BAHRAIN PARKS: A CLOUD PLATFORM FOR PLANNING, DESIGNING, AND MANAGING PUBLIC PARKS

JOAO PINELO SILVA*

Assistant Professor
Department of Architecture and Interior Design
College of Engineering
University of Bahrain
Kingdom of Bahrain
PO Box 32038

Email: jpinelo@uob.edu.bh

Mailing Address: Southern Governorate, UOB Isa Town Campus, Building 35, Room 229

ABSTRACT

Purpose: The paper introduces an innovative, open source cloud platform for the efficient, knowledge-based, sustainable and inclusive development of urban parks - Bahrain Parks.

Design/Methodology/Approach: We describe the local context at the genesis of the initiative, its components, and the rationale for the creation of a comprehensive database. We summarise the data available and explain how the application leverages it to useful information for decision-making on park maintenance, planning, and design.

Findings: We share our initial findings regarding the official acceptance and attempt to integrate the application in workflows at several agencies.

Originality/Value: To our knowledge, the project is unique as it is being developed to incorporate an evidence-based design process comprehensively. Furthermore, it embeds explicit and implicit public participation in decision-making, through park rating and usage. We argue that the project belongs to the sphere of e-governance.

*Corresponding Author

Keywords: Bahrain; sustainable development; built environment; e-government; urban park; cloud computing; public participation; park design; smart city; evidence-based design

INTRODUCTION

In this article, we describe the case study Bahrain Parks, an initiative for the creation of a cloud platform within an e-governance approach, for the efficient and participative management of the infrastructure of public urban parks in the Kingdom of Bahrain. The project aims for the knowledge-driven, sustainable and socially inclusive development of the built environment. We start by referring to the local context for the creation of the project Bahrain Parks. This is followed by a description of the advantages of creating a comprehensive dataset of the park stock. We then describe the main foreseeable advantages that the project brings directly to stakeholders, followed by the components of the initiative, technology, the quantification of usage, and user stated-preferences. We briefly discuss the beginning of the attempts to include the tool in existing workflows, and we debate the initiative within an e-governance framework, before referring to future work.

The Kingdom of Bahrain has an area of approximately 765 square kilometres and a stock of over 200 public urban parks. Parks are planned and built directly by either the Ministry of Works or one of the four Municipalities. Therefore, there are several teams, in different institutions, that are responsible for planning, designing and maintaining the public parks. This situation is typical in other countries, often with hundreds of teams involved with building and managing the infrastructure of parks.

Overall, in Bahrain, the parks share a mix of successful and unsuccessful results, measured by their use. Although park usage is not systematically measured, it is empirically evident that some parks are vastly more successful than others. While some are underused, others are so popular that they often generate more traffic than the infrastructure is programmed to support. Both extreme cases are undesirable. Underuse indicates a waste of both the resources and the opportunity to create a successful park. Often, such cases seem to lead to a lack of maintenance and the decay of the infrastructure, which has the potential to spread through the urban environment and has been known to attract improper use. On the other hand, overuse penalises both the users of the parks and the surrounding areas. While both situations are detrimental, frequently, the reasons behind success and failure are not clear. These situations cause concern to the governance and the institution that is ultimately responsible for the built environment; the Ministry of Works wishes to address it to optimise the use of resources in serving the population.

Apart from providing users with socialising and leisure opportunities, the parks play other key roles in the sustainability of the urban environment. For example, the porosity of pavements facilitates aquifer recharge and the management of stormwater (Andersson-Skold et al., 2015). The abundance of vegetation breaks down noise pollution (Pathak et al., 2008), and contributes to the quality of the atmosphere by reducing airborne particles (Janhäll, 2015). Through its relatively low mass (in comparison with concrete and other traditional building materials) and albedo, the vegetation decreases the temperature at the local scale thus playing a part in the mitigation of the heat island effect (Susca et al., 2011). Vegetation also contributes to the local biodiversity by creating a habitat for species that would otherwise not exist within the urban environment (Savard et al., 2000).

