# A Discrepancy-based Framework to Compare Robustness between Multi-Objective Evaluations

Juste Raimbault[1,2] and Julien Keutchayan[3]

[1] UMR CNRS 8504 Géographie-Cités, Paris, France
[2] UMR-T IFSTTAR 9403 LVMT, Champs-sur-Marne, France
[3] Département de mathématiques et de génie industriel,
Ecole Polytechnique de Montréal, Montréal, Canada

juste.raimbault@polytechnique.edu
julien.keutchayan@polytechnique.edu

**Abstract.** Multi-objective evaluation is a necessary aspect when managing complex systems, as the intrinsic complexity of a system is generally closely linked to the potential number of optimization objectives. In the case of spatial socio-technical systems (territorial and urban systems), the recent example of *eco-districts* [1] is a typical example where such an evaluation is central to the development of sustainable innovative concepts. However, an evaluation makes no sense without its robustness being given (in the ill-defined sense of reliability). Statistical robustness computation methods are highly dependent of underlying statistical models and a model-independent generic framework is still lacking. We propose a generic formulation of such a framework in the case of integrated aggregated indicators, that allows to define a relative measure of robustness taking into account data structure and indicator values. We implement and apply it to virtual data-driven cases of urban systems for Paris districts. First numerical results show the potentialities of this new method. Furthermore, its relative independence of system type and system model appears as an advantage compared to classical statistical robustness methods.

**Keywords:** Multi-objective Evaluation, Model-Independent Robustness, Urban System, Discrepancy

## 1 Introduction

### 1.1 General Context

Multi-objective problems are organically linked to the complexity of underlying systems. Indeed, either in the field of *Complex Industrial Systems*, in the sense of engineered systems, where construction of Systems of Systems (SoS) by coupling and integration always leads to contradictory objectives [2], or in the field of *Natural Complex Systems*, in the sense of physical, biological or social system that present emergence and self-organization properties, where objectives can

e.g. be the result of heterogeneous interacting agents (see [3] for a large survey of systems concerned by this approach), multi-objective optimization can be explicitly introduced to study or design the system but is often already implicitly ruling the internal mechanisms of the system. The case of socio-technical Complex Systems is particularly interesting as, following [4], they can be seen as hybrid systems embedding social agents into "technical artifacts" (sometimes to an unexpected degree creating what PICON describes as *cyborgs* [5]), and thus cumulate propensity to be at the origin of multi-objective issues[1]. The new notion of *eco-districts* [1] is a typical example where sustainability implies contradictory objectives. The example of transportation systems, which conception shifted during the second half of the 20th century from cost-benefit analysis to multi-criteria decision-making, is also typical of such systems [6]. Geographical system are now well studied from such a point of view in particular thanks to the integration of multi-objective frameworks within Geographical Information Systems [7]. As for the micro-case of eco-districts, meso and macro urban planning and design may be made sustainable through indicators evaluation [8].

A crucial aspect of an evaluation is a certain notion of its reliability, that we call in a fuzzy way *robustness*. Various definitions of robustness are possible in different frames, and it will have a precise definition in our framework. Statistics naturally include such a notion since the construction and estimation of statistical models give diverse indicators of the consistence of results [9]. The first example that comes to mind is the application of the law of large numbers to obtain the *p-value* of a model fit, that can be interpreted as a confidence measure of estimates. Besides, confidence intervals and *beta-power* are other important indicators of statistical robustness. Concerning multi-objective optimization, in particular through heuristic algorithms (for example genetic algorithms, or operational research solvers), the notion of robustness of a solution concerns more the stability of the solution on the phase space of the corresponding dynamical system. Recent progresses have been done towards unified formulation of robustness for a multi-objective optimization problem, such as [10] where robust Pareto-front as defined as solutions that are insensitive to small perturbations. In [11], the notion of degree of robustness is introduced, formalized as a sort of continuity of other solutions in successive neighborhood of a solution.

However, there still lack a generic method to estimate robustness of an evaluation that would be model-independent, i.e. that would be extracted from data structure and indicators but that would not depend on the method used. Some advantages could be for example an *a priori* estimation of potential robustness of an evaluation and thus to decide if the evaluation is worth doing. We propose here a framework answering this issue. It is data-driven and not model-driven in the sense that robustness estimation does not depend on how indicators are

---

[1] We design by *Multi-Objective Evaluation* all practices including the computation of multiple indicators of a system (it can be multi-objective optimization for system design or multi-objective evaluation of an existing system ; our examples being in the second case but the overall method staying rather generic).

computed, as soon as they respect some assumptions that will be detailed in the following.

