Urban Forestry & Urban Greening 19 (2016) 88–94

Contents lists available at ScienceDirect

Urban Forestry & Urban Greening

journal homepage: www.elsevier.com/locate/ufug

Normalized difference vegetation index (NDVI) as a marker of surrounding greenness in epidemiological studies: The case of Barcelona city

Mireia Gascon[a,b,c,d,*], Marta Cirach[b,c,d], David Martínez[b,c,d], Payam Dadvand[b,c,d], Antònia Valentín[b,c,d], Antoni Plasència[a], Mark J. Nieuwenhuijsen[b,c,d]

[a] ISGlobal, Barcelona Ctr. Int. Health Res. (CRESIB), Hospital Clínic – Universitat de Barcelona, Barcelona, Spain
[b] ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
[c] Universitat Pompeu Fabra (UPF), Barcelona, Spain
[d] CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain

ARTICLE INFO

Article history:
Received 2 October 2015
Received in revised form 27 May 2016
Accepted 1 July 2016
Available online 12 July 2016

Keywords:
Built environment
Environmental epidemiology
Green spaces
NDVI
Urban

ABSTRACT

The normalized difference vegetation index (NDVI) is often used as a marker of surrounding greenness in epidemiological studies aiming to evaluate the health effects of green space in urban settings. However, it is not clear the relationship between built environment characteristics, including green space, and NDVI. We aimed to evaluate the relationship between built environment characteristics, based on land-use and land-cover maps, and NDVI as a marker of surrounding greenness in the city of Barcelona. We used data from an already existing cohort of pregnant women in Barcelona (N = 8402). NDVI was derived and averaged within buffers of 100 m and 300 m for each participant, and categories of the built environment (m^2) were derived from land-use and land-cover maps of Barcelona. We conducted ANOVA models to calculate the contribution (R^2) of each land-use (or land-cover) category. The variability in NDVI in Barcelona was mainly explained by urban green (R^2 between 0.32 and 0.53) and natural green areas (R^2 between 0.19 and 0.52), although for the latter less than 4% of the participants were exposed to this. Both land-use and land-cover maps explained NDVI at 300 m better (full models explaining 76% and 78%, respectively) than at 100 m buffers (full models explaining 55% and 54%, respectively). Results of the present study indicate that NDVI can be a useful greenness metric depending on the hypothesis and area of study. However, for certain sizes of study areas (buffers smaller than 100 m), NDVI might have a lower predictive value. Results of the present study should be replicated in studies from other cities with different urban characteristics and climate conditions.

1. Introduction

Improved mental health (Gascon et al., 2015), better pregnancy outcomes (Dadvand et al., 2012a; Grazuleviciene et al., 2015), as well as a reduction of the risk of mortality (Lachowycz and Jones, 2014), cardiovascular diseases (Pereira et al., 2012), asthma related symptoms (Dadvand et al., 2014a) or obesity (Dadvand et al., 2014a; Lachowycz and Jones, 2011) have been reported in relation to residential green spaces (James et al., 2015). So far, in such studies two main approaches have been used to define exposure to residential green spaces: (1) distance to minor and/or major green spaces as a surrogate of access to green spaces and (2) amount of surrounding greenness within a certain buffer as a surrogate of general greenness of home, workplace, or school. Surrounding greenness has been mainly characterized as the percentage of green spaces measured based on land-cover or land-use maps, or as the amount of photosynthetically-active greenness measured by the satellite-derived normalized difference vegetation index (NDVI) (James et al., 2015).

NDVI is the most common and easy obtainable vegetation index to detect live green plant canopies using multispectral remote sensing data based on spectral reflectance measurements acquired in the visible (red band) and near-infrared regions, respectively

* Corresponding author at: Parc de Recerca Biomèdica de Barcelona (PRBB) – ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Doctor Aiguader, 88, 08003 Barcelona, Catalonia, Spain.
E-mail addresses: mgascon@creal.cat (M. Gascon), mcirach@creal.cat (M. Cirach), dmartinez@creal.cat (D. Martínez), pdadvand@creal.cat (P. Dadvand), avalentin@creal.cat (A. Valentín), antoni.plasencia@isglobal.org (A. Plasència), mnieuwenhuijsen@creal.cat (M.J. Nieuwenhuijsen).

http://dx.doi.org/10.1016/j.ufug.2016.07.001
1618-8667/

(James et al., 2015). NDVI values range from −1 to 1; in general terms, very low values of NDVI (0.1 and below) correspond to barren areas of rock, sand, water or snow. Moderate values (0.2–0.3) represent shrubs and grassland, while high values (0.6–0.8) indicate temperate and tropical rainforests (Weier and Herring, 2015). However, at urban scale it is not very clear which urban characteristics are actually explaining the NDVI values obtained within a certain area. Moreover, NDVI is limited in capturing the high spatial heterogeneity that exists in an urban setting, and therefore, depending on the resolution, NDVI can underestimate urban green space variability and availability. However, existing land-use and land-cover maps in the city of Barcelona are based on orthophotos of a high resolution (pixel size of 0.5 m) (Burriel Moreno et al., 2006; CREAF, 2013) and, therefore, we can consider them as the gold-standard to determine the distribution of different types of urban uses or urban areas.

