

## Article

# Moduli of Continuity in Metric Models and Extension of Livability Indices

Roger Arnau , Jose M. Calabuig , Álvaro González  and Enrique A. Sánchez Pérez \* 

Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain; ararnnot@posgrado.upv.es (R.A.); jmcabalu@mat.upv.es (J.M.C.); agoncor@posgrado.upv.es (Á.G.)

\* Correspondence: easancpe@mat.upv.es

**Abstract:** Index spaces serve as valuable metric models for studying properties relevant to various applications, such as social science or economics. These properties are represented by real Lipschitz functions that describe the degree of association with each element within the underlying metric space. After determining the index value within a given sample subset, the classic McShane and Whitney formulas allow a Lipschitz regression procedure to be performed to extend the index values over the entire metric space. To improve the adaptability of the metric model to specific scenarios, this paper introduces the concept of a composition metric, which involves composing a metric with an increasing, positive and subadditive function  $\phi$ . The results presented here extend well-established results for Lipschitz indices on metric spaces to composition metrics. In addition, we establish the corresponding approximation properties that facilitate the use of this functional structure. To illustrate the power and simplicity of this mathematical framework, we provide a concrete application involving the modeling of livability indices in North American cities.

**Keywords:** Lipschitz; index; metric; extension; model; livability

**MSC:** 62J02; 26A16; 54E35



**Citation:** Arnau, R.; Calabuig, J.M.; González, Á.; Sánchez Pérez, E.A. Moduli of Continuity in Metric Models and Extension of Livability Indices. *Axioms* **2024**, *13*, 192. <https://doi.org/10.3390/axioms13030192>

Academic Editor: Anna Maria Fino

Received: 20 February 2024

Revised: 11 March 2024

Accepted: 12 March 2024

Published: 14 March 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Metric models, constructed from the aggregation of various variables, serve as useful frameworks that enable prospective approaches in fields such as the social sciences. One version of these models is given by so-called index spaces (see [1] and the related literature), which consist of a triple  $(M, d, I)$ , where  $(M, d)$  represents a metric space and  $I$  denotes an index (essentially, a non-negative Lipschitz function satisfying additional regularity properties). Such models prove advantageous in practical contexts, as they allow the extension of significant indices defined within a metric subspace of  $(M, d)$  to the entire space with the same Lipschitz norm, as seen in the classical scenario of the Lipschitz regression. This methodology is justified when attempting to model special circumstances that do not conform to linear constraints. The usefulness of these results has been demonstrated in recent years by a large number of research papers in various disciplines. For example, concrete applications can be observed in machine learning, where this conceptual framework is employed [2]. Analogous concepts are also widely used in other scientific fields [3,4].

However, there are many cases in which the metric of the problem is not established and has to be defined at some stage in the modeling process. This choice can be critical, as different metrics can lead to very different results. This article attempts to redefine this methodology by introducing a novel approach to obtain, in a straightforward manner, from a simple metric, another metric that better fits the problem. This newly proposed metric, denoted  $d_\phi$ , is chosen from the set of composition metrics  $\phi \circ d$ , where  $d$  is the original metric and  $\phi$  is a continuity modulus. In particular, we show that this adaptation of the initial metric space yields significant improvements over using the original metric structure

$(M, d)$ . To this end, we provide an example that consists of extending the AARP livability index, known for certain US cities, to a wider range of cities, thereby evaluating and analyzing the errors incurred under different considerations. Furthermore, we introduce a new category of indices, called standard indices, aimed at increasing the effectiveness of the proposed approximation technique. In all of these cases, the Lipschitz constant, together with other normalization properties related to the Katetov condition, serves as a central tool for controlling the resulting extension over all scenarios.

The ideas presented here are not new. The notion of modulus of continuity can be found in many works, usually in the context of mathematical analysis (see [5] and the references therein). However, it is not easy to find applied works in the literature that use this tool, which is essentially of a theoretical nature. The concept of composition metrics can be found, under various names, in several works, including some articles from the early part of the twentieth century. The notion of continuous metric transformations (metric-preserving functions in the literature) is first introduced in Wilson's 1935 article [6]. A paper by Borsík and Doboš in 1981 [5] compiles some of the conditions under which a given function composed with a metric will return a metric, and vice versa. Furthermore, the article by Valentine [7] already introduces a continuity modulus in the Lipschitz condition.

Therefore, our interest is to carry out all of this study in the framework of metric models related to the extension of Lipschitz functions, specifically by introducing a class of Lipschitz maps that aims to give a broader definition by incorporating an alternative metric for improved extension capabilities. This involves a more skillful adaptation to the dataset and an improved extension of the indices. Our research will explore the instances of this new space that are relevant to us. We will explore its relevance to the field of interest, in particular to the index theory, while analyzing its effectiveness and contrasting it with established methods. To achieve this, we have organized this paper as follows. In Section 2, we will present the main concepts on real Lipschitz functions and metric spaces and re-introduce the generalization proposal (composition metrics and other mathematical tools). Section 3 deals with the formal context of what we call indices. We will analyze spaces whose elements are these objects, exposing two methods of extension to larger domains, whose results, in practical contexts, will also be presented at the end of the paper. Thus, finally, in Section 4, we will explore the practical application of the above findings to algorithms that numerically execute the described extension procedures. We will delve into specific examples in which we examine their functionality, paying particular attention to performance differences compared to the original alternative and other potential procedures.

## 2. Basic Definitions and Concepts

Let us first explain some elementary facts on metric spaces and Lipschitz functions (see [8,9]). Consider a set  $D$  that is non-empty and a real map from  $d: D \times D \rightarrow \mathbb{R}^+$  (being the positive reals). It is said that  $d$  is a metric (distance) in the case that  $d(a, b) = 0$  if  $a = b$ , symmetric (that is,  $d(a, b) = d(b, a)$ ) and transitive ( $d(a, b) \leq d(a, c) + d(c, b)$ ) for  $a, b, c \in D$ . The metric  $d$  is bounded if  $d(x, y) < K$  for a constant  $K$ . A metric subspace of  $D$  is any subset  $D_0$  endowed with the same distance. A metric space as above is compact. For each sequence  $\{s_n\}_n \subseteq D$ , there is a subsequence  $\{s_{n_k}\}_k$  that tends to some element  $a \in D$ . This property is sometimes called the sequential characterization of compactness in metric spaces.

The maps we are interested in in this paper are those that satisfy the Lipschitz condition. A metric space valued function  $f$  from a metric space  $(D, d)$  to another metric space  $(R, r)$  is Lipschitz if there is a number  $K \in \mathbb{R}$  that satisfies:

$$r(f(x), f(y)) \leq Kd(x, y), \quad x, y \in D.$$

The best constant  $K$  above is what is called the Lipschitz norm (or constant), and sometimes it is said that  $f$  is  $K$ -Lipschitz. The symbol  $Lip(f)$  is used for the Lipschitz norm of  $f$ . Clearly,

$$\sup_{x,y \in D, x \neq y} r(f(x), f(y)) \cdot (d(x, y))^{-1} = Lip(f).$$

Suppose  $f$  is a real function defined on  $D$  but known only in  $S \subset D$ . Several classical results concern the extension of  $f$  to  $D$  while preserving some initial properties of  $f$  in  $S$ . The main result in this direction is when the property we want to preserve is the Lipschitz condition and is called the McShane–Whitney theorem [10,11]. Because of the useful extension formulas that can be obtained, this result has been the starting point for much research on the subject.

**Theorem 1.** *Let  $L > 0$ , and let  $f: S \subset D \rightarrow \mathbb{R}$  be an  $L$ -Lipschitz map. Then, there is always an extension of  $f$  preserving the Lipschitz constant  $L$ , that is, a function  $F: D \rightarrow \mathbb{R}$  satisfying  $F|_S = f$ . This extension can be given by any of the following formulas (sometimes called the Whitney and McShane extensions),*

$$F^W(x) = \inf_{y \in S} \{f(y) + Ld(x, y)\}, \quad F^M(x) = \sup_{y \in S} \{f(y) - Ld(x, y)\}$$

for every  $x \in D$ .

