339

Tuning in

Investigating OSMnx RCL model preparation methods for Angular Segment Analysis

LUCY DONEGAN & FELIPE TAVARES DA SILVA

UNIVERSIDADE FEDERAL DA PARAÍBA (UFPB), UNIVERSIDADE FEDERAL DA PARAÍBA (UFPB)

ABSTRACT

Automated street network models' advantages mirror the complex and constantly transforming nature of cities. While Open Street Maps (OSM) increased open Road Centre Lines (RCL) coverage and accuracy, and work for Angular Segment Analysis (ASA), these models are not ready for accurate results and efficient processing. OSM RCL includes roads not open to multi-modal movement and too many segments affecting ASA. The OSMnx python library offers an automated way of filtering RCL multi-modal routes. For representation adjustments, publications mentioned Douglas-Peucker simplification with different thresholds, from which oversimplification cuts corners too short losing urban form accuracy. Newer tools can include other preparation steps with an angular threshold however without documentation or results comparison. As well-functioning OSM RCL preparation remains an unclear issue, this paper contrasts and compares automated RCL ASA preparation methods searching for best suited models. In order to do this, a city OSMnx driveable RCL is prepared with (i) this study GIS-CAD interoperability algorithm (ii) default and adjusted Space Syntax Toolkit (SST) network cleaner; (iii) default and adjusted GRASS QGIS generalization. Steps to prepare RCL are described with resulting system sizes, representations and ASA centrality values. The default SST and GRASS parameters were not ideal; modified parameter and our algorithm reached finer tuned models, of which Grass adjusted seems ideal. Findings indicate automated open steps whose adjusted and rather refined street network model point to a preparation standard that can contribute to updated space syntax city analysis elsewhere.

KEYWORDS

OSMnx, RCL preparation, ASA, GRASS generalize, SST network cleaner

1 INTRODUCTION

As cities are problems in organized complexity (Jacobs, 1992) urban informatics have evolved as means by which to understand the urban environment (Batty, 2021). Advantages for automated models for space syntax urban analysis coincide with the very complex and evolving nature of cities. Automated steps facilitate modelling and model updating, as well as overcome problems of manually drawn models, more prone to different movement permeabilities interpretations (Batty and Rana, 2002). This can be an issue especially as ways in which urban models are drawn and prepared for analysis affect results (Ratti, 2004; Marshall et al., 2018). Understanding the use of well-functioning RCL for space syntax urban analysis is a current challenge (Van Nes and Yamu, 2021, p. 83); this paper describes steps for automated OSM RCL download and preparation, comparing and discussing results for one city street network.

Space Syntax ASA can be performed both with segment maps coming from axial maps or with RCL models (Turner, 2007). Axial maps are manually drawn models consisting of the smallest set of the longest lines that reach and connect all spaces open to movement within a system (Hillier and Hanson, 1984). Axial analysis quantifies accessibilities and centralities for each line focusing on topologic measures and radii. For axial analysis every turn in the street representation system equals one step depth; the more turns needed to access a particular place or line in the system, the more difficult it is to access it, and the less accessible the place is. To transform an axial map into a segment map (for ASA), each axial line is broken at the intersection with other lines, and small stubs are removed.

ASA quantifies accessibilities and centralities for each segment focusing on angular measures and processing radii which can be adjusted metrically. The main centrality measures are integration, or closeness centrality, that relates to potential movement to places, and choice, or betweenness centrality, that relates to potential movement through places (Hillier and Iida, 2005). As ASA centrality measures can yield very different values according to system sizes, normalized measures of integration and choice were proposed to help compare different urban systems (Hillier, Yang and Turner, 2012).

Besides the fact that drawing axial maps for cities and, even more, metropolitan areas, is a monotonous and time-consuming task to undertake (Van Nes and Yamu, 2021), model representation can suffer individual interpretation inconsistencies and can be more prone to human mistake. As the task is time-consuming, it is difficult to maintain updated models in rapidly changing urban environments. Thus, more automated street network representations (Batty & Rana, 2012) are desirable.

