

## Methodology note

## An integrated approach to create a spatial database of parks for urban health research

Mario Fontán-Vela<sup>a,b,\*</sup>, Roberto Valiente<sup>a,c</sup>, Manuel Franco<sup>a,d</sup>, Pedro Gullón<sup>a,e</sup><sup>a</sup> Grupo de Investigación en Salud Pública y Epidemiología, Facultad de Medicina, Universidad de Alcalá, Alcalá de Henares, Madrid, España<sup>b</sup> Servicio de Medicina Preventiva, Hospital Universitario Infanta Leonor, Madrid, España<sup>c</sup> Departamento de Geología, Geografía y Medio Ambiente, Universidad de Alcalá, Alcalá de Henares, Madrid, España<sup>d</sup> Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA<sup>e</sup> School of Global, Urban and Social Studies, RMIT University, Melbourne, Australia

## ARTICLE INFO

## Article history:

Received 6 May 2021

Accepted 29 July 2021

Available online 6 October 2021

## Keywords:

Madrid

Parks

Urban health

GIS

Databases

## ABSTRACT

The role of urban parks has gained interest in urban health research. However, there are currently no clear definitions of these urban assets as exposure variable across the literature, which do have implications for the analysis. In this study, we exhibit an approach to design a comprehensive spatial database of urban parks in cities, using the example of Madrid. This approach consisted of the combination of different spatial databases containing information about parks and green spaces in the city of Madrid, and then a spatial processing to keep only those that met the selected criteria for urban park definition: accessible for pedestrians (to have at least one path around or through the area) and with the minimum size suitable (bigger than 0.50 ha) for physical activity. This approach can be helpful to design spatial databases that enable conducting epidemiological studies on the relationship between urban parks and population health.

© 2021 SESPAS. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Un enfoque integrado para crear una base de datos espacial de parques para la investigación en salud urbana

## RESUMEN

El papel de los parques ha cobrado interés en la investigación sobre salud urbana. Sin embargo, actualmente no existen definiciones claras de estos activos urbanos como variable de exposición, lo que tiene implicaciones para el análisis. En este estudio se muestra un enfoque para diseñar una base de datos espacial integral de parques en ciudades, utilizando el ejemplo de Madrid. Se combinaron diferentes bases de datos espaciales con información sobre parques y espacios verdes de la ciudad de Madrid, que luego se procesaron espacialmente para mantener aquellos que cumplían los criterios de definición de parque: accesible (disponer de al menos un camino alrededor o a través del área) y un tamaño mínimo adecuado para realizar actividad física (mayor de 0,5 ha). Este enfoque puede ser útil para diseñar bases de datos espaciales que permitan realizar estudios epidemiológicos sobre la relación entre parques urbanos y salud poblacional.

© 2021 SESPAS. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Palabras clave:

Madrid

Parques

Salud urbana

GIS

Bases de datos

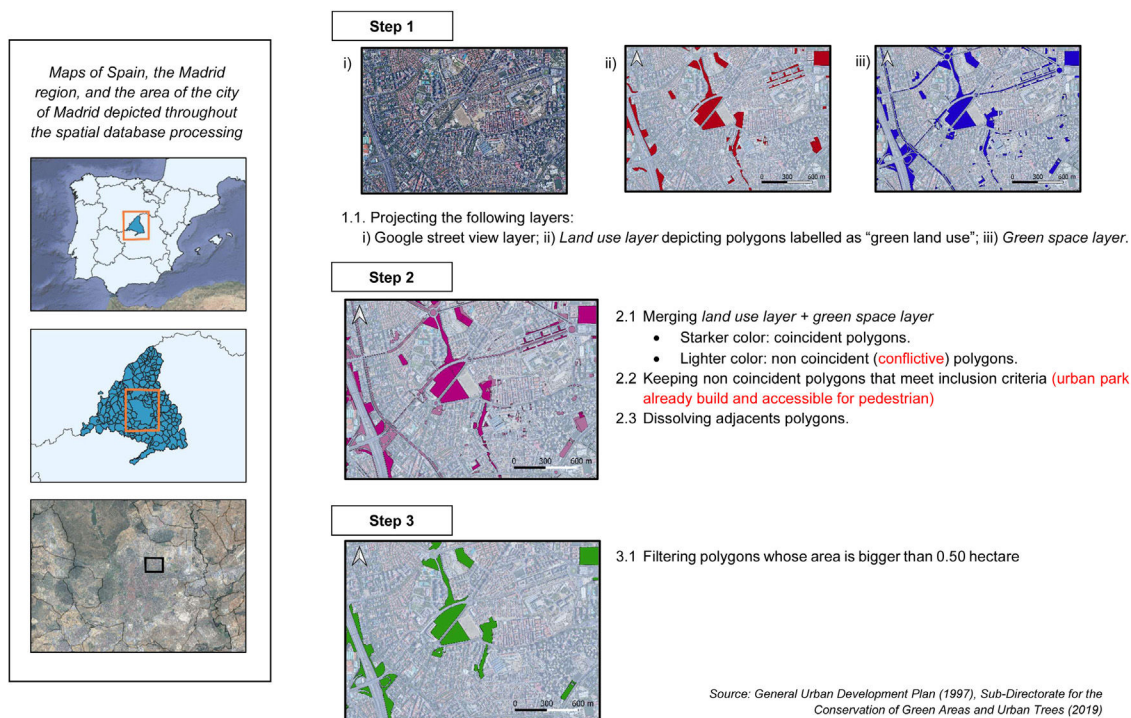
## Introduction

Green spaces have become an important field of study in urban health research focused on the relationship between built environment and health outcomes. Several studies have stressed the beneficial impact of green spaces on the population's health by reducing harm (i.e., exposure to air pollution), restoring capacities (i.e., psychophysiological stress recovery), and building capacities (i.e., encouraging physical activity and facilitating social cohesion).<sup>1,2</sup>

Different categories of green spaces have been used in urban research: “green land use” usually refers to land areas in urban planning allocated for green use, such as vegetation or parks, and “urban parks” usually refers to green spaces within cities with vegetation and accessible for pedestrians.<sup>3</sup> Urban parks have been studied as the main pathway between green spaces and physical activity within cities. However, a standard and operative definition of what constitutes an urban park is not yet established, although a set of characteristics have been proposed as those required to promote physical activity.<sup>3</sup> The methods conducted to estimate and define the extension of urban parks have widely differed among studies: while some used processed cartographic layers of information of green land-uses provided by public institutions, other used satellite images to create their own layers or to estimate the level of

\* Corresponding author.

E-mail address: [mario.fontan@edu.uah.es](mailto:mario.fontan@edu.uah.es) (M. Fontán-Vela).



**Figure 1.** Workflow for the spatial processing of databases selected for parks and urban health research analysis: the example of Madrid.

greenness by performing complex operations like the NDVI.<sup>4</sup> The former represent an easy-to-implement methodology with ready-to-use data, but with limited capacity to define what is considered as an urban park, while the use of satellite images allows a more flexible definition of urban park, but it implies a complex methodology using Geographic Information Systems (GIS). Different urban park definitions can yield different results in public health studies: while some found an association between urban parks and mental health using land-use layers, others found no association using obesity as the outcome variable.<sup>5,6</sup> Alongside this complexity, the need of relying on secondary data could difficult these analysis, since each database was created to respond to a specific purpose and it could gather different information.

This paper aims to present an approach to create a comprehensive spatial database of urban parks for public health research purposes using the example of Madrid.

## Methodology

### Study setting

The proposed approach is for the study of urban parks using the example of Madrid, the capital of Spain, with 3,334,730 inhabitants, according to the National Statistics Institute for the 2020 year, and an area of 604,45 km<sup>2</sup>.

### Database selection

Two specific databases offering data on green spaces and/or urban parks in geospatial vector format (shapefile) were gathered: 1) the land uses reported in the General Urban Development Plan of the Regional Government (GUDP) of the year 1997 —land use layer—,<sup>7</sup> and 2) the green spaces for the year 2019 whose maintenance is the responsibility of the City Council —green space layer—, provided by the Sub-Directorate for the Conservation of Green Areas and Urban Trees. Both were collected by public institutions. Despite the year of the GUDP is 1997, this is the main tool to guide

the urban development in the Spanish context, and some further updates until 2016 of the plan were included in the database.

### Spatial database processing

The land-use layer comprises information on the land use distribution, in which the extent of the city is divided into polygons with different land-use (e.g., residential, industrial, etc.). This layer was filtered to keep those polygons labelled as any "green land use". In contrast, the green space layer is a more specific database representing only those areas that constitute green spaces within the city. We used GIS to merge both layers into a new one —final layer— depicting all the areas classified as "green space" by the land-use or the green space layer. We identified some polygons in which both layers were overlapped. Both sources classify these polygons as "green space", and we automatically included them in the final layer. However, we found polygons that were not coincident (i.e., areas depicted as "green space" in the land use layer, but not in the green space layer, and vice versa).

To solve this conflict, two researchers conducted a two-step analysis. Firstly, one of the researchers assessed, using photo-interpretation with Google Street View imagery, whether each conflictive polygon complied the following criteria:

- As the GUDP might account as "green spaces" for some polygons that are not currently built, but it is expected to be converted into a green space in the future, the park should be already built and verifiable using Google Street View imagery.
- Urban parks should be accessible for pedestrians, considered as the presence of at least one walking path around the perimeter or crossing the area.