These extensions are in a sense optimal: if  $F$  is any other Lipschitz extension, we have  $F^M \leq F \leq F^W$ . Note also that  $F := tF^W + (1-t)F^M$  is a Lipschitz extension if  $t \in (0, 1)$ , and the Lipschitz norm is also  $L$ .

The explanation and proofs of the general results on Lipschitz maps can be found in [9].

### Composition Metrics and Modulus of Continuity

We now focus on the possibility of generalizing the Lipschitz condition to include a broader class of functions, satisfying a more relaxed criterion and thereby preserving the extension theorems. This paper will address this in a positive way with what we call  $\phi$ -Lipschitz maps. First, we will define the family of maps  $\Phi$ , which will serve as a form of the generalized continuity module [12]. However, our main focus will be on metrics that can be formulated as compositions of a pre-existing metric with these functions, rather than on the functions themselves. Our specific interpretation of such a continuity module is outlined below:

**Definition 1.** *A map  $\phi: \mathbb{R}^+ \rightarrow \mathbb{R}^+$  belongs to  $\Phi$  when, for every  $x, y \in \mathbb{R}^+$ , we have:*

- (i).  $\phi(x + y) \leq \phi(x) + \phi(y)$  and  $\phi(0) = 0$ .
- (ii).  $\phi(x) < \phi(y)$  when  $x < y$ .
- (iii).  $\phi$  is a continuous function.

As we said, such functions are often called moduli of continuity. Examples of functions belonging to  $\Phi$  are  $\phi(x) = x^\alpha$  for  $0 < \alpha < 1$  or the sigmoid function  $\sigma(x) = \frac{1}{1+e^{-x}}$ . Thus, a modulus of continuity is defined to be an increasing function  $\omega: [0, +\infty) \rightarrow [0, +\infty)$  which satisfies that  $\omega(0) = 0$  is continuous at 0. Sometimes, other requirements are imposed. For example, in [9], (Chapter 6, Section 4) continuity modules are related to metrics in a similar way to the one we present here, although restricted to the context of the study of uniform continuity.

Now, following the scheme of the definition of the Lipschitz map, we will introduce a function  $\phi \in \Phi$  of the distance. The definition is as follows.

**Definition 2.** If  $(D, d)$  and  $(R, r)$  are metric spaces, take a map  $f: D \rightarrow R$ . Suppose that  $\phi \in \Phi$ . Then,  $f$  is a  $\phi$ -Lipschitz map whenever we have the existence of a constant  $K > 0$  that satisfies that, for all  $x, y \in D$ ,

$$r(f(x), f(y)) \leq K\phi(d(x, y)).$$

As in the standard case, the constant  $K > 0$  is a Lipschitz constant of the  $\phi$ -Lipschitz map  $f$ . We will consider a new function  $\phi \circ d = d_\phi$  as the composition of the distance  $d$  with  $\phi \in \Phi$ .

We observe that every  $L$ -Lipschitz map is also a  $\phi$ -Lipschitz map for  $\phi(x) = x$  and  $K = L$ . In addition, consider the function  $f: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , defined as  $f(x) = \sqrt{x}$ . While this function is not Lipschitz, it meets the criteria of being  $\phi$ -Lipschitz for  $\phi = f \in \Phi$  and  $K = 1$ . Thus, this category of mappings effectively generalizes the notion of the  $L$ -Lipschitz map. We will see further improvements motivated by this construction in Section 3 by means of a different map,  $f(x) = \log(x + 1)$ , for example.

We propose redefining the distance to encompass new maps satisfying the proposed condition alongside the Lipschitz maps within the resulting metric space. By adjusting  $\phi$ , we gain direct control over modifying the original metric to better suit the posed problem. Subsequently, we will demonstrate that  $d_\phi$  indeed constitutes a valid distance.

**Proposition 1.** If  $(D, d)$  is a metric space, and supposing that  $\phi \in \Phi$ , then we have that  $d_\phi$  is a distance too.

**Proof.** Because  $\phi$  is an injective function (due to its strict monotonicity) and  $\phi(0) = 0$ , it is evident that  $d_\phi(x, y) = 0$  if and only if  $d(x, y) = 0$ . Since  $d$  is a metric, this occurs if and only if  $x = y$ , establishing the identity of indiscernibles for  $d_\phi$ . The symmetry of  $d$  implies directly the symmetry of  $d_\phi$ . Considering the triangle inequality, monotonicity and subadditivity of  $\phi$ , we have

$$d_\phi(x, z) \leq \phi(d(x, y) + d(y, z)) \leq d_\phi(x, y) + d_\phi(y, z).$$

Consequently,  $d_\phi$  is a distance.  $\square$

The reciprocal of this result is usually false. For example, if  $D = \mathbb{R}$  and we let  $\phi \in \Phi$  be  $\phi(x) = \sqrt{x}$  for  $x \geq 0$ , and  $d(x, y) = |x - y|^2$ , we know that  $d_\phi = |x - y|$  is a metric on  $D$ , while  $d$  is not. To illustrate this, consider taking  $x = 0$ ,  $y = 1$  and  $z = 3$ , and observe that  $d(x, z) > d(x, y) + d(y, z)$ , which violates the triangle inequality. However, there exists a sort of converse implication. Assuming that for any metric space  $(D, d)$ , the function  $d_\phi$  constitutes a metric, we can infer certain conditions that  $\phi$  must meet. For instance, it is straightforward to verify that  $\phi(0) = 0$ . Moreover, we will demonstrate next that  $\phi$  must exhibit subadditivity.

Let  $x, y \in \mathbb{R}^+$  and define a metric space  $D = \{a, b, c\}$  such that  $d(a, b) = x$ ,  $d(b, c) = y$  and  $d(a, c) = x + y$ . Such a  $d$  is evidently a metric on  $D$ . Since  $d_\phi$  is a metric, applying the triangle inequality yields  $\phi(x + y) = d_\phi(a, c) \leq d_\phi(a, b) + d_\phi(b, c) = \phi(x) + \phi(y)$ , thus establishing the subadditivity of  $\phi$ . Thus, if a function  $\phi: \mathbb{R}^+ \rightarrow \mathbb{R}^+$  satisfies that  $d_\phi$  is a metric for all metrics  $d$ , then  $\phi$  is subadditive and  $\phi(0) = 0$ . However, there might be instances where the required conditions of monotonicity and continuity for  $\phi$  are not fulfilled. In other words,  $d_\phi$  can qualify as a metric for every metric space  $(D, d)$  without  $\phi$  being strictly monotonically increasing or continuous. For instance, consider  $\phi$  defined as follows:  $\phi(x) = 0$  if  $x = 0$ , and  $\phi(x) = 1$  if  $x > 0$ . For each metric space  $(D, d)$ , we observe the following:

- For any  $a, b \in D$ , the condition  $d_\phi(a, b) = 0$  holds if and only if  $d(a, b) = 0$  because  $\phi$  is only null at 0. Since  $d$  is a metric, this holds if and only if  $a = b$ , demonstrating that  $d_\phi$  satisfies the identity of indiscernibles.
- Given that  $d(a, b) = d(b, a)$  for all  $a, b \in D$ , it is evident that  $d_\phi(a, b) = d_\phi(b, a)$ , indicating the symmetry of  $d_\phi$ .

- (c) Let  $a, b, c \in D$  with corresponding distances  $x = d(a, b)$ ,  $y = d(b, c)$  and  $z = d(a, c)$ . If  $a, b, c$  are distinct, then  $x, y, z > 0$  since  $d$  is a metric, leading to  $\phi(x) = \phi(y) = \phi(z) = 1$ .

Consequently,  $d_\phi$  satisfies the triangle inequality for  $a, b, c$ . In cases where some elements of  $D$  coincide, the triangular inequality for  $d_\phi$  trivially holds for them.