Although RCL models can be used to perform Space Syntax ASA (Turner, 2007), these models are not immediately ready to perform this analysis accurately. Issues addressed earlier for using

Road Centre Line models for network analysis mentioned topological errors such as disconnected elements (Dalton, Peponis and Dalton, 2003; Kolovou et al., 2017; Krenz, 2018), and different OSM RCL coverage for different areas (Dhanani et al., 2012). However, more recently OSM RCL has grown in representation coverage and precision worldwide (Minaei, 2020), made viable by collaboration and by developing techniques including georeferencing and remote sensing (Husain and Vaishya, 2018).

Other OSM RCL issues recur. First, OSM RCL includes roads not open to multi-modal movement and has route entities that need to be cleaned for the studied permeability analysis. Second, and more importantly, RCL models' representation following the centre of roads are made up of polylines with too many vertices. The many vertices not necessarily needed for route representation impact negatively on ASA processing and scores, limiting system sizes analysis and comparison of values between different systems. Some earlier studies used RCL model without simplification steps, adjusting ASA values by weighing segment lengths (Turner, 2007; Dhanani et al., 2012). Other studies used some Space Syntax Toolkit functions (Gil et al., 2015) to address OSM topology errors and proposed Douglas-Peucker simplification and remodelling steps to approximate RCL models to axial maps (Kolovou et al., 2017; Krenz, 2018).

Remodelling steps included joining close parallel lines and eliminating roundabouts that have no buildings in the middle. The argument to join parallel lines is that an axial map is not directional (Kolovou et al., 2017; Krenz, 2018). The first OSM RCL road types issue has been tackled with the OSMnx python library (Boeing, 2017), that downloads and filters OSM RCL in an automated manner, rather than obtaining data from Geofabrik website and erasing features by filtering each type classification (Kolovou et al., 2017). However, OSMnx does not tackle the second RCL ASA preparation issue as it does not include steps to clean vertices.

OSMnx was used for space syntax ASA studies in Italy, that was then coupled with Douglas-Peucker RCL simplification via the GRASS (Geographic Resources Analysis Support System) generalisation tool at QGIS (Pezzica, Cutini and Souza, 2019; Altafini and Cutini, 2020). However, Cutini and colleagues did not explain different GRASS simplification calibrations and street network centralities outcomes.

Although modes of handling OSM RCL for ASA have been presented - including cleaning routes by type classification, adjusting ASA values by weighing segment lengths and some simplification and remodelling steps - there is no clear standardized appropriate option, thus there is still room for improvement. For OSM RCL representation adjustments, publications mention Douglas-Peucker simplification with different radius thresholds (Kolovou et al., 2017), where oversimplification starts cutting corners too short losing urban form accuracy. Newer steps for RCL preparation can achieve a finer street network representation, using the latest 2020 Space Syntax Toolkit (SST) version and the GRASS version 7.8.5 QGIS vector generalize tool. These tools' recent versions include complementary steps besides the Douglas-Peucker simplification. However, these steps are not yet clearly presented as to what they entail, how different inputs change models' configurations and how they impact on city space syntax ASA.

To investigate more ideal automated RCL space syntax city ASA preparation methods, this paper gathers, tests, compares and shares data from different RCL preparation steps. The paper uses a city OSMnx driveable RCL base comparing models transformed by the following methods: (i) our proposed algorithm using GIS-CAD interoperability; (ii) default and adjusted SST network cleaner; (iii) default and adjusted GRASS QGIS generalization. Aiming to contribute to space syntax updated city analysis elsewhere, the following section describes all OSM RCL sourcing and preparation steps, from graph creation with OSMnx, different layer preparation steps and ASA processing. The results section exhibits main models results and comparisons, discussed in the following section highlighting current best cost benefit fit for RCL ASA preparation.

2 DATASETS AND METHODS

This section shares automated steps to capture and prepare an updated city model, and perform space syntax ASA with Depthmap and Space Syntax Toolkit graph analysis module.

