

Fuzzy Set Theory and Principal Component Analysis: complementary aggregation solutions

José Manuel Roche

J.M.Roche@sussex.ac.uk

PhD Student at the University of Sussex, UK
Research Fellow at CISOR, Venezuela

General interest:

Monitoring inequality in capabilities between social groups

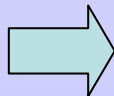
Technical problem:

How to measure and aggregate qualitative data?

Methodological proposal:

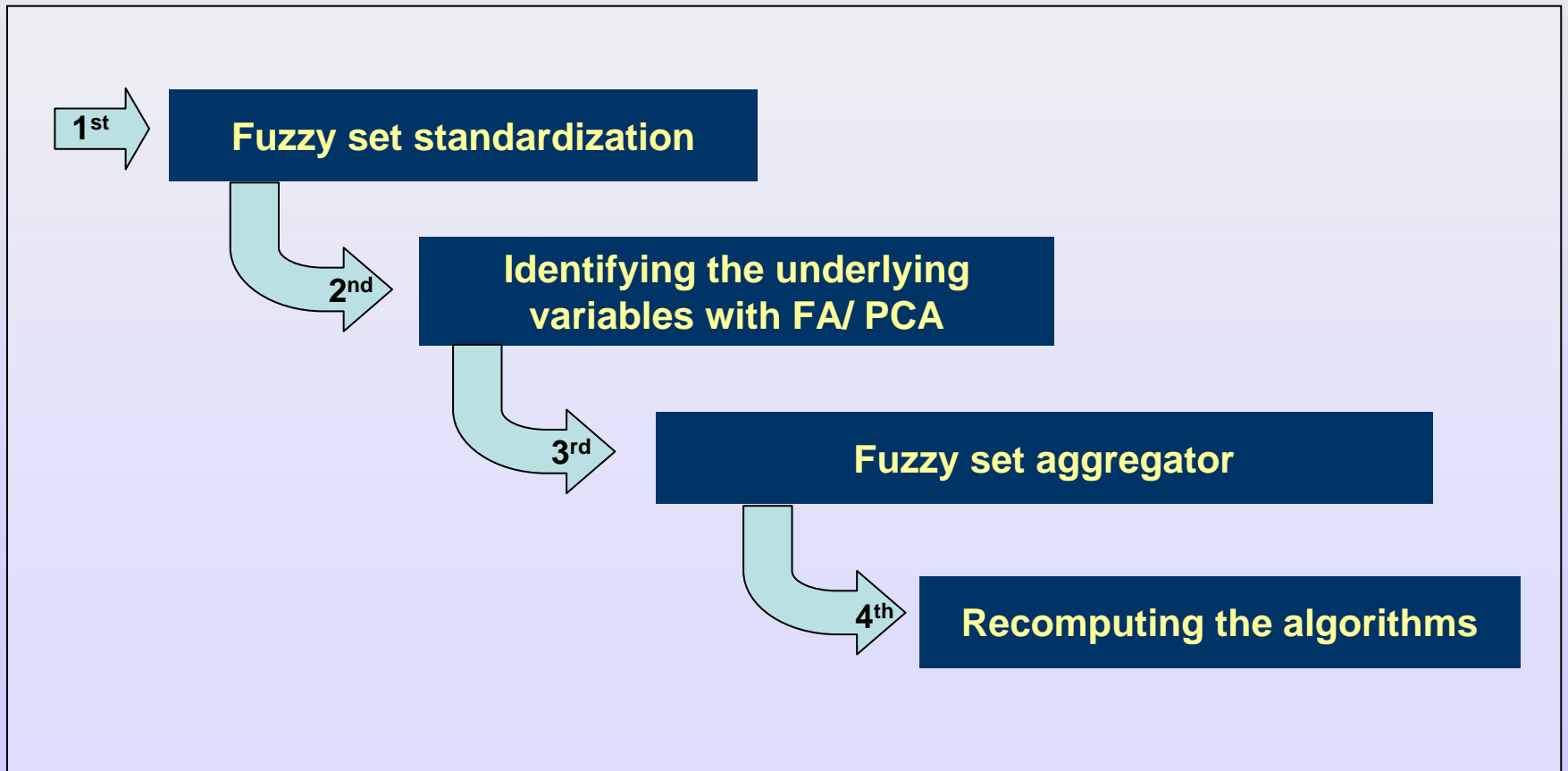
A combined use of **Fuzzy Set Theory (FST)**, and **Principal Component Analysis (PCA)**