In epidemiological studies, in order to provide appropriate evidence of the health effects of a particular exposure, it is important to properly characterize this exposure and to minimize exposure misclassification (Burns et al., 2014). Many studies evaluating the health effects of living near green spaces have used NDVI as a proxy to define exposure to urban green spaces, but none of them has actually assessed the relationship between built environment characteristics, including green space, and NDVI. Therefore, in the current study we aimed to evaluate the relationship between built environment characteristics, based on land-use and land-cover maps, and NDVI as a marker of surrounding greenness in the city of Barcelona (Catalonia, Spain).

2. Materials and methods

2.1. The city of study

Various epidemiological studies have been conducted in the city of Barcelona in relation to green spaces and health outcomes (Dadvand et al., 2014a,b; Triguero-Mas et al., 2015; Dadvand et al., 2015); for this reason we conducted the current study in this city located in Catalonia, Spain. It has an area of approximately 100 km^2 and is limited by the coastline in the south-east and by the Collserola hills (the highest point, Tibidabo, is 516 m) in the north-west. Currently 1.6 million inhabitants live in Barcelona, which has a high population density of 16.056 inhabitants/km^2. Barcelona has a Mediterranean climate, with relatively humid and mild winters and warm and dry summers. The rainiest seasons are autumn and spring.

2.2. The study participants

We based the current study on data obtained from a cohort of pregnant women recruited during 2001–2005 from the obstetrics department of the Hospital Clinic of Barcelona, a major university hospital covering Barcelona city (Figueras et al., 2008). We limited the analyses to those participants residing in the city of Barcelona with known residential address at the time of delivery (N = 8402).

2.3. NDVI

To determine the surrounding greenness we used NDVI derived from the Landsat 8 OLI/TIRS data at 30 m × 30 m resolution. The Landsat 8 imagery data was acquired for April 16th 2007 covering Barcelona city area. Surrounding greenness was abstracted as the average of NDVI in buffers of 100 m and 300 m around each maternal place of residence which was geocoded according to the address at delivery time.

2.4. Land use and land cover maps

In order to define factors of the built environment that explain NDVI we used two types of land maps (land-use map and land-cover map). Although both of them are based on orthophotos of a resolution of a pixel size of 0.5 m (Burriel Moreno et al., 2006; CREAF, 2013), the categories created for each map provide different types of information regarding the characteristics of the area of study. Therefore, we considered that, in order to answer our research question, it was important to use both maps.

The first map that we used was the "Ecologic map of Barcelona", a land-use map prepared by the Centre for Ecological Research and Forestry Application (CREAF) (Burriel Moreno et al., 2006). This map displays the ecological areas that compose the urban green infrastructure (natural areas, semi-natural areas and built-up areas). For the current study we used the edition of 2004.The map is organized in a hierarchical legend of three levels. Level 3 (the most detailed classification) contains a total of 58 categories, level 2 groups the original categories into 31 and level 1 groups them into 10, which were the categories used to conduct the current study. These 10 categories are: natural waters, urban waters, crops, abandoned crops, dense urban area, non-dense urban area, natural green, urban green, non-green natural areas, non-built area (Burriel Moreno et al., 2006). Table 1 describes the land-uses indentified in the 8402 buffers analyzed.

The second map was the "Map of Land Covers of Catalonia (2005–2007)" (CREAF, 2013), which contains a total of 233 simple covers, hierarchically grouped into different levels. To conduct this study we used the level grouping these 233 categories into 24, which are: landfills, continental waters, deciduous forests, conifer forests, sclerophyllous forests, arboreal crops, herbaceous crops, sea, shrubs, cliffs, riverbank plantations, beaches, meadows, rocks and screes, wetland vegetation, riverbank vegetation, roads and parking areas, vineyards, areas under construction, burned areas, industrial areas, barren areas, urban areas, urban green. Table 1 describes the land-covers indentified in the 8402 buffers analyzed.