2.1 Capturing OSM RCL

The free open-source collaborative Open Street Map Road data was captured via python language library (OSMnx) and can be easily be performed for other places and contexts worldwide. If OSM road files were to be captured by geofabrik.de website as used elsewhere (Kolovou et al, 2017), cleaning steps would have to be taken type by type to erase routes not open to vehicular movement, as erasing service roads, cycle and pedestrian paths. This would also demand larger files to be downloaded then cropped. The OSMnx library with the osmnx.graph module (Boeing, 2017) can automatically get street networks graphs from the OSM database by type. For this paper the João Pessoa city (Paraíba state, Brazil) drivable street network was downloaded by place name[1], almost entirely streets openly accessible[2] in the municipal boundaries.

The next step involved adjusting the coordinate reference system. The OSM street network layer employs a geographic coordinate system (GCS WSG 1984 EPSG 4326). For RCL preparation functions and ASA Depthmap processing to read metric distances this OSM layer was transformed into a planar coordinate system. For João Pessoa location SIRGAS 2000 / UTM zone 25S EPSG 31985 system was used.

---

[1] Downloaded in August 2021 using python language at Spyder integrated development environment (IDE).

[2] No service roads appeared, but some private condominium routes remained.

This unprepared RCL OSMnx layer was used to perform ASA, however with too many nodes impacting on centrality values and processing time, as highlighted by the literature and exemplified in our results.

2.2 Preparing OSMnx RCL for ASA

Different automated preparation steps were performed for João Pessoa RCL OSMnx. To understand better possible RCL adjustments, preparation steps were performed with the Computer Aided Design (CAD) Rhinoceros program and the visual programming environment Grasshopper (GH). This clarified the RCL street network model nature in terms of form, topology and tabular information. This study was aided by GH libraries Shrimp add-on for GIS-CAD interoperability and DeCoding Spaces add-on to build a graph model from polylines to simplify the RCL model using the Douglas-Peucker algorithm (Figure 1).

The Douglas-Peucker is a well-known algorithm which performs a polyline simplification reducing the number of vertices by setting a radius distance threshold in which some vertices should be erased, overridden by setting snaps to maintain some vertices. There are two different issues concerning RCL street network representation: (i) close to corners there are many un-needed vertices; (ii) straight long streets and roads many times exhibit un-needed vertices as this does not signify any route change. The first issue could be addressed by a Douglas-Peucker distance limitation approximating softer urban block curves, such as setting a 10-metre threshold. This step, however, does not solve the second issue for vertices set on longer straight roads. The solution found to address both RCL street network representation issues was to establish criteria for vertices that should remain in the model, so a larger Douglas-Peucker simplification threshold could be performed without losing significant street network information. To maintain vertices that entailed some break in linearity, an angular criterion was set to fix vertices between segments with three degrees of angle deviation or more. Additionally, as only one local point is needed to represent a position, a Cull point function was set to select and fix one point within a 10-metre threshold radius. So that real street network crossings would not be lost, polyline end-points and graph intersections were fixed. The intersections were found using a graph built from segments, filtering the vertices that had two or more segments connected. With all these points fixed a large Douglas-Peucker simplification radius was set (~ 2000 m) reducing unnecessary vertices and segments while maintaining significant street network form information.

To perform this simplification task more efficiently, same street name polylines were joined in the same branch at a data-tree structure. Accordingly, polylines were grouped by street name in an identical data-tree structure to simplify each polyline separately. Figure 1 exhibits an algorithm flowchart that explains the proposed simplification method implemented on GH using Shrimp and Decoding Spaces add-ons.