In the particular context of Bahrain, with relatively warm temperatures for three to four months of the year, parks, with their vegetation, shade, evapotranspiration, and low mass, are the coolest outdoor places, although this has not yet been quantified.

Measuring the contribution of parks to the urban ecosystem on the aspects mentioned above would require detailed information regarding the parks. The Department for Desert Farming at the Arabian Gulf University approached us for the inclusion of data beyond our initial plan. For example, the inventory and layout of the vegetation of a park would allow for estimating its role regarding the filtration of air pollutants. Quantifying the green waste per park (and therefore for the whole stock) could generate information to support the first national strategy for the transformation of waste into a resource, eventually creating an industry. We share the view of Scott Campbell in believing that environmental protection, economic development, and social justice, three goals of planning, can be perspectives of the same reality and not mutually exclusive aims (Campbell, 1996).

A COMPREHENSIVE GRASP OF THE PARK STOCK

The study of parks and their contribution to a sustainable urban ecosystem, whether social relevance or environmental impact, could take two complementary forms: an holistic study of all park characteristics on a case-by-case basis, or the study of the features of the park stock in aggregate. Both avenues of inquiry require data: as is often the case, this is not readily available. Furthermore, the second avenue of enquiry, the aggregate study of parks, requires data that are consistent across the sample. The consistency of data means that the same data are available for all parks, measured and processed through the same methods. With the initiative Bahrain Parks, we are creating such a database not of a sample of parks, but of all the public urban parks in Bahrain. Furthermore, the initiative aims to make the full dataset publicly available.

To open the database in a way that has visibility and is meaningful, we make it accessible through a Web application. The application delivers the data in a straightforward fashion to each of the different user groups of stakeholders. While experts can export the whole dataset and analyse it at will, the application also offers users basic tools to explore data. For users who do not wish to study the data in an active way, algorithms create listings of information ready to be used. The general public can, for example, consult a geographic list of the top ten family-oriented parks. Nonetheless, the application is being prepared to be used by planners, designers, and maintainers of parks in their daily activities, with minimal effort, largely outweighed by benefits.

One other aspect of the application is that it also allows for the creation of data. For example, park ratings, user counts, vegetation surveys and logging of green waste are made directly through the application. Such an approach contributes towards data availability, the ease of creating data, and the standardisation of methods and procedures.

The availability of data from all parks, its homogeneity, and integration in one single database creates unique conditions for analysis. The readiness of the data through an application that guides users to simple analysis leverages the intervention of planners, designers and park managers. The application is created to be easy to use and provide immediate feedback. Moreover, since all stakeholders are familiar with the database and the application, this facilitates the transfer of knowledge between planning and design teams at the different institutions. Similarly, when researchers publish the results of elaborate studies not carried out by the stakeholder agencies, the teams at these organisations are familiar with the data used, and can eventually assimilate the interpretation of the results more quickly. Furthermore, being able to access the data directly, the teams have the possibility of quickly checking the situation of their stock by referring to the indicators provided by research.

The comprehensive and consistent dataset containing the full stock of parks will allow for the statistical analysis of the role of design elements in the success of parks based on the number and frequency of users and ratings. Statistical techniques, such as multiple regression analysis, can be applied to the dataset to create information for the planning and design of parks. Such information can then be used to improve the parks continuously and iteratively. By establishing baselines of usage and ratings and stimulating continuous measurements, the dataset makes it possible to predict and then measure the impact of design iterations on usage and ratings. As the knowledge about design increases, the differences between predicted and measured effects will diminish, reflecting the growth in the effectiveness of the design process. The method for such development must not be a blind search for correlational phenomena but the development of theoretical propositions (Batty, 2013; West, 2013). The data provides the opportunity for the testing of the new hypothesis. A caveat of such a process, as highlighted by Batty (2013) about big data, relates to the fact that despite the potential of covering a broad range of time-spans, to begin with, it lacks long-term data, therefore excluding most existing urban planning theories. Simultaneously, this opens the possibility to focus on the short-term, which must be carefully considered (Batty, 2013).

