding ***exact solutions*** has been proposed by Munda (2005).

- linear median orders are computed by using their ***theoretical equivalence with maximum likelihood rankings***
- ***outranking matrices*** are used as a starting computational step.

The Knowledge Economy Dataset (a continuously updated dataset…)

- Series A gives indicators for the four/five main drivers (see WP 1.1)
- Series B gives indicators for two types of outcomes: economic and social
- Series C gives additional indicators that could be useful in the scenario analyses.

A Component

27 Indicators

A1. Production and diffusion of ICT

- Investment in ICT (A1a4)
- # patent applic. to the EPO (A1a5)
- Broadband penetr. rate (A1c4)

A2. Human resources, skills and creativity:

- Pisa mathematical literacy of 15 year olds (A2a1)
- New PhDs per thousand population aged 25-34 (A2a4)
- Participation in LLL per working age population (25-64) (A2c3)
- Job to job mobility by NACE (A2e2)

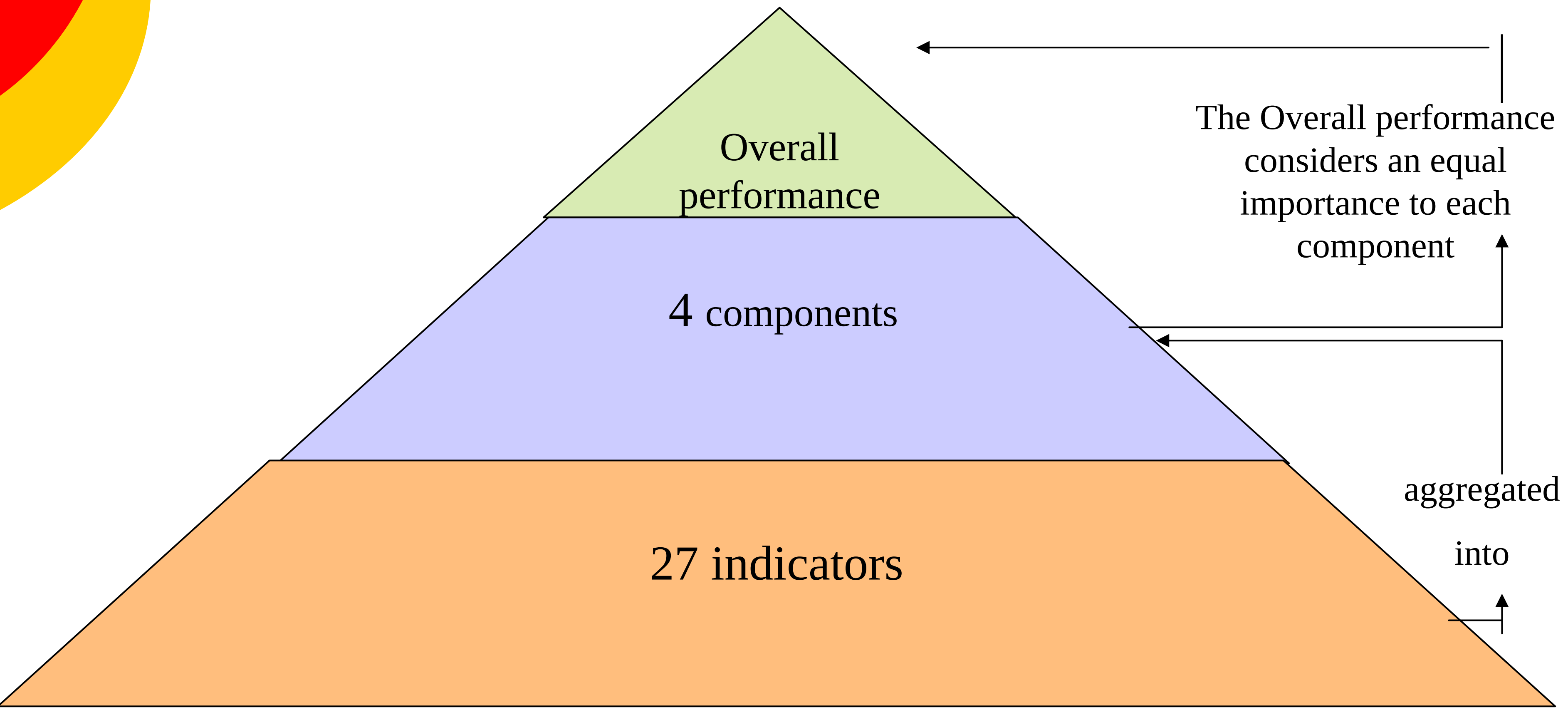
A3. Knowledge production and diffusion:

