Un Cadre Bas sur la Discrpance pour une Comparaison de la Robustesse entre Evaluations Multi-attributs

Juste Raimbault[1,2]

[1] UMR CNRS 8504 Géographie-Cités, Paris, France
[2] UMR-T IFSTTAR 9403 LVMT, Champs-sur-Marne, France
juste.raimbault@polytechnique.edu

Abstract. Les valuations multi-objectifs sont un aspect essentiel de la gestion de systmes complexes, puisque la complexit intrinsque d'un systme est gnralement troitement lie au nombre d'objectifs d'optimisation potentiels. Cependant, une valuation ne fait pas sens si sa robustesse, au sens de sa fiabilit, n'est pas donne. Les mthodes statistiques usuelles fournissant une mesure de robustesse sont trs dpendantes des modles sous-jacents. Nous proposons une formulation d'un cadre indpendant du modle, dans le cas d'indicateurs intgrs et agrgs (valuation multi-attributs), qui permet de dfinir une mesure de robustesse relative prenant en compte la structure des donnes et les valeurs des indicateurs. La mthode est teste sur donnes urbaines synthtiques associes aux arrondissements de Paris, et des donnes relles de revenus pour l'valuation de la sgrgation urbaine dans la rgion mtropolitaine du Grand Paris. Les premiers rsultats numriques montrent les potentialits de cette nouvelle mthode. De plus, sa relative indpendance au type de systme et au modle pourrait la positionner comme une alternative aux mthodes statistiques classiques d'valuation de la robustesse.

Keywords: Evaluation Multi-attributs, Robustesse Non-modle-dpendante, Systmes Urbains, Discrpance

1 Introduction

1.1 Contexte Gnral

Les problmes multi-objectifs sont organiquement lis la complexit des systmes sous-jacents. En effet, que ce soit dans le champ des *Systmes Complexes Industriels*, dans le sens de systmes conus par ingnierie, o la construction de Systmes de Systmes (SoS) par couplage et intgration induit souvent des objectifs contradictoires [**?**], ou dans le champ des *Systmes Complexes Naturels*, au sens de systmes non dsigns, physiques, biologiques ou sociaux, qui prsentent des proprits d'mergence et d'auto-organisation, pour lesquels les objectifs peuvent e.g. tre le rsultat de l'interaction d'agents htrognes (voir [**?**] pour une revue tendue des types de systmes concerns par cette approche), l'optimisation multi-objectifs peut tre explicitement introduite pour tudier ou dsigner le systme, mais rgit gnralement dj implicitement les mcanismes internes du systme. Le cas des Systmes Complexes Sociaux-techniques est particulirement intressant puisque selon Haken [?], ils peuvent tre vus comme des systmes hybrides embarquant des agents sociaux dans des “artefacts techniques” (parfois jusqu' un niveau inattendu, crant ce que PICON dcrit comme *cyborgs* [?]), et cumulent ainsi la potentialit d'tre l'origine de problmes multi-objectifs[1]. La notion rcente d'*co-quartier* [?] est un exemple typique pour lequel la durabilit implique des objectifs contradictoires. L'exemple des systmes de transport, dont la conception a gliss durant la seconde moiti du 20me sicle d'analyses cot-bnfices la price de dcision multi-critres, est galement typique de tels systmes [?]. Les systmes gographiques sont prsent bien tudis d'un tel point de vue, en particulier grce l'intgration des cadres multi-objectifs au sein des Systmes d'Information Gographiques [?]. Comme dans le cas microscopique des co-quartiers, la planification et le design urbains msoscopiques et macroscopiques peuvent tre rendus durables grce aux valuations par indicateurs [?].

Un aspect crucial de l'valuation est une certaine notion de sa fiabilit, que nous nommerons ici *robustesse*. Les mthodes statistiques incluent naturellement cette notion puisque la construction et l'estimation de modles statistiques donne divers indicateurs de la consistence des rsultats [?]. Le premier exemple venant l'esprit est l'application de la loi des grands nombres pour obtenir la *p-valeur* d'une estimation de modle, qui peut tre interprt comme une mesure de confiance en les valeurs estimes. D'autre part, les intervalles de confiance et le *beta-power* sont d'autres indicateurs importants de robustesse statistique. L'infrence baysienne fournit galement des mesures de robustesse quand la distribution des paramtres est estime de manire squentielle. Concernant les optimisations multi-objectifs, en particulier par des algorithmes heuristiques (comme par exemple les algorithmes gntiques, ou les solveurs de recherche oprationelle), la notion de robustesse d'une solution consiste plus en la stabilit de la solution dans l'espace des phases du systme dynamique correspondant. Des progrs rcents ont t faits vers une formulation unifie de la robustesse pour les problmes d'optimisation multi-objectifs, comme dans [?] o les fronts de Pareto robustes sont dfinis comme des solutions insensibles aux petites perturbations. Dans [?], la notion de degr de robustesse est introduite, formalise comme une sorte de continuit des autres solutions dans des voisinages successifs d'une solution.