## 1.2   Proposed Approach

*Objectives as Spatial Integrals* Our approach needs to restrict the scope of application. We assume that objectives can be expressed as spatial integrals, so it should apply to any territorial system and our application cases are urban systems. It is not much restrictive in terms of possible indicators if one uses suitable variables and integrated kernels : in a way analog to the method of geographically weighted regression [12], any spatial variable can be integrated against regular kernels of variable size and the result will be a spatial aggregation which sense depends on kernel size. The example we use in the following such as conditional means or sums suit well the assumption. Even an already spatially aggregated indicator can be interpreted as a spatial indicator by using a Dirac distribution on the centroid of the corresponding area.

*Linearly Aggregated Objectives* A second assumption we make is that the multi-objective evaluation is done through linear aggregation of objectives. 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 many domains apply such a method that can be called *Multi-Criteria Decision Analysis* [13]. 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 [14] 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.

The rest of the paper is organized as follows : section 2 describes intuitively and mathematically the proposed framework ; section 3 then details implementation, data collection for case studies and numerical results including sensitivity analysis ; section 4 finally discuss limitations and potentialities of the method.

## 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 [15] besides a more general benchmarking study on indicator sense and relevance in a sustainability context. Intuitively, it relies on the same simple empirical base which results from the following statements :

- 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 a will to translate particular aspects of the system, that is to capture a realization of an "urban fact" (*fait urbain*) in the sense of Mangin [16] - 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 themselves. In that space, data will represent more or less well real systems, depending e.g. on initial scale, precision of data, missing data. We precisely capture that through the notion of point cloud discrepancy, which is a mathematical tool coming from sampling theory which express how a dataset is well distributed in the space it is embedded in [17]. Coupling discrepancies with appropriate weights depending on indicator importance allows to introduce a relative ratio of robustness between two evaluations.

## 2.2   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 : $\mathcal{X}_{i,k} = \mathbb{R}^{n_{i,k}^X p_{i,k}^X}$ (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 [18], 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

$$
\begin{aligned}
(\mathcal{X}, \mathcal{Y}) \underset{def}{=} & \left( \prod \tilde{\mathcal{X}}_c \right) \times \left( \prod \tilde{\mathcal{Y}}_c \right) \\
= & \left( \prod_{\mathcal{X}_{i,k} \in \mathcal{D}_{\mathcal{X}}} \mathbb{R}^{p_{i,k}^X} \right) \times \left( \prod_{\mathcal{Y}_{i,k} \in \mathcal{D}_{\mathcal{Y}}} \mathbb{R}^{p_{i,k}^Y} \right)
\end{aligned}
\tag{1}
$$

, with $\mathcal{D}_{\mathcal{X}} = \{\mathcal{X}_{i,k} | I(i,k) \text{ distincts}, n_{i,k}^X \text{ maximal}\}$ (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.

## 2.3   Aggregation

As described in introduction, the multi-objective nature of the problem is tackled through linear aggregation. Although it may be seen as a reductive approach, it becomes an asset for the application of integral approximation theorems as the next step of our method. 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.

Weights result of two components, one deterministic and one "subjective" part permitting more flexibility and the inclusion of judgement in the decision process. For $i, c$ and $h_c \in H_c$ given, the weight $w_{i,c}$ is a mean of :

- objective weight, corresponding to relative importance of the indicator

$$w_{i,c}^L = \frac{\hat{q}_{i,c}}{\sum_c \hat{q}_{i,c}}$$

  where $\hat{q}_{i,c}$ is an estimator of $q_c$ for data $\mathbf{X}_{i,c}$ (i.e. the effectively calculated value)
- subjective weight : various potential methods (notations, importance order, etc.), reviewed in [13], are available to give a normalized subjective weight. In practice, these will be set to zero in our first implementation as we test the framework on virtual cases.

## 2.4   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 [19], 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

$$D_{i,c} = Discp_{\tilde{\mathcal{X}}_c, \infty}(\mathbf{X}_{i,c})$$

the discrepancy of data points cloud [20].