- GERD expenditure/GDP (A3a1)
- GERD per capita (A3a3)
- Estimated Civil GERD as % of GDP (A3a4)
- GOVERD (I) (calculated) and HERD (II) as % of GDP (A3a5 I, II)
- GBAORD as % of GDP (calculated) (A3a6)
- BERD as a percentage of GDP (A3a7)
- BERD as a percentage of value added in industry (A3a9)
- BERD/GERD (calculated) (A3a10)
- Triadic patent families by priority year (A3b5 I, II)
- Share of all firms reporting public research (universities & institutes) as a high value information source (A3e1)
- Share of all firms reporting public research (universities & institutes) as a cooperation partner (A3e2)
- High tech exports/total exports (A3f2)

A4. Innovation, Entrepreneurship and creative destruction:

- Firm entries (A4a2)
- Firm exits (A4a3)
- Share of firms introducing new-to-market products (A3d1)
- Share of total sales from new-to-market products (A3d2)
- Share of total sales from new-to-firm products (A3d4)
- (A5a1)

Aggregation of indicators

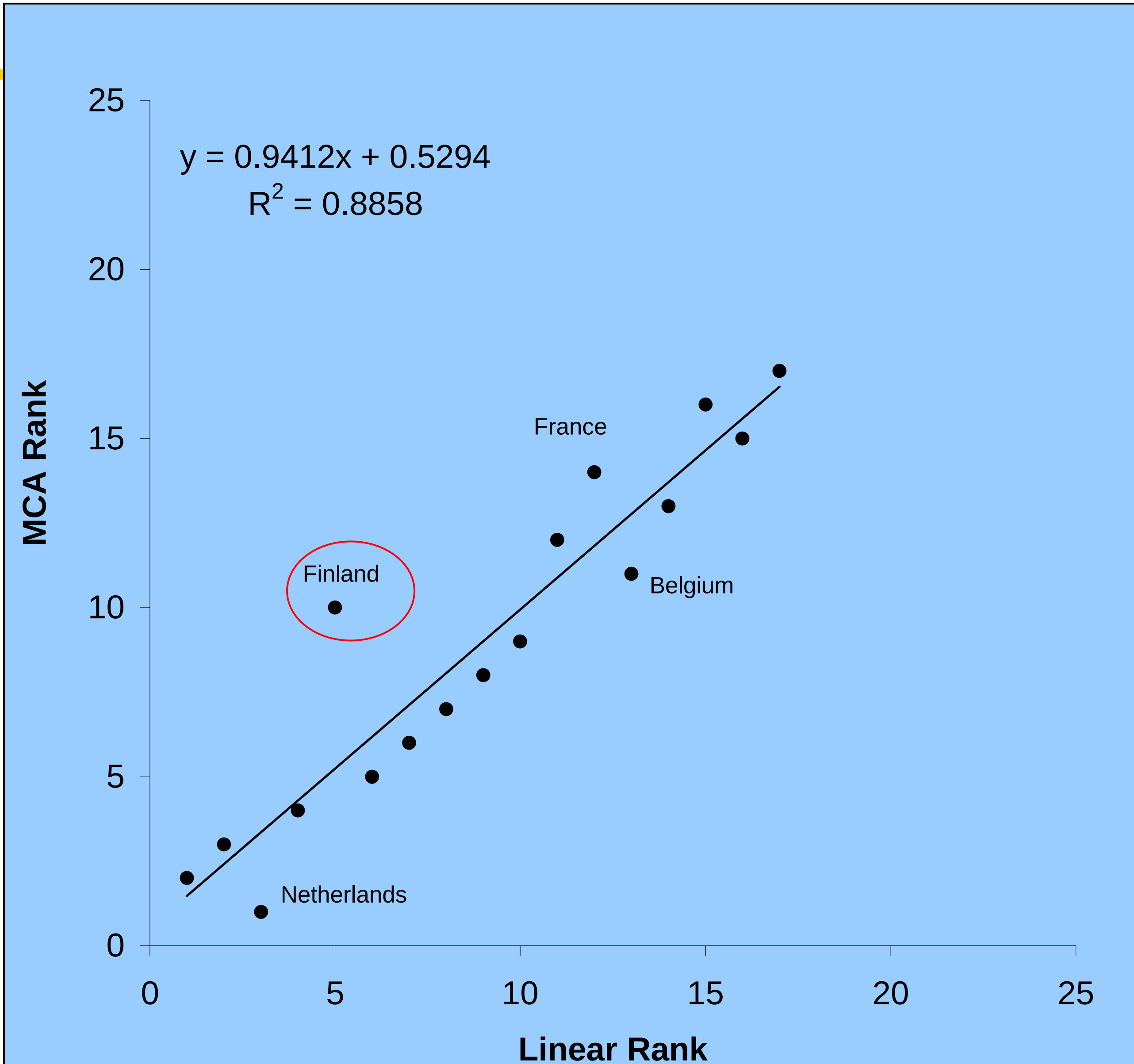


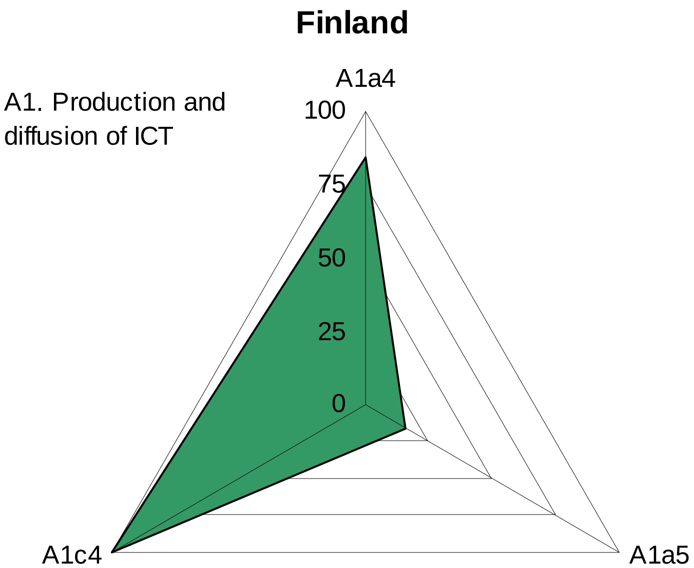
Comparison of aggregation methods

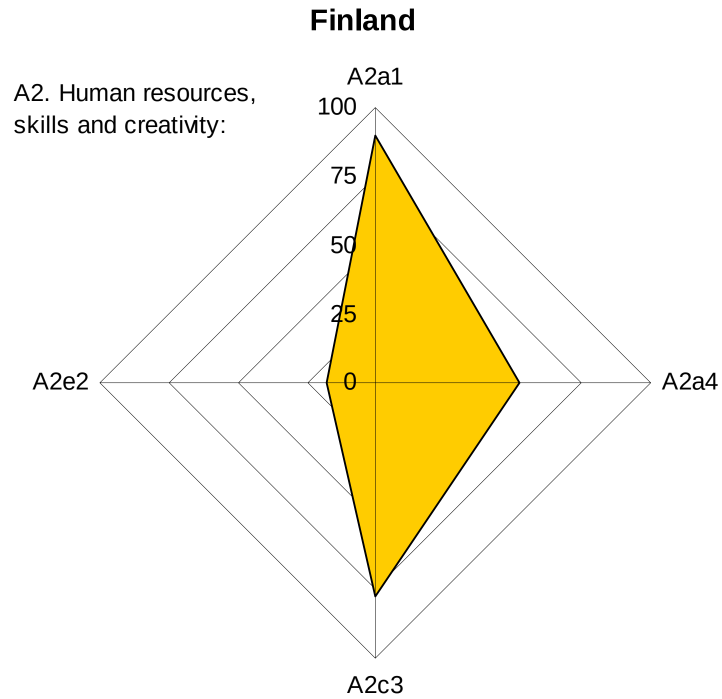
• aggregation method affects principally the middle-of-the-road countries

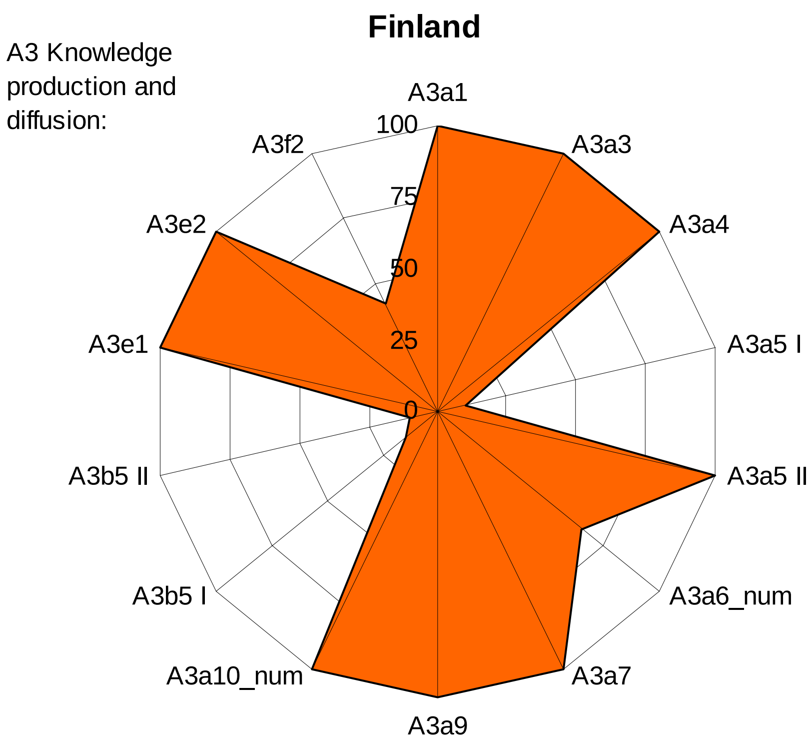
• both aggregation schemes produce comparable rankings

