Brazilian Journal of Cartography (2016), Nº 68/6, Special Issue GEOINFO 2015: 1157-1169
Brazilian Society of Cartography, Geodesy, Photgrammetry and Remote Sense
ISSN: 1808-0936

SMALL AREA HOUSING DEFICIT ESTIMATION: A SPATIAL MICROSIMULATION APPROACH

Estimativa do Deficit Habitacional em Pequenas Áreas: Uma Abordagem Baseada em Microssimulação Espacial

Flávia da Fonseca Feitosa[1], Thiago Corrêa Jacovine[1] & Roberta Guerra Rosemback[2]

[1]Universidade Federal do ABC – UFABC
Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas – CECS
Rua Arcturus, 3, Jardim Antares. CEP: 09606-070. São Bernardo do Campo, SP - Brasil
{flavia.feitosa, jacovine.thiago}@ufabc.edu.br, @ufabc.edu.br

[2]Universidade Federal de Minas Gerais – UFMG
Centro de Desenvolvimento e Planejamento Regional – CEDEPLAR
Av. Antônio Carlos, 6627. CEP: 31270-901. Belo Horizonte, MG - Brasil
rosemback@cedeplar.ufmg.br

Received on October 14, 2015/ Accepted on March 23, 2016
Recebido em 14 de Outubro, 2015/ Aceito em 23 de Março, 2016

ABSTRACT

This paper presents a new methodology for measuring housing deficit at small areas. It combines the advantages of two types of census data: (a) individual-level sample data, which are very useful for depicting many dimensions of the housing deficit, but do not present detailed geographic information; and (b) universal data with detailed spatial resolution (census tracts), but aggregated. For that, we explore an approach based on spatial microsimulation. We simulate spatial microdata by using aggregate data as constraints to expand and allocate individual-level data to census tracts. This procedure allowed us to estimate a particular dimension of the housing deficit (housing cost) at higher spatial resolution.

Keywords: Housing Deficit Estimation, Census Data, Spatial Microsimulation.

RESUMO

Este artigo apresenta uma nova metodologia para a mensuração do déficit habitacional em pequenas áreas. Ela combina as vantagens proporcionadas por dois tipos de dados censitários: (a) microdados da amostra, que são úteis para capturar muitas dimensões do déficit habitacional, mas não apresentam informação geográfica detalhada; e (b) dados do universo com resolução espacial detalhada (setores censitários), mas agregados. Para tanto, foi explorada uma abordagem baseada em microssimulação espacial. Microdados espaciais foram simulados a partir da utilização dos dados censitários agregados como restrições para expandir e alocar microdados em setores censitários. Este procedimento permitiu a estimativa de uma dimensão particular do déficit habitacional (custo da moradia) em resolução espacial mais detalhada.

Palavras chaves: Estimativa do Déficit Habitacional, Dados Censitários, Microssimulação Espacial.

1. INTRODUCTION

The lack of decent housing for families with limited means is a major social problem in Brazil and other developing countries. To support appropriate housing policies and programs, it is essential to characterize and quantify the multiple dimensions of housing deficit. This is particularly important in the Brazilian context where, according to the National Housing Policy (PNH - Política Nacional de Habitação), federal funding towards municipal social housing projects will only be allocated to municipalities with a Local Plan of Social Interest Housing (PHLIS - Plano Local de Habitação de Interesse Social), which includes a local diagnosis of the current housing situation.

The PNH assigns to the municipalities a role that was previously attributed to the federal and state levels. This leading role of municipalities in the development of housing diagnosis raises new methodological demands, which include a refinement of scale and a thorough discussion on the potential of available data sources (ROSEMBACK *et al.* 2014).

Building a methodology for estimating housing deficit requires the definition of which data sources will be used and the identification of their limitations and potentialities. The Census is the most comprehensive statistical survey carried out in Brazil. It collects data on the composition and characteristics of population, households, dwellings and their surroundings. It is available to all municipalities and is therefore a unique data source for understanding the Brazilian housing conditions.

The census survey relies on two types of questionnaires: a sample one, with 108 items, which is applied to a fraction of the surveyed households (about 11% of the population), and a simplified one (37 items), applied to the whole population. As a result, the Census provides different types of data for public use, two of which are of particular interest for this work: (a) individual-level sample data (microdata), which are very useful for depicting many dimensions of the housing deficit, but are not universal and, for confidentiality reasons, do not present detailed geographic information; and (b) universal data with detailed spatial resolution, but aggregate by small areal units known as census tracts.

Based on these two types of census data (individual and aggregated), two different approaches have been adopted to estimate the Brazilian housing deficit: household-based and place-based approach. On the one hand, household-based studies use individual-level census data and have the advantage of considering more variables and, therefore, providing a better representation of the multiple dimensions of housing deficit. On the other hand, place-based studies use universal and aggregated census data, which provides a detailed spatial resolution and facilitates the integration with spatial data from other sources.

Considering the fact that each of these types of studies has advantages and shortcomings, this work represents an effort to develop a new hybrid approach for measuring housing deficit at small areas. For that, spatial microsimulation is used to combine the advantages of the two types of census data (sample microdata and universal aggregate data). Spatial microsimulation, in this work, is understood as the process of generating spatial microdata by taking data at the individual level and using aggregated level constraints to allocate these individuals to small areas (LOVELACE, 2014).

The possibilities opened by spatial microsimulation are wide. By allocating individual data in small areas it becomes possible, for instance, to combine this microdata with other spatial data that are relevant to characterize the housing deficit, such as information about land use and land cover (including the location of precarious and irregular settlements) or areas subject to landslides and floods.

In the next Section, we present seven conceptual dimensions of the housing deficit and additional information about the place-based and household-based approaches. Along the description of these two approaches, we point out the potential and shortcomings of census data for capturing the multiple dimensions of the housing deficit. Afterwards, an overview of spatial microsimulation and its methods is introduced in Section 3. This Section includes the description of the spatial microsimulation method explored in this work, the "Iterative Proportional Fitting" (IPF).

To illustrate how the IPF method can

address limitations identified in the current housing deficit estimation approaches, Section 4 presents an experiment with census data from a region of the city of São Bernardo do Campo, Brazil. In this experiment, we estimate one particular dimension of the housing deficit ("housing cost") at higher spatial resolution by using universal aggregate data as constraints to expand and allocate individual-level sample data to small zones (census tracts). The dimension "housing cost" was chosen due to its relevance in the Brazilian context.