Cependant, il n'existe pas de mthode gnrique qui permettrait une valuation de la robustesse de faon indpendante au modle, i.e. qui serait extraite de la structure des donnes et des indicateurs mais ne dpendrait pas de la mthode utilise. Un avantage serait par exemple une estimation *a priori* de la robustesse potentielle d'une valuation et de dcider ainsi si elle vaut la peine d'tre faite. Nous proposons un cadre rpondant cette contrainte dans le cas particulier des valuations multi-attributs, i.e. quand le problme est rendu unidimensionnel par

[1] Nous dsignons ici par *Evaluation Multi-objectifs* toutes les pratiques incluant le calcul de multiples indicateurs d'un systme (il peut s'agir d'optimisation multi-objectif pour un design de systme, une valuation multi-objectif d'un systme existant, une valuation multi-attributs ; notre cadre particulier correspondra au dernier cas).

agrgation des objectifs. Il est bas sur les donnes et non sur les modles, au sens ou l'estimation de la robustesse ne dpendra pas de la manire dont les indicateurs sont calculs, tant qu'ils respectent certaines hypothses dtailles par la suite.

1.2 Approche Propose

Objectifs comme Intgrales Spatiales Nous supposons que les objectifs peuvent tre exprims comme intgrales spatiales, ce qui devrait s'appliquer tout systme territorial, et nos cas d'application sont des systmes urbains. Ce n'est pas si restrictif en terme d'indicateurs possibles si l'on utilise les bonnes variables et noyaux intgrs : de faon analogue la mthode de Regression Gographique Pondre [**?**], toute variable spatiale peut tre intgre contre des noyaux rguliers de taille variable et le rsultats sera une agrgation spatiale dont la signification dpendra de l'tendue du noyau. Les exemples utiliss par la suite comme des moyennes conditionnelles ou des sommes vrifient parfaitement cette hypothse. Mme un indicateur dj agrg dans l'espace peut tre interprt comme une intgrale spatiale en utilisant une distribution de Dirac au centrode de la zone correspondante.

Objectifs Agrgs Linairement A second assumption we make is that the multi-objective evaluation is done through linear aggregation of objectives, i.e. that we are tackling a multi-attribute optimization problem. If $(q_i(\boldsymbol{x}))_i$ are values of objectives functions, then weights $(w_i)_i$ are defined in order to build the aggregated decision-making function $q(\boldsymbol{x}) = \sum_i w_i q_i(\boldsymbol{x})$, which value determines then the performance of the solution. It is analog to aggregated utility techniques in economics and is used in many fields. The subtlety lies in the choice of weights, i.e. the shape of the projection function, and various approaches have been developed to find weights depending on the nature of the problem. Recent work [**?**] proposed to compare robustness of different aggregation techniques through sensitivity analysis, performed by Monte-Carlo simulations on synthetic data. Distribution of biases where obtained for various techniques and some showed to perform significantly better than others. Robustness assessment still depended on models used in that work.

Le reste de cette monographie est organis de la faon suivante : la section 2 dcrit intuitivement puis mathmatiquement le cadre propos ; la section 3 dtaille ensuite l'implmentation, la collecte des donnes pour les cas d'tude et les rsultats numriques pour une valuation intra-urbaine synthtique et un cas rel mtropolitain ; la section 4 discute finalement les limitations et les potentialits de la mthode.

2 Framework Description

2.1 Intuitive Description

We describe now the abstract framework allowing theoretically to compare robustnesses of evaluations of two different urban systems. Our framework is a generalization of an empirical method proposed in [?] besides a more general benchmarking study on indicator sense and relevance in a sustainability context. Intuitively, it relies on empirical base resulting from the following axioms :

- Urban systems can be seen from the information available, i.e. raw data describing the system. As a data-driven approach, this raw data is the basis of our framework and robustness will be determined by its structure.
- From data are computed indicators (objective functions). We assume that a choice of indicators is an intention to translate particular aspects of the system, i.e. to capture a realization of an "urban fact" (*fait urbain*) in the sense of Mangin [?] - a sort of stylized fact in terms of processes and mechanisms, having various realizations on spatially distinct systems, depending on each precise context.
- Given many systems and associated indicators, a common space can be built to compare them. In that space, data represents more or less well real systems, depending e.g. on initial scale, precision of data, missing data. We precisely propose to capture that through the notion of point cloud discrepancy, which is a mathematical tool coming from sampling theory expressing how a dataset is distributed in the space it is embedded in [?].