The dataset and the application presented here do not constitute big data. For now, it does not have the tabular dimensions typically associated with big data (loosely defined as containing over one million rows), nor is it continuously fed by sensors in real-time. However, it has the possibility to become so. First, the application registers all events with a timestamp, which allows for endless time-span aggregations (a typical characteristic of big datasets). Second, we can automate the manual component of most semi-automated features with the use of sensors (another hallmark of big datasets). We see Bahrain Parks as a means of a gradual transition between virtually no data to a big data scenario. The process seems rich in opportunities for learning how to deal with previously unavailable data, on the selection and use of data, and on methods to link different data streams.

Tailored and Augmented Information for Stakeholders

The dedicated nature of the data towards specific goals, its digital format, and centralisation in a unique database allow for ease of use and simple delivery of content. It also permits the provision of augmented information. Through automated methods of analysis and sophisticated operations, users without formal statistical knowledge can get access to information that is more reliable than the data, per se. For example, when a park manager consults park usage, the application can deliver results by eliminating outliers, and in the case of small sampling, produce Monte Carlo simulations to improve the reliability of averages. For example, a designer can, with a few clicks, consult the design features of one park or any chosen set, or quickly plot the relationship between any two (numerical) variables. Such insights are helpful during design but were not readily available.

The Components of Bahrain Parks

The project has three domains: a) database, b) web application, and c) social media, which we describe below. The database holds the data related to the parks. Dasu and Johnson (2003) report that 80% of the time spent doing data analysis is used to clean and prepare data. This would not be acceptable for the purpose of the project, and therefore we follow the tidy data protocol as defined by Wickham (2014), where each variable forms a column, each observation forms a row, and each type of observational unit creates a table. Tables structure the data as they relate to each other through the key ‘park ID’, forming a relational database, as described by Codd (1982).

The main table consists of manually gathered park features that are relevant to the planning and design of parks. On this table, each row represents one park and has a unique key (identifier). Each column represents one feature (such as the total area of the park, the area shaded, picnic area, playground area, access by public transport, parking, the length of walkways, the presence of toilets, food and beverage facilities, latitude, longitude, location, municipality, data open, and name). The list of variables is to be expanded based on research outputs and requests from stakeholders.

The other four tables follow a similar structure but hold data collected semi-automatically through the application. The four tables are a) user rating b) user counts c) vegetation and d) green waste as follows:

- The table ‘user rating’ consists of users’ votes on five categories: vegetation, amenities, conservation, and schedule and follows the tidy data protocol. Each row consists of one vote, and the columns refer to a) the park ID b) the rating value (0–5, where 5 is best) c) the category to which the vote refers and d) a timestamp;
- The table ‘user counts’ is a log of the users spotted at each park. On this table, each row corresponds to one user. The columns refer to the a) park IDb) user category (alphabetically: blind, child, man, senior, wheelchair and woman) and c) timestamp. We address the rationale behind counting and categorising users below;
- The table ‘vegetation’ consists of an account of all such elements of each park. In this table, each row corresponds to one element, such as an individual tree, shrub, planter or grass patch. The columns contain information regarding a) park ID b) type c) species d) subtype e) condition f) timestamp;
- The table `green waste’ is a log of categorised green waste disposed at each park. Each row corresponds to one load (for example, a bin bag). Columns contain information for each load, namely the park ID, the weightand a category (grass, green leaves, and dry leaves).

The inclusion of timestamps together with data logging, such as user counts, waste removed or ratings, allows for a time-wise analysis as well as for the aggregation of data with different time-spans. This flexibility is advantageous because it makes the data useful to study both different phenomena based on time references, and the same phenomenon in different time-scales.

The web application is publicly accessible online at www.bahrainparks.org/home/app/. We organised it regarding its primary target audiences. A description of the nine tabs available at the time of writing (December 2016) follows.

