#51

ON THE CORRELATION OF PEDESTRIAN FLOWS TO URBAN ENVIRONMENT MEASURES:

A Space Syntax and Walkability Analysis comparison case

PAULO CAMBRA
CERIS, CESUR, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001, Lisboa, Portugal
paulo.cambra@tecnico.ulisboa.pt

FILIPE MOURA
CERIS, CESUR, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001, Lisboa, Portugal
fmoura@tecnico.ulisboa.pt

ALEXANDRE GONÇALVES
CERIS, CESUR, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001, Lisboa, Portugal
alexandre.goncalves@tecnico.ulisboa.pt

ABSTRACT

Today, urban planners face the challenge of making cities more walkable. A range of methodologies has been developed for "walkability" assessment, a wide range of actors are involved in the evaluation of relations between urban environment and pedestrian behaviour, and different definitions on how to measure this relation arise. Space Syntax (SSyntax) research has been a key contributor to the study of the relation between urban environment and pedestrian movement providing solid evidence on their correlation. International SSyntax research has focused on the development and comparison of built environment indicators and models, whilst the methods used for pedestrian data collection and processing used to validate such models, have somehow been overlooked. This study puts forward operational issues regarding pedestrian counts and their correlation to urban environment measures. A series of pedestrian counts was performed in a sample of 60 streets in Lisbon (approx.10.000 counts) to validate the results of two built environment assessment methods – space syntax and walkability analysis. A very high degree of spatial-temporal variation in pedestrian flows and a very significant proportion of people sojourning was observed.

Addressing consistency, this paper presents and discusses a set of tests on the correlation strength between the pedestrian data and built environment measures. Findings point out the influence of the aggregation of pedestrian data on the correlation coefficients, suggesting that data "taming" using average values of the pedestrian flows may have implications on the interpretation of results. The results also suggest that future research on the integration of SSyntax and Walkability approaches may improve the explanatory power of pedestrian behaviour.

KEYWORDS

Pedestrian movement; Analytics; Urban metrics; Walkability; Pedestrian behaviour

INTRODUCTION

Walking is gaining growing attention as a key factor in the promotion of healthier, environmental friendly and socially active communities. In the past decades, researchers from various science fields, namely architecture, urban planning, transportation and public health have been developing tools and measures to provide an objective answer to the question "how walkable is my street/neighbourhood/city?".

Space Syntax (henceforth SSyntax) concerns the relation between the built environment configuration and urban dynamics. It is a concept built upon the works of Hillier and Hanson, developed in the 1970s at the University College London, establishing relations for the understanding of the impact of urban morphology on people's movement. (Koohsari, Owen, Cerin, Giles-corti, & Sugiyama, 2016) (Pereira, Holanda, Medeiros, & Barros, 2012). Within SSyntax concept, streets can be represented as axial lines, corresponding to straight lines of sight, whose topological relations form the base of an extensive set of indicators.

Likewise, the influence of the urban built environment on pedestrian travel behaviour has been the focus of walkability assessment research. Many environmental factors have been associated with walking behaviour but agreement on the relative importance of each factor is still contentious. These factors are broadly related to street connectivity, accessibility, land use, urban design and facilities. Numerous tools and methods have been put forward to measure walkability, including audit tools, checklists, inventories, level-of-service scales, surveys, questionnaires and indices.

Albeit of different nature, Space Syntax and Walkability tools share common grounds in attempting to understand and explain pedestrian behaviour and movement. Both methods address the influence of the urban morphology in walking. SSyntax regards connectivity as a central piece of its rationale whilst a walkability tool considers it as one of the key factors amongst others.

In walkability analysis the network connectivity plays an important role but so does convenience (of reaching different land uses), comfort (of the pedestrian infrastructure), conviviality (chances of meeting and socializing), conspicuousness (finding the way around), coexistence (with other transport modes) and policy level commitment (Moura, Cambra, & Gonçalves, 2017).

At the operational level SSyntax analysis are somewhat more straightforward to perform as they use cartographic data as main input whilst walkability analysis demand more extensive datasets with attributes related to environmental features and land uses that, in many cases, need to be obtained by means of street auditing. The approximation of both approaches has been suggested by Cambra (2012) and more recently by Koohsari et al. (2016).

In order to trust the outcomes of these tools as accurate estimators of pedestrian movement, some kind of validation should be performed. Various methods can be applied, namely travel surveys, household surveys, interviews, questionnaires and pedestrian counts.

Within the transportation and urban planning research, the use of pedestrian counts for validation of urban design measures has been addressed by Ewing and Connors (2013) but in practical terms, such relation has been more extensively addressed within public space research. The work of Gehl and Svarre (2013) provides a comprehensive guide to various manual and automated registration methods, namely counting (often registering the pedestrian flow), mapping (more oriented to record the locations where people stay), tracing (registering movement patterns), and tracking (to address walking speeds and route events). Although more sophisticated and automated methods have been recently developed and tested, such as real time video counters or public access image databases the most used method to obtain pedestrian flows is still the "gate" method in which a standing observer counts the pedestrians that cross an imaginary line in front of him (Vaughan & Grajewski, 2001). Using this straightforward method, the Space Syntax research community has been a significant contributor to the understanding of the relation between built environment attributes and pedestrian volumes, providing consistent evidence of a positive and significant relation between the network connectivity and the pedestrian activity (B. Hillier, Penn, Hanson, Grajewski, & Xu, 1993)(Hillier & Iida, 2005)(Kalakou & Moura, 2014) (Hajrasouliha, 2015).

Space Syntax theory suggests that the amount of movement occurring on each street is influenced by its configuration and by the relation that each street establishes with the other streets in the urban system. The spatial distribution of flows in the system is therefore essentially morphological and topological, being a functional result of the urban configuration (Pereira et al., 2012).