Synthesizing these requirements, we propose a notion of *Robustness* of an evaluation that captures both, by combining data reliability with relative importance,

1. *Missing Data* : an evaluation based on more refined datasets will naturally be more robust.
2. *Indicator importance* : indicators with more relative influence will weight more on the total robustness.

2.2 Formal Description

Indicators Let $(S_i)_{1\leq i\leq N}$ be a finite number of geographically disjoints territorial systems, that we assume described through raw data and intermediate indicators, yielding $S_i = (\mathbf{X}_i, \mathbf{Y}_i) \in \mathcal{X}_i \times \mathcal{Y}_i$ with $\mathcal{X}_i = \prod_k \mathcal{X}_{i,k}$ such that each subspace contain real matrices : $\mathcal{X}_{i,k} = \mathbb{R}^{n^X_{i,k} p^X_{i,k}}$ (the same holding for $\mathcal{Y}_i$). We also define an ontological index function $I_X(i,k)$ (resp. $I_Y(i,k)$) taking integer values which coincide if and only if the two variables have the same ontology in the sense of [?], i.e. they are supposed to represent the same real object. We distinguish "raw data" $\mathbf{X}_i$ from which indicators are computed via explicit deterministic functions, from "intermediate indicators" $\mathbf{Y}_i$ that are already integrated and can be e.g. outputs of elaborated models simulating some aspects of the urban system. We define the partial characteristic space of the "urban fact" by

$$(\mathcal{X}, \mathcal{Y}) \underset{def}{=} \left(\prod \tilde{\mathcal{X}}_c\right) \times \left(\prod \tilde{\mathcal{Y}}_c\right) = \left(\prod_{\mathcal{X}_{i,k} \in \mathcal{D}_\mathcal{X}} \mathbb{R}^{p^X_{i,k}}\right) \times \left(\prod_{\mathcal{Y}_{i,k} \in \mathcal{D}_\mathcal{Y}} \mathbb{R}^{p^Y_{i,k}}\right) \quad (1)$$

with $\mathcal{D}_{\mathcal{X}} = \{\mathcal{X}_{i,k} | I(i,k) \text{ distincts}, n^X_{i,k} \text{ maximal}\}$ (the same holding for $\mathcal{Y}_i$). It is indeed the abstract space on which indicators are integrated. The indices c introduced as a definition here correspond to different indicators across all systems. This space is the minimal space common to all systems allowing a common definition for indicators on each.

Let $\mathbf{X}_{i,c}$ be the data canonically projected in the corresponding subspace, well defined for all i and all c. We make the key assumption that all indicators are computed by integration against a certain kernel, i.e. that for all c, there exists H_c space of real-valued functions on $(\tilde{\mathcal{X}}_c, \tilde{\mathcal{Y}}_c)$, such that for all $h \in H_c$:

1. h is "enough" regular (tempered distributions e.g.)
2. $q_c = \int_{(\tilde{\mathcal{X}}_c, \tilde{\mathcal{Y}}_c)} h$ is a function describing the "urban fact" (the indicator in itself)

Typical concrete example of kernels can be :

– A mean of rows of $\mathbf{X}_{i,c}$ is computed with $h(x) = x \cdot f_{i,c}(x)$ where $f_{i,c}$ is the density of the distribution of the assumed underlying variable.
– A rate of elements respecting a given condition C, $h(x) = f_{i,c}(x)\chi_{C(x)}$
– For already aggregated variables $\mathbf{Y}$, a Dirac distribution allows to express them also as a kernel integral.