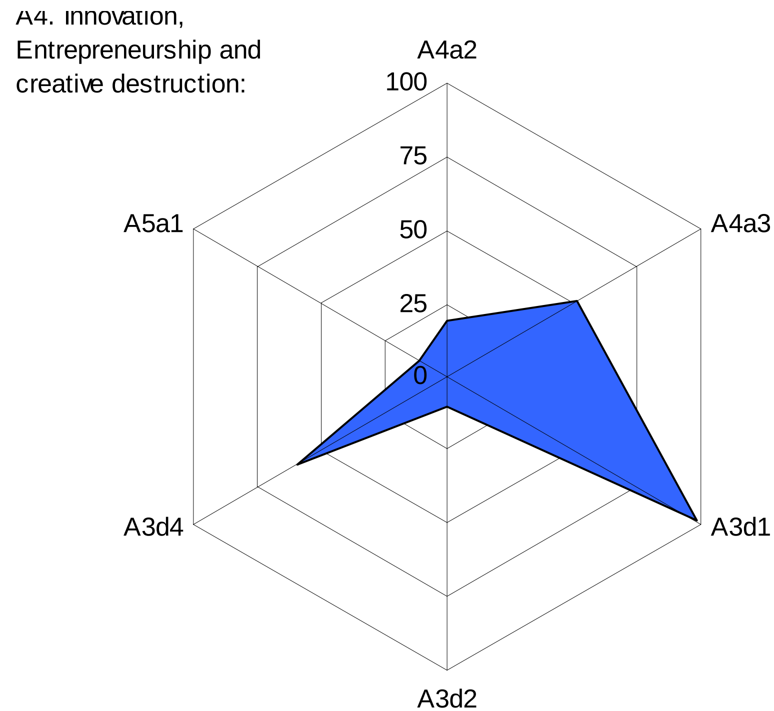
• when compensability is not allowed, countries performing very poorly on some indicators, such as Finland or France see their rank lowered with respect to the linear aggregation, whereas countries that have less extreme values improve their situation, such as Belgium




Finland
A1. Production and
diffusion of ICT
A1a4
100
75
50
25
0
A1c4
A1a5

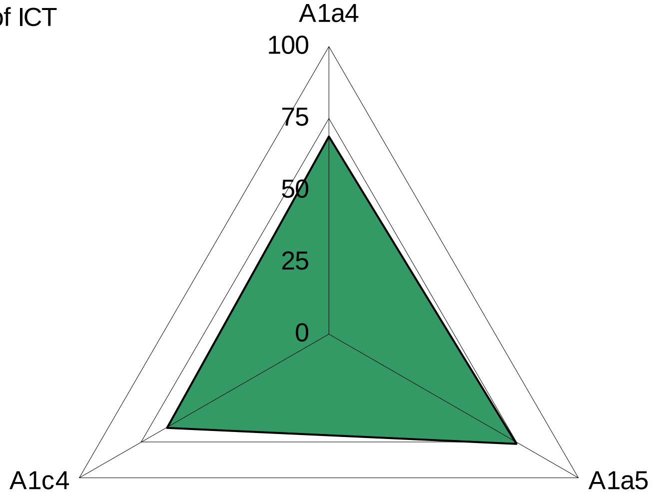

Finland
A2. Human resources,
skills and creativity:
A2a1
100
75
50
25
0
A2e2
A2a4
A2c3


Finland
A3 Knowledge
production and
diffusion:
A3a1
100
75
50
25
0
A3a3
A3a4
A3a5 I
A3a5 II
A3a6_num
A3a7
A3a9
A3a10_num
A3b5 I
A3b5 II
A3e1
A3e2
A3f2