**Complementary
solutions**



Their combined use contributes with empirical evidence and operational solutions for the design of synthetic indices

Core of the proposal



An illustrative application

Focus: Housing conditions
(the capability of 'being well sheltered')

Context: Venezuela

Data: Household Survey (2001)
Census ('71, '81, '90, 2001)

Selected Indicators

Sewage system

Water

Electricity

Fuel

Floors

Roofs

Walls

Housing Overcrowding Index

What difference does it make?
Does it matter?



Indicator transformation and fuzzy set standardization

1st

Variables recoding
(standardization '71, '81, '90, 2001)

**Rearrangement of categories
in a meaningful ordinal scale**

**Assign a fuzzy set
membership function**

$$\mu_A : X \rightarrow [0, 1]$$

Identifying the underlying variables with PCA

Unrotated, Varimax-rotated common components matrix

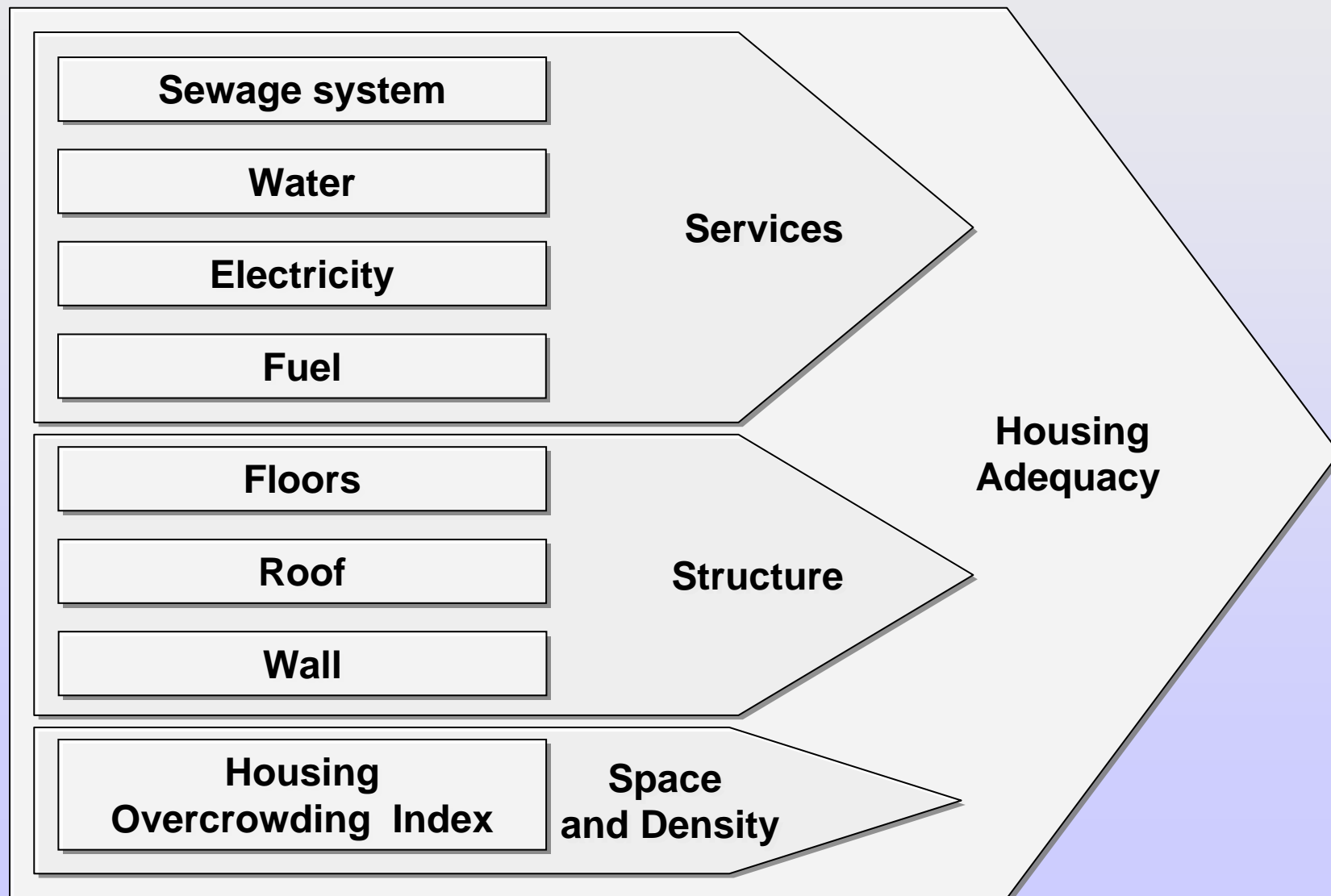
	Unrotated Component			VARIMAX-rotated Component		
	1	2	3	1	2	3
Sewage	0.734	0.120	-0.010	0.518	0.418	0.331
Water	0.565	0.435	0.144	0.695	0.100	0.190
Electricity	0.420	0.529	0.138	0.687	-0.014	0.061
Fuel used for cooking	0.401	0.495	-0.088	0.620	0.147	-0.087
Floors	0.752	-0.208	-0.310	0.226	0.752	0.297
Roofs	0.597	-0.312	-0.595	0.018	0.897	0.070
Walls	0.692	-0.228	0.345	0.258	0.250	0.721
Housing Overcrowding Index	0.495	-0.513	0.513	-0.064	0.101	0.870