Aggregation Weighting objectives in multi-attribute decision-making is indeed the crucial point of the processes, and numerous methods are available (see [?] for a review for the particular case of sustainable energy management). Let define weights for the linear aggregation. We assume the indicators normalized, i.e. $q_c \in [0,1]$, for a more simple construction of relative weights. For i, c and $h_c \in H_c$ given, the weight $w_{i,c}$ is simply constituted by the relative importance of the indicator $w^L_{i,c} = \frac{\hat{q}_{i,c}}{\sum_c \hat{q}_{i,c}}$ where $\hat{q}_{i,c}$ is an estimator of q_c for data $\mathbf{X}_{i,c}$ (i.e. the effectively calculated value). Note that this step can be extended to any sets of weight attributions, by taking for example $\tilde{w}_{i,c} = w_{i,c} \cdot w'_{i,c}$ if $\mathbf{w}'$ are the weights attributed by the decision-maker. We focus here on the relative influence of attributes and thus choose this simple form for weights.

Robustness Estimation The scene is now set up to be able to estimate the robustness of the evaluation done through the aggregated function. Therefore, we apply an integral approximation method similar to methods introduced in [?], since the integrated form of indicators indeed brings the benefits of such powerful theoretical results. Let $\mathbf{X}_{i,c} = (\boldsymbol{X}_{i,c,l})_{1 \leq l \leq n_{i,c}}$ and $D_{i,c} = Disc_{\tilde{\mathcal{X}}_c, L^2}(\mathbf{X}_{i,c})$ the discrepancy of data points cloud[2] [?]. With $h \in H_c$, we have the upper bound on the integral approximation error

[2] The discrepancy is defined as the $L2$-norm of local discrepancy which is for normalized data points $\mathbf{X} = (x_{ij}) \in [0,1]^d$, a function of $\mathbf{t} \in [0,1]^d$ comparing the number of points falling in the corresponding hypercube with its volume, by $disc(\mathbf{t}) = \frac{1}{n}\sum_i \mathbb{1}_{\prod_j x_{ij} < t_j} - \prod_j t_j$. It is a measure of how the point cloud covers the space.

$$\left\| \int h_c - \frac{1}{n_{i,c}} \sum_l h_c(\boldsymbol{X}_{i,c,l}) \right\| \leq K \cdot |||h_c||| \cdot D_{i,c}$$

where K is a constant independent of data points and objective function. It directly yields

$$\left\| \int \sum w_{i,c} h_c - \frac{1}{n_{i,c}} \sum_l w_{i,c} h_c(\boldsymbol{X}_{i,c,l}) \right\| \leq K \sum_c |w_{i,c}| \, |||h_c||| \cdot D_{i,c}$$

Assuming the error reasonably realized ("worst case" scenario for knowledge of the theoretical value of aggregated function), we take this upper bound as an approximation of its magnitude. Furthermore, taking normalized indicators implies $|||h_c||| = 1$. We propose then to compare error bounds between two evaluations. They depend only on data distribution (equivalent to *statistical robustness*) and on indicators chosen (sort of *ontological robustness*, i.e. do the indicators have a real sense in the chosen context and do their values make sense), and are a way to combine these two type of robustnesses into a single value.

We thus define a *robustness ratio* to compare the robustness of two evaluations by

$$R_{i,i'} = \frac{\sum_c w_{i,c} \cdot D_{i,c}}{\sum_c w_{i',c} \cdot D_{i',c}} \tag{2}$$

The intuitive sense of this definition is that one compares robustness of evaluations by comparing the highest error done in each based on data structure and relative importance.

By taking then an order relation on evaluations by comparing the position of the ratio to one, it is obvious that we obtain a complete order on all possible evaluations. This ratio should theoretically allow to compare any evaluation of an urban system. To keep an ontological sense to it, it should be used to compare disjoints sub-systems with a reasonable proportion of indicators in common, or the same sub-system with varying indicators. Note that it provides a way to test the influence of indicators on an evaluation by analyzing the sensitivity if the ratio to their removal. On the contrary, finding a "minimal" number of indicators each making the ratio strongly vary should be a way to isolate essential parameters ruling the sub-system.

3 Results

Implementation Preprocessing of geographical data is made through QGIS [?] for performance reasons. Core implementation of the framework is done in R [?] for the flexibility of data management and statistical computations. Furthermore, the package `DiceDesign` [?] written for numerical experiments and sampling purposes, allows an efficient and direct computation of discrepancies. Last but not least, all source code is openly available on the `git` repository of the project[3] for reproducibility purposes [?].

3.1 Implementation on Synthetic Data

We propose in a first time to illustrate the implementation with an application to synthetic data and indicators, for intra-urban quality indicators in the city of Paris.