To better comprehend the relationship between  $d$  and  $d_\phi$ , we depict in Figures 1 and 2 a visual representation of the behavior of two specific metrics: the Euclidean metric  $d$  in  $\mathbb{R}$  and its composition with  $\phi(x) = \log(1 + x)$ , denoted as  $d_\phi$ . In Figure 1, it is evident that  $d_\phi$  “smoothens” the distance between two real numbers concerning  $d$ . More precisely,  $d$  and  $d_\phi$  exhibit similar behavior when  $x$  and  $y$  are “close” to each other, but, for more distant values,  $d_\phi$  dampens the growth in comparison to  $d$ . Figure 2 offers a comparison of the triangular inequality behavior between  $d$  and  $d_\phi$ . Once again, we observe how the logarithmic growth of  $\phi$  influences  $d_\phi$  and its triangular inequality,

$$d_\phi((x, y), (z, t)) = \log\left(1 + \sqrt{(x - z)^2 + (y - t)^2}\right). \quad (1)$$

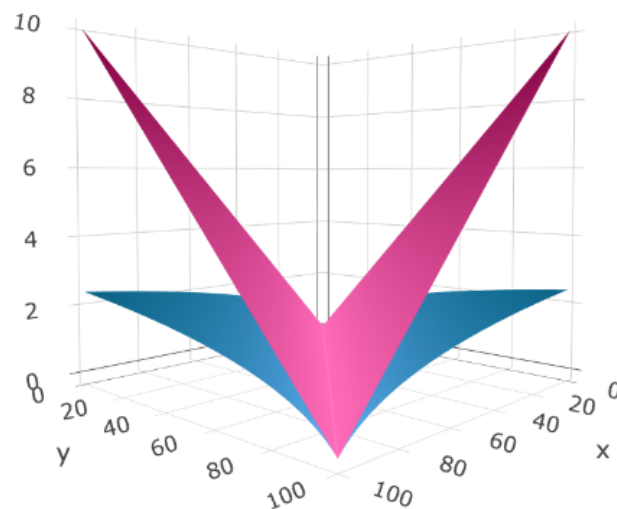


Figure 1. Comparison:  $d(x, y)$  (pink) and  $d_\phi(x, y)$  (blue).

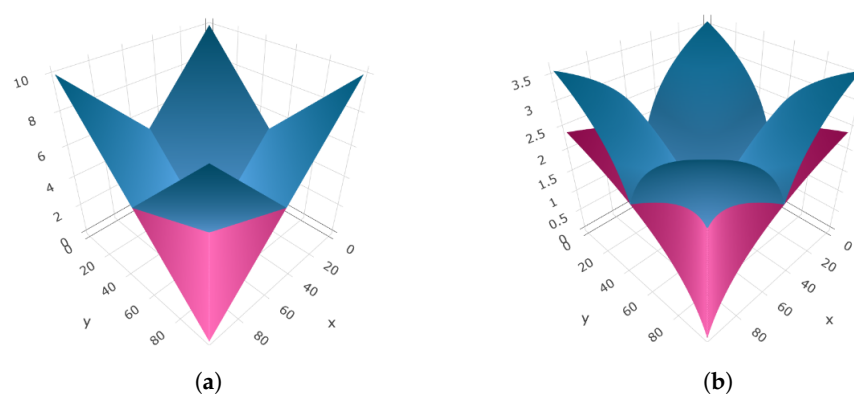


Figure 2. Visualization of the triangular inequality of both  $d$  and  $d_\phi$ : (a)  $d(x, y)$  (pink) versus  $d(x, 0) + d(0, y)$  (blue) and (b)  $d_\phi(x, y)$  (pink) versus  $d_\phi(x, 0) + d_\phi(0, y)$  (blue).

**Example 1.** Additional examples of  $\phi$ -Lipschitz maps are provided below:

- A mapping  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is termed  $\alpha$ -Hölder continuous if there exist constants  $C > 0$  and  $\alpha > 0$  satisfying for all  $x, y \in D$

$$r(f(x), f(y)) \leq Cd(x, y)^\alpha.$$

These mappings qualify as  $\phi$ -Lipschitz maps for  $K = C$ . Specifically, if  $\alpha \in (0, 1]$ , then defining  $\phi(x) = x^\alpha$  places  $\phi$  within  $\Phi$ , and  $f$  is  $\phi$ -Lipschitz with  $K = C$ . If  $\alpha > 1$ , it can be demonstrated that  $f$  becomes a constant map, rendering it  $\phi$ -Lipschitz for any  $\alpha \geq 0$ .

- Consider  $\mathbb{R}$  equipped with its standard metric, and let  $\phi \in \Phi$ . Subadditivity implies  $\phi(x) - \phi(y) \leq \phi(x - y)$  for  $x \geq y$ ; hence,  $d(\phi(x), \phi(y)) \leq \phi(d(x, y))$ . Thus, every  $\phi \in \Phi$  qualifies as a  $\phi$ -Lipschitz function with  $K = 1$ . Examples of  $\phi$  functions, in addition to those already mentioned, include  $\phi(x) = \arctan(x)$ ,  $\phi(x) = x(x + 1)^{-1}$  or  $\phi(x) = x(x^2 + 1)^{-1/2}$ .

The compactness for subsets of the metric spaces  $(D, d_\phi)$  in terms of the compactness of  $(D, d)$  is straightforward. For example, in  $(\mathbb{R}^2, d_\phi)$  with  $d_\phi$ , as in (1), any closed ball is compact.

**Proposition 2.** Suppose that  $(D, d)$  is compact. Then, for every function,  $\phi \in \Phi$   $(D, d_\phi)$  is also compact.

### 3. Index Extension

The aim of this section is to adapt the concepts of index and index space in the terms that were introduced in [1] to the context of the composition metrics. We will also explain the extension methods available to obtain an approximation framework for Lipschitz regression. The first is supported by what are known in [1] as standard indices, which are essentially defined by the distance to a reference point. The second one is an application of the already explained McShane and Whitney extension formulas.

Our models essentially consist of a metric space  $(D, d)$  and an index  $I$  (a real Lipschitz map) defined on it. These indices give meaningful values to the elements of  $D$ , which are the objects of the models. The Lipschitz condition of the index is the tool that introduces a certain concordance of the metric with the nature of the distance in  $D$ . The essence of the research we propose in this paper is that we introduce another tool to improve models based on index spaces: the construction of composition metrics  $d_\phi$  that improve the fit of the metric to the situation we want to model.

The term index space is defined in [1] as a triplet  $(D, d, I)$  where  $I: D \rightarrow \mathbb{R}$  serves as the index function, which is considered bounded if  $\sup_{a \in D} |I(a)| \leq C$  for some  $C > 0$ . The infimum of all such constants  $C$  satisfying this inequality is denoted as  $B(I)$ , represented by  $B(I) = \sup_{a \in D} |I(a)|$ . When discussing the normed space structure of the space of functions with the uniform norm, we use  $|\cdot|_\infty$  in place of  $B(\cdot)$ .

We will need two more boundedness-type properties to characterize the set of indices we are interested in.

**Definition 3.** Consider a function  $I: D \rightarrow \mathbb{R}$  to be an index. If  $Q > 0$ , we say that it is  $Q$ -normalized when for all  $a, b \in D$

$$d_\phi(a, b) \leq Q(|I(a)| + |I(b)|).$$

This kind of map is sometimes called a Katetov function. The normalization constant  $N(I)$  for  $I$  is the optimal  $Q$ . If  $\phi \in \Phi$ , the index  $I$  is said to be  $\phi$ -coherent for  $K > 0$  if for all  $a, b \in D$

$$|I(a) - I(b)| \leq K d_\phi(a, b),$$

which means that it is  $\phi$ -Lipschitz for  $K$ . The coherence constant  $C(I)$  for  $I$  is the infimum of all  $K$ 's.

In practical scenarios involving the construction of models using index spaces, the index value might not be readily accessible for every element within the metric space. This situation can arise when acquiring all the necessary data for the model's context is either costly or infeasible. Hence, it becomes advantageous to possess tools enabling the approximation of the index from non-indexed elements. As previously mentioned, this section proposes two methods to address this challenge: firstly, by identifying or



approximating the desired index using what we term as standard indices, and, secondly, by approaching the question as a Lipschitz regression problem using the classical extension formulas. Both techniques can be adapted to the case of composition metrics, as we will see below.