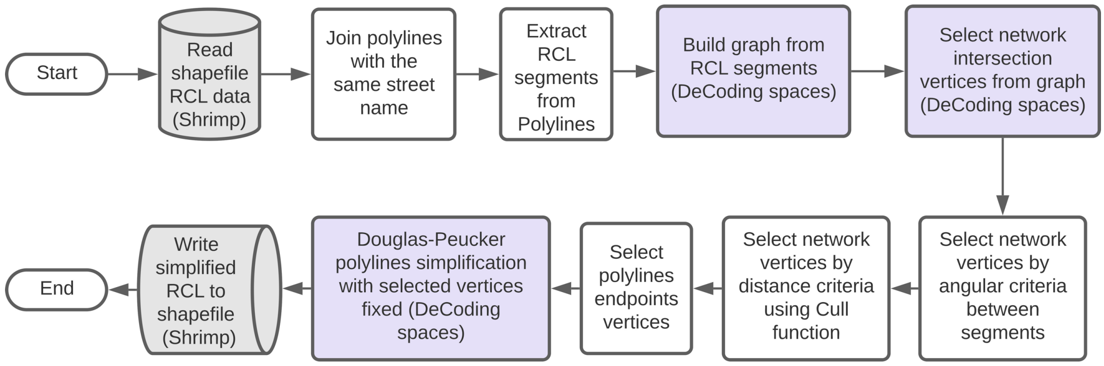

Figure 1: Algorithm CAD-GIS interoperability and performed polyline simplifications flowchart.

The proposed algorithm helped understand other simplification tools available on QGIS platform, once documentation was not found to clarify how to obtain a better-tuned RCL simplification. Other preparation steps were then tested with QGIS tools. The Space Syntax Toolkit (Gil et al., 2015) is a QGIS plugin that allows an interface with Depthmap program. This toolkit's various modules have been updated throughout the years. The RCL-topology-cleaner module at one point involved a cleaning process to validate geometries to become suitable for GIS analysis, such as correcting RCL maps' topological errors and line overlaps (Kolovou et al., 2017). This was noted by Kolovou and colleagues (2017) to apply mostly to OSM as this dataset was not found to be rigorously validated. Douglas-Peucker simplification as described by Kolovou et al (2017) was performed elsewhere. However, the newest SST version 0.3.9 (March 2021) currently has a cleaner module that, amongst other functions, involves "snap endpoints threshold" with 10 set as default and "simplify angular changes" also with 10 set as default. As this default setting seemed to oversimplify and distort the street network, other adjustments were also tested for these functions. This paper will present models resulting from both the default setting and a better adjusted setting. If the SST Cleaner only had its previous topologic error functions, there would be no need to use it for current RCL OSM, as no topologic errors were found for Joao Pessoa OSM.

QGIS GRASS also features a preparation tool for RCL. This function has already been used for ASA elsewhere (Pezzica, Cutini and Souza, 2019; Altafini and Cutini, 2020), with Douglas-Peucker simplification. The current QGIS version 3.16.14 GRASS vector generalize tool exhibits, amongst some other functions, simplification algorithm set as Douglas as default with a maximum distance tolerance value of 1. As the default setting seemed to under simplify the street network, other adjustments were made increasing Douglas tolerance. Models were tested and will be presented with both default and adjusted settings.

The current SST tool functions includes steps that resemble this study's GIS-CAD performed functions, as it allows for an angular and distance thresholds together with Douglas-Peucker simplification.

To understand and compare the resulting OSM street network models preparations, the following adjustments were performed and will be presented in this paper results with the acronyms:

- OSMnx: OSM RCL drivable as downloaded by OSMnx without simplification steps, and served as base for the subsequent preparation steps;
- This Study: the model resulting from the steps performed by the paper authors, set at 10 metres threshold distance and 3 degrees for angular change, synthesized in Figure 1;
- SST_10_10: the Space Syntax Toolkit preparation default configuration with 10 set at snap endpoint threshold and 10 at simplify angular changes;
- SST_5_5: the Space Syntax Toolkit preparation adjusted to 5 set at snap endpoint threshold and 5 at simplify angular changes;
- GRASS_1: GRASS generalisation function with default set at 1 Douglas Simplification;
- GRASS_5: GRASS generalisation function adjusted to 5 Douglas Simplification

All these OSM RCL João Pessoa city models were processed for ASA centralities via Depthmap with SST plugin graph analysis module communicating with DepthmapXnet 0.35. Integration and choice, and their following normalized values (respectively NAIN and NACH), were processed for global (n) and metric radii: 400, 1200, 3000 and 5000. These radii encompass either short or longer walkable and vehicular distances, as well as the whole global centrality. These centralities average and maximum values were visualized and compared with all models by linear graphs. Centralities integration and choice global distribution maps were visualized with values separated with Jenks natural breaks (six classes) using Depthmap classic colour band.

