# Social polarization in the metropolitan area of Marseille. Modelling uncertain knowledge with probabilistic and possibilistic networks

Giovanni Fusco, Cristina Cao, Didier Dubois, Henri Prade, Floriane Scarella, Andrea G. B. Tettamanzi

# Social polarization in the metropolitan area of Marseille. Modelling uncertain knowledge with probabilistic and possibilistic networks[1].

Giovanni Fusco[*], Cristina Cao[*], Didier Dubois[**], Henri Prade[**], Floriane Scarella[*], Andrea Tettamanzi[***]

## Abstract

A Bayesian Network and a Possibilistic Network are used to produce trend scenarios of social polarization in the metropolitan area of Marseille (France). Both scenarios are based on uncertain knowledge of relationships among variables and produce uncertain evaluations of future social polarization. We show that probabilistic models should not be used just to infer most probable outcomes, as these would give a fallacious impression of certain knowledge. The possibilistic model produces more uncertainty-laden results which are coherent with model uncertainties and respect elicited values of possibilities. Results of the two models converge when probability values are "degraded".

## Keywords

Social polarization, Uncertainty, Bayesian networks, Possibilistic networks, Marseille.

[*] UMR ESPACE, CNRS / University of Nice Sophia Antipolis
[**] UMR IRIT, CNRS / University of Toulouse Paul Sabatier
[***] UMR I3S, CNRS / University of Nice Sophia Antipolis

## Introduction

The metropolitan area of Aix-Marseille in southern France has experienced ongoing social polarization since the 1980s. The geography of unemployment, on the one hand, and the concentration of high-skilled professionals, on the other, contribute considerably to the structuring of a contrasted metropolitan social morphology (Centi, 1996; Fusco and Scarella, 2011). Future continuation of metropolitan logics, reinforcing social polarization, raises questions on the social cohesion. For the purpose of our research, the metropolitan area is defined as encompassing 439 municipalities and more than 3 million inhabitants, around Marseille, Aix-en-Provence, Toulon and Avignon.
Knowledge of factors inducing social polarization of the municipalities in the study area is nevertheless uncertain. Several factors contribute to the valorisation or to the de-valorisation of residential space. But these factors have "soft", uncertain impacts on the phenomena under investigation: the same causes can lead to different effects. A probabilistic model of these socio-spatial mechanisms in the metro area of Marseille has already been proposed (Scarella, 2014) in the form of a Bayesian network (Jensen, 2001). More particularly, this model was used to investigate several scenarios of future evolution of spatial polarization in the metro area. Nevertheless, the uncertainty content of model results has not been completely explored. Moreover, alternative theoretic frameworks exist to model uncertain knowledge. Possibility theory (Dubois and Prade, 1988, 2001) seems particularly appropriate to model uncertainties in the knowledge of geographic phenomena and of their causal relations. Thanks to recent advancements in the implementation of possibilistic networks (Caglioni *et al.*, 2014), a new possibilistic model has thus been developed by the authors of this paper (Dubois *et al.*, 2015).

## Two Uncertainty-based Models

In both models, the Bayesian Network (BN) and the Possibilistic Network (PN), social polarization is described through the overrepresentation of two target populations: in valorised municipalities executives and professionals are overrepresented; in de-valorised ones the unemployed are overrepresented. The models have the same structure and include 26 variables covering different factors, namely position in the metropolitan area, migration flows, presence/absence of environmental amenities, nature of the housing stock, planning policies and path-dependence of social specialization. Other common features are:

- They are expert-based. Only BN parameters were actually elicited; a least committing probability-to-possibility preference preserving transformation (Dubois *et al.*, 1993) was then used to obtain PN parameters.
- They mix observable and non-observable variables.
- Probabilistic/Possibilistic relations are modelled through noisy/uncertain logical gates (Or, And, Max), reducing considerably the number of parameters to be elicited.
- They use uncertain relations, include leak parameters (taking into account the impact of factors not included in the model) and produce uncertain results.