With $h \in H_c$, we have the upper bound on the integral approximation error

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

where $K$ is a constant regarding 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

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

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

We apply our framework to a "toy-model", as elaborated indicator definition nor raw data collection was not the primary aim of this work, which was more

to theoretically introduce the framework. It is however essential for such a data-driven paradigm to demonstrate through simulation the potentialities of the method, what we do here on semi-synthetic data.

### 3.1 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 [21] that provides accurate high definition data for many urban features. We use the street network and position of buildings within the city of Paris[2]. Limits of *arrondissements*, used to overlay and extract features when working on single districts, are also extracted from the same source. An example of visualization of used data is shown in Fig. 1.
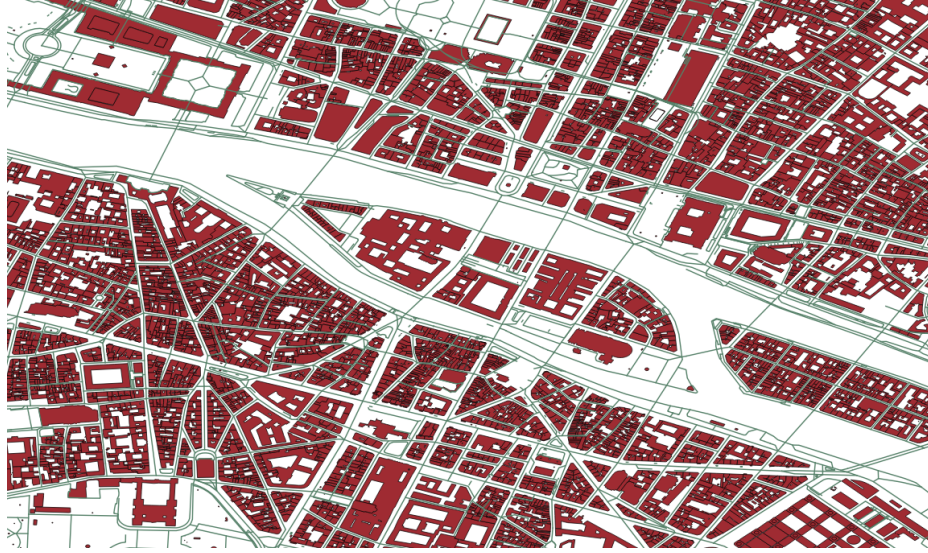


**Fig. 1.** Example of a visualization of used data for center area of Paris. We use centroids of buildings which polygons are filled in red, and segments of street network drawn in green. Dataset overall consists of around 200000 building features and 100000 road segments.

### 3.2 Implementation

Preprocessing of geographical data is made through QGIS [22] for performance reasons. Core implementation of the framework is done in R [23] for the flexibility of data management and statistical computations. Furthermore, the package

---

[2] for which the third-party website `https://mapzen.com/metro-extracts/` provides already packaged extracts as for many metropolitan areas of the world

`DiceDesign` [24] 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 [25].

### 3.3   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 $\alpha_m$ 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,

$$\bar{\varphi} = 1 - \frac{1}{|s \in A(a)|} \cdot \sum_{s \in A(a)} \varphi(s) \cdot \frac{l(s)}{\max(l(s))}$$

 – 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.

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

---

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

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.

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

**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.

# 4    Discussion

## 4.1    Applicability to 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 [26] which calculates various indices of urban segregation.

*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 administrative city level, smaller units have 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 [27] 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 [28], 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 overly 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 [29].

## Conclusion

We have proposed a model-independent framework to compare the robustness of multi-objective 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. Limiting assumption are however the need of linear aggregation and of indicators that can be expressed through spatial kernel integrals. We propose a toy implementation based on real data for the city of Paris and numerical results confirm general expected behavior. Further work should be oriented towards sensitivity analysis of the method, application to real cases and theoretical assumptions relaxation.