3 RESULTS

OSM studied models encompass all drivable routes inside Joao Pessoa (Figure 2). Two clips were set to help visualize and analyse models and respective centralities. The first local clip (Clip 1) is an urban fraction exhibiting urban blocks close to highway BR 230 and Bancários Avenue[3] (marked in light blue in Figure 2). The second larger frame (Clip 2) captures Joao Pessoa Old City Centre and main expansion routes. Epitácio Pessoa Avenue is João Pessoa's longest linear route and, together with Ruy Carneiro Avenue, form the city main expansion to the sea. Other main expansions from the old city centre direct south: (i) around the east of the Atlantic Forest preservation via Pedro II Avenue, BR 230 highway and Bancários avenue; (ii) west of the Atlantic Forest Preservation via Cruz das Armas Avenue. Highway BR 230 helps connect João Pessoa with neighbouring municipalities to the north and east, and distributing flows in Joao Pessoa transversely to main expansion routes from the Old City centre.

[3] Bancários neighbourhood central avenue, officially Empresário João Rodrigues Alves Street and continuing streets.

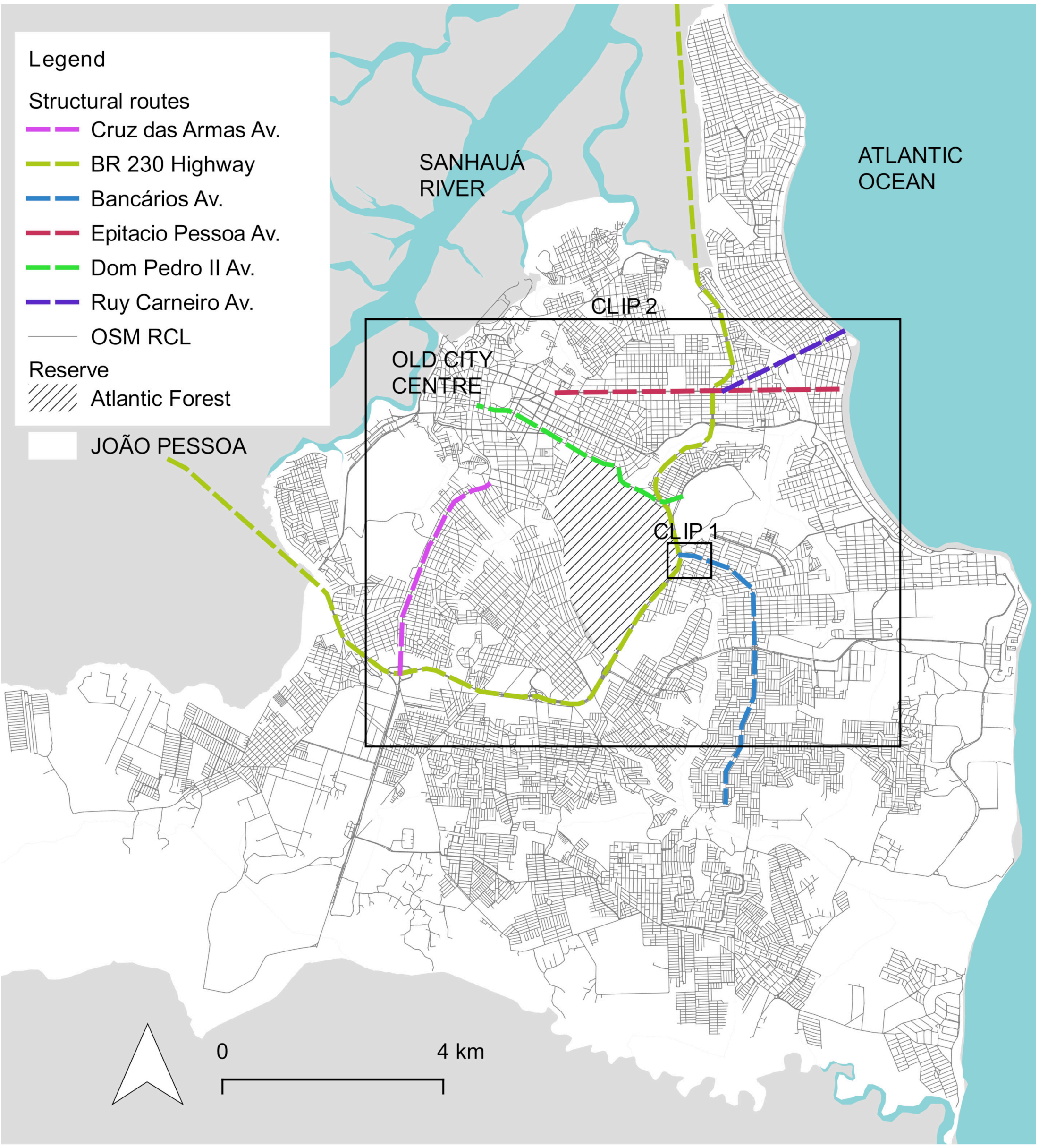

Figure 2: Joao Pessoa OSMnx map exhibiting main routes and clips.

OSM road centre lines representation usually represents one line at route centre. This happens when streets have one traffic flow or two traffic flows but no separating curbs. Examples of streets in João Pessoa with two traffic flows without physical separation and represented with one centre line by OSM can be seen on Figure 3 right-hand cases. On the other hand, routes with two traffic flows with curbs as physical separations are represented with two parallel centre lines by OSM, one for each side. This is the case for such structural routes in João Pessoa as Highway BR-230 and avenues Ruy Carneiro, Epitácio Pessoa and the beginning of Bancários avenue; the latter ones can be viewed on Figure 3 left-hand cases. Curbs separating sides of traffic are usually accompanied by other street furniture, such as light and sign posts, and sometimes trees.