2. THE HOUSING DEFICIT: DIMENSIONS AND MEASUREMENT APPROACHES

The development of a diagnostic that is suitable for supporting social housing policies demands a multidimensional view of the housing issue. According to Rosemback *et al.* (2014), it is possible to recognize at least seven (7) dimensions of adequacy that must be considered in the assessment of housing needs. These dimensions are summarized in Table 1.

Building a methodology that is able to measure these multiple dimensions of the housing deficit remains as an important challenge to be faced. In Brazil, it is possible to identify two different measurement approaches to address this issue. The first is a place-based approach that relies on the identification of so-called "squatter settlements", i.e., inadequate human settlements occupied by low-income residents. The second approach, household-based, adopts the concept of "housing needs" and addresses more explicitly some of the dimensions shown in Table 1.

In the place-based approach, the housing deficit is estimated by counting the families living in areas demarcated as squatter settlements. It demands, therefore, data with detailed spatial information. These diagnostics often use local data provided by the municipalities and/or census data. In the latter case, the census data used is obtained from the simplified questionnaire and aggregated by census tracts, which is the smallest spatial unit of analysis available for public use.

The Brazilian Institute of Geography and Statistics (IBGE) identifies a set of census tracts that corresponds to cluttered and dense poor settlements, most lacking basic infrastructure and services. These special census tracts have been called "subnormal agglomerates" since the 2010 survey and are often used to place-based estimations of the housing deficit. According to the 2010 Census, 323 municipalities registered 6,329 subnormal agglomerates, which corresponds to about 6% of Brazil's population (11,425,644 people) and 5.6% of its permanent dwellings (3,224,529 dwellings). The Southeast Region, which is the richest in the country, concentrates almost half (49.8%) of these subnormal agglomerates.

Table 1: Housing needs: dimensions of housing adequacy (ROSEMBACK *et al.*, 2014).

Dimension	Description
1. Housing Cost	The household spending on housing should not severely compromise the total household income.
2. Physical Suitability of the Dwelling Unit	Dwellings should be made of materials that permanently ensure weather protection, the health, privacy, and security of their residents.
3. Dwelling Unit Suitability to the Household	The household density in a building should not be excessively high. Families should not cohabit for lack of choice.
4. Environmental Safety	Dwellings should not be located in areas of environmental risks, including risks of flooding or landslides, contaminated areas, etc.
5. Legal Security	Households must have legal security of tenure.
6. Infrastructure and Public Services	Dwellings should be served by sewage, water supply, electricity network, street lighting, paving, trees, curb, sidewalk, etc.
7. Location and Accessibility	The location of dwellings should promote the integration into the city, including appropriate access to employment options, efficient public transportation, health services, school, culture and leisure;

Despite its importance, many authors have been pointing out the underestimation of the housing deficit obtained from the IBGE's subnormal agglomerates (TASCHNER, 2008; CARDOSO *et al.*, 2009; DENALDI, 2013). To address this issue, the Center for Metropolitan Studies (CEM) used statistical and GIS techniques to identify the census tracts whose characteristics are similar to subnormal agglomerates, but which are not identified by IBGE as such. The census tracts identified by this method are called "precarious tracts". This study advanced towards an enhanced detection of squatter settlements and was conducted for

the municipalities of selected metropolitan regions. In the case of São Bernardo do Campo, the housing deficit estimation for 2010 increased from 43,052 to 45,846 households living in subnormal or precarious tracts.

These place-based estimations for the housing deficit (IBGE and CEM) rely on universal data aggregated by census tracts. Due to its high spatial resolution, this data can be more easily combined with auxiliary data on dimensions that are not covered by the census survey, such as risk areas maps ('Environmental Safety') and municipal data about the legal status of land tenure ('Legal Security'). The detailed spatial resolution of these units (census tracts) allows analysis at the intra-urban scale, which is particularly important for housing policies.

Nevertheless, the variables obtained from the simplified census questionnaire (universal) can explicitly depict only one of the seven housing adequacy dimensions presented in Table 1 - 'Infrastructure and Public Services'. Another disadvantage of this data comes from the fact that it is aggregated. In this case, it is possible to know how many dwellings without connection to water supply networks or without garbage collection service can be found within the area of a certain census tract. However, it is not possible to know how many dwellings are not served by both services. Thus, the identification of squatter settlements and the use of aggregated data prove to be insufficient for the characterization and quantification of the housing deficit.

The household-based approach demands individual-level data. The João Pinheiro Foundation (FJP) adopted this approach and developed a method for estimating the Brazilian housing deficit that is considered a reference among social housing experts. This method relies on sample data from the 2010 Census and quantifies the housing deficit of all municipalities in the country (FJP, 2013). In fact, the data obtained from the complete census questionnaire represents a valuable source of information for measuring five of the seven dimensions presented in Table 1. Only the dimensions 'Environmental Safety' and 'Legal Security' cannot be represented by any of the variables provided by this dataset.

Available as microdata, this data allows us to consider the household as an analytical category and also provides a richer set of information about their living conditions. It includes many variables that are not available in the aggregated data and admits multiple combinations among them. The dimension 'Housing Cost', for instance, demands a combination between information on the household spending on housing and the total household income, which can only be obtained from individual-level data.

However, the census microdata does not provide detailed spatial information. The housing deficit estimates calculated from this dataset are only made available for the municipality as a whole or, at best, for weighting areas (áreas de ponderação, in Portuguese). Weighting areas are large geographical partitions used in the census sampling weighting procedures and formed by a mutually exclusively set of contiguous census tracts, with no less than 400 dwellings sampled (IBGE, 2012). To incorporate the two dimensions that demand auxiliary data into these household-based estimates ('Environmental Safety' and 'Legal Security'), it is essential to improve the spatial resolution of the results. In other words, to be able to integrate data from different sources, it is important to obtain more detailed information on the location of the sampled households.