The seminal work of Hillier et al. (1993) analysed the relation between the configuration of the King's Cross area in London and the observed pedestrian movement. A positive and significant degree of correlation was found between the pedestrian flow and SSyntax indicators (connectivity, integration, control and global choice), of which "integration" was found to show the strongest association. Using simple linear regression between pedestrian movement rates and integration, a coefficient of determination (R2) of 0.238 was obtained. A higher R2 was obtained (0.547) when using the natural logarithm of pedestrian movement rates. The degree of correlation also increased when each of the 10 subareas of the global study area was analysed individually, ranging from 0.617 to 0.798, meaning that almost 60 to 80% of the variation in the pedestrian movement rates could be explained by configurational aspects only.

Similar tests in different places have reached similar results in the demonstration of positive and significant associations between pedestrian movement and urban configuration, being the explained variation of pedestrian flow consistently in the range of 60-70% (Lerman, Rofè, & Omer, 2014).

At this stage it is worth noticing that, albeit pedestrian movement has been the key for the calculation of correlation indices by the SSyntax approach, sojourning pedestrians (people sitting, standing, socializing) can be a significant part of the pedestrian activity on the streets. Hence we can pose the question of how the degree of correlation would be affected if both moving and static pedestrians would be considered in pedestrian counts. In fact, the observations of Hillier et al (1993) reported "encounter rates" with moving and static pedestrians but only movement rates were used in the study. Another interesting take drawn from this work was the use of a "moving observer" technique in order to capture movement and static pedestrians. An adaptation of the "moving observer" technique, which we used for collecting the pedestrian counts, is described in the next section.

In this sense, adding the walkability perspective, it is expected that a friendlier walking environment, where connectivity plays a major role (SSyntax measures) but where attractiveness factors are also relevant (walkability analysis), will be associated to higher pedestrian activity (moving and sojourning pedestrians).

A series of pedestrian counts was performed in order to validate the walkability assessment tool developed by Moura et al (2017) and a series of consistency issues relating to pedestrian counting and correlation analysis were identified. This paper focus on the consistency issues identified in collecting and processing the pedestrian counts, and in its potential implications in calculating correlation indices. In addition, by performing SSyntax analysis to the studied area, we were able to compare the correlation between pedestrian volumes, SSyntax measures and Walkability scores.

2. BUILT ENVIRONMENT MEASURES: SPACE SYNTAX AND WALKABILITY SCORES

We used two methods for measuring the built environment in relation to pedestrian behaviour: the SSyntax methodology and a walkability assessment framework.

The studied areas were located in Lisbon, Portugal Arroios (Study Area 1) and Avenidas Novas, (Study Area 2), both located in the city central area.

Study Area 1, Arroios: It is a mixed use area with predominant residential occupation formed by 4-6 story buildings. Its urban development was initiated in the early 20th century, having registered a major progress within the years 1919-1945. It is composed by a mix of regular and irregular street pattern of one to two lane road spaces. It is a hilly setup with a noticeable lack of open spaces and green areas.

Study Area 2, Avenidas Novas: Resulted from a planned urban intervention in the late 19th century. It is a mixed use area formed by medium rise buildings with a significant office occupation. The street design incorporated the French boulevard style, being characterized by a regular grid of large avenues with tree alignments in the medians.

Figure 1 - Lisbon Axial map showing Integration (HH) and study area locations

2.1 WALKABILITY SCORES

A number of walkability assessment methodologies have been developed for application in urban planning, used by planning professionals and policymakers to understand the scope and extent of local pedestrian conditions as well as to identify possible interventions for its promotion. The walkability of the environment was assessed by a tool, previously developed by the research team. This walkability framework was named IAAPE -Indicators of Accessibility and Attractiveness of the Pedestrian Environment (www.iaape.org). The IAAPE framework stands out from the majority of similar tools due to its participatory nature, by involving the main stakeholders into the selection and ranking of the indicators that structure the backbone of the pedestrian environment assessment. It also allows the measurement to meet distinct pedestrian groups and trip purposes. Similarly to other walkability tools, the IAAPE framework assesses the pedestrian environment by performing street audits. A set of indicators related to 7 key dimensions (Connectivity; Convenience; Comfort; Conviviality; Conspicuousness; Coexistence; Commitment) is used to score the street environment qualities. Data is collected and stored in a GIS platform where a pedestrian network was previously built, representing the extent of the pedestrian realm (sidewalks, crossings, footpaths and corresponding attributes – for instance, curbed sidewalks for inclusive accessibility) in a more accurate way than commonly used road centreline networks. The digital pedestrian network is also used to compute the Connectivity indicator by using the Network Analyst tool of ArcMap © GIS software. The Connectivity indicator uses the formulation of Cambra (2012) consisting in the ratio of the network distance and the straight-line distance between specific origins and destinations.

The IAAPE tool uses a simple multicriteria compensatory model, where the walkability score of a street segment (link) is obtained by adding the 7 C's indicators multiplied by their relative weight, being represented on a 0 to 100 scale. For a more comprehensive description on the framework refer to Moura et al. (2017).

The IAAPE framework was used to assess the walkability of Arroios (Study Area 1) and Avenidas Novas, (Study Area 2), providing walkability scores for approximately 800 street segments. Comparing the two areas in terms of aggregated walkability scores (classified in 5 categories from A do E) significant differences are found: In Arroios, the majority (67%) of the pedestrian network falls in an average classification (C: score between 40 and 60), with 10% of the network scoring below 40 (categories D and E). Only 22% of the network is in category B, whilst streets with high walkability scores were not present in the area. In the other study area, Avenidas Novas, the differences in the walking environment is quite noticeable. The majority (70%) of the pedestrian network scores B and there are more streets of higher walkability (2% in category A) than there are of lower walkability (less than 1% in categories D and E). The remainder 27% score C.