Figure 3: Images of routes with two traffic flows in João Pessoa, on the left those represented by two OSM centre lines; on the right those represented by one OSM centre line. Source: the authors, 2021

OSMnx models are compared first by their segment form for Clip 1, then their system's segment sizes and numbers according with models' transformation into segment graphs via Depthmap Space Syntax Toolkit plugin. Then ASA centrality values for different metric radii are presented, and finally, centrality distribution visualization for Joao Pessoa Clip 2.

To illustrate models' route segments nodes characteristics, Highway BR 230 joining Bancários Avenue enlarged (Figure 4), location in the city viewed as City clip 1 in Figure 2. RCL models' representations are set at the centre of routes made of polylines following routes' curves. Although the resulting street representation does not vary much for the studied models, the number of nodes varies significantly. The untreated RCL OSMnx exhibits many nodes and segments for curved routes, especially roundabouts. For GRASS_1 with a smaller simplification threshold, some nodes on smoother curves vanish yet not significantly; roundabouts still carry unnecessary large numbers of nodes. Route representation seems more unbundled for the models on Figure 4 right-hand side. This Study, SST_5_5 and GRASS_5. This Study and GRASS_5 approximate more and capture roundabouts similarly.

On the other hand, SST_10_10 oversimplifies the street network cutting some corners, as at highway BR-230, the western route passes over the Atlantic reserve. SST_10_10 joins some close parallel routes elsewhere in the system, while separations are maintained for other more distant parallel lines, as the two sides of BR-230. However, at some routes this happens inconsistently, as there are places of separation and some joining points; this happens at Bancários Avenue heading west east at Figure 4 clip and elsewhere, changing street network representation inconsistent with actual traffic continuity.