But the models also implement two different approaches to epistemic uncertainty. The BN is based on Bayesian probabilities (Pearl, 2000; Jensen, 2001); its sum/product compositions are well suited to an "exact" knowledge of probabilities. The PN is based on possibility theory (Dubois and Prade, 1988); its max/min compositions are better suited to the qualitative knowledge of possibilities.

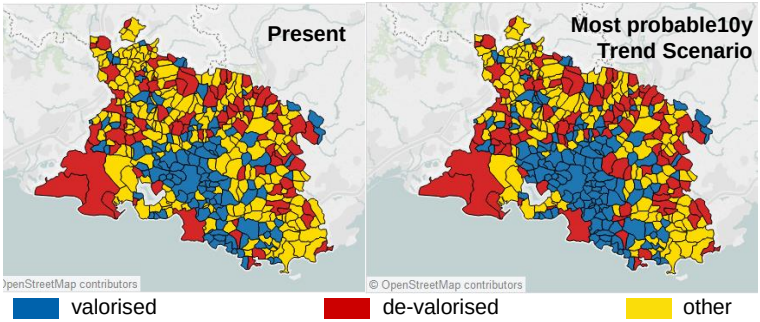


Figure 1 – Naïve use of the probabilistic model.

The two models have been compared in the way they infer a 10-years trend scenario for the social polarization within the study area. Model results are compared using appropriate interactive visualizations on the Tableau® data-viz platform (https://public.tableau.com/profile/fusco#!/vizhome/RepresentingUncertainFutures/Story1).
Representing knowledge of complex phenomena, like social polarization, is achieved through a dashboard system, with dynamic links between maps and diagrams. Particular attention is given to visualizing uncertain knowledge. Only appropriate visualizations can make uncertain knowledge useful for scientific understanding and for planning policies (Harrower 2003). We convey uncertainty through interactive visualizations: by choosing a given certainty threshold, the user is presented with different sets of results. Fig. 1 shows a naïve use of the probabilistic model. Present state of social polarization is compared to the most probable outcomes inferred for the 10 years trend scenario. The future state of each unit can be valorised, de-valorised or other. Most probable outcomes show an increase of valorisation around Aix, in the centre, and relative stability in the rest of metro area. But this simple map of most probable outcomes gives a false impression of certainty.

Indeed, the certainty/uncertainty content of probabilistic inferred values is not fully exploited. Fig. 2-a thus includes the second most probable outcome of the BN model. Moreover, a slider can filter the outcomes by their estimated probability. Fig. 2-b/c show most probable and second most probable outcomes only when they are inferred with probability more than 0.5 and 0.7, respectively. Grey colour represents "uncertain" units, for which a most probable outcome cannot be inferred with the given probability threshold. The reading of the resulting maps is not straightforward. By requiring higher probability values (i.e. more certainty) many units "become" uncertain.

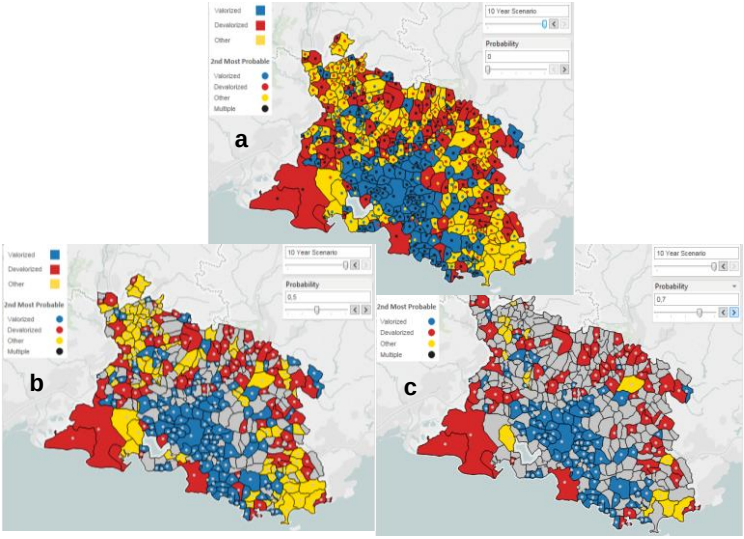


Figure 2 – Filtering uncertainty levels of the BN results.