In Table 2 we present a comparison between both datasets and approaches used to characterize and quantify housing needs. The list of advantages and disadvantages of the place-based and household-based approaches reveals their complementarity and the importance of developing a hybrid method, able to capture the best of each approach and dataset. This hybrid method should also be able to provide an estimation of the housing deficit assigned to areas inside and outside of squatter settlements. These separate assessments are particularly important for the development of housing policies and programs, as they subsidize different types of actions and investments.

To address this challenge, a recent attempt has been made in the "Diagnóstico Habitacional da Região do Grande ABC: informações municipais" (DENALDI *et al.*, 2016), which proposes a method that conciliates household- and place-based approaches to calculate the housing deficit of the Great ABC Region (São Paulo, Brazil).

Table 2: Place-based and household-based approaches with census data: advantages and disadvantages

PLACE-BASED APPROACH	HOUSEHOLD-BASED APPROACH
Uses aggregated data. It is not possible to capture details or cumulative inadequacies presented by a household. **(–)**	Uses microdata. This allows combinations between two or more variables (i.e., renters with excessive housing cost burdens) and the identification of cumulative inadequacies of a household. **(+)**
Reduced number of variables (37 items – basic questionnaire). Therefore, it is not possible to capture most of housing inadequacies dimensions, including "Housing Cost", "Physical Suitability of the Dwelling Unit", "Dwelling Unit Suitability to the Household", etc. (-)	Larger number of variables (108 items – sample questionnaire). It is possible to represent five out of seven housing inadequacies dimensions considered in this work. **(+)**
Does not capture inadequacies that exist outside of squatter settlements. (-)	Captures inadequacies that exist outside of squatter settlements **(+)**, but does not discriminate which inadequacies occur inside or outside of these settlements. (-)
More detailed spatial information (census tract). **(+)**	Less detailed spatial information (weighting areas – in municipalities with more than 190.000 hab.). **(–)**
Easy integration with other spatial data, such as satellite images, risk maps, land use information, etc. **(+)**	Difficult to integrate with other data sources. **(–)**
(+) Advantages	**(-) Disadvantages**

First, this study estimates the housing deficit of the region, separately, according to household-based and place-based strategies. Afterwards, based on the Set Theory, it considers that these two estimations present regions of intersection and difference (Figure 1). The region of intersection (Region 1 in Figure 1) covers households with housing inadequacies that are captured by the household-based approach (blue rectangle in Figure 1) and the place-based approach (yellow rectangle in Figure 1).

Two regions of difference are identified. The first (Region 2 in Figure 1) represents households with housing inadequacies that are only captured by the place-based approach. The second region of difference (Region 3 in Figure 1) is of particular importance to the study as it represents households with housing inadequacies that are not located in squatter settlements. This deficit is only captured by the household-based approach.

To estimate the deficit assigned to areas outside squatter settlements, the study computed a factor that represents housing inadequacies often observed in these areas (i.e., excessive housing cost and excessive density in a dwelling) and multiplied it by the total number of households living outside squatter settlements. Once the housing deficit is estimated for Region 3 (Figure 1), Set Theory rules are applied to estimate the deficit in Regions 1 and 2.

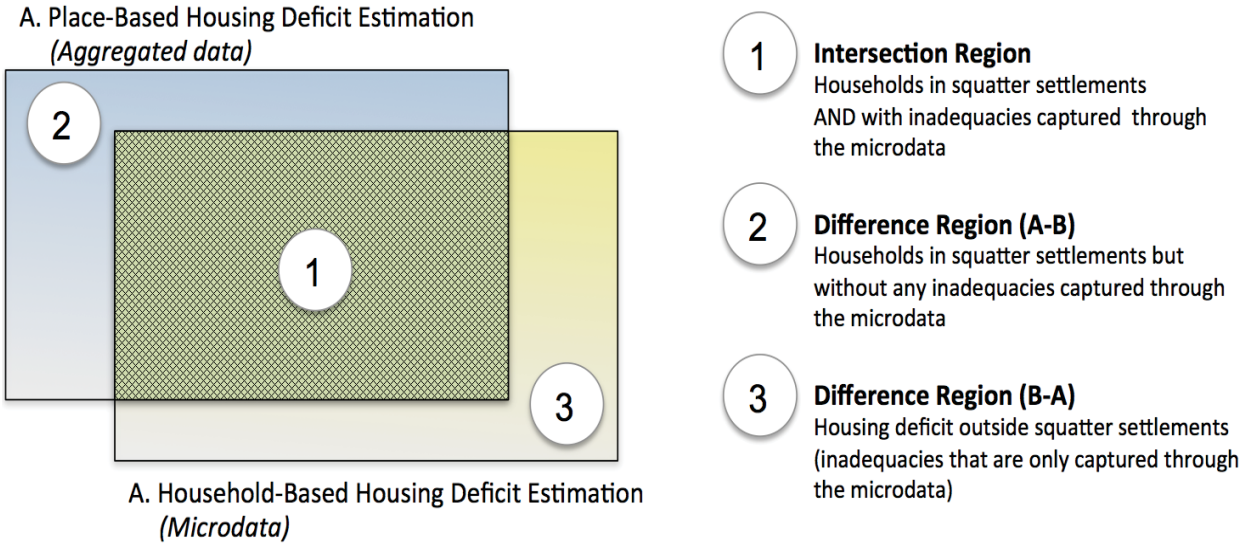

Fig. 1 – Compatibilization of place-based and household-based approaches for estimating the housing deficit.

This methodological effort led to an improvement in the measurement and characterization of the housing deficit in Brazilian cities. However, the method adopted for estimating the deficit in areas outside squatter settlements faces the difficulties of defining an appropriate multiplicative factor for each reality and, therefore, presents a relatively high degree of imprecision.

In addition, the method computes estimates for the whole study area, without the possibility of presenting a detailed spatial resolution. It is not possible, for instance, to estimate the deficit of a particular neighborhood.

Another important limitation is that the census universal data (used in the place-based approach) presents a limited number of variables, which makes difficult a multidimensional characterization of the housing deficit. As stated in Table 2, while the sample survey covers 108 questions, the universal survey presents only 37. Therefore, despite the merit of reconciling both approaches (place-based and household-based), the method does not improve the number of variables that characterizes the squatter settlements.