2.2 SPACE SYNTAX MEASURES

For the SSyntax measures for the study area we used the axial map of the city of Lisbon (previously developed by the SSyntax research group at Instituto Superior Tecnico – University of Lisbon) and Depthmap software v.10. We obtained 3 common spatial measures: connectivity, global integration and local integration (R=3). The connectivity index, which reflects the number of axial lines directly intersecting each line in a network, was much higher in Avenidas Novas (10,42) than in Arroios (5,97), reflecting its grid pattern. The integration index, which reflects the average topological distance, that is to say the number of direction changes from each line to all the other lines in a network, was relatively similar in both areas (Arroios=0,56; Avenidas Novas=0,65) relating to their central position in the system.

The local integration index (with a radius of 3)which reflects the topological proximity of a line to its nearby axial lines (in this case two turns away from a given axial line) was also relatively similar and slightly higher in Avenidas Novas (2,86) than Arroios (2,86). Both areas show some of the highest values for local integration found in the city of Lisbon.

3. PEDESTRIAN DATA

3.1 COUNTING METHOD AND ACCURACY ISSUES

Two observation methods were considered two collect pedestrian data to validate the built environment measures: the gate method and the moving observer method. For the validation of the built environment measures we wanted to address different travel purposes of walking. Considering utilitarian travel to be more related to movement and social/recreational travel to be related to sojourning, we ought to collect data on moving and static pedestrians. According to the Space Syntax Observation Manual (Vaughan & Grajewski, 2001), "the gate method is the workhorse of spatial observing techniques ". It is a simple and widely adopted method consisting of observing the people (or vehicles) passing through an imaginary screen line crossing the street space at a right angle, being suitable for recording moving people but not for stationary people.

The moving observer method, on the other hand, is suitable to record moving and static pedestrians, having been used with slight variations in SSyntax research (B. Hillier et al., 1993) (Choi & Koch, 2015). In this method the observer walks at a regular pace along the street, counting people in movement (in the same direction of the observer, opposite direction or both) and static people (seating, standing).

In order to increase accuracy and reliability, the moving observer method demands several passes in each way. Sample size was determined on preliminary observation of the pilot study whereby every street segment of 4 blocks was analysed in the same area, with 4 counts each. Based on the results (mean; variance) 30 observations per street segment was found to bind the standard error to a 10% range relative to the mean, attending to the high variance of pedestrian flows. The number of needed observations is in line with Hillier's method, which required routes to be observed "between twenty to thirty times".

Given the high variance of pedestrian flows, even a large number of observations per street (30) did not assure a standard error of 10% for each street of our study area. Figure 2 illustrates the error in the relation to pedestrian count results, showing that for pedestrian flows lower than 80 people per 15 minutes the standard error tends to be higher than 10% and higher than 20% for a pedestrian flow lower than 20 people/15 mins.