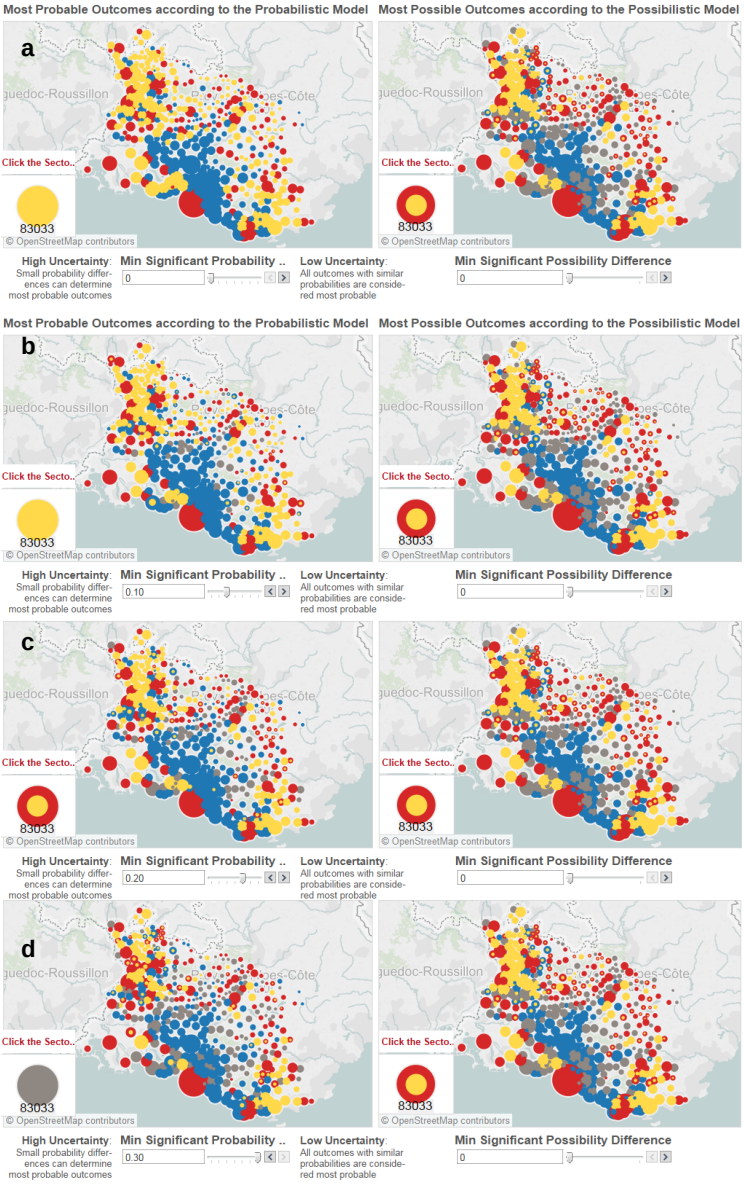


Figure 3 – Comparing probabilities and possibilities.

In reality, we are eliminating the false sense of certainty given by naïve representation of most probable outcomes. Municipalities for which a given future state can be predicted are now characterised by higher levels of confidence. The geographic reading of these more "robust" results is also different. Instead of predicting a spread-out of valorisation around Aix-en-Provence, we can only say that valorised municipalities around Aix will very probably stay valorised and that those which are de-valorised in more peripheral areas will very probably remain de-valorised. Of course, we cannot say much for units represented in grey. Compared to the map of the most probable outcomes, we have a much more uncertainty-aware picture.