### 3.1. Standard Indices and Approximation

A distance function  $d$  allows us to define indices that are inherently natural, known as standard indices. For a given point  $a \in D$ , we can define an index  $I_a$  as  $I_a(b) := d(a, b)$ , where  $b \in D$ . These functions are called standard indices. Usually, the criterion for selecting  $a$  involves identifying the element that minimizes specific properties, although any other element could be chosen instead. The aim of this section is to introduce an additional tool, the composition metric  $d_\phi$ , which allows us to use the associated standard indices  $b \mapsto d_\phi(a, b)$ . Our interest, therefore, is in showing a method for approximation by means of standard indices or any other index. We will first present some results about the compactness of these function spaces to establish the framework. We will then turn our attention to the central issue of approximation. The main goal is to show that a robust approximation method is possible, at least in theory.

Note that uniformly bounded sets of indices (without Lipschitz-type requirements) are pointwise compact. Indeed, let  $C > 0$  and consider the set of functions  $\mathcal{F}_C := \{I : D \rightarrow \mathbb{R} : B(I) \leq C\}$ . There are two topologies that can be defined here: the norm topology (for  $B(I) := \sup_{a \in D} |I(a)|$ ) and the one of the pointwise convergence.

If a constant  $k$  satisfies  $|k| \leq C$ , take  $\mathcal{F}_{C,k} := \{I : D \rightarrow \mathbb{R} : k \leq B(I) \leq C\}$ . It is a compact space with the topology of pointwise convergence since each element of  $\{I_\eta\}_{\eta \in \Lambda}$  can be seen as a member of  $\prod_{a \in D} [k, C]$ , which is a product of compact sets in a space with the topology of the product. By Tychonoff's theorem, it is compact.

However, although we cannot expect compactness with respect to the topology of uniform convergence of the entire space of uniformly bounded functions, the space of standard indices is compact.

**Proposition 3.** *Let  $\phi \in \Phi$ . Let  $(D, d)$  be a compact metric space. Then there exists  $C > 0$  such that the space  $S := \{d_\phi(a, \cdot) : a \in D\}$  of standard indices associated with the metric  $d_\phi$ , is included in  $\mathcal{F}_C$ , which is compact in the uniform topology.*

**Proof.** First, note that due to the compactness of  $(D, d)$ , the metric  $d$  is bounded by a certain constant  $R > 0$ . Then,  $S$  is bounded by  $C := \phi(R)$  and so  $S \subset \mathcal{F}_C$ . In addition, let  $\{d_\phi(a_n, \cdot)\}_n \subset S$ . By the compactness of  $(D, d)$  again and by Proposition 2, we already know that  $(D, d_\phi)$  is compact, so the sequence  $a_n$  has a subsequence  $a_{n_k}$  that converges to  $a_0 \in D$  with respect to  $d_\phi$ . For any  $\varepsilon > 0$ , there exists  $k_0 \in \mathbb{N}$  such that  $d_\phi(a_{n_k}, a_0) < \varepsilon$  for all  $k \geq k_0$ . Then, for any  $b \in D$ , using the triangular inequality, we have for  $k \geq k_0$

$$\sup_{b \in D} |d_\phi(a_{n_k}, b) - d_\phi(a_0, b)| \leq \sup_{b \in D} d_\phi(a_{n_k}, a_0) = d_\phi(a_{n_k}, a_0) < \varepsilon.$$

Therefore,  $d_\phi(a_{n_k}, \cdot)$  converges uniformly to  $d_\phi(a_0, \cdot)$ .  $\square$

This allows us to obtain that there is always a best approximation for each Lipschitz index  $I$  by the elements of  $S$ , as stated in Corollary 1.

**Corollary 1.** *Let  $\phi \in \Phi$  and  $(D, d)$  be a compact metric space. For every function  $I$ , which is a Lipschitz index acting in the compact space  $(D, d_\phi)$ , there is an element  $a_0 \in S$  such that the standard index  $I_{a_0, \phi}$  is the best approximation to  $I$  in  $S$ , that is,*

$$\inf\{\|I - I_{a, \phi}\|_\infty : a \in D\} = \|I - I_{a_0, \phi}\|_\infty.$$

**Proof.** A direct compactness argument gives the proof. The function  $F : S \rightarrow \mathbb{R}$  defined as

$$F(I_{a,\phi}) = \|I - I_{a,\phi}\|_\infty, \quad a \in D,$$

is continuous (and even Lipschitz), since

$$\begin{aligned} |F(I_{a,\phi}) - F(I_{b,\phi})| &= \left| \|I - I_{a,\phi}\|_\infty - \|I - I_{b,\phi}\|_\infty \right| \leq \|I_{a,\phi} - I_{b,\phi}\|_\infty \\ &= \sup_{c \in D} |\phi \circ d(a, c) - \phi \circ d(b, c)| \leq \phi \circ d(a, b) = d_\phi(a, b), \end{aligned}$$

for all  $a, b \in D$ . The compactness of  $S$  (Proposition 3) proves the result.  $\square$

Let us now focus our attention on the approximation bound for general Lipschitz indices. In particular, we will show that we can always obtain an estimate of the error made by this approximation procedure in terms of the generalization of the normalization and the coherence constants. For technical reasons, we also need to introduce a specific type of sequence that will feature the result we are seeking.

A sequence  $a_{nn} \subset D$  is termed *pointwise Cauchy* if, for each  $b \in D$ , the sequence  $\lim_n d(a_n, b)$  exists. Every convergent sequence is pointwise Cauchy, but the converse is not necessarily true. For instance, the sequence  $a_n = 1/n$  in  $D = (0, 1]$  with the Euclidean metric serves as an example of a non-convergent pointwise Cauchy sequence. The definition above can directly be translated for the composition metrics  $d_\phi, \phi \in \Phi$ .

The following result will finally give us the tool to approximate an index using the standard indices, in the sense of an improvement to [1], (Th. 1). We will also see that this approximation depends on the product of the normalization and coherence constants. Specifically, we will see these results for the following set of indices to be approximated:

$$\mathcal{R}_{K,Q,C} := \left\{ I \geq 0 : \begin{array}{l} |I(a) - I(b)| \leq Kd_\phi(a, b), \\ \frac{1+KQ}{K} \inf(I) + d_\phi(a, b) \leq Q(I(a) + I(b)), \\ B(I) \leq C \end{array} \right\}.$$

The proof of the following result is similar to the theorem discussed before. We include the proof for the aim of completeness.

**Theorem 2.** Let  $K, Q > 0$  such that  $KQ \geq 1$ . For every  $I \in \mathcal{R}_{K,Q,C}$ , there is a sequence  $\{a_n\}_n$  that is pointwise Cauchy and  $I(b) \leq \inf(I) + \lim_n Kd_\phi(a_n, b) \leq KQI(b)$ ,  $b \in D$ .

**Proof.** Take  $b \in D$  and fix  $n \in \mathbb{N}$ . Then, for every  $n \in \mathbb{N}$ , there is an element  $a_n \in D$  such that  $I(a_n) - \frac{1}{n} \leq \inf(I)$  and so

$$\begin{aligned} \inf(I) + Kd_\phi(a_n, b) &\leq KQI(a_n) + KQI(b) - KQ\inf(I) \\ &\leq KQI(b) + KQI(a_n) - KQ\left(I(a_n) - \frac{1}{n}\right) = KQI(b) + \frac{KQ}{n}. \end{aligned}$$

In addition,

$$I(b) - I(a_n) \leq |I(b) - I(a_n)| \leq Kd_\phi(a_n, b),$$

and therefore

$$I(b) \leq Kd_\phi(a_n, b) + I(a_n) \leq Kd_\phi(a_n, b) + \inf(I) + \frac{1}{n} \leq KQI(b) + \frac{1+KQ}{n},$$

for all  $n \in \mathbb{N}$  and  $b \in D$ . Now, observe that  $d_\phi(a_n, \cdot) = I_{a_n}(\cdot) \in S$ . Thus, leveraging the compactness of  $S$ , as demonstrated in Proposition 3, there exists a subsequence  $a_{n_k}$  such



that  $\lim_k d_\phi(a_{n_k}, b) = d_\phi(a_0, b)$  for each  $b \in D$  and some  $a_0 \in D$ . Consequently,  $a_{n_k}$  forms a pointwise Cauchy sequence, and, based on the last inequality, we conclude that

$$I(b) \leq \inf(I) + K \lim_k d_\phi(a_{n_k}, b) \leq KQI(b), \quad b \in D.$$

□

So, the function  $\tilde{I} := \inf(I) + K \lim_k d_\phi(a_{n_k}, \cdot)$  can be seen as an approximation for  $I \in \mathcal{R}_{K,Q,C}$ . The maximum error committed is bounded by

$$\sup_{b \in D} |\tilde{I}(b) - I(b)| \leq \sup_{b \in D} |KQI(b) - I(b)| = (KQ - 1)C.$$

If  $KQ \approx 1$ , the comparison between  $\tilde{I}$  and  $I$  is reasonable. However, as these constants increase, the approximation may deteriorate.