These limitations could be overcome if detailed spatial information related to the survey microdata were available, which does not happen due to confidentiality reasons. To address this challenge, this work proposes a hybrid approach based on the generation of spatial microdata through techniques of spatial microsimulation. In the next Section, an overview of spatial microsimulation methods is presented.

3.SPATIAL MICROSIMULATION: AN OVERVIEW

The use of methods and techniques of spatial microsimulation is not new in the social sciences. Throughout the last fifty years, spatial microsimulation has evolved and been applied to problems from a wide range of domains, including demography, health, regional development, land use planning, etc. (O'DONOGHUE, MORRISEY, LENNON, 2014).

Defined as an approach to "the creation, analysis and modeling of individual level data allocated to geographic zones" (LOVELACE, 2014), spatial microsimulation is part of the microsimulation field. Its origins, according to different authors (HERMES & POULSEN, 2012; O'DONOGHUE, MORRISSEY & LENON, 2014; TANTON & EDWARDS, 2013), date back to the 1950's-1960's with the pioneering works of Hägerstrand (1952; 1957; 1967) and Orcutt (1957). Nevertheless, this approach has gained greater prominence in the last three decades.

Characterized by the inclusion of the spatial dimension as a crucial element for the data simulation process, the idea behind the spatial microsimulation relies on: i) estimating how representative microdata observations (e.g., households) are for each small area or place of interest, and, ii) allocating them to these areas. The result of this modeling is called spatial microdata. Figure 2 illustrates this process considering the two different types of Brazilian Census data (survey microdata and universal data aggregated by census tracts).

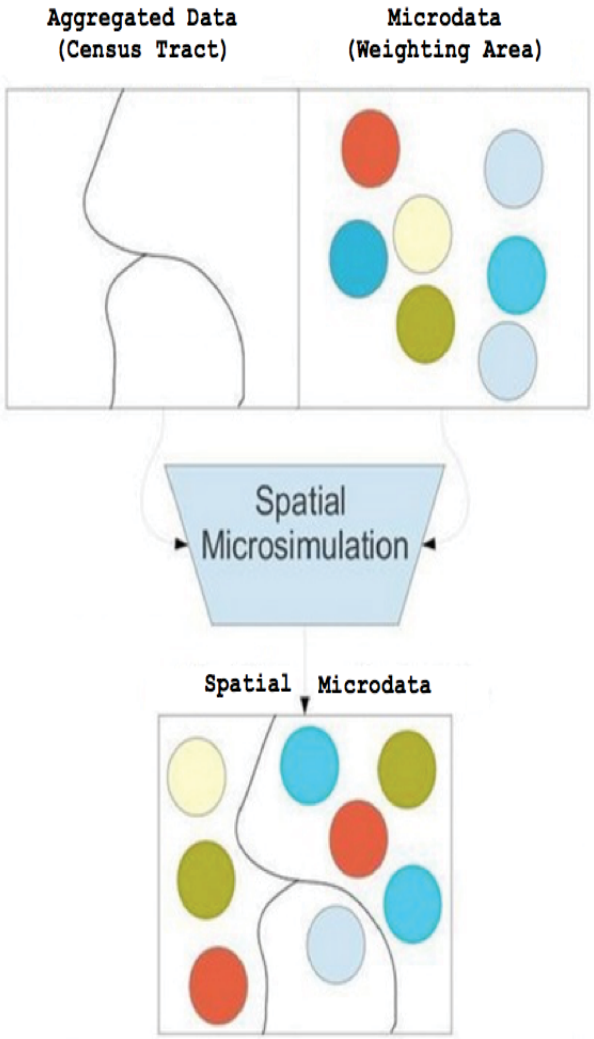

Fig. 2 - Spatial microsimulation process. Adapted from Lovelace (2014).

Spatial microsimulation has been applied to many different problems, including those related to health issues (CLARKE *et al.*, 1984), water demand (CLARKE *et al.*, 1997), socioeconomic analysis (CALDWELL *et al.*, 1998), demographic projections (VAN IMHOFF & POST, 1998) and estimates of dwelling characteristics (WILLIAMSON *et al.* 1998; BALLAS & CLARKE, 1999; BALLAS *et al.*, 1999). To handle the issues that these different types of applications demanded, several spatial microsimulation techniques were developed.

According to the conceptual framework proposed by Tanton (2014), these techniques can be classified into two major groups: dynamic and static (Figure 3). The main difference between them is the ability to model an event over the time. Using probability rules, the dynamic methods predict certain life events (deaths, births, etc.) and produce estimates with these predictions, while static methods are not designed to provide this type of estimates.

Because the purpose of this work is to develop a method for housing deficit estimation that uses data from a specific moment in time, we considered static techniques of spatial simulation as the most appropriate for the case.

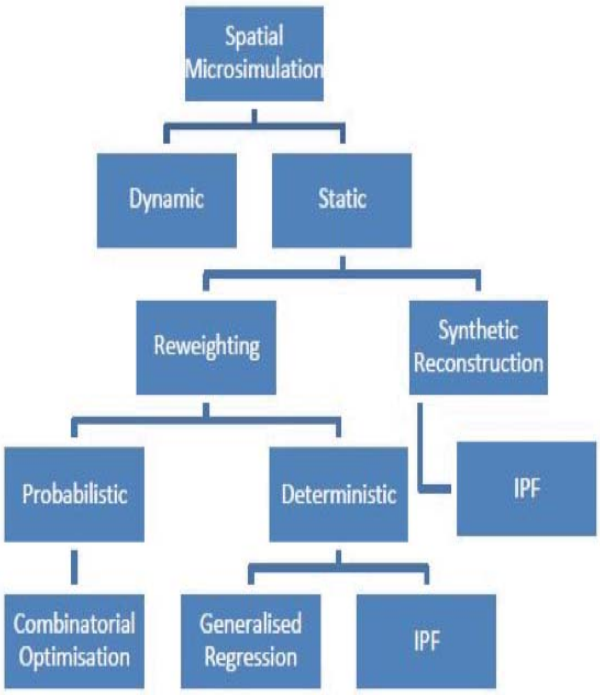

Fig. 3 - Conceptual framework for spatial microsimulation techniques. Source: Tanton (2014)

Static methods of spatial microsimulation can be classified in two different groups: reweighting and synthetic reconstruction. With similar structure and principles, the methods belonging to both groups involve individual-level data and data aggregated by small areas (e.g., census tracts). The main difference between them is related to the individual-level data (microdata). On the one hand, synthetic reconstruction methods artificially generate individual-level observations according to known distributions of characteristics from aggregated data (HERMES & POULSEN, 2012). On the other hand, reweighting methods demands an existing survey microdata.