A4. Innovation,
Entrepreneurship and
creative destruction:
A4a2
100
75
50
25
0
A4a3
A3d1
A3d2
A3d4
A5a1

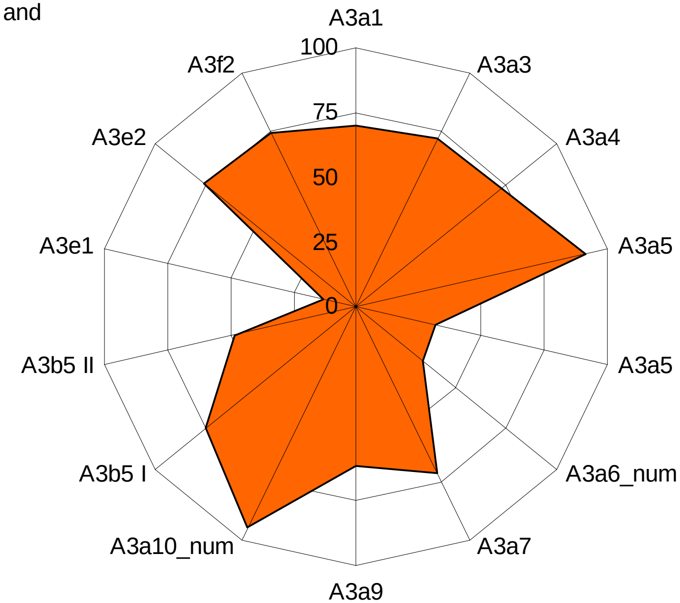
Japan

A1. Production and diffusion of ICT



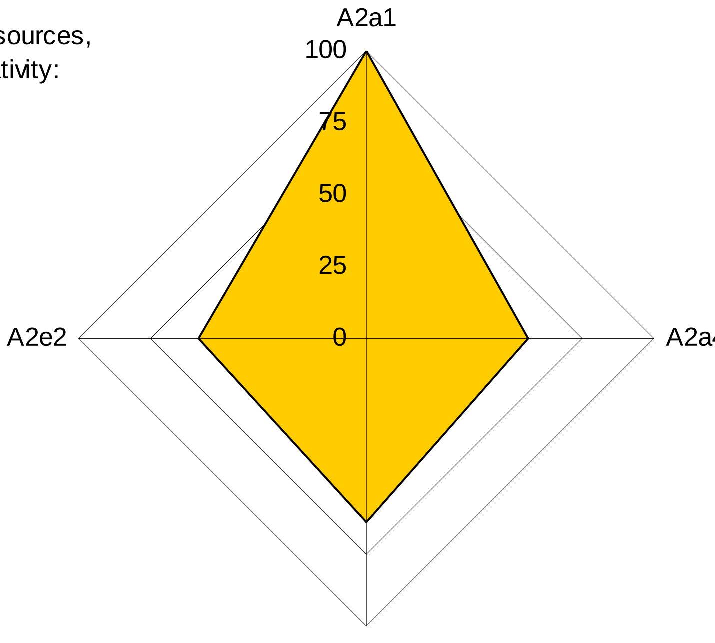
Japan

A3 Knowledge production and diffusion:

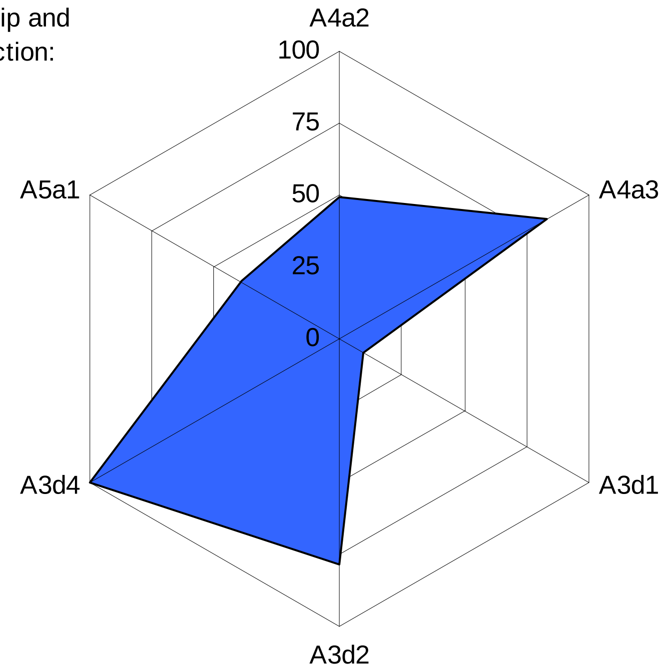


Japan

A2. Human resources, skills and creativity:



A4. Innovation, Entrepreneurship and creative destruction:



B Component

B1. Economic outputs:

- GDP per capita in PPS (B1a1)
- Real GDP growth rate (B1a2)
- Labour productivity per hour worked (B1b1)
- Gross fixed capital formation as % of GDP (calculated) (B1b3 I, II)
- Total employment growth (B1c1)

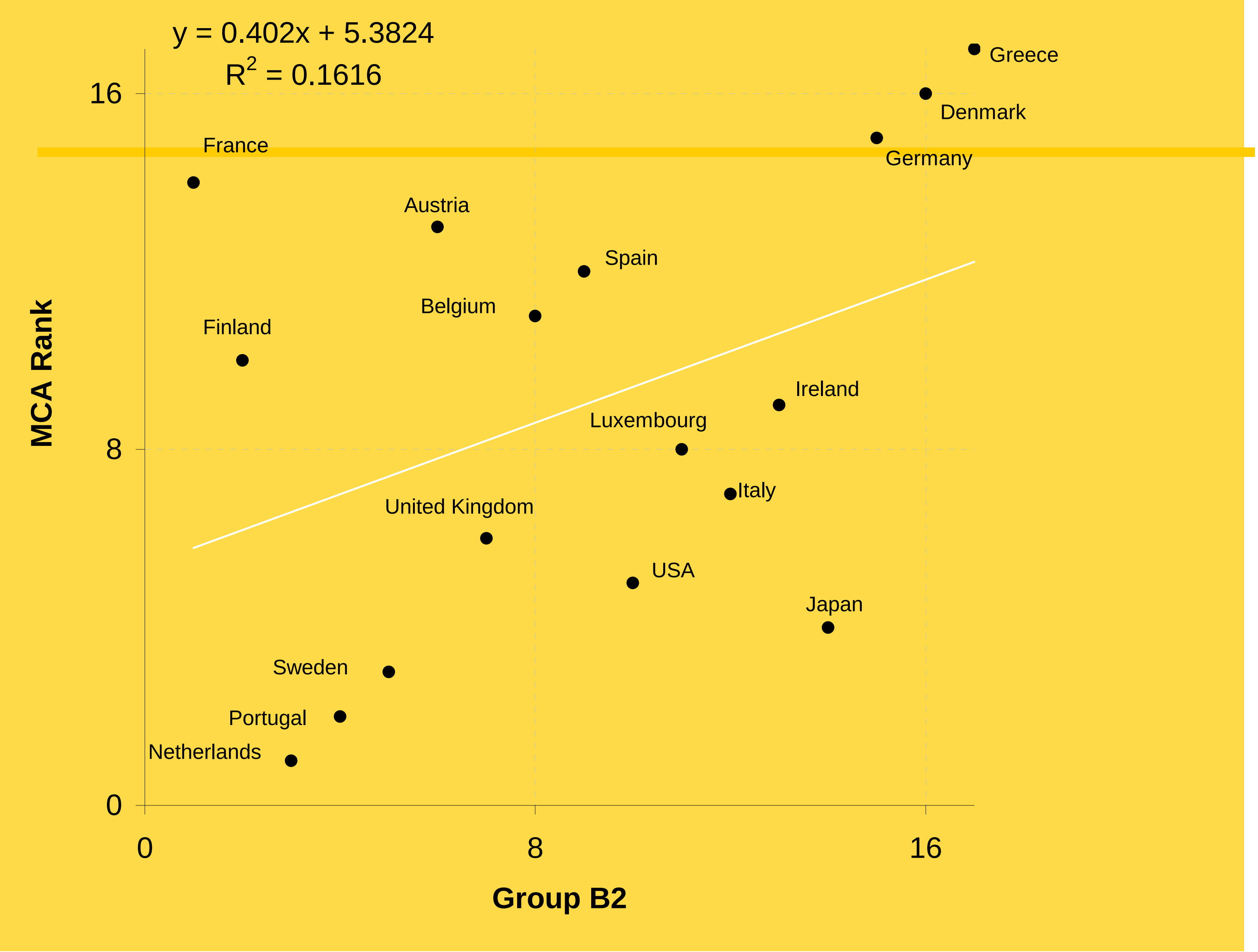
B2. Social performance:

- Energy intensity of the economy (B2a2)
- Employment rate of older workers (B2b1)
- Total employment rate (B1c2)
- Long term unemployment rate (B2b2)
- Inequality of income distribution (B2b4)