Furthermore, by setting  $b = a_0$  in the result of Theorem 2, we obtain

$$I(a_0) \leq \inf(I) + K \lim_k d_\phi(a_{n_k}, a_0) = \inf(I) + Kd_\phi(a_0, a_0) = \inf(I).$$

Therefore,  $I(a_0) = \inf(I)$ . Hence, we can express  $\tilde{I}(\cdot) = I(a_0) + d_\phi(a_0, \cdot)$ , where  $a_0 \in D$  is a point at which  $I$  attains its minimum. Additionally, for the indices we work with, we can assume  $\inf(I) = 0$ , allowing us to make this approximation for any  $Q$ -normalized and  $\phi$ -coherent index for  $K$ . In this scenario,  $K\tilde{I}(\cdot) = d_\phi(a_0, \cdot)$ .

**Remark 1.** The condition  $QK \geq 1$  stated in Theorem 2 is, in a way, universal in index spaces. This is because  $I \geq 0$ , and  $I$  reaches its minimum at  $b \in D$  (which occurs, for example, when  $D$  is finite or compact), and this minimum is 0. Hence,

$$\begin{aligned} I(a) - I(b) &= |I(a) - I(b)| \leq Kd_\phi(a, b) \\ &\leq KQ(I(a) + I(b)) = KQI(a). \end{aligned}$$

Consequently, if there is  $b \in D$  with  $I(b) = 0$ , we also have  $KQ \geq 1$ . Moreover, even if such a scenario does not arise, in the case of  $I$  within a compact space, we have  $\inf(I) = I(b)$  for a  $b \in D$ , implying that  $I_0(a) := I(a) - I(b)$  constitutes a non-negative index that retains the characteristics of  $I$ , with  $I_0(b) = 0$ .

### 3.2. McShane and Whitney Formulas as Approximation Tools

Another approach to the extension problem is to formulate an estimate by means of the expressions for extending Lipschitz functions outlined in Section 2. While there are alternative expansion techniques, such as those discussed in [13], which may be more suitable for specific purposes, for our overarching application it proves more advantageous to use the classical McShane and Whitney formulas. The method follows the procedure explained in [1], adapted to the case of the composition metric. A full explanation can be found in Section 4 of that article and in the example given in Section 5 of the same article.

**Remark 2.** To ensure that the composition  $d_\phi$  defines a metric, in the proof of Proposition 1, we only considered properties (i)–(iii) of  $\Phi$ . That is, the continuity property is not necessary, although it was needed to ensure the compactness results for standard indices and to obtain the result of Theorem 2. However, this is not the case for these results on the McShane–Whitney extensions, since all that is required is that  $d_\phi$  defines a metric. Thus, the functions of  $\Phi$  can be more general than a continuity modulus for this technique.

Convex combinations of the classical extension formulas can also be used. Furthermore, as in the case of standard indices, the product of the normalization and coherence constants is related to the accuracy of the approximation. Note that

$$\begin{aligned} I^W(a) &= \inf_{b \in D} \{I(b) + Kd_\phi(a, b)\} \\ &\leq \inf_{b \in D} \{I(b) + KQ(I(a) + I(b))\} \\ &= (1 + KQ) \inf_{b \in D} I(b) + KQI(a), \end{aligned}$$

for any  $a \in D$ . So,

$$\sup_{a \in D} |I^W(a) - I(a)| \leq (1 + KQ) \left| \inf_{b \in D} I(b) \right| + |KQ - 1| \sup_{a \in D} |I(a)|.$$

If we suppose that  $\inf_{b \in D} I(b) = 0$  and  $KQ \geq 1$ , as in the case of standard indices, the last bound is reduced to

$$\sup_{a \in D} |I^W(a) - I(a)| \leq (KQ - 1)C,$$

what indicates that the approximation improves when  $KQ \rightarrow 1$ . Similar bounds can be found for the McShane formula and then for any convex combination of both.

#### 4. Applications: The Livability Index for Cities

In this section, we outline a methodological approach for practically implementing the theoretical concepts discussed in this paper to extend a given index. We begin with a finite set of elements characterized by certain real variables, where the value of a specific index  $I$  needs to be determined. For some of these elements, the value of the index is already known. Our objective is to extrapolate this known information to estimate the value of the index for elements where it is not yet defined. Mathematically, this set forms a metric space  $D$ , equipped with a suitable distance metric tailored to the nature of the data or the problem under consideration. Additionally, there exists a known index within a subset  $D_0 \subset D$  that we aim to extend to the entire space  $D$ . To accomplish this objective, we propose the next procedure.

##### 4.1. Methodology

The first matter we analyze is whether the diverse characteristics of the variables could distort the distance under consideration, given the heterogeneity of their scales. To mitigate this potential problem, we suggest normalizing the variables to a common scale by the following method: we subtract the minimum and divide by the interval in the range. That is, let  $D = \{y_j\}_{j=1}^n$  and  $y_j = (x_1^j, \dots, x_m^j)$ . Let  $a_k := \max_j x_k^j$  and  $b_k := \min_j x_k^j$  for each  $k = 1, \dots, m$ . We then transform

$$x_k^j \sim \frac{x_k^j - b_k}{a_k - b_k},$$

for all  $j$  and  $k$ , ensuring that the new variables are restricted to the interval  $[0,1]$ , maintaining the same scale.

In order to evaluate the accuracy of the upcoming approximation, we require a measure of the error incurred, the Root Mean Square Error (RMSE). This metric provides the expected error that is given by

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{j=1}^n (\tilde{I}(a_j) - I(a_j))^2},$$

where  $a_1, \dots, a_n$  represent the observations for which we aim to estimate the error, and  $\tilde{I}$  denotes the approximation to  $I$ . However, given the absence of information about the index

values we seek to approximate, we require a strategy to estimate this error. In our approach, we partition  $D_0$ , the subset of observations with known indices, into two subsets: a training set and a test set. Seventy percent of the total observations are randomly selected for the training set, which we use to perform the expansion. The remaining observations from the test set are then utilized to calculate the Root Mean Square Error (RMSE). However, the inherent randomness in this process may influence the resulting error, rendering it potentially non-representative. To address this issue, we employ the method of cross-validation. This implies the repetition several times (twenty) and considers the respective errors to give a more consistent result in terms of accuracy.

In order to select a  $\phi$  function that best fits the model, we engage in an optimization process aimed at minimizing the error bound on the test set. According to Theorem 2, this means minimizing the product of the coherence and normalization constants. We would partition our dataset into three subsets: the training and test subsets mentioned earlier, along with a validation subset used for fitting. However, due to the potentially limited number of available observations, the resulting subsets may not be sufficiently significant for this study. Therefore, we will utilize the values obtained from the test set as a reference for our optimization process. To accomplish this, we will consider that the linear combination of functions in  $\Phi$  with positive scalars is another function in  $\Phi$ . We will first select a set of elementary functions  $(\phi_j)_{j=1}^n \subset \Phi$ , and then examine for which values  $\lambda_1, \dots, \lambda_n \geq 0$  the function  $\phi := \lambda_1 \phi_1 + \dots + \lambda_n \phi_n$  ensures that the metric  $d\phi$  is optimal in terms of the bound. For this purpose, we will employ the particle swarm optimization algorithm available in the R library “pso”. Unlike algorithms based on the gradient of the function, this type of algorithm explores the entire possible set of parameters, thereby avoiding convergence to non-absolute minima.