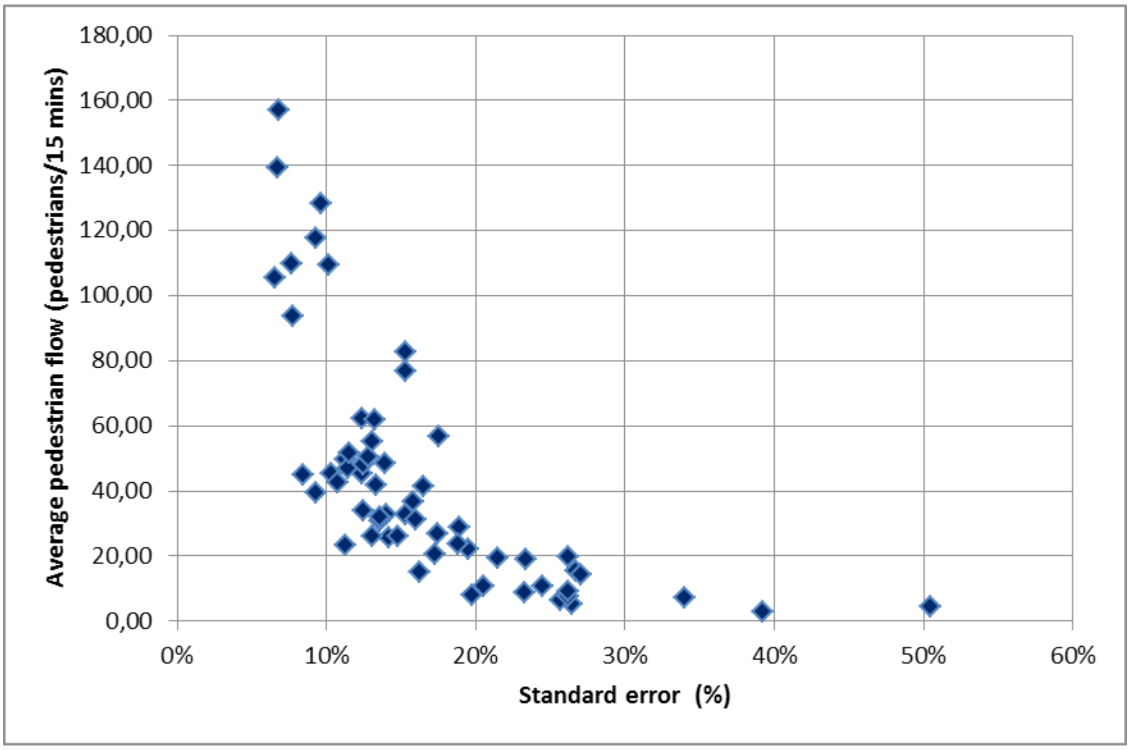

Figure 2: Observation error related to pedestrian volume

The location of the counting spots is another critical subject. The Space Syntax Observation Manual recommends a minimum of 25 "gate" positions placed in a manner to cover a range of pedestrian volumes ("well- used, moderately used and poorly used spaces in and around the area of study"). To capture the area's environmental variety thus covering a full range of pedestrian volumes, the pedestrian network would have to be typified prior to performing the sample selection. In order to obtain a somewhat more refined street typology set, the option was to group streets according to their attributes, in terms of the 7 key dimensions, and their global quality, in terms of their walkability score. Using the statistical software SPPS "n-cluster analysis" we obtained 4 typology clusters comparable to a "street hierarchy". A sample size of 60 street segments (approx.5% of the total) was considered sufficient to ensure the representation of each "hierarchy". The counting locations were randomly selected for each street typology at first, but, as we used the moving observer method, the physical effort needed to perform the observations had to be considered (street length, street to street distances and slope). Given that all the observations made by a single person were required to be made within a 90 minute time period it was necessary to refine the counting locations by manual screening.

The Arroios study area was divided in 6 sections (36 streets to be observed) and Avenidas Novas study area was divided in 4 sections (24 streets), totalizing 60 streets to be observed. The team of counting auditors consisted of 10 students from the University of Lisbon who received training prior to the field work, and were supervised by a coordinator on the field.

In short, the counting procedure consisted on the simultaneous counting of a sample of 60 street segments, by a team of 10 auditors, during 5 week days plus a Saturday. Each street segment was observed 6 consecutive times in 5 different time periods (3 time periods on the Saturday). Concerning observation periods, each street segment would be observed at 5 distinct daily periods consisting of the morning peak (8:00-9:30 AM); morning off-peak (10:00-11:30 AM) lunch peak (12:30-14:00 PM); afternoon off-peak (15:00-16:30 PM) and afternoon peak (17:00-18:30 PM). This method would provide 10.080 distinct pedestrian counts (60 streets*6 counts*5 periods*5 weekdays + 60 streets*6 counts *3 periods*1 Saturday). The survey was performed from May 7th to 13th, 2015, (Thursday to Wednesday, excluding Sunday). The weather conditions were stable and consistent to the time of the year - sunny days with average temperatures of 20º C, with the maximum temperature reaching 30º C in the last 3 days. The location of the streets where the counting took place is provided in Figure 3.

Figure 3: Street counting locations in Area 1 – Arroios and Area 2 – Avenidas Novas

After performing data consistency tests, we obtained a total of 9.136 valid counts from the 10.080 collected observations, resulting in 4.568 count values (1 count value is made of 2 way and return observations). One of the reasons dealt with human error related to accuracy when recording the travel times. Another reason dealt with human fatigue. Although the observations circuits were designed to minimize walking distances there were cases where the counting auditor was not able to complete the observation circuits within the planned time limits, especially during the warmer periods. The counting auditors walked an average of 20km per day. The results for the built environment measures and pedestrian counts for each study areas are presented in Tables 1 and 2.

Street Ref.	Walkability Score	Average Moving Flow (ped/15 mins)	Average proportion of sojourning people (%)	Connectivity	Integration [HH]	Integration [HH] R3
10700	41,92	32,62	10,53	6	0,59	1,98
10713	60,52	31,08	28,95	4	0,59	2,04
10716	68,41	47,02	17,04	9	0,62	2,52
10725	37,40	6,44	9,46	3	0,56	1,81
10733	49,58	30,70	7,43	9	0,53	2,36
10793	48,13	82,80	16,18	12	0,53	2,75
10806	39,54	4,29	3,16	3	0,51	1,88
10861	52,41	7,00	16,26	3	0,56	1,73
10866	39,42	26,96	5,28	4	0,56	2,21
10904	64,31	33,75	5,72	10	0,62	2,63
10914	45,81	48,11	12,22	4	0,66	2,82
10923	41,21	10,85	3,45	6	0,62	2,23
10924	59,71	50,39	18,65	8	0,66	2,99
10976	69,81	55,09	45,47	4	0,66	2,77
10978	74,69	109,76	34,33	9	0,62	2,52
11021	18,99	7,67	12,64	3	0,55	1,39
11042	48,84	9,13	9,15	2	0,56	1,50
11056	70,38	105,45	26,63	33	0,70	3,87
11092	34,01	20,66	22,48	4	0,65	2,74
11093	59,90	19,74	18,24	5	0,63	2,52
11109	45,89	41,65	10,36	3	0,65	2,67
11119	49,28	32,14	42,00	6	0,62	2,33
11130	58,01	48,47	48,24	6	0,63	2,45
11347	70,27	51,51	13,03	3	0,56	1,73
11395	47,50	93,86	24,21	2	0,51	1,37
11404	25,00	23,76	49,07	4	0,48	1,60
11415	51,12	42,55	15,83	7	0,49	2,52
11443	25,06	2,81	4,29	4	0,45	1,60
11479	33,82	14,22	3,54	9	0,47	2,56
11511	34,61	8,80	14,74	3	0,45	1,52
11524	33,30	7,88	14,38	3	0,45	1,80
11532	34,61	10,67	4,84	5	0,47	1,86
11549	41,39	15,04	28,70	9	0,47	2,52
11651	74,51	157,23	26,80	5	0,46	2,10
11652	41,29	19,43	19,86	4	0,45	1,86
11913	71,42	23,40	29,01	1	0,51	1,02
Average	**48,95**	**37,03**	**18,67**	**5,97**	**0,56**	**2,19**

Table 1 - Built environmental measures and pedestrian counts per street segment, Area 1 – Arroios