In addition to the existence of microdata and data aggregated by small areas, reweighting methods require at least one common variable between the two datasets, whose function is to connect them. Depending on the method, these common variables may be known as linking variables or constraint variables.

In this work, we opted for the use of reweighting methods for two reasons: i) the existence of public and accessible census microdata, and ii) comparative studies demonstrating that reweighting methods are the most efficient (HERMES & POULSEN, 2012).

Considering the available reweighting methods, this work explores the most commonly used, the "Iterative Proportional Fitting" (IPF) (HERMES & POULSEN, 2012), The IPF method determines how representative an observation (e.g., household) is for each small area. For that, the IPF algorithm uses weights that are calculated by the following formula:

$$Wn = Wi * Agreg_{var} / Micro_{var} \qquad (1)$$

where,
Wn = New weight
Wi = Original weight (microdata)
$Agreg_{var}$ = Observation's value of the constraint variable (gender, i.e.) in the small area benchmark (aggregated data)
$Micro_{var}$ = Observation's value of the microdata for the same variable used as a constraint (gender, i.e.)

During this process, the weights are iteratively calculated and adjusted for each individual observation (individual-level data) in every census tract (aggregated data) until the known marginal distribution of the census tract population is

matched (HERMES & POULSEN, 2012).

The output is a "spatial microdata", which is a dataset that contains a single row per individual and also an additional variable that indicates the small area (i.e., census tract) where the individual was allocated (Figure 4).

In the example presented in Figure 4, the results obtained from the IPF method show us that the individual 1 from the sample microdata received a weight equal to '3' for census tract 1. This means that the individual 1 represents three individuals of the census tract 1 (that is why individual 1 appears 3 times at the spatial microdata table). Meanwhile, a weight equal to zero was assigned to the individual 3 because its characteristics are untypical for this small area. Thus, the individual 3 is not allocated to this census tract.

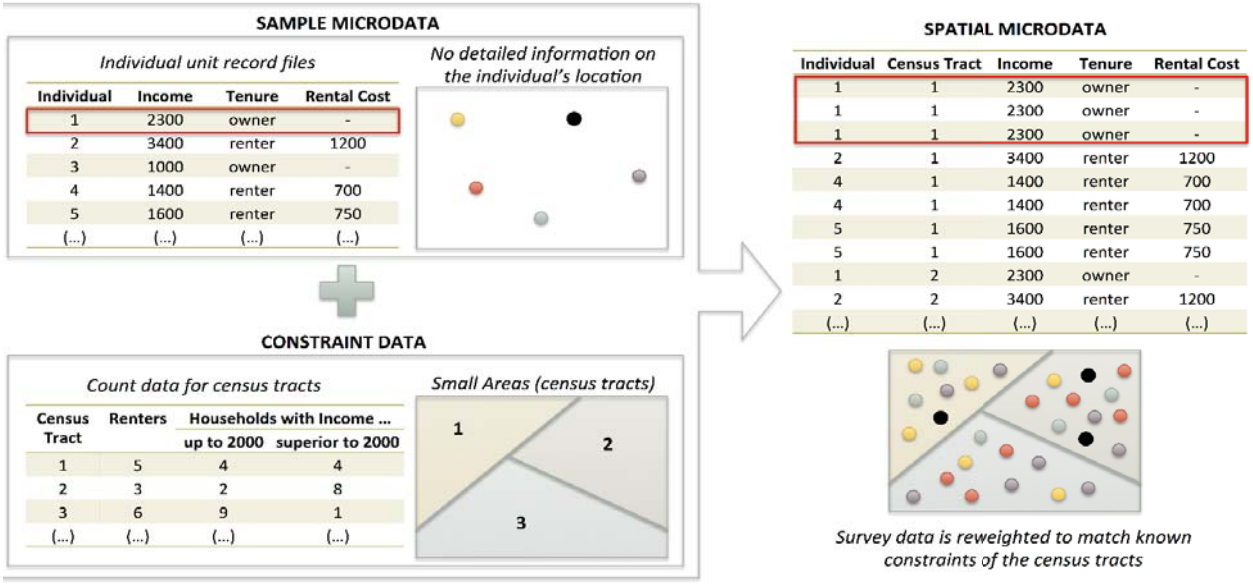

Fig. 4 - Creating spatial microdata using the IPF method. Adapted from Hermes & Poulsen (2012).

4.SPATIAL MICROSIMULATION FOR HOUSING DEFICIT ESTIMATION: AN EXPERIMENT

The lack of individual-level data at fine spatial resolution remains as a major obstacle to capture the multiple dimension of the housing problem in an appropriate manner. Since detailed spatial information about individual-level data cannot be provided due to confidentiality reasons, we advocate that spatial microsimulation represents a valuable resource to address the problem of estimating the housing deficit. To demonstrate that, the spatial microsimulation method known as IPF is used to estimate one of the housing adequacy dimensions at a small area: the "housing cost".

The experiment was conducted for region of São Bernardo do Campo that corresponds to a weighting area. The main contribution of the proposed method is to combine the place-based and household-based approaches, benefiting from the advantages provided by the two types of census data presented in this paper – the sample microdata (more variables and individual-level, but lower spatial resolution) and universal aggregated data (higher spatial resolution and easier integration with auxiliary data sources, but less variables and aggregated). A simplified version of this experiment was presented in Feitosa *et al.* (2015).

Situated at the Metropolitan Region of São Paulo, São Bernardo do Campo is the largest city of the ABC's Region in area (409,532 km2) and population (765,463 inhabitants, according to 2010 census, and 816,925, estimated to 2015). With a high GDP per capita (R$ 59,149.8) – higher than the GDP of the city of São Paulo, R$ 48,275.45 –, the city presented a continuous improvement in the Municipal Human Development Index (IDHM) during the last three decades. In a scale from 0 to 1, São Bernardo do Campo's IDHM was 0.642 in 1991, and 0.805 in 2010.