- The tab ‘Home’ introduces the application by briefly describing public parks and mapping the stock on the database. It also provides insight on the social media feeds;

- The tab 'Families' just lists the top 10 parks that have the right attributes to be considered adequate for families. The criteria are car access, parking, walkways, children's playground, picnic area, shade, food and beverage facilities, and toilet facilities. The application algorithmically calculates the list of parks to be included in the suggestion for families based on the criteria described. At the time of writing, only three of the parks available in the database met the criteria;
- The tab 'Rating' allows any user to rate a park, and illustrates the scores in a ranking that aims at instigating healthy competitiveness between park managers and municipalities. Other displays of ratings are planned but need further evaluation data before being implemented. The application also makes the ratings available on a table on the relevant sub-tab;
- The tab 'Explorer' allows any user to filter the parks based on six criteria such as name and amenities such as a walkway, playground, WC, picnic area, food, and drink.
- The tab 'Professional' is aimed at the planners and designers of parks: it provides the most comprehensive access to the database. It supports the filtering/grouping of parks based on features such as municipality (governorate) and location, among others. It also offers tools for instant analysis such as mapping, accessing tabular data, plotting variables such as XY, and quick histograms of individual variables. The features available are not comprehensive and are under constant development. They are not, however, 'researcher' tools in the proper sense of the word, but a device to provide immediate access to the data in an insightful manner, and to stimulate the emergence of research questions;
- The tab 'Documentation' provides general information about the project and the data, such as the variables available, sources of data, and units. This tab also contains information about the how the application creates (algorithmically) the other variables based on data input. A full list of parks available in the dataset completes the tab, and is organised both alphabetically and by 'park ID';
- The tab 'About' explains the genesis of the project as well as the main goals and objectives; it also names the team and other contributors to the project;
- The tab 'Get Involved' provides the user with the opportunity to express their willingness to contribute to the project, as well as simply provide feedback or point out inaccuracies in the data;
- The tab 'Park Manager' allows park managers to carry out categorised counts of park users.

Social Media

The social media streams are part of the project in the sense that they stimulate user engagement with both the parks and the application. We strive to keep the feeds educational as well as informative. Examples of educational topics are health benefits of exercise and the social and environmental benefits of parks. We have been hosting a photo competition with the parks as the theme, which has been running via Instagram. In the medium-term, we aim to amplify the role of the social media by analysing its data, semantically and geographically.

Explicit User Feedback – Park Rating

Any person with access to the application can rate a park: no registration is needed. We wanted to make it clear that rating submissions are anonymous; this aligns well with making the process as simple as possible. We plan to keep the process anonymous, but we would like to make the rating more informative. In the future, we plan to offer the voter the opportunity to, voluntarily, provide extra information about themselves. We plan to implement this at a later stage once the voting has spread and become a habit. This approach makes sense because only then will the application will be getting sufficient data that is statistically significant when disaggregated by user groups.

To rate, the user must select one park by name, and then proceed to rate the park on five criteria: vegetation, amenities, conservation, cleanliness, and schedule. The criterion ‘vegetation’ is described as relating to the quality of the vegetation in the park. The aspect ‘amenities’ is described as relating to the presence of facilities, such as sitting areas, shading, WC, playground, sports fields, and the existence of busy as well as quiet areas. The criterion ‘conservation’ relates to the state of preservation of the park in general. ‘Cleanliness’ refers to the presence or absence of litter. The aspect ‘schedule’ reflects the adequacy of the opening hours of the park, making it convenient or not for daily use. The system logs each rating with a timestamp, making it possible, in the future, to access the all-time score, but also the rating for a given period, for example, the past weekend. We aim to discuss the expansion of the criteria further with all stakeholders and eventually expand it.

From a governance perspective, the criteria aim to provide information for both the planning/design of parks and their maintenance. Data on the criterion ‘vegetation’ will provide data, which, once treated and analysed throughout the whole set of parks, may be informative on users preferences on vegetation types, species, density, spatial arrangement, and spatial layout. Data on the criterion ‘amenities’, once treated and analysed within the complete dataset of parks, will be informative to park planners and designers.