Street Ref.	Walkability Score	Average Moving Flow (ped/15 mins)	Average proportion of sojourning people (%)	Connectivity	Integration [HH]	Integration [HH] R3
1003	53,39	45,42	23,60	9	0,65	3,04
1006	80,14	117,71	47,39	16	0,66	3,30
1021	47,39	76,66	18,90	27	0,69	3,70
1026	74,63	139,35	18,68	17	0,68	3,29
1063	80,44	45,27	23,06	6	0,65	2,89
1139	59,49	25,49	43,93	4	0,62	2,47
1919	70,20	109,25	15,72	17	0,68	3,29
2071	59,44	41,40	18,44	8	0,64	2,94
2075	79,44	45,13	21,64	21	0,68	3,39
2623	80,42	49,68	33,05	5	0,64	2,59
2674	40,24	36,51	23,51	4	0,65	2,55
2683	50,29	26,02	40,46	10	0,66	3,08
2708	44,89	22,18	28,49	4	0,62	2,11
2725	74,18	56,71	27,25	11	0,67	2,99
2765	75,26	19,12	20,08	10	0,66	2,92
2822	79,14	128,34	37,09	17	0,69	3,48
2864	70,88	28,88	19,09	5	0,65	2,67
2874	74,05	62,42	47,03	9	0,65	2,66
2890	59,88	61,98	39,84	21	0,68	3,39
2924	60,92	26,02	45,32	5	0,64	2,36
2945	64,92	5,27	18,84	3	0,63	2,17
3021	65,37	32,79	26,56	4	0,64	2,14
3258	51,55	39,42	13,20	11	0,67	2,99
3649	51,68	15,38	1,63	6	0,60	2,27
Average	64,51	52,35	27,20	10,42	0,65	2,86

Table 2 - Built environmental measures and pedestrian counts per street segment, Area 2 – Avenidas Novas

3.2 SPATIAL-TEMPORAL VARIABILITY

The above tables underline the spatial variability of pedestrian flows, as very different movement patterns are found not only in streets located within the same study area but in different segments along the same street. Temporal variability is easily noticeable throughout the different day periods.

The spatial-temporal variability of pedestrian flows results in a high volatility of pedestrian counts at individual locations, which can be an issue when performing correlations between pedestrian volumes and built environmental measures. This issue has been addressed within SSyntax research by Park (2013), who analyzed an extensive pedestrian count database of about 10.000 locations in Seoul. Park found that 29% of the total variations in pedestrian counts were temporal, independently of count locations with the remaining 71% of the variation being spatial, across different counting locations, independently of count occasions. Park also noted the strong skewness of the distribution of pedestrian counts over time. In order to obtain a clear pattern of pedestrian volumes, a log transformation was considered necessary but not sufficient to "tame" the distribution. Only when using average values a daily routine pattern of pedestrian movement was noticeable.

The spatial-temporal variability and the "taming" of pedestrian counts are illustrated in figures 4 and 5. Figure 4 shows the recorded pedestrian flows for one of the sampled streets (Ref.1006 in study area 2) during a week. It can be seen that each observed day shows a different behavior and no discernible pattern can be found. When these values are averaged (Figure 5) it is possible to notice a daily routine where most people are found walking during the lunch hour peak. This observation makes sense within the area's land use context: there are many offices and people leave the offices to eat out during lunch time. The pedestrian volume at the morning and afternoon peak hours is lower than the one at lunch time probably because many office buildings have their own car parking so people do not walk to access their workplace. Of course, it can be questioned is if the averaged volume value can be considered a "typical" daily pattern as it is not observable in more than 1 of the 5 week days. Likewise, the flow pattern of a street attained from the values observed in its constituent parts (segments) is very likely an amalgamation of dissimilar travel patterns that do not follow a particular trend unless all segments' observations are combined and averaged values.

Analyzing the 60 sampled streets and respective 4.568 counts, we could conclude that using average values allowed having a clearer notion of the flow patterns but disregards the intrinsic and characteristic spatial-temporal variation of pedestrian movement, possibly inducing bias in the correlation analysis. Location specific spatial-temporal variation of the pedestrian volumes should be taken into consideration when designing the pedestrian counting procedure: where to count, for how long and how many different days to count.