It is interesting to compare the results of the two models in terms of most plausible predictions. When different outcomes are equally plausible (probable/possible), the trend scenario for a given unit becomes more uncertain. For the most uncertain cases all three outcomes are equally probable/possible. Possibilistic results seem more uncertain than their probabilistic counterparts. For each municipality, a most probable outcome is always found, whereas several outcomes can be completely possible (Fig. 3-a). However, probability differences among outcomes can be very small and eventually not significant. When a minimum difference is required, several outcomes can become "equally most probable" (Fig. 3-b/c/d). BN results become thus increasingly uncertain and comparable to those of the PN. The best agreement between the two models is found by requiring a minimal difference of 0.25 for probability values. 77.2% of outcomes are then equal in the two models, possibilistic outcomes are more uncertain in 17.1% of cases, and probabilistic ones in 5.7%. Both models show that the outcome for several municipalities is not totally uncertain, but hesitates between two equally probable/possible states. Complete uncertainty is still the

case for 130 spatial units for the PN and 85 for the BN, which are neither too peripheral nor too central.

## Conclusion

Uncertain knowledge of social polarization needs uncertainty-based models for spatial strategic foresight. Methodologically, probabilistic models should not be used just to infer most probable outcomes, as these would give a false impression of certainty. Possibility theory is an interesting alternative for inference in graphical models. It produces more uncertainty-laden results which are coherent with model uncertainties and respect elicited values of possibilities. The results of the two models converge when probability values are degraded, but this is *a posteriori* bricolage and epistemic uncertainty is more coherently dealt with in the possibilistic model.
Geographically, integrating uncertainty gives different results than the naïve use of probabilities. Valorised municipalities around Aix should stay valorised, peripheral de-valorised municipalities should remain de-valorised, some municipalities could have two different outcomes and several intermediate municipalities seem completely undetermined. This is not just a different picture, but a more sincere one, highlighting defaulting knowledge.

[1] Research carried out with CNRS grant - PEPS GéoIncertitude (2014-15)

## References

Caglioni M., Dubois D., Fusco G., Moreno D., Prade H., Scarella F., Tettamanzi A. (2014), *Mise en Œuvre Pratique de Réseaux Possibilistes, LFA 2014*, Toulouse, Cépadues Edition, pp. 267-274

Centi C. (1996), *Le Laboratoire Marseillais: Chemins d'Intégration Métropolitaine et de Segmentation Sociale*, Paris, L'Harmattan.

Dubois D., Fusco G., Prade H., Tettamanzi A. (2015), *Uncertain Logical Gates in Possibilistic Networks. An Application to Human Geography, LNAI*, 9310, Springer, pp. 249–263.

Dubois D., Prade H., Sandri S. (1993), On possibility/probability transformations, in Lowen R., Roubens M. (eds.) *Fuzzy Logic: State of the Art*, Dordrecht, Kluwer, pp. 103-112.

Dubois D., Prade H. (1988), *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, New York, Springer.

Dubois D., Prade H. (2001), Possibility theory, probability theory and multiple-valued logics: A clarification, *Annals of Mathematics and Artificial Intelligence*, 32, pp. 35–66

Fusco G., Scarella F. (2011), Métropolisation et ségrégation sociospatiale. Les flux des migrations résidentielles en PACA, *L'Espace Géographique*, 40(4), pp. 319-336

Jensen F. (2001), *Bayesian Networks and Decision Graphs*, New York, Springer.

Harrower M. (2003), Representing uncertainty: Does it help people make better decisions, *UCGIS Workshop: Geospatial Visualization and Knowledge Discovery Workshop*, Lansdowne (VA), November 18-20.

Pearl J. (2000), *Causality: Models, Reasoning and Inference*, Cambridge, University Press.

Scarella F. (2014), *La Ségrégation Résidentielle dans l'Espace-Temps Métropolitain.*, PhD thesis, Université de Nice Sophia Antipolis.