Despite these positive indicators, São Bernardo do Campo presents high levels of social

segregation, with an acute spatial isolation of low-income households (LISBOA & FEITOSA, 2016), and the housing deficit in the city remains as an important challenge to be faced. According to the census data on subnormal agglomerates, the city presented 43,072 households in census tracts classified as squatter settlements in 2010 (Figure 5b). This number represented 18% from the total of 239,337 households in the city (IBGE, 2011). It is relevant to emphasize, however, that this housing deficit estimate is restricted to the squatter settlements (place-based approach) and is very limited to represent the multiple dimensions of the issue. It is not able to depict, for instance, important dimensions such as the housing cost and the dwelling unit suitability to the household.

The household-based approach, which uses the census microdata, provides a much better picture of the multiple dimensions of the housing deficit in the city as a whole. Nevertheless, it is not capable of estimating the deficit in a detailed spatial resolution, as the only spatial information available for this dataset is the weighting area where each household is located (Figure 5a).

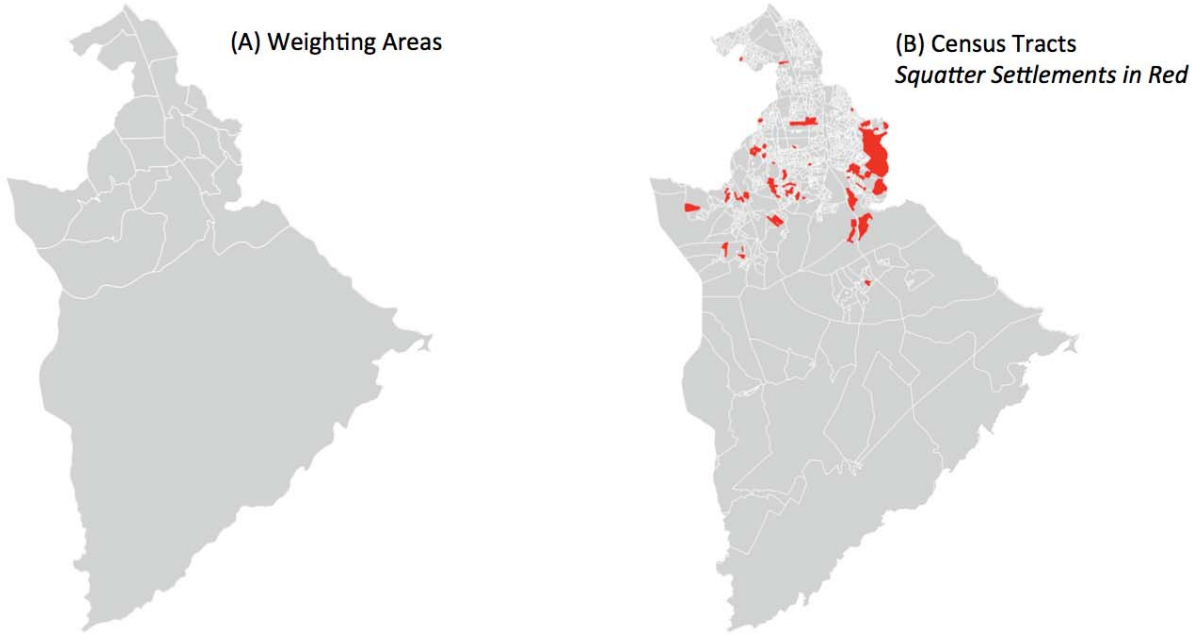

Fig. 5 – Areal Units: (A) Weighting areas and (B) Census tracts (squatter settlements in red).

According to the household-based estimation presented in Denaldi *et al.* (2016), the housing deficit in São Bernardo do Campo is equivalent to 90,881 households (37% of the total).

Among the different housing deficit dimensions captured by this estimation, the "housing cost" deserves particular attention. According to the study, 11,991 households in São Bernardo do Campo (4.9% of the total) are poor and commit more than 30% of their income to rental costs (DENALDI *et al.*, 2016).

Given the relevance of this housing inadequacy, we explore the spatial microsimulation to estimate the amount of households with excessive housing cost at the census tract scale. This choice, which can be replicated to other housing inadequacy dimensions, is also justified by the impossibility of estimating this inadequacy with aggregated data, as it demands individual-level data.

Considering the variables available in the sample census microdata (2010), the inadequacy "housing cost" can be represented as the "number of households with total income up to 3 minimum wages that spend more than 30% of their income with rental costs". Considering the total of low-income households with excessive housing costs in São Bernardo do Campo (11,991), 8.8% (1,055) are located in the weighting area that was chosen for the experiment (Figure 6).

The selected weighting area is located in a central area of the city and contains 61 census tracts. Since there is no information about the housing cost at this spatial level, the aim of the experiment is to estimate this housing inadequacy for each census tract using the IPF

spatial microsimulation method.

The IPF method demands the definition of constraint variables and target variables. The variables “income per capita” (8 intervals of income) and “tenure status” (6 classes) were selected as constraint variables, since they are available in both datasets (individual-level and aggregate). The constraint variables link both databases and allow the expansion and allocation process of individual observations (individual-level sample data) into the aggregated data (census tract). As target variables, we selected the variables of interest, which are only available in the sample microdata by weighting areas: "total household income" and "rental cost". The combination of these two variables allows us to identify the low-income households with excessive housing costs (housing cost dimension).