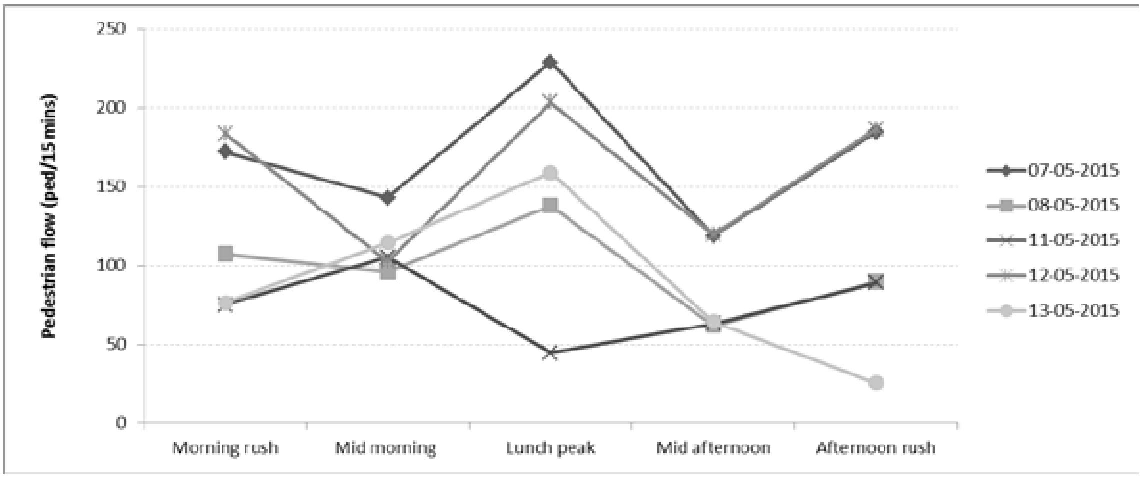

Figure 4 - Pedestrian flow of street segment 1006 over a week

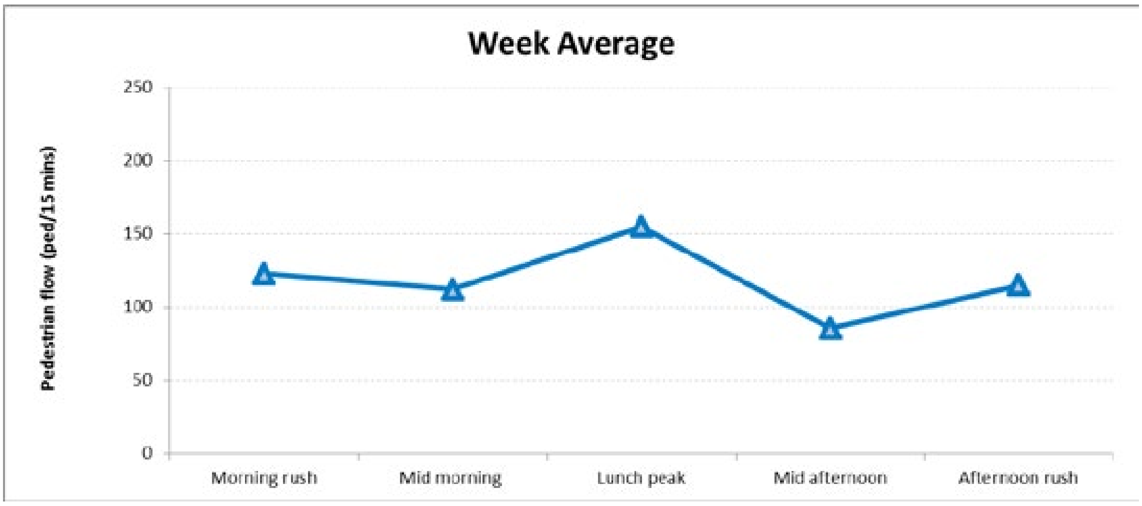

Figure 5 - Average pedestrian flow of street segment 1006 over a week

4. CORRELATING BUILT ENVIRONMENT MEASURES AND PEDESTRIAN DATA

Given the high spatial-temporal variability of pedestrian flows and the significant proportion of sojourning pedestrians (not moving) we run a series of tests, correlating the pedestrian data with built environment measures (SSynt ax indices and Walkability scores).

In the first test, we addressed spatial variability. We used the pedestrian data obtained for the sample of 60 streets. In Arroios, 36 streets were observed for a week in a total of 2700 observations (resulting from 5400 individual observations in each direction). In Avenidas Novas, 24 streets were observed for a week in a total of 1382 observations (resulting from 2764 individual observations in each direction). In this test the pedestrian data was not linearized (using a log transformation) nor summarized (using average values). The obtained Pearson Correlation coefficients (r) show significant differences for each area and for the SSyntax and Walkability approaches, being the highest correlation value observed between the pedestrian flow and walkability scores (r=0,461; R2=0.212).

We note three interesting findings here: i) the highest correlation value to SSyntax measures is related to connectivity, not to integration; ii) the correlation values to SSyntax measures are significantly higher in the area that shows a regular grid pattern and high pedestrian volumes; iii) the opposite happens with the correlation values to Walkability scores (higher correlation values in the area that shows an irregular pattern and lower pedestrian volumes)

		Pearson Correlation (r)			
		Connectivity	Integration	Integration R3	Walkability Score
Flow (pedestrians/15 min)	Arroios (N=2700)	,321**	,180**	,305**	,484**
	Avenidas Novas (N=1382)	,515**	,506**	,533**	,342**
	Global (week days, N=4082)	,430**	,268**	,404**	,461**

**. Correlation is significant at the 0.01 level (2-tailed).

Table 3 - Spatial variability test (per study area)

For the second test we aggregated the pedestrian counts of both areas by counting period in order to address temporal variability during the day. We can observe that the correlation coefficient varies along the day, being highest at the lunch hour peak period and that the temporal variation is lower than spatial variation of the previous test (much in line with Park results for Seoul (Park, 2013). Again, the highest correlation coefficient is obtained for the walkability scores (r=0,525; R2=0.275).

		Pearson Correlation (r)			
		Connectivity	Integration	Integration R3	Walkability Score
Flow (pedestrians/15 min)	P1 (08:00-09:30; N=826)	,419**	,254**	,371**	,416**
	P2 (10:00-11:30; N=822)	,389**	,226**	,386**	,449**
	P3 (12:30-14:00; N=809)	,471**	**,335****	**,473****	**,525****
	P4 (15:00-16:30; N=813)	**,478****	,279**	,422**	,477**
	P5 (17:00-18:30; N=812)	,420**	,249**	,382**	,452**
	All Periods (week days; N=4082)	,430**	,268**	,404**	,461**

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4 - Temporal variability test (per counting period)

The third test was similar than the previous one, with the difference that the temporal aggregation was made by counting day, including the Saturday. We note two interesting findings here: i) pedestrian flow correlations vary significantly along the week days (more than 30% in the case of Integration) showing that picking a "typical day" to perform pedestrian counts may introduce distortion the correlation analysis results;. ii) It is noticeable the change in r between any week day and the Saturday. However, the walkability scores' r remains more consistent over the week (less than 15% variation) than the SSyntax's r (more than 40% variation).