Data Collection We base our virtual case on real geographical data, in particular for *arrondissements* of Paris. We use open data available through the OpenStreetMap project [?] that provides accurate high definition data for many urban features. We use the street network and position of buildings within the city of Paris. Limits of *arrondissements*, used to overlay and extract features when working on single districts, are also extracted from the same source. We use centroids of buildings polygons, and segments of street network. Dataset overall consists of around $200k$ building features and $100k$ road segments.

Virtual Cases We work on each district of Paris (from the 1st to the 20th) as an evaluated urban system. We construct random synthetic data associated to spatial features, so each district has to be evaluated many time to obtain mean statistical behavior of toy indicators and robustness ratios. The indicators chosen need to be computed on residential and street network spatial data. We implement two mean kernels and a conditional mean to show different examples, linked to environmental sustainability and quality of life, that are required to be maximized. Note that these indicators have a real meaning but no particular reason to be aggregated, they are chosen here for the convenience of the toy model and the generation of synthetic data. With $a \in \{1 \ldots 20\}$ the number of the district, $A(a)$ corresponding spatial extent, $b \in B$ building coordinates and $s \in S$ street segments, we take

- Complementary of the average daily distance to work with car per individual, approximated by, with $n_{cars}(b)$ number of cars in the building (randomly generated by associated of cars to a number of building proportional to motorization rate α_m 0.4 in Paris), d_w distance to work of individuals (generated from the building to a uniformly generated random point in spatial extent of the dataset), and d_{max} the diameter of Paris area, $\bar{d}_w = 1 - \frac{1}{|b \in A(a)|} \cdot \sum_{b \in A(a)} n_{cars}(b) \cdot \frac{d_w}{d_{max}}$
- Complementary of average car flows within the streets in the district, approximated by, with $\varphi(s)$ relative flow in street segment s, generated through the minimum of 1 and a log-normal distribution adjusted to have 95% of mass smaller than 1 what mimics the hierarchical distribution of street use (corresponding to betweenness centrality), and $l(s)$ segment length, $\bar{\varphi} = 1 - \frac{1}{|s \in A(a)|} \cdot \sum_{s \in A(a)} \varphi(s) \cdot \frac{l(s)}{\max(l(s))}$

[3] at `https://github.com/JusteRaimbault/RobustnessDiscrepancy`

– Relative length of pedestrian streets $\bar{p}$, computed through a randomly uniformly generated dummy variable adjusted to have a fixed global proportion of segments that are pedestrian.

Arrdt	$< \bar{d}_w > \pm\sigma(\bar{d}_w)$	$< \bar{\varphi} > \pm\sigma(\bar{\varphi})$	$< \bar{p} > \pm\sigma(\bar{p})$	$R_{i,1}$
1 th	0.731655 ± 0.041099	0.917462 ± 0.026637	0.191615 ± 0.052142	1.000000 ± 0.000000
2 th	0.723225 ± 0.032539	0.844350 ± 0.036085	0.209467 ± 0.058675	1.002098 ± 0.039972
3 th	0.713716 ± 0.044789	0.797313 ± 0.057480	0.185541 ± 0.065089	0.999341 ± 0.048825
4 th	0.712394 ± 0.042897	0.861635 ± 0.030859	0.201236 ± 0.044395	0.973045 ± 0.036993
5 th	0.715557 ± 0.026328	0.894675 ± 0.020730	0.209965 ± 0.050093	0.963466 ± 0.040722
6 th	0.733249 ± 0.026890	0.875613 ± 0.029169	0.206690 ± 0.054850	0.990676 ± 0.031666
7 th	0.719775 ± 0.029072	0.891861 ± 0.026695	0.209265 ± 0.041337	0.966103 ± 0.037132
8 th	0.713602 ± 0.034423	0.931776 ± 0.015356	0.208923 ± 0.036814	0.973975 ± 0.033809
9 th	0.712441 ± 0.027587	0.910817 ± 0.015915	0.202283 ± 0.049044	0.971889 ± 0.035381
10 th	0.713072 ± 0.028918	0.881710 ± 0.021668	0.210118 ± 0.040435	0.991036 ± 0.038942
11 th	0.682905 ± 0.034225	0.875217 ± 0.019678	0.203195 ± 0.047049	0.949828 ± 0.035122
12 th	0.646328 ± 0.039668	0.920086 ± 0.019238	0.198986 ± 0.023012	0.960192 ± 0.034854
13 th	0.697512 ± 0.025461	0.890253 ± 0.022778	0.201406 ± 0.030348	0.960534 ± 0.033730
14 th	0.703224 ± 0.019900	0.902898 ± 0.019830	0.205575 ± 0.038635	0.932755 ± 0.033616
15 th	0.692050 ± 0.027536	0.891654 ± 0.018239	0.200860 ± 0.024085	0.929006 ± 0.031675
16 th	0.654609 ± 0.028141	0.928181 ± 0.013477	0.202355 ± 0.017180	0.963143 ± 0.033232
17 th	0.683020 ± 0.025644	0.890392 ± 0.023586	0.198464 ± 0.033714	0.941025 ± 0.034951
18 th	0.699170 ± 0.025487	0.911382 ± 0.027290	0.188802 ± 0.036537	0.950874 ± 0.028669
19 th	0.655108 ± 0.031857	0.884214 ± 0.027816	0.209234 ± 0.032466	0.962966 ± 0.034187
20 th	0.637446 ± 0.032562	0.873755 ± 0.036792	0.196807 ± 0.026001	0.952410 ± 0.038702