Extraction Method: Principal Component Analysis. 3 components extracted.

VARIMAX: Rotation converged in 4 iterations.

Oblimin: Rotation converged in 9 iterations.

Identifying the underlying variables with PCA



Identify the best fuzzy set aggregator that capture the underlying variables

Fuzzy Set Aggregator

It could capture the component structure
Less obscure than the PCA

Housing Structure

	Strong Intersection	Weak Intersection	Arithmetic Average
Component 2 VARIMAX	0.610	0.784	0.700

Housing Services

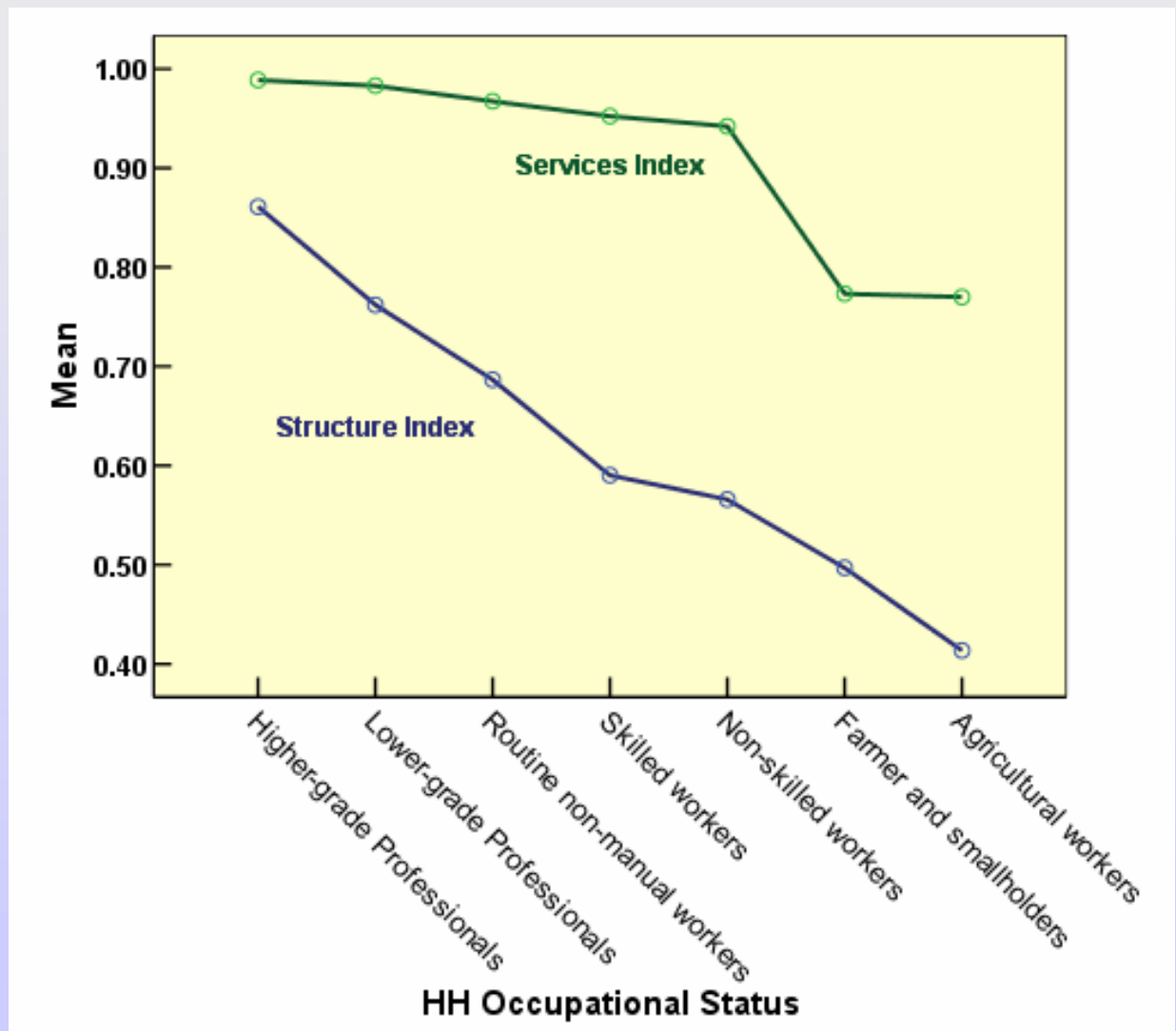
	Strong Intersection	Weak Intersection	Arithmetic Average
Component 1 VARIMAX	0.654	0.655	0.869