## References

1. Taoufik Souami. *Ecoquartiers: secrets de fabrication.* Scrineo, 2012.
2. R Timothy Marler and Jasbir S Arora. Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*, 26(6):369–395, 2004.
3. MEJ Newman. Complex systems: A survey. *arXiv preprint arXiv:1112.1440*, 2011.
4. Herman Haken and Juval Portugali. The face of the city is its information. *Journal of Environmental Psychology*, 23(4):385–408, 2003.
5. Antoine Picon. *Smart cities: théorie et critique d'un idéal auto-réalisateur.* B2, 2013.
6. Jean-Jacques Bavoux, Francis Beaucire, Laurent Chapelon, and Pierre Zembri. *Géographie des transports.* Paris, 2005.
7. Stephen J Carver. Integrating multi-criteria evaluation with geographical information systems. *International Journal of Geographical Information System*, 5(3):321–339, 1991.
8. Anne Jégou, Vincent Augiseau, Cécile Guyot, Cécile Judéaux, François-Xavier Monaco, Pierre Pech, et al. L'évaluation par indicateurs: un outil nécessaire d'aménagement urbain durable?. réflexions à partir de la démarche parisienne pour le géographe et l'aménageur. *Cybergeo: European Journal of Geography*, 2012.
9. Robert L Launer and Graham N Wilkinson. *Robustness in statistics.* Academic Press, 2014.
10. Kalyanmoy Deb and Himanshu Gupta. Introducing robustness in multi-objective optimization. *Evolutionary Computation*, 14(4):463–494, 2006.
11. C. Barrico and C.H. Antunes. Robustness analysis in multi-objective optimization using a degree of robustness concept. In *Evolutionary Computation, 2006. CEC 2006. IEEE Congress on*, pages 1887–1892, 2006.
12. Chris Brunsdon, Stewart Fotheringham, and Martin Charlton. Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(3):431–443, 1998.

13. Jiang-Jiang Wang, You-Yin Jing, Chun-Fa Zhang, and Jun-Hong Zhao. Review on multi-criteria decision analysis aid in sustainable energy decision-making. *Renewable and Sustainable Energy Reviews*, 13(9):2263–2278, 2009.
14. Melissa J Dobbie and David Dail. Robustness and sensitivity of weighting and aggregation in constructing composite indices. *Ecological Indicators*, 29:270–277, 2013.
15. A. Ali, I. Carneiro, L. Dussarps, F. Guédel, E Lamy, J. Raimbault, L. Viger, V. Cohen, T. Aw, and S. Sadeghian. Les eco-quartiers lus par la mobilité : vers une évaluation intégrée. Technical report, Ecole des Ponts ParisTech, June 2014.
16. David Mangin and Philippe Panerai. *Projet urbain*. Parenthèses, 1999.
17. Josef Dick and Friedrich Pillichshammer. *Digital nets and sequences: Discrepancy Theory and Quasi–Monte Carlo Integration*. Cambridge University Press, 2010.
18. Pierre Livet, Jean-Pierre Muller, Denis Phan, and Lena Sanders. Ontology, a mediator for agent-based modeling in social science. *Journal of Artificial Societies and Social Simulation*, 13(1):3, 2010.
19. Suzanne Varet. *Développement de méthodes statistiques pour la prédiction d'un gabarit de signature infrarouge*. PhD thesis, Université Paul Sabatier-Toulouse III, 2010.
20. H Niederreiter. Discrepancy and convex programming. *Annali di matematica pura ed applicata*, 93(1):89–97, 1972.
21. Jonathan Bennett. *OpenStreetMap*. Packt Publishing Ltd, 2010.
22. DT QGis. Quantum gis geographic information system. *Open Source Geospatial Foundation Project*, 2011.
23. R Core Team. R language definition, 2000.
24. Jessica Franco, Delphine Dupuy, Olivier Roustant, and Astrid Jourdan. Dicedesign-package. *Designs of Computer Experiments*, page 2, 2009.
25. Karthik Ram. Git can facilitate greater reproducibility and increased transparency in science. *Source code for biology and medicine*, 8(1):7, 2013.
26. Mihai Tivadar, Yves Schaeffer, André Torre, and Frédéric Bray. Oasis–un outil d'analyse de la ségrégation et des inégalités spatiales. *Cybergeo: European Journal of Geography*, 2014.
27. Mei-Po Kwan. Space-time and integral measures of individual accessibility: a comparative analysis using a point-based framework. *Geographical analysis*, 30(3):191–216, 1998.
28. Arnaud Banos. Pour des pratiques de modélisation et de simulation libérées en géographie et shs. *Thèse d'Habilitation à Diriger des Recherches, UMR CNRS 8504 Géographie-Cités, ISCPIF*, Décembre 2013.
29. SS Dragomir. The ostrowski's integral inequality for lipschitzian mappings and applications. *Computers & Mathematics with Applications*, 38(11):33–37, 1999.