Table 1. Numerical results of simulation for each district with $N = 50$ repetitions. Each toy indicator value is given by mean on repetitions and associated standard deviation. Robustness ratio is computed relative to first district (arbitrary choice). A ratio smaller than 1 means that integral bound is smaller for upper district, i.e. that evaluation is more robust for this district. Because of the small size of first district, we expected a majority of district to give ratio smaller than 1, what is confirmed by results, even when adding standard deviations.

As synthetic data are stochastic, we run the computation for each district $N = 50$ times, what was a reasonable compromise between statistical convergence and time required for computation. Table 1 shows results (mean and standard deviations) of indicator values and robustness ratio computation. Obtained standard deviation confirm that this number of repetitions give consistent results. Indicators obtained through a fixed ratio show small variability what may a limit of this toy approach. However, we obtain the interesting result that a majority of districts give more robust evaluations than 1st district, what was expected because of the size and content of this district : it is indeed a small one with large administrative buildings, what means less spatial elements and thus a less robust evaluation following our definition of the robustness.

3.2 Application to a Real Case : Metropolitan Segregation

The first example was aimed to show potentialities of the method but was purely synthetic, hence yielding no concrete conclusion nor implications for policy. We propose now to apply it to real data for the example of metropolitan segregation.

Data We work on income data available for France at an intra-urban level (basic statistical units IRIS) for the year 2011 under the form of summary statistics (deciles if the area is populated enough to ensure anonymity), provided by INSEE[4]. Data are associated with geographical extent of statistical units, allowing computation of spatial analysis indicators.

Indicators We use here three indicators of segregation integrated on a geographical area. Let assume the area divided into covering units $\mathcal{S}_i$ for $1 \leq i \leq N$ with centroids (x_i, y_i). Each unit has characteristics of population P_i and median income X_i. We define spatial weights used to quantify strength of geographical interactions between units i, j, with d_{ij} euclidian distance between centroids : $w_{ij} = \frac{P_i P_j}{\left(\sum_k P_k\right)^2} \cdot \frac{1}{d_{ij}}$ if $i \neq i$ and $w_{ii} = 0$. The normalized indicators are the following

- Spatial autocorrelation Moran index, defined as weighted normalized covariance of median income by $\rho = \frac{N}{\sum_{ij} w_{ij}} \cdot \frac{\sum_{ij} w_{ij}\left(X_i - \bar{X}\right)\left(X_j - \bar{X}\right)}{\sum_i \left(X_i - \bar{X}\right)^2}$
- Dissimilarity index (close to Moran but integrating local dissimilarities rather than correlations), given by $d = \frac{1}{\sum_{ij} w_{ij}} \sum_{ij} w_{ij} \left|\tilde{X}_i - \tilde{X}_j\right|$ with $\tilde{X}_i = \frac{X_i - \min(X_k)}{\max(X_k) - \min(X_k)}$
- Complementary of the entropy of income distribution that is a way to capture global inequalities $\varepsilon = 1 + \frac{1}{\log(N)} \sum_i \frac{X_i}{\sum_k X_k} \cdot \log\left(\frac{X_i}{\sum_k X_k}\right)$

Numerous measures of segregation with various meanings and at different scales are available, as for example at the level of the unit by comparison of empirical wage distribution with a theoretical null model [**?**]. The choice here is arbitrary in order to illustrate our method with a reasonable number of dimensions.