2.5. The analyses

We conducted the current study using the buffers of 100 m and 300 m, which are buffer sizes commonly used in epidemiological studies. The buffer of 500 m has also been used in such studies, but because of the size of the city of Barcelona we considered that conducting the analyses using this buffer would not be appropriate, as the variability between subjects would be minor. For each category of land-use or land-cover we obtained the total amount of m^2, respectively, within the buffer of interest.

We first calculated spearman correlations between each land-use and each land-cover category, respectively. In order to understand the degree of correlation between land-use and land-cover categories, we also calculated the correlation between the categories of both types of land maps. Secondly, we estimated the association between NDVI and the several land-use categories, on one hand, and the land-cover categories, on the other hand, by including the respective categories in linear regression models. In order to observe changes, the estimates are presented for each one hectare of increase of each land-use (or land-cover) category by multiplying by 10000 (1 ha) the original estimates, which were for each 1 m^2. Thirdly, also separately for land-use and land-cover maps, we conducted ANOVA models to calculate the contribution (R^2, %) of each land-use (or land-cover) category. We first included each variable alone (one single variable model) and, secondly, we included all of them in one single model (full model). The existence of collinearity between variables included in the full model could lead to an overrepresentation of some of these variables and to an underrepresentation of others on the variability (%) explained. In

Table 1
Description of each land-use and land-cover categories identified in the current study.

Land use map	
Urban waters	Swimming pool; artificial pond
Crops	Herbaceous crop; vineyard; woody crop; greenhouse
Dense urban areas	Buildings with several floors with garden; highly densely built with or without green spaces; industry; warehouse; port area; big commercial area; buildings of public use (hospitals, schools, markets, etc).
Non-dense urban areas	Houses with 1–2 families with small or big gardens; big properties with isolated buildings
Natural green	Forests; shrubs; riverbank vegetation; meadows
Urban green	Parks; gardens; flower beds; grass from sport areas
Non-green natural areas	Beach; rocky area; naked forest area; natural course
Non-built areas	Parking area; roads and highways; pedestrian area; train line; landfill; unbuilt land
Land cover map	
Continental waters	Natural course; lakes; natural and artificial ponds; rivers; reservoir; coastal lagoons
Deciduous forests	If at least 5% of deciduous species; buffer strips
Conifer forests	If at least 5% of conifer species; buffer strips
Sclerophyllous forests	If at least 5% of sclerophyllous species; buffer strips
Arboreal crops	Arboreal crops
Herbaceous crops	Herbaceous crops
Shrubs	Shrubs, including abandoned crops and shrubs in firebreaks
Riverbank plantations	*Platanus and populus species*
Beaches	Beaches
Meadows	Meadows
Rocks and screes	Rocks and screes
Riverbank vegetation	Riverbank vegetation (shrubs and deciduous and evergreen forests)
Roads and parking areas	Roads and parking areas
Areas under construction	Urban areas under construction
Industrial areas	Warehouses
Barren areas	Soil eroded by natural agent or human action; forest firebreaks
Urban areas	Isolated buildings; buildings; unbuilt urban areas; pedestrian areas without vegetation
Urban green	Artificial green areas and urban woodland

order to avoid or reduce collinearity problems we used the variance inflation factor (VIF) (Lin et al., 2011) to observe collinearity between variables (VIF > 5, data not shown). We therefore excluded "dense urban areas" from the land-use model and "urban areas" from the land-cover model, as these showed the highest VIF, and once excluded the VIF of the other variables was <5. To calculate the contribution of each land-use (or land-cover) category in the model including all the variables, we first calculated the residuals of the model including all the variables (full model). Afterwards we excluded each category from the full model and calculated the new residuals. Using the following formula, based on the explained variance of the model (sum of squares of the residuals), we derived the contribution of each variable included in the full model: [R^2 of the land-use (or land-cover) category "X" = (Sum of squares of the residuals excluding the category "X"- Sum of squares of the residuals of the full model)/Sum of squares of the residuals excluding the category "X"]. All analyses were conducted with STATA 12.0.

3. Results

3.1. Land uses and land covers variability

Table 2 shows that for certain types of uses and covers, mainly forests, urban waters and crops, the percentage of participants with at least 1 m^2 was very small (less than 5%). Meanwhile, for other types of uses and covers, the percentage of participants with at least 1 m^2 was moderate (e.g. urban green) or high (e.g. [dense] urban areas) in both maps (Table 2). The percentage of participants with at least 1 m^2 of certain land-use and land-cover categories substantially increased when the buffer increased up to 300 m; for instance, around 95% of the study participants had at least 1 m^2 of urban green in both types of maps (Table S1).