The need for amenities reflects the character and use of the park as well as its local context. In some cases, amenities might not be provided in the park but in the immediate vicinity. For example, one can speculate that a small park within a residential neighbourhood might not need toilets on site as most of its regular users are within walking distance from home. The understanding of parks within their local context versus their classification regarding amenities is relevant to plan the location of parks and their design (relating the creation of amenities), which has significant budgeting consequences. Planning a park within the appropriate surroundings could have a positive impact on its use while decreasing the need for amenities and therefore maximising the reach of the budget.

Detailed rating data will guide planners and designers in the medium- and long-terms. Ratings will also provide useful information for park maintenance, both in real-time and in short (daily/weekly) timeframes. The application automatically produces the algorithmic computation of the ranking of parks, from the average of ratings across all features. The application prominently displays a live plot of the top 15 of the overall ranking. The full national ranking/rating is also made accessible, live, on the application in a tabular format. Other readily available displays of rating and ranking are planned, as well as the integration on rating and ranking on queries.

Park Users

As discussed above, parks play a broad set of roles, from which the social aspects are of relevance. In simple terms, parks afford the population opportunities for leisure and informal socialising. Therefore, the quantification of users is at the base of any measurement of social impact. Furthermore, a categorised account of users extends the detail of information and creates opportunities for planners and designers of parks to create facilities that integrate a broad range of users. Currently, the application offers park managers the possibility of categorising users into six groups: blind, child, man, senior, wheelchair and woman. The categories reflect classes of users with different needs from the built environment. The planners and designers of parks have agency over such features and can, therefore, use the information to make the parks more inclusive. Establishing usage baselines grants the ability to observe long-terms trends. For example, an increase in the number of blind or wheelchair users reflects advancement towards a more inclusive society. In the short-term, regular user surveys can provide insight on the efficacy of design solutions. The delta of usage before and after interventions provides the means for a quick and accurate assessment of the success of an intervention, a post-occupancy evaluation. The knowledge obtained from the process can be fed back into new designs quickly and with confidence.

TECHNOLOGY

Providing all stakeholders access to one database, in real-time, implies the database is online. Furthermore, the database must be continuously available to get data, not only serve data; we therefore hosted the database on an online server, 'the cloud'. To serve the data and facilitate data creation, we created the application, hosted at the domain www.bahrainparks.org/home/app/. We developed the application in R (R Development Core Team, 2016), and R-Shiny (Chang et al., 2015). R is a computer programming language for statistical computing, and therefore ideal for data handling and advanced statistical analysis (Chambers, 2008). R-Shiny is a Web framework for the online implementation of programs written in the R language (Chang et al., 2015).

Park Stock and Data Available

Although we do not have an accurate estimate of the total number of parks, the database has data on more than 60 parks. The data is inserted by hand through the measurement of park features on digital cartography, building plans (facilitated by the Ministry of Works and the municipalities), and on site. A second team carries out data verification. On finding errors, the team double-checks the data and updates the database. We are developing automatic methods of verifying the data that is inserted directly into the application.

Timeline and Engagement

We imagined the initiative in Spring 2016 and started developing the database and the application in parallel during the summer break. In October 2016, we initiated contact with the Ministry of Works, which was interested in the project. We aim to meet on a regular basis to discuss the implementation of features and the data deemed necessary. The process feels slower than the ambitious team excitedly and naïvely expected it to be, but the adherence rate has been 100%. All stakeholders contacted have shown a high interest in being involved. The public is responding as well, and less than 30 days after revealing our project, over 200 users provided park ratings. We are continuously adding parks to the database. We aim to have 100% of the parks in the database by June 2018. We have just started to verify the data. By June 2018, we expect to have completed 95% of the work on the database. This estimate does not account for the vegetation inventory, which an affiliated institution, the Department of Desert Farming at the Arabian Gulf University, will help us to complete.