[4] `http://www.insee.fr`

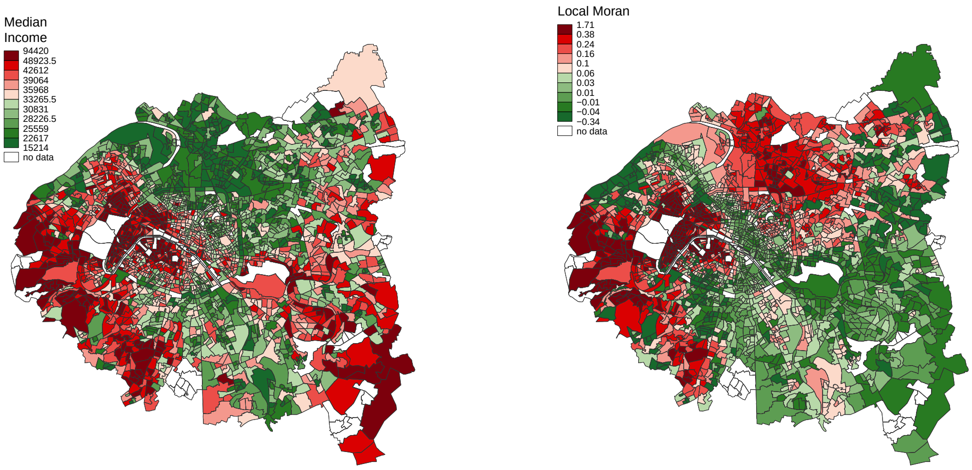

Fig. 1. Maps of Metropolitan Segregation. Maps show yearly median income on basic statistical units (IRIS) for the three departments constituting mainly the Great Paris metropolitan area, and the corresponding local Moran spatial autocorrelation index, defined for unit i as $\rho_i = N/\sum_j w_{ij} \cdot \frac{\sum_j w_{ij}(X_j - \bar{X})(X_i - \bar{X})}{\sum_i (X_i - \bar{X})^2}$. The most segregated areas coincide with the richest and the poorest, suggesting an increase of segregation in extreme situations.

Results We apply our method with these indicators on the Greater Paris area, constituted of four *départements* that are intermediate administrative units. The recent creation of a new metropolitan governance system [**?**] underlines interrogations on its consistence, and in particular on its relation to intermediate spatial inequalities. We show in Fig. 1 maps of spatial distribution of median income and corresponding local index of autocorrelation. We observe the well-known West-East opposition and district disparities inside Paris as they were formulated in various studies, such as [**?**] through the analysis of real estate transactions dynamics. We then apply our framework to answer a concrete question that has implications for urban policy : *how are the evaluation of segregation within different territories sensitive to missing data ?* To do so, we proceed to Monte Carlo simulations (75 repetitions) during which a fixed proportion of data is randomly removed, and the corresponding robustness index is evaluated with renormalized indicators. Simulations are done on each *department* separately, each time relatively to the robustness of the evaluation of full Greater Paris. Results are shown in Fig. 2. All areas present a slightly better robustness than the reference, what could be explained by local homogeneity and thus more fiable segregation values. Implications for policy that can be drawn are for example direct comparisons between areas : a loss of 30% of information on 93 area corresponds to a loss of only 25% in 92 area. The first being a deprived area, the inequality is increased by this relative lower quality of statistical information. The study of standard deviations suggest further investigations as different response regimes to data removal seem to exist.

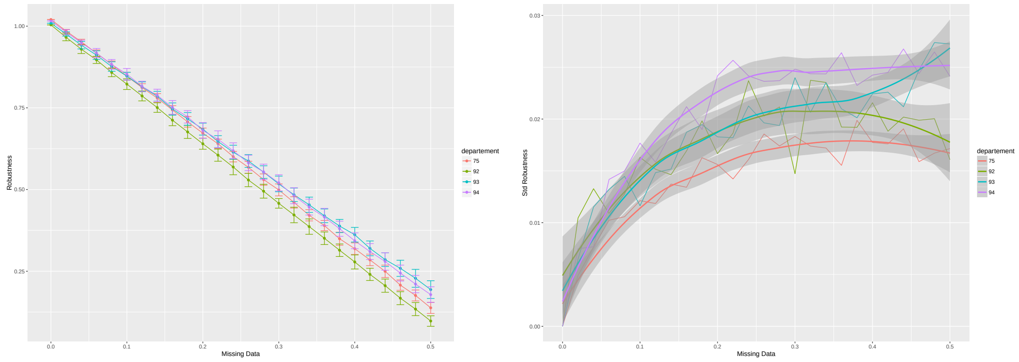

Fig. 2. Sensitivity of robustness to missing data. *Left.* For each department, Monte Carlo simulations (N=75 repetitions) are used to determine the impact of missing data on robustness of segregation evaluation. Robustness ratios are all computed relatively to full metropolitan area with all available data. Quasi-linear behavior translates an approximative linear decrease of discrepancy as a function of data size. The similar trajectory of poorest departments (93,94) suggest the correction to linear behavior being driven be segregation patterns. *Right.* Corresponding standard deviations of robustness ratios. Different regimes (in particular 93 against others) unveil phase transitions at different levels of missing data, meaning that the evaluation in 94 is from this point of view more sensitive to missing data.