3.2. Correlations between land-uses and land-covers

In buffers of 100 m correlations between land-use categories were very weak for most of the combinations, and moderate between natural green and dense urban area (r = −0.28), non-dense urban area (r = 0.28), urban green (r = 0.14) and non-green natural areas (r = 0.15), as well as for other combinations (Table 3). Correlations were negative and strong between dense urban areas and urban green (r = −0.61) and non-built area (r = −0.71) (Table 3).

Correlations between land-cover categories were in general very weak. The strongest correlations were between sclerophyllous and conifer forests and between these two types of forests and shrubs (r between 0.34 and 0.53). Also, between urban areas and roads and parking areas (r = −0.33), industrial areas (r = −0.34), barren areas (r = −0.28) and urban green (r = −0.75), where the correlations were negative (Table 4).

Correlations between the land-use and land-cover categories are shown in Table S2. For buffers of 300 m correlations followed a similar pattern as with buffers of 100 m, although in some cases the correlations became stronger and in others weaker (Tables S3–S5).

3.3. The built environment explaining NDVI

In the models including only one land-use category at a time, those variables of the built environment explaining most of the NDVI variability in buffers of 100 m were the dense urban areas [R^2 = 0.34, β (95%CI) = −0.06 (−0.06, −0.05)], the natural green [R^2 = 0.26, β (95%CI) = 0.16 (0.16, 0.17)], the urban green [R^2 = 0.23, β (95%CI) = 0.08 (0.08, 0.09)] and, to a lesser extent, non-dense urban areas [R^2 = 0.14, β (95%CI) = 0.10 (0.10, 0.11)] (Table 5). Once all land-use categories were included in the full model (except dense urban areas due to collinearity) urban green was the variable explaining most of the NDVI variability [R^2 = 0.32, β (95%CI) = 0.08 (0.08, 0.08)], followed by natural green [R^2 = 0.28, β (95%CI) = 0.14 (0.14, 0.15)] and non-dense urban areas [R^2 = 0.11, β (95%CI) = 0.07 (0.06, 0.07)] (Table 5). The full model containing all the land-use categories explained 55% of the variability in NDVI. Similar results were obtained for buffers of 300 m. However, in this case, once all variables were included in the full model, the natural green

Table 2
Description of NDVI, land uses and land covers within buffers of 100 m around study participants (N = 8402).

Indicator or map (year of capture/realization)	Mean	Standard deviation	% of participants with at least 1 m^2 of each use or cover
NDVI (year 2007) – ratio	0.2	0.1	NA
Land use map (year 2004) – m^2			
Urban waters	10.8	130.0	1.20
Crops	3.2	85.6	0.29
Dense urban areas	27123.9	5663.6	99.56
Non-dense urban areas	317.7	1928.8	6.43
Natural green	304.6	1716.8	7.83
Urban green	1451.0	3073.4	47.68
Non-green natural areas	7.9	191.3	0.25
Non-built areas	2046.1	3199.6	56.89
Land cover map (years 2005–2007) – m^2			
Continental waters	15.9	150.8	1.93
Deciduous forests	0.5	42.0	0.01
Conifer forests	75.1	1000.9	1.19
Sclerophyllous forests	11.3	276.2	0.30
Herbaceous crops	4.2	110.9	0.38
Shrubs	56.3	631.9	2.12
Beaches	1.3	86.4	0.04
Meadows	37.1	322.1	2.73
Riverbank vegetation	1.0	94.3	0.01
Roads and parking areas	290.4	1152.7	13.08
Industrial areas	643.1	2524.2	18.58
Barren areas	122.3	768.9	6.59
Urban areas	28538.6	4465.4	99.96
Urban green	1603.6	2705.1	55.18

NA = not applicable.

Table 3
Correlation between land-use categories within buffers of 100 m around study participants (N = 8402).

	Urban waters	Crops	Dense urban areas	Non-dense urban areas	Natural green	Urban green	Non-green natural areas	Non-built areas
Urban waters	1							
Crops	−0.01	1						
Dense urban areas	−0.06	−0.07	1					
Non-dense urban areas	0.00	0.01	−0.25	1				
Natural green	0.01	0.14	−0.28	0.28	1			
Urban green	0.05	0.07	−0.61	0.05	0.14	1		
Non-green natural areas	0.02	0.00	−0.08	0.05	0.15	0.01	1	
Non-built areas	0.00	0.02	−0.71	0.00	0.00	0.21	0.05	1

[R^2 = 0.52, β (95%CI) = 0.02 (0.02, 0.02)] explained more than urban green [R^2 = 0.41, β (95%CI) = 0.01 (0.01, 0.01)]. Non-dense urban areas explained again 11% of the variability [β (95%CI) = 0.01 (0.01, 0.01)] (Table S6). Additionally, in 300 m buffers, the full model containing all the land-use categories explained 76% of the variability in NDVI.