		Pearson Correlation (r)			
		Connectivity	Integration	Integration R3	Walkability Score
Flow (pedestrians /15 min)	Monday (N=813)	**,497****	**,325****	**,447****	,466**
	Tuesday (N=810)	,394**	,223**	,367**	,489**
	Wednesday (N=810)	,394**	,212**	,375**	,423**
	Thursday (N=839)	,413**	,264**	,405**	,451**
	Friday (N=810)	,443**	,313**	,417**	,479**
	Saturday (N=486)	**,292****	**,123****	**,234****	**,422****
	Week days (N=4082)	,430**	,268**	,404**	,461**
	All days (N=4568)	,415**	,254**	,388**	,456**

**. Correlation is significant at the 0.01 level (2-tailed).

Table 5 - Temporal variability test (per day of the week)

Next, we addressed pedestrian activity in the streets – that is moving and sojourning people - in comparison to pedestrian flows. In the following tests we applied a natural log transformation to the pedestrian flow and analyzed each study area independently. The observations for all days were used (week days and Saturday).

Area 1 - Arroios				
N = 3.024; 36 streets	Connectivity	Integration	Integration R3	Walkability Score
Pedestrian flow (per 15 min)	,322**	,177**	,300**	,491**
Ln (PedFlow15min)	,292**	,202**	,307**	,456**
PedActivity (moving+sojourning)	,344**	,129**	,229**	,527**
Area 2 – Avenidas Novas				
N = 1.544; 24 streets	Connectivity	Integration	Integration R3	Walkability Score
Pedestrian flow (per 15 min)	,488**	,489**	,515**	,334**
Ln (PedFlow15min)	,475**	,465**	,510**	,298**
PedActivity (moving+sojourning)	,355**	,305**	,393**	,358**
Aggregated areas				
N = 4.568; 60 streets	Connectivity	Integration	Integration R3	Walkability Score
Pedestrian flow (per 15 min)	,415**	,254**	,388**	,456**
Ln (PedFlow15min)	,377**	,236**	,368**	,404**
PedActivity (moving+sojourning)	,409**	,274**	,377**	,496**

**. Correlation is significant at the 0.01 level (2-tailed).

Contrary to the findings of other studies, using the natural log transformation did not result in an improvement of the correlation coefficient for the SSyntax measures (an exception was found in Arroios area). But, when using the full week dataset and the Ln transformation for each area independently then the local integration (R3) becomes the best spatial predictor for pedestrian movement.

When considering moving and static pedestrians –pedestrian activity- a lower r is found for SSyntax measures, whilst a higher r is observed regarding walkability scores. At this stage we can consider SSyntax to be better suited for explaining movement than to explain a broader street use, which includes moving, standing, sitting and socializing pedestrians, translated in the concept of pedestrian activity.

Overall the best fit to pedestrian counts is obtained by correlating pedestrian activity and walkability scores (r=0.496; R2=0.246); which is quite lower than the results found in SSyntax research studies, and, in statistical terms, considered a weak correlation . The low explanatory power of the built environment measures (either SSyntax or Walkability scores) in relation to pedestrian volumes was striking, especially when other studies point to a coefficient of determination (R2) higher than 60%, meaning that 60% of the variation of pedestrian volumes could be explained by the spatial/environmental attributes of the urban space. In our case, no more than 25% of the pedestrian volume variation could be explained.

Such weak correlation is linked to the high volatility of pedestrian flows. Recalling the previous section of this paper, the pedestrian flows observed on a single street segment do not allow distinguishing a clear pattern, but, when the observation values are summarized in average values, patterns become discernible. On the other hand, when using an average value we are not only discarding relevant information about variation but also reducing the sample size, thus lessening statistical significance.

We tested using the average values of the pedestrian flow for each street. This means all the 4.568 observations were summarized into 60 records prior to the correlation analysis.

The findings were considerably different: first, the best spatial predictor for pedestrian movement was now Local Integration, for both study areas; second, using the Ln transformation increased the r value to figures closer to similar studies (namely in Area 2, where the local integration r=0.737); third, the correlation coefficient of SSyntax measures and walkability scores was now very similar at the aggregated area analysis (Local integration r=0.619; Walkability score r =0.634) and fourth, some of the correlations for the individual areas were not statistically significant.

According to these values, the power of the built environment measures to explain the (average) pedestrian flow variation in the aggregated study area was now between 38% -40%, which is a significant improvement compared to the previous model specifications.

Area 1: Arroios				
N = 36; 36 streets	Connectivity	Integration	Integration R3	Walkability Score
Pedestrian flow (per 15 min)	,422*	,232	,394*	,644**
Ln (PedFlow15min)	,402*	,419*	**,513***	**,693****
Area 2: Avenidas Novas				
N = 36; 36 streets	Connectivity	Integration	Integration R3	Walkability Score
Pedestrian flow (per 15 min)	,660**	,671**	,696**	,428*
Ln (PedFlow15min)	,660**	,684**	**,737****	,358
Aggregated areas				
N = 60; 60 streets	Connectivity	Integration	Integration R3	Walkability Score
Pedestrian flow (per 15 min)	,560**	,346**	,524**	,582**
Ln (PedFlow15min)	,537**	,502**	**,619****	**,634****

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 7 - Correlation coefficients of average pedestrian flow values in Area 1

5. CONCLUSIONS

In summary, our tests showed that the spatial-temporal variations inherent of pedestrian flows can induce relevant changes in the correlation analysis between pedestrian volumes and built environment measures.

Also, the use of average values can introduce potential bias in the correlation analysis. This is a very pertinent issue when realizing that within SSyntax research there is a considerable body of evidence built upon the correlation of spatial predictors to pedestrian volume. The omission or the unclear description of the pedestrian counting procedures at the observation stage and at the processing stage may result in unfeasible comparisons between models and unfeasible transferability of findings.