4 Discussion

4.1 Applicability to Real situations

Implications for Decision-making The application of our method to concrete decision-making can be thought in different ways. First in the case of a comparative multi-attribute decision process, such as the determination of a transportation corridor, the identification of territories on which the evaluation may be flawed (i.e. has a poor relative robustness) could allow a more refined focus on these and a corresponding revision of datasets or an adapted revision of weights. In any case the overall decision-making process should be made more reliable. A second direction lays in the spirit of the real application we have proposed, i.e. the sensitivity of evaluation to various parameters such as missing data. If a decision appears as reliable because data have few missing points, but the evaluation is very sensitive to it, one will be more careful in the interpretation of results and taking the final decision. Further work and testing will however be needed to understand framework behavior in different contexts and be able to pilot its application in various real situations.

Integration Within Existing Frameworks The applicability of the method on real cases will directly depend on its potential integration within existing framework. Beyond technical difficulties that will surely appear when trying to couple or integrate implementations, more theoretical obstacles could occur, such as fuzzy formulations of functions or data types, consistency issues in databases, etc. Such multi-criteria framework are numerous. Further interesting work would be to attempt integration into an open one, such as e.g. the one described in [**?**] which calculates various indices of urban segregation, as we have already illustrated the application on metropolitan segregation indexes.

Availability of Raw Data In general, sensitive data such as transportation questionnaires, or very fine granularity census data are not openly available but provided already aggregated at a certain level (for instance French Insee Data are publicly available at basic statistical unit level or larger areas depending on variables and minimal population constraints, more precise data is under restricted access). It means that applying the framework may imply complicated data research procedure, its advantage to be flexible being thus reduced through additional constraints.

4.2 Validity of Theoretical Assumptions

A possible limitation of our approach is the validity of the assumption formulating indicators as spatial integrals. Indeed, many socio-economic indicators are not necessarily depending explicitly on space, and trying to associate them with spatial coordinates may become a slippery slope (e.g. associate individual economic variables with individual residential coordinates will have a sense only if the use of the variable has a relation with space, otherwise it is a non-legitimate artifact). Even indicators which have a spatial value may derive from non-spatial variables, as [**?**] points out concerning accessibility, when opposing integrated accessibility measures with individual-based non necessarily spatial-based (e.g. individual decisions) measures. Constraining a theoretical representation of a system to fit a framework by changing some of its ontological properties (always in the sense of real meaning of objects) can be understood as a violation of a fundamental rule of modeling and simulation in social science given in [**?**], that is that there can be an universal "language" for modeling and some can not express some systems, having for consequence misleading conclusion due to ontology breaking in the case of an over-constrained formulation.

4.3 Framework Generality

We argue that the fundamental advantage of the proposed framework is its generality and flexibility, since robustness of the evaluations are obtained only through data structure if ones relaxes constraints on the value of weight. Further work should go towards a more general formulation, suppressing for example the linear aggregation assumption. Non-linear aggregation functions would require however to present particular properties regarding integral inequalities. For example, similar results could search in the direction of integral inequalities for Lipschitzian functions such as the one-dimensional results of [**?**].

Conclusion

We have proposed a model-independent framework to compare the robustness of multi-attribute evaluations between different urban systems. Based on data discrepancy, it provide a general definition of relative robustness without any assumption on model for the system, but with limiting assumptions that are the need of linear aggregation and of indicators being expressed through spatial kernel integrals. We propose a toy implementation based on real data for the city of Paris, numerical results confirming general expected behavior, and an implementation on real data for income segregation on Greater Paris metropolitan areas, giving possible insights into concrete policy questions. Further work should be oriented towards sensitivity analysis of the method, application to other real cases and theoretical assumptions relaxation, i.e. the relaxation of linear aggregation and spatial integration.

Acknowledgments

The author would like to thank Julien Keutchayan (Ecole Polytechnique de Montréal) for suggesting the original idea of using discrepancy, and anonymous reviewers for the useful comments and insights.