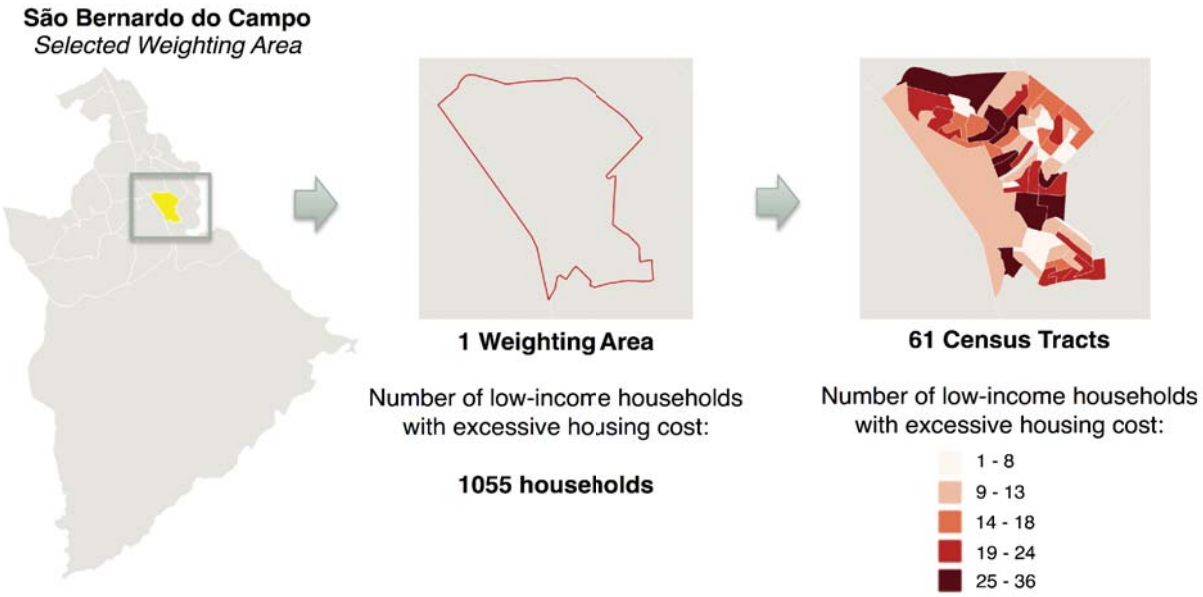

Fig. 6 - Experiment results: Spatial distribution of low-income households with excessive rental cost, based on simulated data.

The spatial microssimulation procedure was done at the statistical and modeling software R (package 'ipfm' - LOVELACE, 2014). The results are summarized in Figure 6. The simulation expands and allocates the original data in census tracts and allows a much more detailed spatial distribution of the households with excessive rental costs.

Figure 6a (before simulation) presents the total number of low-income households with excessive housing costs in the selected weighting area (1,055). It is not possible, with the original data, to observe how these households are spatially distributed inside this area.

The spatial microsimulation process generates spatial microdata and allows estimating the deficit for each census tract (Figure 6b - after simulation). In the selected weighting area, all the 61 census tracts presented at least one household with excessive rental costs. It is possible to observe that this housing inadequacy presents no clear pattern of clustering.

Due to its higher spatial resolution, the simulated microdata can be more easily integrated with other spatial data. This fact provides new possibilities for overcoming a limitation faced by the methods for housing deficit estimation, which is the difficulty to differentiate the housing needs inside and outside of squatter settlements.

Figure 7 illustrates how the simulated spatial data can be integrated with information about the limits of squatter settlements in order to provide information about the total number of households with excessive housing costs who live inside and outside of these settlements. The results indicate that this housing inadequacy is usually more intense outside the squatter settlements. Considering the 10 census tracts with the highest number of households living

under this inadequate condition, only one of them represents a precarious settlement.

From the total number of low-income households with excessive rental in this study area (1,055), 882 (88.6%) are not located in squatter settlements.

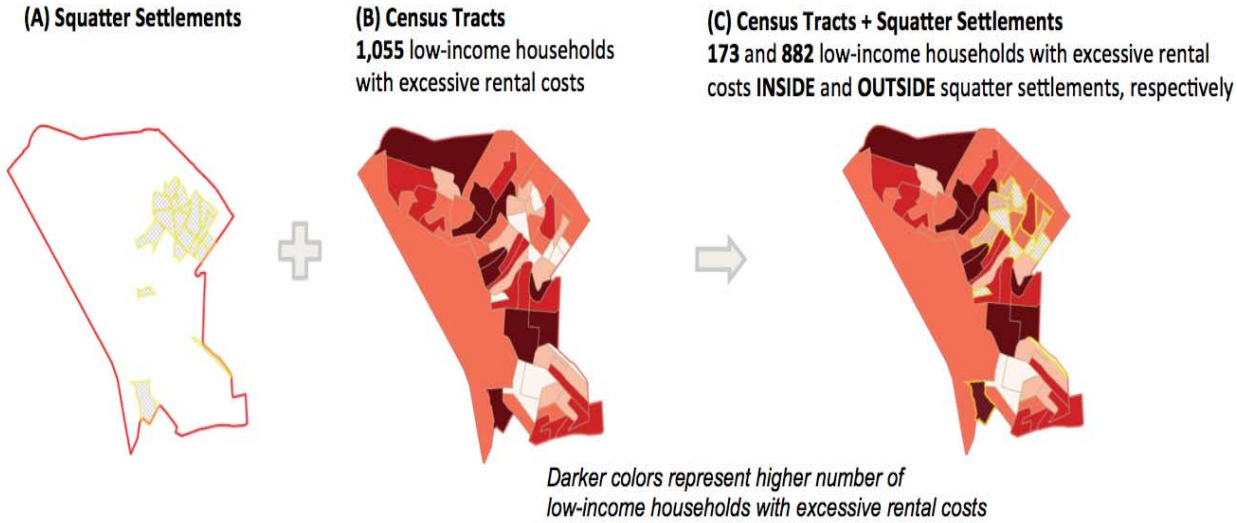

Fig. 7 – Experiment results: Spatial distribution of low-income households with excessive rental cost - Inside and Outside the Squatter Settlements.

5. CONCLUDING REMARKS

In this paper we advocate that spatial microsimulation techniques introduce new possibilities for the development of a method for measuring housing deficit at small areas. While the data aggregated by census tracts presents a fine spatial resolution, but lacks details on the households, the census sample microdata presents a richer dataset that is suitable for capturing different dimensions of the housing deficit, but lacks information on the spatial location of households.

By addressing the shortcomings of both data types, and therefore both place-based and household-based approaches for measuring housing deficit, we presented a hybrid approach that is able to better depict the housing deficit in Brazilian cities. Such approach should be able to explore not only the full capability of census data, which is available for all municipalities, but also allow the integration with auxiliary data that may be available at the local level, such as natural hazard/risk assessments and municipal data on legal status of land tenure.