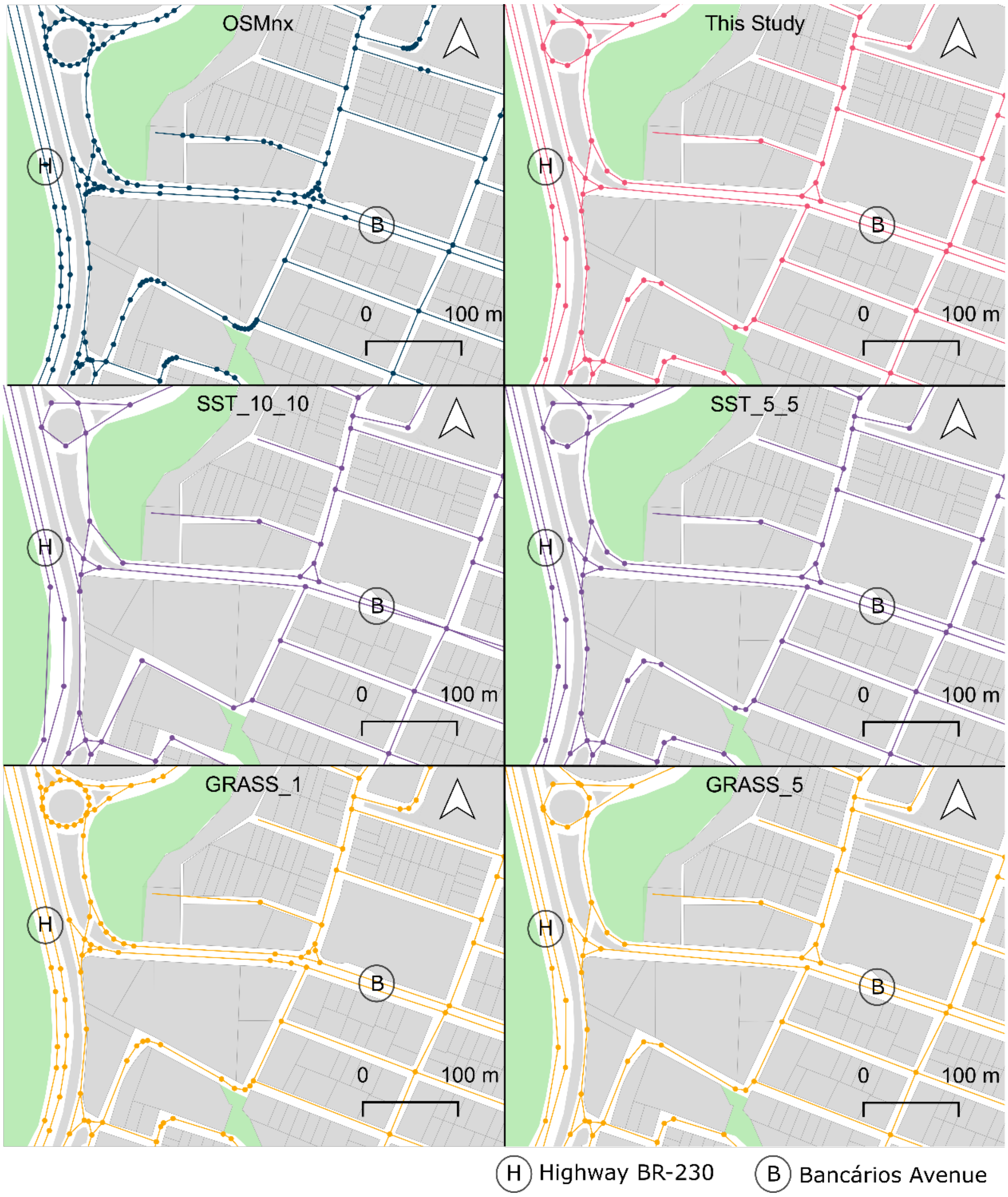

Figure 4: RCL models exhibiting polyline segments (Clip 1).

SST-10_10 higher simplification and more route joining is expressed by the model's lowest number of segments and smaller system length (Figure 5). For other models, total route length varies little while the number of segments decreases steadily, to close to 73% of OSMnx for Grass_1, slightly over half for SST_5_5 (56%), Grass_5 (56%) and This study (57%) models, practically half for SST_10_10 (51%). Figure 5 segments are the models graph nodes.

In Figure 4 - exhibiting nodes and segments at a local clip - This Study and Grass_5 seem almost identical. Yet for Figure 5 - especially regarding distribution of number of segments by segment lengths -, the models that overlap are Grass_5 and SST_5_5, although This Study is still very close to these models. The segment lengths histogram (Figure 5) shows that for OSMnx the largest incidence of segments has between 5 and 10 metres length. For Grass_1 most recurrent segment size is between 5 and 10 metres, with a close peak just over 50 metres, and a third peak with segments around 75 metres. For other models' main peaks happen for these values, either slightly over 50 metres or around 75 metres. These values seem to relate more with the city urban form, as representing block sizes or distances between crossings formed by urban blocks. Again, SST_10_10 differs slightly from other models with less smaller segments (up to 50 metres), and has the highest mean segment size.