The researcher classified the conflictive polygons into three categories depending on whether the polygon should be included, excluded, or to review. As the database information was previously validated by public institutions, we considered that only a double-check was needed for the conflictive polygons. Thus, the

other researcher checked those polygons categorized as 'excluded' or 'to review', and classified them following the previous criteria. Discrepancies were discussed between the researchers, and the final decision was based on their agreement. Both databases were comprised of 8124 polygons, of which 545 (6.7%) were non-coincident, and 97 of these (17.8%) were finally included at the end of the process. This database had an area of 93.03 km<sup>2</sup>.

Then, adjacent polygons were dissolved to be considered as a unique urban park. Subsequently, a polygon size filter was applied to remove those polygons smaller than 0.5 ha, a threshold posed by the World Health Organization as the minimum size for a park to promote physical activity.<sup>3</sup> The final area of the layer was 70.1 km<sup>2</sup>, which supposes a 75.35% of the previous one.

At the end of the process, the final layer was composed of those polygons corresponding to urban parks of the city of Madrid, accessible for pedestrians and whose area is bigger than 0.5 ha. [Figure 1](#) illustrates the methodological design of this study.

All the spatial processing of databases was conducted using QGIS software (A Coruña 3.10 version).

### Practical use

The final spatial database of green spaces could be used for several purposes. Although we focus on the relationship between urban parks and health, environmental studies have stressed the importance of accessibility and availability of parks in terms of environmental justice.<sup>8</sup> Other researchers placed parks at the center of long-term urban dynamics such as gentrification and try to understand the potential side effects of these areas in terms of economic, social, and health impact.<sup>9</sup> Although both databases came from a different public institution, we considered that this approach enabled us to obtain a more comprehensive database with information related to urban parks that could be replicable in other settings where unified databases for parks are not available.

### Conclusions

The study of the impact of urban parks in different research areas has gained interest in the urban research field. A combination of different databases may be useful to gather comprehensive information on the location and characteristics of green spaces, thus overcoming the problems of definition and heterogeneity of parks databases, and allowing for a better understanding of the relationship of these urban assets with different outcomes in public health studies.

### Editor in charge

Javier García Amez.

### Transparency declaration

The corresponding author on behalf of the other authors guarantee the accuracy, transparency and honesty of the data and

information contained in the study, that no relevant information has been omitted and that all discrepancies between authors have been adequately resolved and described.

### Authorship contributions

M. Fontán-Vela, P. Gullón and M. Franco conceptualized and designed the study. M. Fontán-Vela and P. Gullón conducted the spatial database processing, and R. Valiente contributed with the software part of the spatial processing. M. Fontán-Vela wrote the first draft. P. Gullón, R. Valiente and M. Franco critically revised the paper for intellectual content. P. Gullón and M. Franco performed project supervision and administration, and managed funding acquisition.

### Funding

This project was funded by the Instituto de Salud Carlos III, Subdirección General de Evaluación y Fomento de la Investigación, Government of Spain (PI18/00782), and by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013/ERC Starting Grant Heart Healthy Hoods Agreement no. 623 336893).

### Conflicts of interest

The authors declare no conflict of interest.

### References

1. Markevych I, Schoierer J, Hartig T, et al. Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environ Res.* 2017;158:301–17. <http://dx.doi.org/10.1016/j.envres.2017.06.028>.
2. James P, Banay RF, Hart JE, et al. A review of the health benefits of greenness. *Curr Epidemiol Rep.* 2015;2:131–42. Available at: <http://link.springer.com/10.1007/s40471-015-0043-7>.
3. WHO Regional Office for Europe. Urban green spaces and health. Copenhagen: WHO; 2016.
4. Fong KC, Hart JE, James P. A review of epidemiologic studies on greenness and health: updated literature through 2017. *Curr Environ Health Rep.* 2018;5:77–87. Available at: <http://link.springer.com/10.1007/s40572-018-0179-y>.
5. Klompaker JO, Hoek G, Bloemsma LD, et al. Green space definition affects associations of green space with overweight and physical activity. *Environ Res.* 2018;160:531–40. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0013935117310691>.
6. Zhang L, Tan P. Associations between urban green spaces and health are dependent on the analytical scale and how urban green spaces are measured. *Int J Environ Res Public Health.* 2019;16:578. Available at: <http://www.mdpi.com/1660-4601/16/4/578>.
7. Dirección General de Planificación Estratégica. Ayuntamiento de Madrid PGOU 97. Plano de ordenación. 1997 [cited 2021 Apr 16]. Available at: <https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbc4b2e4b284f1a5a0/?vgnnextoid=98fca5b5cfbac510VgnVCM2000001f4a900aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
8. Wolch JR, Byrne J, Newell JP. Urban green space, public health, and environmental justice: the challenge of making cities 'just green enough'. *Landsc Urban Plan.* 2014;125:234–44. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0169204614000310>.
9. Cole HVS, Lamarca MG, Connolly JJT, et al. Are green cities healthy and equitable? Unpacking the relationship between health, green space and gentrification. *J Epidemiol Community Health.* 2017;71:1118–21.