Ranks - different components A, B

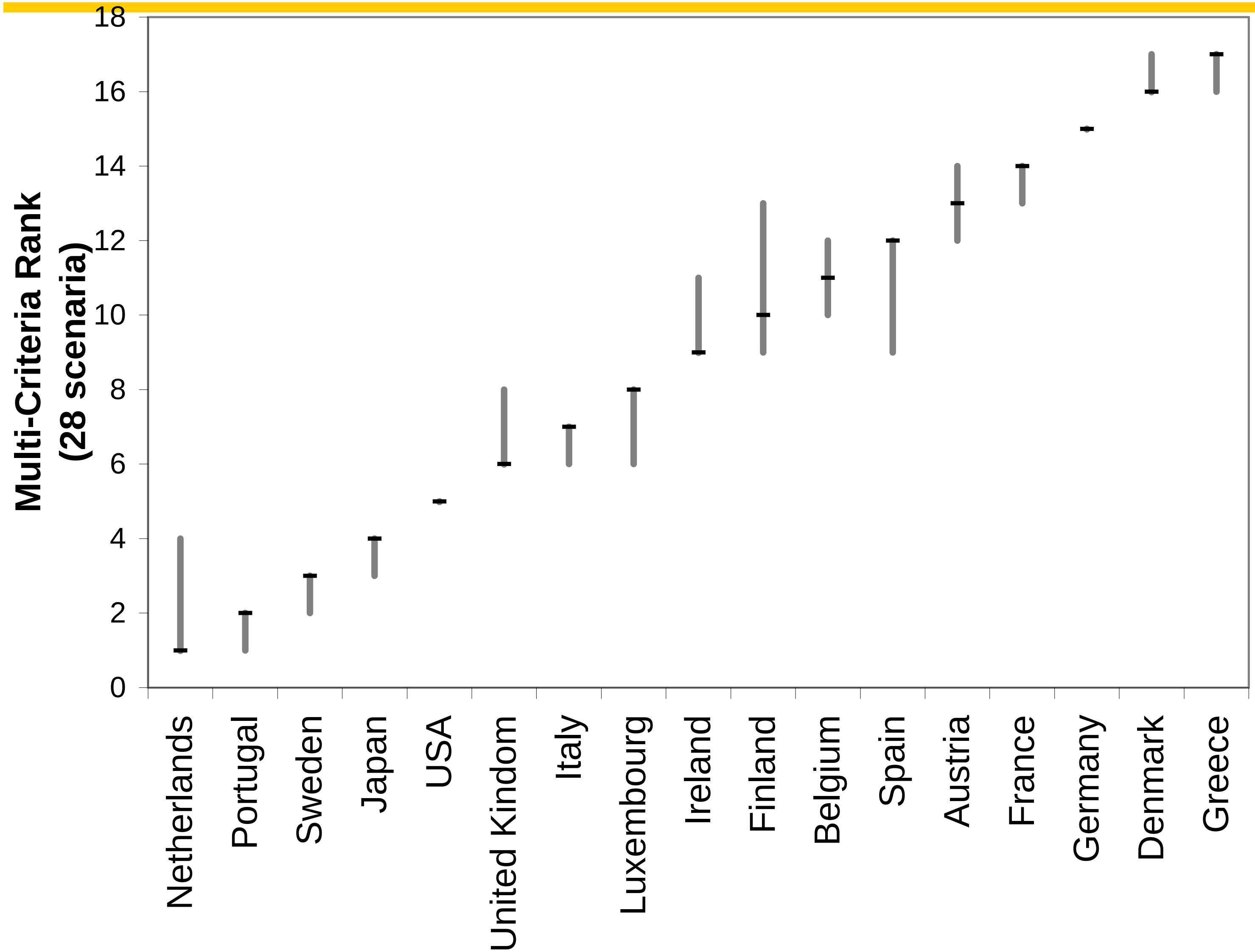
	A	A1	A2	A3	A4	B1	B2
Austria	13	11	13	12	9	4	6
Belgium	11	4	10	13	12	11	8
Denmark	16	16	17	17	11	14	16
Germany	15	15	16	15	14	16	15
Greece	17	17	15	16	17	17	17
Spain	12	8	3	10	15	5	9
Finland	10	1	14	1	13	7	2
France	14	9	11	9	16	12	1
Ireland	9	10	4	14	8	8	13
Italy	7	12	2	11	1	9	12
Luxembourg	8	14	1	8	10	6	11
Netherlands	1	5	8	3	2	1	3
Portugal	2	2	5	2	3	3	4
Sweden	3	3	12	4	4	2	5
United Kindom	6	13	9	7	5	10	7
USA	5	7	7	6	6	13	10
Japan	4	6	6	5	7	15	14