Models including only one land-cover category at a time showed similar results to those obtained using land-use categories; in 100 m buffers, urban green areas [R^2 = 0.30, β(95%CI) = 0.11 (0.11, 0.11)], urban areas [R^2 = 0.22, β(95%CI) = −0.06 (−0.06, −0.05)] and conifer forests [R^2 = 0.18, β(95%CI) = 0.23 (0.22, 0.24)] explained most of the NDVI (Table 5). Once all the land-use categories were included in the full model (except urban areas due to collinearity) urban green remained the variable explaining most of the variability [R^2 = 0.39, β(95%CI) = 0.11 (0.11, 0.11)], followed by conifer forests [R^2 = 0.19, β(95%CI) = 0.19 (0.18, 0.20)] and to a much lesser extent shrubs [R^2 = 0.05, β(95%CI) = 0.14 (0.13, 0.16)] (Table 5). The full model containing all the land-cover categories explained 54% of the variability in NDVI. Results were similar for buffers of 300 m; again, in the full model, urban green areas explained most of the variability [R^2 = 0.53, β (95%CI) = 0.01 (0.01, 0.01)], followed by conifer forests [R^2 = 0.22, β (95%CI) = 0.02 (0.02, 0.02)] and in a much lesser extent shrubs [R^2 = 0.09, β (95%CI) = 0.01 (0.01, 0.01)] (Table S6). In this case, the full model containing all the land-cover categories explained 78% of the variability in NDVI.

4. Discussion

Results of the present study show that although the percentage of participants with natural green areas within 100 m and 300 m was quite low compared to the percentage of participants with urban green areas, both types of green spaces had a similar impact on NDVI variability. Within natural areas, conifer forests explained a greater part of NDVI variability than other natural areas.

In the current study we used land-use and land-cover maps as the gold-standard to determine the distribution of different types of urban uses or urban areas. The category “natural green” in the land-use map includes forests, shrubs, riverbank vegetation and meadows, while in the land-cover map natural green areas are split into separate categories showing that conifer forests is the category with a greater impact on NDVI variability in the city of Barcelona, although few participants (less than 4%) had a conifer forest within the buffers of study. Regarding urban green, in the land-use map this category includes parks, gardens, flower beds and grass from sport areas, whereas in land-cover maps it includes artificial green areas and urban woodland. Another land-use category that had a relevant role in explaining NDVI variability was the category known

Table 4
Correlation between land-cover categories within buffers of 100 m around study participants (N = 8402).

	Continental waters	Deciduous forests	Conifer forests	Scleroph-yllous forests	Herbaceous crops	Shrubs	Beaches	Meadows	Riverbank vegetation	Roads and parking areas	Industrial areas	Barren areas	Urban areas	Urban green
Continental waters	1													
Deciduous forests	0.00	1												
Conifer forests	0.01	0.10	1											
Sclerophyllous forests	0.01	0.00	0.46	1										
Herbaceous crops	−0.01	0.00	0.01	0.00	1									
Shrubs	0.01	0.00	0.53	0.34	0.06	1								
Beaches	0.00	0.00	0.00	0.00	0.00	0.00	1							
Meadows	−0.02	0.00	0.16	0.07	0.04	0.13	0.00	1						
Riverbank vegetation	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00	1					
Roads and parking areas	0.00	0.00	0.06	0.07	0.01	0.07	−0.01	0.09	0.03	1				
Industrial areas	−0.04	−0.01	−0.05	−0.03	0.07	−0.01	−0.01	0.02	−0.01	0.03	1			
Barren areas	0.01	0.00	0.05	0.06	0.14	0.13	−0.01	0.14	0.04	0.09	0.06	1		
Urban areas	−0.11	−0.02	−0.16	−0.09	−0.08	−0.18	−0.01	−0.16	−0.02	−0.33	−0.34	−0.28	1	
Urban green	0.10	−0.01	0.01	−0.02	0.07	0.04	−0.01	0.08	0.01	0.11	−0.10	0.18	−0.75	1

as "non-dense urban areas". This category includes houses with 1–2 families with small or big gardens. This vegetation in these properties might explain the 11% of NDVI variability associated to these areas.