The introduction of a potential bias in the result is illustrated in Table 10 which summarizes the discussed correlation analysis and shows how the correlation coefficients vary according to the pedestrian dataset used. Looking at the Local Integration correlation values to the log of pedestrian flow it can be seen that the use of "raw" data results in r=0.368, which is considered a weak correlation whereas the use of "tamed" data results in r=0.619, which is already a moderate to strong correlation. In conclusion we can admit the use of both "raw" and "tamed" pedestrian data but this should be a conscious option.

			Pearson Correlation (aggregated area)			
Counting period	Pedestrian volume variable	N observations	Connectivity	Integration	Local Integration (R3)	Walkability Score
Week days (Mon:Fri)	Pedestrian flow / 15 mins	4082	0,430	0,268	0,404	0,461
Week days + Saturday	Pedestrian flow / 15 mins	4568	0,415	0,254	**0,388**	0,456
	LN (pedestrian flow / 15 mins)	4568	0,377	0,236	0,368	0,404
	Pedestrian activity (moving+static)	4568	0,409	0,274	0,377	0,496
	Average pedestrian flow / 15 mins	60	0,560	0,346	0,524	0,582
	LN (average pedestrian flow / 15 mins)	60	0,537	0,502	**0,619**	0,634

			Coefficient of determination (R2)			
Counting period	Pedestrian volume variable	N observations	Connectivity	Integration	Local Integration (R3)	Walkability Score
Week days (Mon:Fri)	Pedestrian flow / 15 mins	4082	18,5%	7,2%	16,3%	21,3%
Week days + Saturday	Pedestrian flow / 15 mins	4568	17,2%	6,5%	15,1%	20,8%
	LN (pedestrian flow / 15 mins)	4568	14,2%	5,6%	13,5%	16,3%
	Pedestrian activity (moving+static)	4568	16,7%	7,5%	14,2%	24,6%
	Average pedestrian flow / 15 mins	60	31,4%	12,0%	27,5%	33,9%
	LN (average pedestrian flow / 15 mins)	60	28,8%	25,2%	**38,3%**	**40,2%**

Table 8 - Summary of the correlation analysis: Pearson correlation

The highest explanatory power of the built environment measures to explain the pedestrian flow variation was found to be in the 40% range, being very similar between the SSyntax approach (38,3%) and the walkability approach (40,2%).

The walkability score approach was found to deliver more consistent correlates over the issue of spatial-temporal variability and over pedestrian travel behavior, dealing with moving and sojourning pedestrians. As walking can be associated to different travel purposes – utilitarian/social/recreation – and different travel purposes can be related to distinct activity patterns, then addressing a full range of pedestrian activity patterns would be of interest of research on the built environment, physical activity and health.

The walkability approach is also more resource intensive. It requires collecting and processing extensive attribute datasets, relying many times in qualitative analyses that are prone to a high degree of subjectivity. From the application of both methods (SSyntax and Walkability) in the same study areas we found a minor difference (2%) between the respective coefficients of determination. Hence we may pose the question – is it worth it?

Given the major strengths of each side – the pragmatism and the decades of collaborative international research of SSyntax; and the robustness and consistency of walkability analysis – further research and experimentation is needed on the integration between SSyntax and Walkability measures.

REFERENCES

Cambra, P. (2012). Pedestrian Accessibility and Attractiveness Indicators for Walkability Assessment. Thesis for the Master Degree (MSc) in Urban Studies and Territorial Management. Retrieved from https://fenix.tecnico.ulisboa.pt/downloadFile/2589873355564/Dissertacao.pdf

Ewing, R., & Connors, M. (2013). Validating Urban Design Measures. In Transportation Research Board 92nd annual meeting (pp. 1–18).

Gehl, J., Svarre, B., Press, I., & Steenhard, K. A. (2013). How to study public life. Washington: Island Press. http://doi.org/10.5822/978-1-61091-525-0

Hajrasouliha, A. (2015). The impact of street network connectivity on pedestrian volume. Urban Studies, 52(13), 2483–2497. http://doi.org/10.1177/0042098014544763

Hillier, B., & Iida, S. (2005). Network and psychological effects: a theory of urban movement. In Spatial information theory (pp. 475–490). Springer.

Hillier, B., Penn, A., Hanson, J., Grajewski, T., & Xu, J. (1993). Natural movement: or, configuration and attraction in urban pedestrian movement. Environment and Planning B: Planning and Design, 20, 29–66.

Kalakou, S., & Moura, F. (2014). Bridging the Gap in Planning Indoor Pedestrian Facilities. Transport Reviews, 34(4), 474–500. http://doi.org/10.1080/01441647.2014.915441

Koohsari, M. J., Owen, N., Cerin, E., Giles-corti, B., & Sugiyama, T. (2016). Walkability and walking for transport : characterizing the built environment using space syntax. International Journal of Behavioral Nutrition and Physical Activity, 1–9. http://doi.org/10.1186/s12966-016-0448-9

Lerman, Y., Rofè, Y., & Omer, I. (2014). Using Space Syntax to Model Pedestrian Movement in Urban Transportation Planning, 392–410. http://doi.org/10.1111/gean.12063

Moura, F., Cambra, P., & Gonçalves, A. B. (2017). Measuring walkability for distinct pedestrian groups with a participatory assessment method: A case study in Lisbon. Landscape and Urban Planning, 157. http://doi.org/10.1016/j.landurbplan.2016.07.002

Park, H.-T. (2013). The spatio-temporal routine of pedestrian movements in large cities : The case of Seoul. Proceedings of the 10th International Space Syntax Symposium, 1–17.

Pereira, R., Holanda, F., Medeiros, V., & Barros, A. P. (2012). THE USE OF SPACE SYNTAX IN URBAN TRANSPORT ANALYSIS : limits and potentials. Proceedings of the 8th International Space Syntax Symposium, 1–25.

Vaughan, L., & Grajewski, T. (2001). Space syntax observation manual. University College London.