Integration on Workflows

We believe that a measure of the success of the application is its integration with workflows. Regarding planning and design teams, we are developing the integration by introducing the application with some examples of unexpected insights revealed by the use of the tool. For example, a quick plot of the area of parks versus their shaded area reveals a subtle correlation. In a country where shade is paramount,

one would expect the larger parks to have larger amounts of shade and therefore a significant correlation; this is absent. Furthermore, we can expand the analysis by isolating municipalities and checking whether some tend to have a different pattern. It turns out that is the case. Such examples seem to instigate curiosity and the use of the application for consultation during planning and design work. We have not yet quantified this. Another aspect of integration is the extent to which it is used to create new data. Much of the data collection is carried out on-site.

We are discussing the possibility of establishing the role of park manager with the Ministry of Works. This would be a person who could become responsible for one (or more) parks and be the point of contact between the park, the users, and the design and maintenance teams. The park manager would be mostly on-site to carry out data collection, and would also guide maintenance through user feedback (user ratings from the application). Without the role of park manager in place, we have as yet been unable to find a systematic way of implementing regular counts.

We are developing two other features for data collection, one to quantify green waste, and another to create an inventory of vegetation per park. Regarding the inventory, the plan is to integrate data collection on post-graduate theses, so it is carried out within a scientific context supervised by experts. The park manager is responsible for the quantification of green waste. Traditionally, the waste is collected in plastic bins and carried from the park in trucks; each bin is weighted and its contents classified during the loading of the truck.

Governance

Demediuk et al. (2014) observe that, “Local governance occurs when a local government give citizens a say in things that really matter to them” (Demediuk et al., 2014). The application Bahrain Parks provides citizens with an electronic and continuous opportunity to participate in the management, design, and planning of urban public parks. Therefore, the initiative falls within the sphere of e-governance, in another step to facilitate and stimulate public participation. Towards the socially inclusive and sustainable, knowledge-based, development of the built environment, the case study does not follow traditional efforts to modernise public management systems. Bahrain Parks creates the potential for a more agile and customer-centred approach, which are the essence of Digital-Era Governance (DEG) (Dunleavy et al., 2006). In particular, it follows several themes and components of the DEG as described by Dunleavy et al. (2006).

Under the ‘Reintegration’ theme, the component is ‘reengineering back-office functions’ (improve productivity based on newer IT). Under the theme ‘Needs-based Holism’, the components are:

a) client-based or function-based reorganisation (re-integrate agencies around specific client groups);
b) interactive and 'ask-once' information seeking (holistic view of people's needs and preferences);
c) end-to-end service reengineering (create different service provision models focussing on the whole process, eliminating artificial boundaries);
d) agile government (quick and flexible responsiveness).

On the third theme, 'Digitization Changes' describes organisational and cultural changes within agencies and behavioural shifts by civil society, such as the habit of rating parks, the component 'facilitating isocratic administration and co-production' (shift from agency-centred to stakeholder-centred processes) (Dunleavy et al., 2006).

Future Work

We keep introducing new parks to the database and verifying all of the data; this is accomplished by independent teams. The work towards the integration of the application in the workflow of planners, designers, and managers of parks has just started. We will push forward to explore among groups of the public administration where it is foreseeable that the database might provide insightful information. In the transition to a scenario of big data, the verification of information obtained from such sources is necessary. Bahrain Parks, with a large georeferenced and timestamped database, and an associated social media stream, is uniquely positioned to explore the relationships between the digital and the physical phenomena. For example, we aim to explore the potential of the semantic analysis of social media streams to create information that has a bearing on the built environment. To do this, we will investigate the relationship between social media data and other (tangible) user inputs such as park rating, usage, and features. As data becomes informative and useful, the development of its scope and depth is likely to grow. For example, we expect to transition some of the existing data collection methods from semi-automated to fully automated through sensor-based input. It is likely that such a transition will follow a logic of cost/benefit; finding a balance between the usefulness of the data versus the costs of implementing a viable network of sensors.

REFERENCES

Andersson-Skold, Y., Thorsson, S., Rayner, D., Lindberg, F., Janhäll, S., Jonsson, A., Moback, U., Bergman, R. and Granberg, M. (2015), An integrated method for assessing climate-related risks and adaptation alternatives in urban areas. *Climate Risk Management*, Vol. 7, pp. 31–50. http://doi.org/10.1016/j.

crm.2015.01.003

Batty, M. (2013), Big data, smart cities and city planning. *Dialogues in Human Geography*, Vol. 3, No. 3, pp. 274–79. http://doi.org/10.1177/2043820613513390.