Taking the Whitney and McShane extensions, we may analyze whether we can consider an intermediate extension that optimizes the error. That is to say, to find the value  $\alpha \in [0, 1]$  for which the formula  $I := (1 - \alpha)I^W + \alpha I^M$  makes the error to be minimum. The better way to obtain this real number would be by considering a validation set, but, as we explained above, we will use the test set instead. So, we determine the parameter  $\alpha$  using the next result.

**Proposition 4.** Let  $I: D_0 \subset D \rightarrow \mathbb{R}$  be an index that is  $\phi$ -coherent,  $K > 0$ . Let  $S_1, S_2 \subset D_0$  such that  $S_1 \cup S_2 = D_0$  and  $S_1 \cap S_2 = \emptyset$ . Consider

$$I^W(b) = \inf\{a \in S_1 : I(a) + Kd_\phi(a, b)\}, \quad I^M(b) = \sup\{a \in S_1 : I(a) - Kd_\phi(a, b)\}.$$

Naming  $I_\alpha^E := (1 - \alpha)I^W + \alpha I^M$ , then

$$\min_{0 \leq \alpha \leq 1} \sum_{b \in S_2} \left( I(b) - I_\alpha^E(b) \right)^2 = \sum_{b \in S_2} \left( I(b) - I_{\alpha_0}^E(b) \right)^2,$$

for

$$\alpha_0 = \frac{\sum_{b \in S_2} (I^W(b) - I(b)) (I^W(b) - I^M(b))}{\sum_{b \in S_2} (I^W(b) - I^M(b))^2}.$$

**Proof.** Consider  $F(\alpha) = \sum_{b \in S_2} (I(b) - I_\alpha^E(b))^2$ . We can see that

$$F(\alpha) = \sum_{b \in S_2} \left( I(b) - I^W(b) + \alpha (I^W(b) - I^M(b)) \right)^2,$$

so

$$F'(\alpha) = 2 \sum_{b \in S_2} \left( I(b) - I^W(b) + \alpha (I^W(b) - I^M(b)) \right) (I^W(b) - I^M(b)) \\ + 2\alpha \sum_{b \in S_2} \left( I^W(b) - I^M(b) \right)^2.$$

The solution of the equation  $F'(\alpha) = 0$  gives the desired value of  $\alpha_0$ , taking into account that

$$F''(\alpha) = 2 \sum_{b \in S_2} \left( I^W(b) - I^M(b) \right)^2 \geq 0.$$

□

We opted to utilize the Root Mean Square Error (RMSE) as the error metric instead of other alternatives because it ensures the differentiability of the function  $F$  under consideration. This facilitates a straightforward analysis of its minimum, as demonstrated in the previous proposition. Opting for other error measures defined in terms of absolute value, such as Mean Absolute Error (MAE) or Symmetric Mean Absolute Percentage Error (SMAPE), would necessitate a more intricate and indirect approach to determine the optimal  $\alpha$  due to the non-differentiable nature of such definitions. If we consider extending by identifying the index with a standard one, we proceed to locate the element  $a_0 \in D$  that minimizes the index. In this scenario, we take  $\tilde{I}(b) := Kd_\phi(a_0, b)$  to find an estimate of  $I(b)$ , taking into account  $\min(I) = 0$  after processing. However, it is also important to note that the choice of how we quantify an error can lead to different extensions. For example, RMSE is known to be sensitive to outliers, so minimizing this measure tends to result in models that are more sensitive to large errors. Conversely, favoring SMAPE minimization leads to models that take into account the proportional accuracy of predictions, making them more robust to outliers and better suited to scenarios in which relative accuracy is important. Thus, the choice ultimately depends on the specific objectives and requirements of the problem domain.

#### 4.2. Extension of a Livability Index

In the following, we will validate our proposed approach using the so-called AARP Livability Index. In 2018, 55 percent of the world's population lived in urban areas, a figure that is expected to rise to 68 percent by 2050, according to [14]. The rapid urbanization emphasizes the increasing significance of examining and quantifying concepts such as quality of life or livability in cities, as evidenced by various indices detailed in [15]. The objectives of these indices are twofold: firstly, to conceptualize livability and determine the parameters that characterize it and, secondly, to offer insights into which cities or neighborhoods provide superior living conditions. Equipped with this knowledge, urban planners can identify areas in need of intervention, while public administrations can target development and investment initiatives in regions with poor living conditions. However, assessing livability can be a complex task due to the large number of factors involved, many of which are difficult to assess. For example, *The Economist's* "Global Liveability Index", one of the most widely used, includes thirty indicators across five categories, with some indicators building on others. In addition, certain factors, such as climate discomfort for travelers or levels of corruption, are subjective and inherently difficult to quantify. In this section, we propose the utilization of the index extension theory developed in the previous section to approximate livability using only indicators related to alternative mobility. The aim is to shift the focus away from subjective or complex social indicators and instead concentrate on factors that can be easily estimated based on existing infrastructure and the connectivity of urban patterns.

Let us elaborate on the datasets we will utilize. Walk Score® <https://www.walkscore.com> (accessed on 12 January 2024) is a website that assesses the walkability performance

of 123 cities in the United States and Canada on a scale from 0 to 100. This assessment is based on various parameters such as intersection density, block length, and access to amenities within a 5 min walk. Additionally, Walk Score employs a similar scoring system to evaluate cities' performances for walking and cycling. Our objective is to use these three indicators to approximate the AARP Livability Index <https://livabilityindex.aarp.org/scoring> (accessed on 11 November 2023). This index evaluates 61 different indicators across seven categories (housing, neighborhood, transportation, environment, health, engagement and opportunity) to gauge the livability of US cities, including factors such as housing costs, crime rates, air quality and income inequality. The resulting score ranges from 0 to 100, with 50 representing an average score, higher scores indicating above-average performance and vice versa.

In mathematical terms, our metric space is  $D \subset [0, 100]^3$ , where each element  $(x, y, z) \in D$  represents a city with walk, transport and bike scores  $x, y$  and  $z$ , respectively, equipped with the canonical metric  $d$  of  $\mathbb{R}^3$ . For 101 US cities, we have defined the index of interest, which we will refer to as  $I$ , and our objective is to define it for 22 Canadian cities as well. Table 1 provides an example of our dataset.

**Table 1.** Examples of scores and indices for some cities.

City	Walk Score	Transit Score	Bike Score	$I$
New York	88	88.6	69.3	63
Los Angeles	68.6	52.9	58.7	49
Chicago	77.2	65	72.2	57
Toronto	61	78.2	61	?
Houston	47.5	36.2	48.6	48
Montreal	65.4	67	72.6	?

We have implemented the extension of our index according to the considerations of Section 4.1. In particular, for illustrative purposes, we study our method by considering two linear combinations of functions of  $\Phi$  as follows:

$$\phi(x) = p_1 x + p_2 \log(1 + x) + p_3 \arctan(x) + p_4 \frac{x}{1 + x}, \quad p_1, \dots, p_4 \geq 0,$$

and

$$\psi(x) = p_1 \sqrt{x} + p_2 \log(1 + \sqrt{x}) + p_3 \arctan(\sqrt{x}) + p_4 \frac{\sqrt{x}}{1 + \sqrt{x}}, \quad p_1, \dots, p_4 \geq 0.$$

Table 2 presents a comparative analysis of the two extension procedures studied (standard indices and McShane–Whitney type approximation) and of the various metrics considered. The results show that writing our index as a standard index provides poor performance (a high mean error, making it unsuitable for consideration). As far as the McShane–Whitney-type formulas are concerned, it preserves the standard deviation, also reducing the expected RMSE a little. However, it should be noted that our technique requires a longer computation time, mainly due to the optimization process involved. This time can vary according to the optimization algorithm chosen and the dimension of the problem, although, as we will see, it can be similar to most techniques used today.