The generation of new data at small in our case) through this new hybrid approach opens new possibilities for analysis and provides important insights for public policies. Through the conducted experiment we were able not only to estimate housing cost inadequacies in small areas, but also to measure this inadequacy inside and outside of squatter settlements.

Additional tests must be conducted to ensure that the simulated spatial microdata is as representative as possible of the aggregate constraints. For that, it is necessary to explore the choice of different constraint variables and validate the resulting estimates. In addition, it is important to test and compare different methods for spatial microsimulation by exploring its main features, variability and validity comparing with external datasets.

ACKNOWLEDGEMENTS

The authors would like to thank the financial support received from CNPq (Grant 443052/2014-0), FAPEMIG and the Intermunicipal Consortium of the Greater ABC Region.

REFERENCES

BALLAS, D.; CLARKE, G. Regional versus local multipliers of economic change? A microsimulation approach. In: **39th Regional Science Association (ERSA) Congress**,

University College, Dublin, p. 1-31, 1999.

BALLAS, D.; CLARKE, G.; TURTON, I. Exploring microsimulation methodologies for the estimation of household attributes. In: **4th International conference on GeoComputation**, Mary Washington College, Fredericksburg, VA, p. 1-46, 1999.

CALDWELL S.; CLARKE G.; KEISTER, A. Modelling regional changes in US household income and wealth: a research agenda, **Environment and Planning C: Government and Policy**, v. 16, n. 6, p. 707-722, 1998.

CARDOSO, A.L; ARAÚJO, R.L.; GUILHARDI, F.H. Necessidades habitacionais. In: BRASIL, Ministério das Cidades, **Curso à distância: Planos Locais de Habitação e Interesse Social**. Brasília: MCidades, 2009. 216p.

CENTRO DE ESTUDOS DA METRÓPOLE (CEM). **Diagnóstico dos Assentamentos Precários nos Municípios da Macrometrópole Paulista**. Primeiro Relatório. São Paulo, 2013. 55p.

CLARKE, M.; FORTE, P.; SPOWAGE, M.; WILSON, A. G. A strategic planning simulation model of a district health service system: The in-patient component and results. In: **Systems science in health care**. Berlin: Springer, p. 949-954, 1984.

CLARKE G.; KASHTI A.; MCDONALD A.; WILLIAMSON P. Estimating small area demand for water: a new methodology, **Journal of the Chartered Institution of Water and Environmental Management**, v. 11, p.186-192, 1997.

DENALDI, R (org.). **Planejamento Habitacional. Notas sobre a precariedade e terra nos Planos Locais de Habitação**. 1ª. São Paulo: Ed. Annablume, 2013. 308p.

DENALDI, R.; FEITOSA, F. F.; MENCIO, M.; FERRARA, L. N.; JESUS, P. M.; BRAJATO, D.; CARVALHO, G. C.; SILVEIRA, K. J.; PETRAROLLI, J. G.; SILVA, P. H.; ROSEMBACK, R. G.; JACOVINE, T. C.; GONÇALVES, G. S.; FALCÃO, K. S.; AGUIAR, R. R. **Diagnóstico Habitacional da Região do Grande ABC: informações municipais** (Relatório). São Bernardo do Campo, 2016. 251p.

FEITOSA, F. F.; ROSEMBACK, R. G.; JACOVINE, T. C. Small Area Housing Deficit Estimation: A Spatial Microsimulation Approach. In: XVI Geoinfo - Brazilian Symposium on Geoinformatics, Campos do Jordão, 2015. **Anais**. p. 131-136, São José dos Campos, MCTI/INPE, 2015.

FUNDAÇÃO JOÃO PINHEIRO (FJP). **Déficit habitacional municipal no Brasil 2010**. Belo Horizonte, Fundação João Pinheiro, Centro de Estatística e Informações, 2013. 80p.

HERMES, K.; POULSEN, M. A review of current methods to generate synthetic spatial microdata using reweighting and future directions. **Computers, Environment and Urban Systems**, v. 36, p. 281-290, 2012.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Censo Demográfico 2010: Aglomerados Subnormais – Primeiros Resultados**. Rio de Janeiro, IBGE, 2011. 259p.

_____. **Censo Demográfico 2010: Microdados da Amostra**. Rio de Janeiro, IBGE, 2012. 235p.

LOVELACE, R. **Introducing Spatial Microsimulation with R: A Practical**. National Centre for Research Methods Working Paper 08/14. Leeds, University of Leeds, 2014. 23p.

LISBOA, F.S.; FEITOSA, F.F. Para Além da Perspectiva Residencial: A Construção de Índices de Segregação dos Espaços de Atividades. **Revista Brasileira de Cartografia**, v. 68, n. 4, p. 797-813, 2016.

O'DONOGHUE, C.; MORRISSEY, K.; LENNON, J. Spatial Microsimulation Modelling: a Review of Applications and Methodological Choices. **International Journal of Microsimulation**, v. 7, n. 1, p. 26-75, 2014.

ROSEMBACK, R.; RIGOTTI, J.; FEITOSA, F.; MONTEIRO, A. As dimensões da questão habitacional e o papel dos dados censitários nos diagnósticos municipais: uma sugestão de análise frente às novas exigências da Política Nacional de Habitação. In: XIX Encontro Nacional de Estudos Populacionais, São Pedro, 2014. **Anais**. Disponível em <http://abep.info/anais/anais.php?id=53#.VsRZNJMrKRs>.

Acesso: 10 out. 2015.

TANTON, R.; EDWARDS, K.L. "**Spatial Microsimulation: a reference guide for users**". London, Springer, 2013. 277p.

TANTON, R. A Review of Spatial Microsimulation Methods. **International Journal of Microsimulation**, V. 7(1), p.4-25, 2014.

TASCHNER, S.P. O desafio da mensuração. **Política habitacional e a integração urbana de assentamentos precários. Parâmetros conceituais, técnicos e metodológicos**. Brasília, MCidades, 2008. 144p.

VAN IMHOFF; E., POST, W. Microsimulation methods for population projection. **Population**, v. 10, n. 1, p. 97-136, 1998.

WILLIAMSON P.; BIRKIN, M.; REES, P. The estimation of population microdata by using data from small area statistics and samples of anonymised records, **Environment and Planning A**, v. 30, n. 5, p. 785-816, 1998.