Both land-use and land-cover maps available for the city of Barcelona are of high resolution (they allow taking into account objects with a minimum size of 10 m width and 50 m length, except roads and train tracks, which needed a minimum width of 8 m), and therefore are considered to be good tools to evaluate NDVI variability. However, these maps have explained NDVI variability better when buffers were of 300 m (0.76% and 0.78%, respectively) than of 100 m (0.54% and 0.55%, respectively). These results might be explained by the fact that to explain NDVI in small buffers, such as 100 m, the resolution needs to be higher than the 30 m × 30 m NDVI resolution of the current study. In this sense, better resolutions can be obtained from commercial satellites such as Digital Globe (WorldView, Geoeye, Ikonos, QuickBird), Astrium (SPOT, Pleiades, TerraSAR), eGeos (Cosmo-Sky-Med, Radarsat) and RapidEye (RapidEye satellites). The use of better resolutions would allow better capturing the high variability between land-uses or land-covers and also reducing the potential underestimation of greenness exposure that occurs at small buffers (i.e. 100 m) with the LandSat data, which has the advantage that can be obtained for free. Furthermore, our results show that in the city of Barcelona the percentage of participants with at least 1 m^2 of natural green areas or forests is low, whereas these areas have an important impact on NDVI variability. In this sense, results of the present study indicate that before running epidemiological studies it is important to understand what NDVI means in a particular area of study according to the exposure of interest (i.e. general greenness – both urban and natural – or only urban greenness) or the hypothesis of the study, as different built environment characteristics or urban green elements can have different implications with regards to a number of health outcomes or mechanisms. For instance, different types of urban green elements may have different implications to asthma and allergy (Dadvand et al., 2014a), or benefits of close-range exposures (e.g window views) might not be captured by using NDVI as exposure metric. Therefore, depending on the objectives of the study NDVI might not be the most appropriate metric to be used.

Although our study is novel in its aims and methodology, there are some methodological limitations that need be discussed. Apart from limitations due to spatial resolution, which have already been discussed, in our study we had to address problems of collinearity between the different land-use and land-cover categories included in the respective models. For that, we applied VIF, which allowed us to identify such problems and forced us to exclude one variable (dense urban areas and urban areas, respectively) from each model. Furthermore, land-use and land-cover maps contain many more categories (58 and 233, respectively) than the ones used in the current study (10 and 24, respectively). The use of more categories would have provided more detailed information, but as we did not have enough study points to evaluate such heterogeneity, we were forced to group these categories and therefore lose refined information. Also, in order to refine our analyses, in the current study we wanted to use the number of trees in an area as a variable to take into account in our analyses; however, the municipality of Barcelona has only information on public trees (i.e. their map does not include wild trees or those in private properties), and therefore the information is incomplete and biased. Finally, the data used in the current study was obtained at different years (NDVI in 2007, land-uses in 2004 and land-covers between 2005 and 2007). However, findings from previous studies support the stability of the NDVI spatial contrast over seasons and years (Dadvand et al., 2012a,b), as well as for land-use and land-cover categories (Burriel Moreno et al., 2006; CREAF, 2013), when considering few years of difference.

Table 5
Association β(95%CI) between NDVI and each land-use and land-cover category, respectively, and contribution (R^2) of each category within buffers of 100 meters around study participants.

	One single category in the model		All categories in the model	
	R^2	β (95%CI)[a]	R^2	β (95%CI)[a]
LAND USE MAP				
Total R^2	NA		0.55[b]	
Urban waters	0.00	−0.094 (−0.183, −0.005)	0.01	−0.206 (−0.265, −0.146)
Crops	0.01	0.469 (0.335, 0.604)	0.00	0.160 (0.065, 0.251)
Dense urban areas	0.34	−0.055 (−0.057, −0.054)	Excluded[c]	
Non-dense urban areas	0.14	0.103 (0.098, 0.109)	0.11	0.069 (0.065, 0.073)
Natural green	0.26	0.162 (0.156, 0.167)	0.28	0.140 (0.136, 0.145)
Urban green	0.23	0.085 (0.081, 0.088)	0.32	0.081 (0.079, 0.084)
Non-green natural areas	0.00	−0.048 (−0.109, 0.012)	0.03	−0.332 (−0.373, −0.291)
Non-built areas	0.01	0.015 (0.012, 0.019)	0.01	0.009 (0.006, 0.011)
LAND COVER MAP				
Total R^2	NA		0.54[b]	
Continental waters	0.00	0.060 (-0.017, 0.137)	0.00	−0.157 (−0.210, −0.105)
Deciduous forests	0.01	1.033 (0.759, 1.307)	0.00	0.507 (0.319, 0.695)
Conifer forests	0.18	0.230 (0.219, 0.240)	0.19	0.191 (0.183, 0.200)
Sclerophyllous forests	0.05	0.454 (0.413, 0.495)	0.03	0.224 (0.194, 0.254)
Herbaceous crops	0.01	0.415 (0.312, 0.519)	0.00	0.103 (0.031, 0.174)
Shrubs	0.09	0.255 (0.238, 0.273)	0.05	0.142 (0.129, 0.155)
Beaches	0.00	−0.106 (−0.240, 0.028)	0.00	−0.073 (−0.164, 0.017)
Meadows	0.01	0.141 (0.106, 0.177)	0.00	0.036 (0.011, 0.061)
Riverbank vegetation	0.00	0.403 (0.281, 0.525)	0.00	0.266 (0.183, 0.349)
Roads and parking areas	0.00	0.023 (0.013, 0.033)	0.00	−0.011 (−0.018, −0.004)
Industrial areas	0.01	−0.018 (−0.023, −0.014)	0.01	−0.014 (−0.017, −0.010)
Barren areas	0.00	0.043 (0.028, 0.058)	0.00	−0.014 (−0.025, −0.003)
Urban areas	0.22	−0.057 (−0.059, −0.054)	Excluded[c]	
Urban green	0.30	0.109 (0.105, 0.112)	0.39	0.110 (0.107. 0.113)