Furthermore, Table 3 provides the result when a convex expression as  $\psi$  is considered in this scenario, alongside the results of the classical technique. In particular, there is a better result when compared with the original metric, particularly noticeable in the case of standard indices. This variance in performance highlights the fact that the choice of the  $\phi$ -metric has a significant impact on the degree of improvement obtained relative to the original metric. Another option would be to use the concept of local Lipschitz to split up our problem and obtain locally better Lipschitz constants. We could see whether it is more efficient to have many fast classical local extensions, or one global but slower one, using an approximation in the context of the  $\phi$ -Lipschitz functions.

**Table 2.** Different procedures for the design of the function  $\phi$ .

Function $\phi$	Standard		McShane–Whitney	
	Lipschitz	$\phi$ -Lipschitz	Lipschitz	$\phi$ -Lipschitz
Mean RMSE	138.43	79.48	5.08	5.04
Median RMSE	140.49	81.25	5.12	5.05
Standard deviation	24.11	8.78	0.61	0.60
Seconds per iteration	$1.039 \times 10^{-3}$	$3.316 \times 10^{-1}$	$1.513 \times 10^{-3}$	$3.330 \times 10^{-1}$

**Table 3.** Different procedures for the design of the function  $\psi$ .

Function $\psi$	Standard		McShane–Whitney	
	Lipschitz	$\psi$ -Lipschitz	Lipschitz	$\psi$ -Lipschitz
Mean RMSE	138.43	16.69	5.08	4.55
Median RMSE	140.49	16.50	5.12	4.47
Standard deviation	24.11	1.52	0.61	0.63
Seconds per iteration	$1.039 \times 10^{-3}$	$3.000 \times 10^{-1}$	$1.513 \times 10^{-3}$	$3.013 \times 10^{-1}$

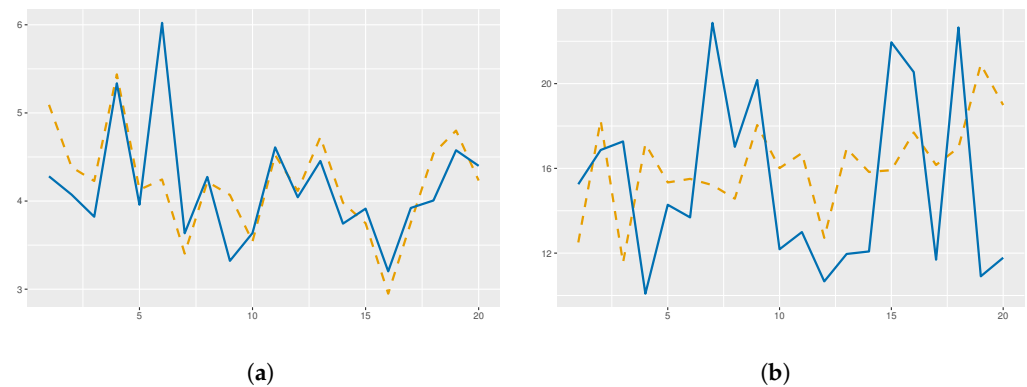
We also provide a comparison of these results with those obtained by applying different regression algorithms widely studied in the literature: neural networks and linear regression. The results of this comparison are summarized in Table 4. It is noteworthy that the results obtained by a neural network are comparable to those obtained by the McShane–Whitney technique, both in terms of the estimated error and its deviation and in terms of execution time. We can also compare the performance of the standard indices with that of a linear regression. For our example, similar results are obtained, except for the computation time, where linear regression proves to be more efficient. To allow a more comprehensive comparison of these behaviors, Figure 3 illustrates the different errors obtained in each iteration of the cross-validation. Each of the  $\psi$  metric models is plotted next to the most similar one discussed above.

**Table 4.** Performance of  $\psi$ -metric models and other regression methods.

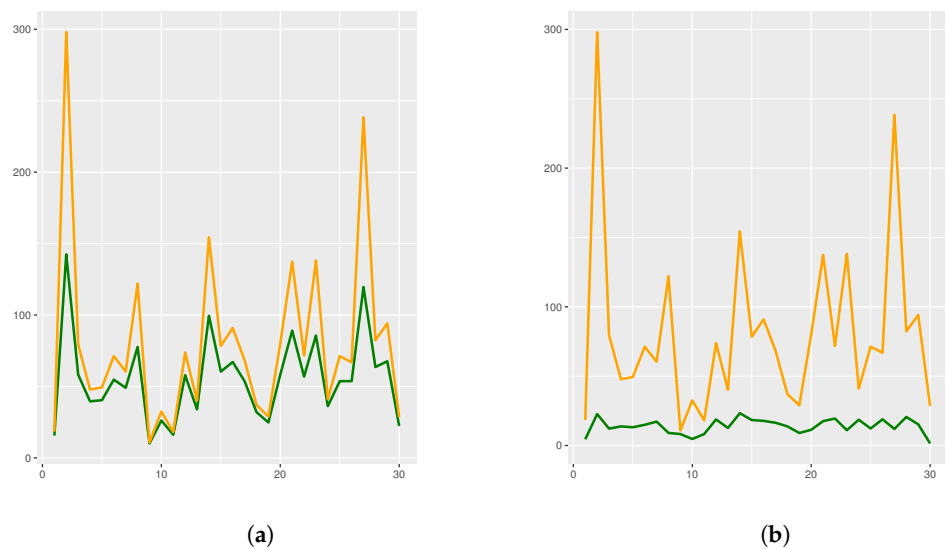
	Standard	McShane–Whitney	Neural Net	Linear
Mean RMSE	16.69	4.55	4.40	13.80
Median RMSE	16.50	4.47	4.42	13.65
Standard deviation	1.52	0.63	0.46	3.25
Seconds per iteration	$3.000 \times 10^{-1}$	$3.013 \times 10^{-1}$	$1.923 \times 10^{-1}$	$4.341 \times 10^{-3}$

It should also be noted that the way in which this improvement is achieved could vary from one method to another and from one metric to another. With this in mind, we have taken a concrete division into training and test sets (the 71 most populated cities will serve as the training and the rest as the test) and applied the results of the extension studied. In the case of standard indices, if we look at Figure 4, we can see that by redefining the metric to  $d_\phi$  or  $d_\psi$ , the improvement is general, i.e., we obtain a lower error for each element following the original trend. However, in the case of the McShane–Whitney method, Figure 5 shows that depending on the new metric, the accuracy can be different for each element, although on average it is lower, as we have seen before.

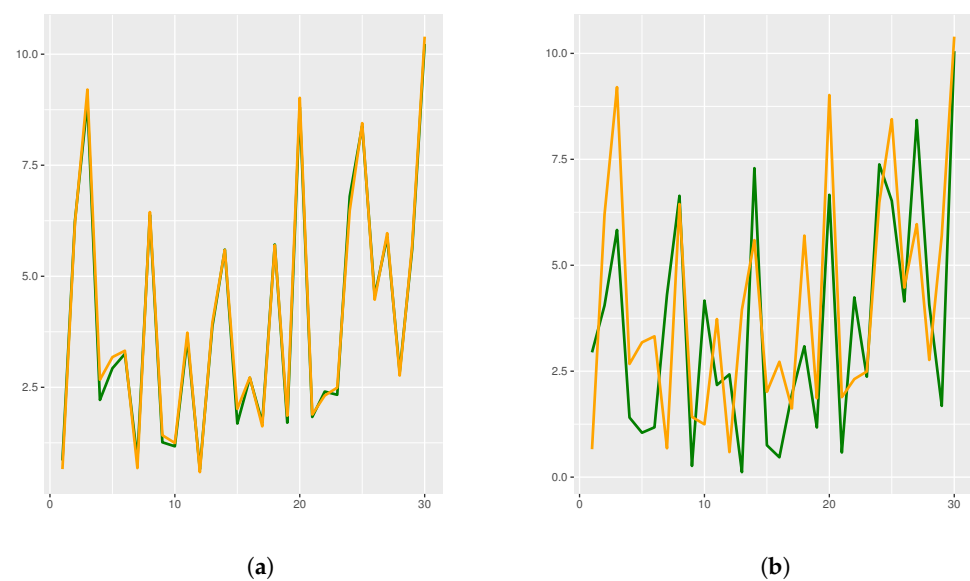




**Figure 3.** Comparison of extension errors using McShane–Whitney formulas for different metrics: (a) results for  $d$  (yellow) and  $d_\phi$  (blue) and (b) results for  $d$  (yellow) and  $d_\psi$  (blue).