Recomputing the set of indicators in other data set

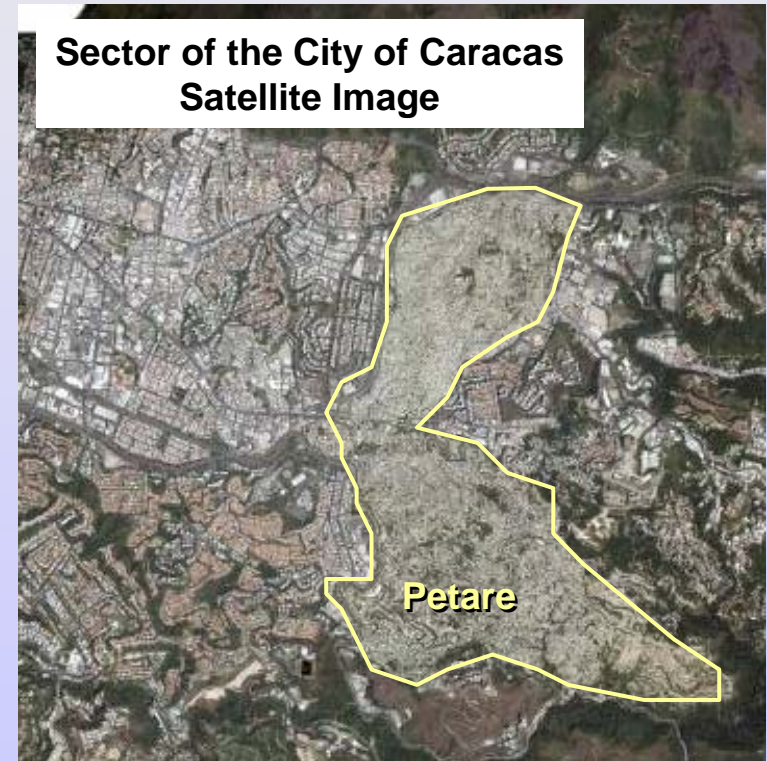
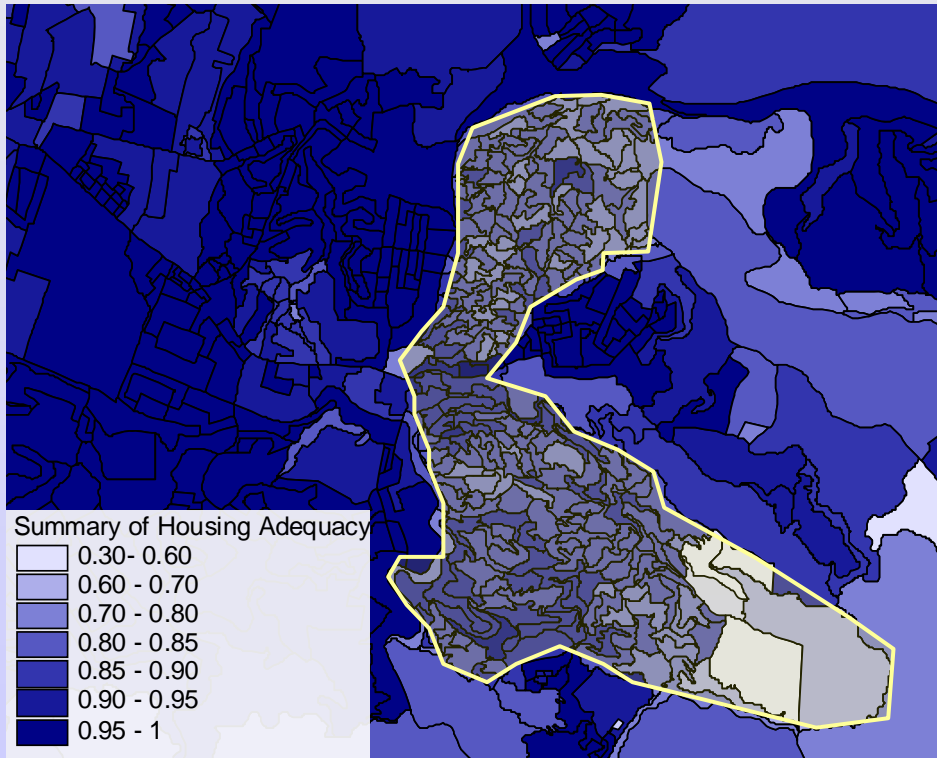
Monitoring progress

Assessing inequality

Overall housing adequacy
Specific area



Monitoring progress and inequalities at a local level Detection of vulnerable groups



CONCLUSIONS

The methodological proposal is enriched by the strengths of both techniques

Principal Component Analysis

- Identifying the underlying or latent variables
- Empirical evidence to support aggregation decisions

Fuzzy Set Theory

- Expressing the initial indicators in terms of degree of achievements
- Arithmetic procedures to capture the underlying variables