Campbell, S. (1996), Green Cities, Growing Cities, Just Cities?: Urban Planning and the Contradictions of Sustainable Development. *Journal of the American Planning Association*, Vol. 62, No. 3, pp. 296–312. http://doi.org/10.1080/01944369608975696.

Chambers, J.M. (2008), *Software for Data Analysis Programming with R*. Springer Science and Business Media. http://doi.org/10.1007/978-0-387-75936-4.

Chang, W., Cheng, J., Allaire, J., Xie, Y. and McPherson, J. (2015), Shiny: web application framework for R. *R package version 0.11, 1*.

Codd, E.F. (1982), Relational database: a practical foundation for productivity. *Communications of the ACM*, Vol. 25, No. 2, pp. 109–117. http://doi.org/10.1145/358396.358400.

Dasu, T. and Johnson, T. (2003), *Exploratory Data Mining and Data Cleaning*. (Vol. 479). John Wiley & Sons.

Demediuk, P., Burgess, S. and Solli, R. (2014), Local government as a democracy actor or a service delivery actor: The supporting roles for e-governance initiatives. In *E-Governance and Social Inclusion: Concepts and Cases* (pp. 127–147). http://doi.org/10.4018/978-1-4666-6106-6.ch008.

Dunleavy, P., Margetts, H., Bastow, S. and Tinkler, J. (2006), New public management is dead - Long live digital-era governance. *Journal of Public Administration Research and Theory*, Vol. 16, No. 3, pp. 467–94. http://doi.org/10.1093/jopart/mui057.

Janhäll, S. (2015), Review on urban vegetation and particle air pollution - Deposition and dispersion. *Atmospheric Environment*, Vol. 105, pp. 130–37. http://doi.org/10.1016/j.atmosenv.2015.01.052.

Pathak, V., Tripathi, B.D. and Mishra, V.K. (2008), Dynamics of traffic noise in a tropical city Varanasi and its abatement through vegetation. *Environmental Monitoring and Assessment*, Vol. 146, Nos 1–3, pp. 67–75. http://doi.org/10.1007/s10661-007-0060-1.

R Development Core Team (2016), R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna Austria, ISBN: 3-900051-07-0. Available online at http://www.R-project.org/.

Savard, J.P.L., Clergeau, P. and Mennechez, G. (2000), Biodiversity concepts and urban ecosystems. *Landscape and Urban Planning*, Vol. 48, Nos 3–4, pp. 131–42. http://doi.org/10.1016/S0169-2046(00)00037-2.

Susca, T., Gaffin, S.R. and Dell'Osso, G.R. (2011), Positive effects of vegetation:

Urban heat island and green roofs. Selected Papers from the Conference Urban Environmental Pollution: Overcoming Obstacles to Sustainability and Quality of Life (UEP2010), 20-23 June 2010, Boston, USA, *Environmental* Pollution, Vol. 159, Nos 8–9, pp. 2119–26. http://doi.org/10.1016/j.envpol.2011.03.007.

West, G.F. (2013), Big data needs a big theory to go with it. Retrieved December 20, 2016, from https://www.scientificamerican.com/article/big-data-needs-big-theory/.

Wickham, H. (2014), Tidy Data. *Journal of Statistical Software*, Vol. 59, No. 10, pp. 1–23. http://doi.org/10.18637/jss.v059.i10.

BIOGRAPHY

Joao Pinelo Silva is an assistant professor of architecture, with interests in the implications of design on the use of buildings and cities. His research focusses on spatial cognition and behaviour of people in the built environment. He follows a science-based and human-focussed approach, and has developed software tools for the analysis of behaviour within spatial frameworks. Joao is a registered architect, who earned a PhD from University College London, and is Evidence-based-design Certified (EDAC USA).