NA = not applicable.
[a] Increase or decrease of NDVI for each hectare of land-use or land-cover, respectively.
[b] The total R^2 does not equal the sum of the individual R^2 because the later are calculated by excluding each land-use or land-cover category from the model with all categories in (see methods section for more information).
[c] These categories were excluded from the model because the variance inflation factor (VIF) was >5, indicating collinearity with other categories.

LandSat NDVI imaging, which has been shown to be associated with different health outcomes in a number of studies (Pereira et al., 2012; Pilat et al., 2012; Gascon et al., 2015; Dadvand et al., 2014a), is a useful tool to define exposure to surrounding greenness because it is a standardized measure of greenness that can be easily and often freely obtained from available on-line depositories. However, results of the present study show that studies need to consider the limitations of using NDVI, particularly LandSat NDVI imaging, of less resolution, as already discussed above. Finally, in the current study we used geocoding data from a well-established pregnancy cohort that has been used in different epidemiological studies (Dadvand et al., 2012b, 2011a,b); therefore, our findings can be relevant to similar epidemiological studies. Nevertheless, results of the present study should be replicated in studies from other cities with different urban characteristics and climate conditions. Improvement of green exposure assessment and of the understanding of built environment determinants of NDVI or similar indexes is not only useful for epidemiologist but also for urban designers and policy makers, as they are in charge to redesign specific urban spots to, among other things, make them "healthier".

Acknowledgments

This project was funded by the CERCA Institutes Integration Program (SUMA 2013) [promoted and managed by the Secretariat for Universities and Research of the Ministry Economy and Knowledge of the Government of Catalonia (SUR), the Agency for Management of University and Research Grants (AGAUR) and the CERCA Institute]. Payam Dadvand is funded by a Ramón y Cajal fellowship (RYC-2012-10995) awarded by the Spanish Ministry of Economy and Competitiveness.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ufug.2016.07.001.