y = 0.5588x + 3.9706
R^2 = 0.3123
MCA Rank
Group B1
16
8
0
0
8
16
Greece
Denmark
Germany
France
Austria
Spain
Belgium
Finland
Ireland
Luxembourg
Italy
United Kingdom
USA
Japan
Sweden
Portugal
Netherlands



y = 0.402x + 5.3824
R2 = 0.1616
MCA Rank
Group B2
16
8
0
0
8
16
Greece
Denmark
Germany
France
Austria
Spain
Belgium
Finland
Ireland
Luxembourg
Italy
United Kingdom
USA
Japan
Sweden
Portugal
Netherlands

Analysis of different scenaria



when to use what?

When using a model or an algorithm to describe a real-world issue formal coherence is a necessary property BUT not sufficient.

The model in fact should fit objectives and intentions of the user, i.e. it must be the most appropriate tool for expressing the set of objectives that motivated the whole exercise.

The choice of which indicators to use, how those are divided into classes, whether a normalization method has to be used (and which one), the choice of the weighting method, and how information is aggregated, all these features stem from a certain perspective on the issue to be modelled.

when to use what?

The absence of an "objective" way of constructing composites should not result in a rejection of whatever type of composite. Composites can meaningfully supply information provided that the relation between the framing of a problem and the outcome in the decision space are made clear.

A backward induction exercise could be useful in this context. Once the context and the modeller's objectives have been made explicit, the user can verify whether and how the selected model fulfils those objectives.

No model is a priori better than another, provided internal coherence is assured. In practice, different models can meet different expectations and stakes. Therefore, stakes must be made clear, and transparency should guide the entire process.

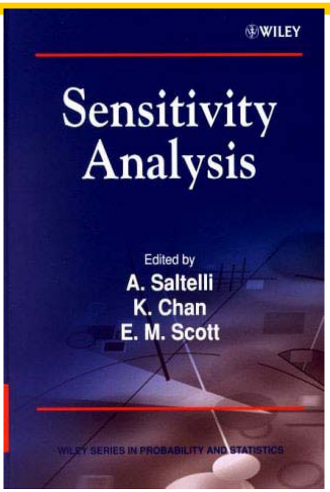
Selected References

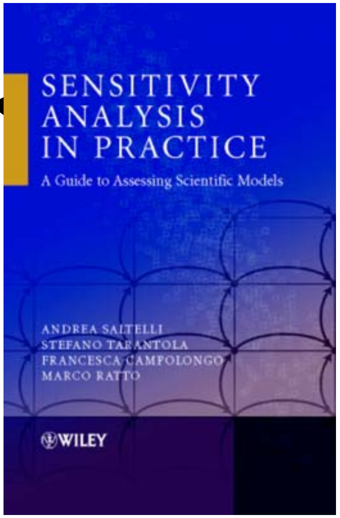
Books

- Saltelli, A., K. Chan and M. Scott, Eds., 2000, *Sensitivity analysis*, John Wiley & Sons publishers, Probability and Statistics series (a multi-authors book - can be used as topical reference).

- Saltelli A. Tarantola S., Campolongo, F. and Ratto, M., 2004, *Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models*, John Wiley & Sons publishers (a primer).

- Saltelli, A., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D. Pennoni, F., Ratto, M., Saisana, M., Tarantola, S., 2007, Sensitivity Analysis of Scientific Models, John Wiley & Sons publishers (a textbook for University students, forthcoming!).





Selected References

References –Recent papers

Saisana M., Saltelli A., Tarantola S. (2005) Uncertainty and Sensitivity analysis techniques as tools for the quality assessment of composite indicators, *Journal of the Royal Statistical Society - A*, **168**(2), 307-323 (application to statistical indicators).

- Zádor J., Zsély, I.G., Turányi, T., Ratto, M., Tarantola, S., and Saltelli, A., 2005, Local and Global Uncertainty Analyses of a Methane Flame Model, *Journal Physical Chemistry A.*, November 2005 (application to chemistry).
- Saltelli, A., M. Ratto, S. Tarantola and F. Campolongo (2005) Sensitivity Analysis for Chemical Models, *Chemical Reviews*, **105(7)** pp 2811 – 2828 (concise review paper).
- Hall, J.W., Tarantola, S., Bates, P.D. and Horritt M.S. (2005) Distributed sensitivity analysis of flood inundation model calibration, *J. Hydraulic Engineering, ASCE,* **131(2)** 117-126 (application to geographically distributed output).