**Figure 4.** Comparison of extension errors using standard indices for different metrics: (a) results for  $d$  (yellow) and  $d_\phi$  (green) and (b) results for  $d$  (yellow) and  $d_\psi$  (green).



**Figure 5.** Comparison of extension errors using McShane–Whitney formulas for different metrics: (a) results for  $d$  (yellow) and  $d_\phi$  (green) and (b) results for  $d$  (yellow) and  $d_\psi$  (green).

Finally, in Table 5, we will present the predictions generated by each method and provide a ranking based on these predictions. Given that we lack a benchmark for our index, we will assess its performance by comparing it with other existing indices when evaluating the resulting rankings. “The Mercer Quality of Living City Ranking” <https://mobilityexchange.mercer.com/Insights/quality-of-living-rankings> (accessed on 3 December 2023) ranks Canadian cities in the following increasing order: Calgary, Ottawa, Toronto and Vancouver. It is noteworthy that this ranking is largely consistent with the one we have derived. Although the standard indices may not offer an exact approximation of the index in question, they still provide a reliable overview for establishing a ranking.

**Table 5.** Ranking of Canadian cities.

Ranking	Standard	McShane–Whitney	Neural Net	Linear
1	Vancouver	Montreal	Vancouver	Vancouver
2	Toronto	Vancouver	Toronto	Toronto
3	Montreal	Longueuil	Montreal	Montreal
4	Burnaby	Toronto	Burnaby	Burnaby
5	Longueuil	Saskatoon	Longueuil	Longueuil
6	Mississauga	Winnipeg	Ottawa	Ottawa
7	Winnipeg	Burnaby	Winnipeg	Winnipeg
8	Ottawa	Mississauga	Mississauga	Surrey
9	Brampton	Ottawa	Brampton	Laval
10	Quebec	Brampton	Quebec	Mississauga
11	Surrey	Surrey	Laval	Kitchener
12	Laval	Quebec	Surrey	Brampton
13	Kitchener	Edmonton	Kitchener	Hamilton
14	Calgary	Kitchener	Calgary	Saskatoon
15	Saskatoon	Windsor	Gatineau	Calgary
16	Markham	Laval	Markham	Quebec
17	Hamilton	Hamilton	London	Windsor
18	Edmonton	Calgary	Hamilton	Edmonton
19	London	London	Edmonton	Vaughan
20	Gatineau	Gatineau	Windsor	Markham
21	Vaughan	Markham	Vaughan	London
22	Windsor	Vaughan	Saskatoon	Gatineau

## 5. Conclusions

We have introduced an innovative procedure for the extension of indices defined on metric subspaces of metric models. The main contribution of our results is that they include a new class of metrics (composition metrics) that give us more flexibility in the approximation tools for finding adapted results of Lipschitz regression processes. While the state-of-the-art Lipschitz extensions are highly dependent on the chosen metric, this methodology provides us with a straightforward way to improve extension results without a complex study of which metrics are well suited to the problem. In particular, the introduction of this wide class of metrics starting from a standard distance (e.g., the Euclidean distance) allows this improvement by reducing the Lipschitz constant. We provide the approximation formulas as well as the error bounds for the proposed procedure. However, the limitation of reducing the use of experimental data has to be taken into account, as some variables (those involved in the definition of the metric) are necessary for the application of the model. Only a subset of the index dataset can be artificially reconstructed using the proposed method; otherwise, the reliability of the results is not guaranteed.

As an applied example, we show how we can extend the so-called livability index for large urban centers in the United States, for which this index is known, from Canadian cities. These results are interesting as they provide more information on a topic of current interest. The results show that our techniques produce similar results to other widely used techniques, with the advantage of better interpretability.

The methodology presented here can be applied to a broad class of problems where classification is performed by one or more specific indices. Lipschitz extensions can be used to obtain a combined technique for analyzing livability indices as the core of our research plan. On the one hand, we measure experimental data to obtain real information. On the other hand, this information is complemented by a new classification of cities based on the extension of the experimental data. By adding both inputs in an optimal way, in this work, we have shown that consistent results can be obtained without the need to increase the experimental work too much, making quality research on the livability of cities possible using fewer resources.

**Author Contributions:** Conceptualization, R.A., Á.G. and E.A.S.P.; formal analysis, R.A. and E.A.S.P.; investigation, J.M.C., Á.G. and E.A.S.P.; methodology, R.A. and Á.G.; supervision, J.M.C.; writing—original draft, Á.G. and E.A.S.P.; writing—review and editing, R.A. and J.M.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** Enrique A. Sánchez Pérez was supported by Grant PID2020-112759GB-I00 funded by MCIN/AEI /10.13039/501100011033. Roger Arnau was supported by a contract of the Programa de Ayudas de Investigación y Desarrollo (PAID-01-21), Universitat Politècnica de València.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** Figures and algorithms are available via GitHub <https://github.com/Algoncor/Composition-metrics> (accessed on 28 February 2024).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Erdoğan, E.; Ferrer-Sapena, A.; Jiménez-Fernández, E.; Sánchez-Pérez, E.A. Index spaces and standard indices in metric modelling. *Nonlinear Anal. Model. Control* **2022**, *27*, 803–822. [\[CrossRef\]](#)
2. Calabuig, J.; Falciani, H.; Sánchez-Pérez, E. Dreaming machine learning: Lipschitz extensions for reinforcement learning on financial markets. *Neurocomputing* **2020**, *398*, 172–184. [\[CrossRef\]](#)
3. Mémoli, F.; Sapiro, G.; Thompson, P.M. Geometric surface and brain warping via geodesic minimizing Lipschitz extensions. In Proceedings of the 1st MICCAI Workshop on Mathematical Foundations of Computational Anatomy: Geometrical, Statistical and Registration Methods for Modeling Biological Shape Variability, Munich, Germany, 9 October 2006; pp. 58–67.
4. Dacorogna, B.; Gangbo, W. Extension theorems for vector valued maps. *J. Math. Pures Appl.* **2006**, *85*, 313–344. [\[CrossRef\]](#)
5. Doboš, J.B. Functions whose composition with every metric is a metric. *Math. Slovaca* **1981**, *31*, 3–12.
6. Wilson, W. On certain types of continuous transformations of metric spaces. *Am. J. Math.* **1935**, *57*, 62–68. [\[CrossRef\]](#)
7. Valentine, F.A. On the extension of a vector function so as to preserve a Lipschitz condition. *Bull. Am. Math. Soc.* **1943**, *49*, 100–108. [\[CrossRef\]](#)
8. Deza, E.; Deza, M.M. *Encyclopedia of Distances*; Springer: Berlin/Heidelberg, Germany, 2009.
9. Cobzaş, Ş.; Miculescu, R.; Nicolae, A. *Lipschitz Functions*; Springer International Publishing: Berlin/Heidelberg, Germany, 2019.
10. McShane, E.J. Extension of range of functions. *Bull. Am. Math. Soc.* **1934**, *40*, 837–842. [\[CrossRef\]](#)
11. Whitney, H. Analytic extensions of differentiable functions defined in closed sets. In *Hassler Whitney Collected Papers*; Springer: Berlin/Heidelberg, Germany, 1992; pp. 228–254.
12. Efimov, A. Modulus of Continuity, Encyclopaedia of Mathematics. 2001. Available online: <https://encyclopediaofmath.org/wiki/Continuity> (accessed on 22 October 2023).
13. Juutinen, P. Absolutely minimizing Lipschitz extensions on a metric space. *Ann. Acad. Sci. Fenn. Math.* **2002**, *27*, 57–67.
14. Ritchie, H.; Roser, M. Urbanization. *Our World Data* **2018**. Available online: <https://ourworldindata.org/urbanization?> (accessed on 4 October 2023).
15. Paul, A.; Sen, J. A critical review of liveability approaches and their dimensions. *Geoforum* **2020**, *117*, 90–92. [\[CrossRef\]](#) [\[PubMed\]](#)

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.