References

Burns, C.J., Wright, J.M., Pierson, J.B., Bateson, T.F., Burstyn, I., Goldstein, D.A., Klaunig, J.E., Luben, T.J., Mihlan, G., Ritter, L., Schnatter, A.R., Symons, J.M., Yi, K.D., 2014. Evaluating uncertainty to strengthen epidemiologic data for use in human health risk assessments. Environ. Health Perspect. 122, 1160–1165.
Burriel Moreno, J., Ibànez Martí, J., Terradas i Serra, J., 2006. El mapa ecológico de Barcelona: los cambios de la ciudad en las últimas tres décadas. Cuad. Geográficos 39, 167–184.
Centre de Recerca Ecològica i Aplicacions Forestals (CREAF) 2013. Legend MCSC-3v1/SIOSE Codification of the covers; Bellaterra (Cerdanyola del Vallès).
Dadvand, P., Basagaña, X., Sartini, C., Figueras, F., Vrijheid, M., de Nazelle, A., Sunyer, J., Nieuwenhuijsen, M.J., 2011a. Climate extremes and the length of gestation. Environ. Health Perspect. 119, 1449–1453.
Dadvand, P., Basagaña, X., Figueras, F., Amoly, E., Tobias, A., de Nazelle, A., Querol, X., Sunyer, J., Nieuwenhuijsen, M.J., 2011b. Saharan dust episodes and pregnancy. J. Environ. Monit. 13, 3222–3228.
Dadvand, P., Sunyer, J., Basagaña, X., Ballester, F., Lertxundi, A., Fernández-Somoano, A., Estarlich, M., García-Esteban, R., Mendez, M.A., Nieuwenhuijsen, M.J., 2012a. Surrounding greenness and pregnancy outcomes in four Spanish birth cohorts. Environ. Health Perspect. 120, 1481–1487.
Dadvand, P., de Nazelle, A., Figueras, F., Basagaña, X., Su, J., Amoly, E., Jerrett, M., Vrijheid, M., Sunyer, J., Nieuwenhuijsen, M.J., 2012b. Green space, health inequality and pregnancy. Environ. Int. 40, 110–115.
Dadvand, P., Villanueva, C.M., Font-Ribera, L., Martinez, D., Basagaña, X., Belmonte, J., Vrijheid, M., Gražulevičienė, R., Kogevinas, M., Nieuwenhuijsen, M.J., 2014a. Risks and benefits of green spaces for children: a cross-sectional study of associations with sedentary behavior, obesity, asthma, and allergy. Environ. Health Perspect. 122, 1329–1335.
Dadvand, P., Wright, J., Martinez, D., Basagaña, X., McEachan, R.R.C., Cirach, M., Gidlow, C.J., de Hoogh, K., Gražulevičienė, R., Nieuwenhuijsen, M.J., 2014b.

Inequality, green spaces, and pregnant women: roles of ethnicity and individual and neighbourhood socioeconomic status. Environ. Int. 71, 101–108.
Dadvand, P., Nieuwenhuijsen, M.J., Esnaola, M., Forns, J., Basagaña, X., Alvarez-Pedrerol, M., Rivas, I., López-Vicente, M., De Castro Pascual, M., Su, J., Jerrett, M., Querol, X., Sunyer, J., 2015. Green spaces and cognitive development in primary schoolchildren. Proc. Natl. Acad. Sci. U. S. A. 112, 7937–7942.
Figueras, F., Meler, E., Iraola, A., Eixarch, E., Coll, O., Figueras, J., Francis, A., Gratacos, E., Gardosi, J., 2008. Customized birthweight standards for a Spanish population. Eur. J. Obstet. Gynecol. Reprod. Biol. 136, 20–24.
Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Forns, J., Plasència, A., Nieuwenhuijsen, M., 2015. Mental health benefits of long-term exposure to residential green and blue spaces: a systematic review. Int. J. Environ. Res. Public Health 12, 4354–4379.
Grazuleviciene, R., Danileviciute, A., Dedele, A., Vencloviene, J., Andrusaityte, S., Uždanaviciute, I., Nieuwenhuijsen, M.J., 2015. Surrounding greenness, proximity to city parks and pregnancy outcomes in Kaunas cohort study. Int. J. Hyg. Environ. Health 218, 358–365.
James, P., Banay, R.F., Hart, J.E., Laden, F., 2015. A review of the health benefits of greenness. Curr. Epidemiol. Rep. 2, 131–142.
Lachowycz, K., Jones, A.P., 2011. Greenspace and obesity: a systematic review of the evidence. Obes. Rev. 12, e183–e189.
Lachowycz, K., Jones, A.P., 2014. Does walking explain associations between access to greenspace and lower mortality? Soc. Sci. Med. 107, 9–17.
Lin, D., Foster, D.P., Ungar, L.H., 2011. VIF regression: a fast regression algorithm for large data. J. Am. Stat. Assoc. 106, 232–247.
Pereira, G., Foster, S., Martin, K., Christian, H., Boruff, B.J., Knuiman, M., Giles-Corti, B., 2012. The association between neighborhood greenness and cardiovascular disease: an observational study. BMC Public Health 12, 466.
Pilat, M.A., McFarland, A., Snelgrove, A., Collins, K., Waliczek, T.M., Zajicek, J., 2012. The effect of tree cover and vegetation on incidence of childhood asthma in metropolitan statistical areas of Texas. Horttechnology 22, 631–637.
Triguero-Mas, M., Dadvand, P., Cirach, M., Martínez, D., Medina, A., Mompart, A., Basagaña, X., Gražulevičienė, R., Nieuwenhuijsen, M.J., 2015. Natural outdoor environments and mental and physical health: relationships and mechanisms. Environ. Int. 77, 35–41.
Weier, J., Herring, D., 2015. NASA Earth Observatory, http://earthobservatory.nasa.gov/Features/MeasuringVegetation/ (accessed 13.7.2015.).