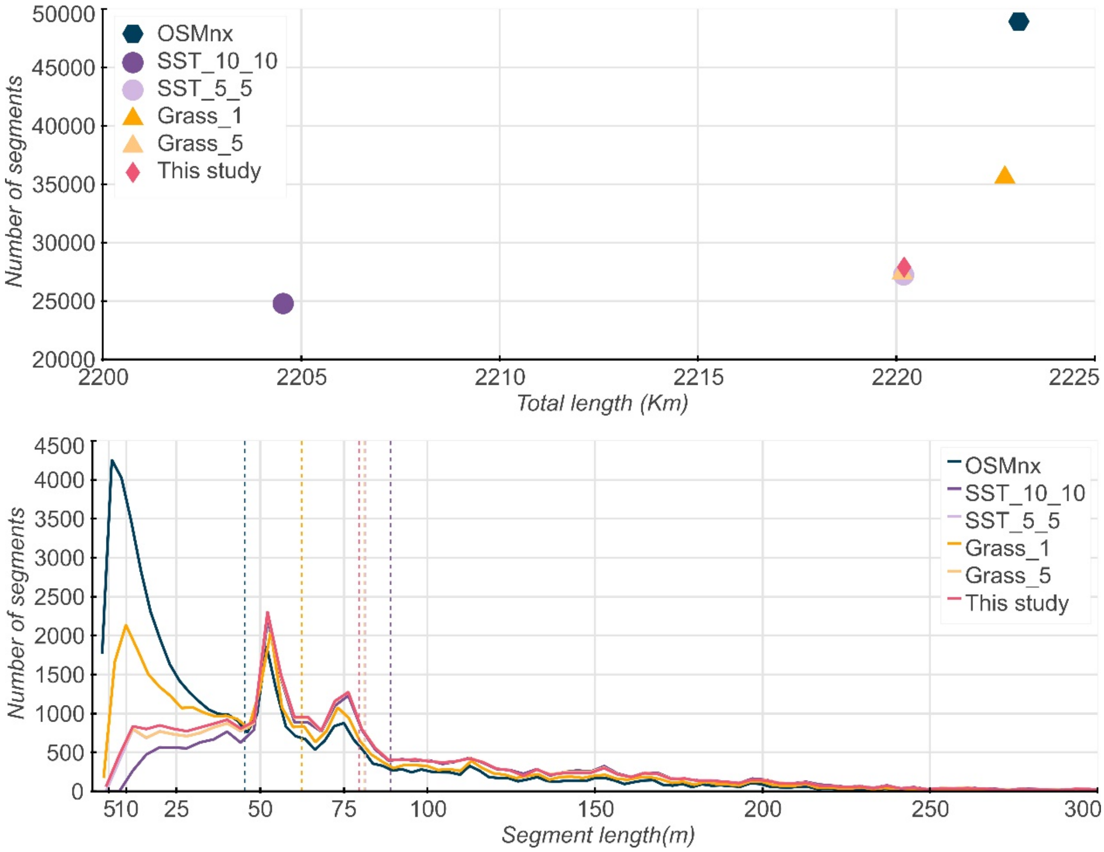

Figure 5: Models number of segments related with their total length in km (up) and number of segments by segment length in metres (down).

Trends between models translate similarly in terms of ASA centrality maximum and mean values (Figure 6). OSMnx and Grass_1 distance themselves more from other models. Overall differences in values are either higher for (i) global integration and, especially, choice values or (ii) for local 400 m maximum integration and normalized integration and choice values. Following more segments, choice mean and maximum values are much higher for OSMnx. A higher discrepancy is also found for local integration and normalized centrality values. Routes without neighbours at a local scale might reach unnecessarily high centrality value as the radius might encompass many consecutive segments without having an actual street legibility hierarchy. As the number of segments decrease with preparation steps of the models, discrepancies in these local values also decrease.

For most centrality values, SST_5_5, Grass_5 and This study models are close. However, there are slight shifts especially for more local NAIN and NACH maximum between and SST_5_5_ Grass_5 and This study, as the latter exhibits higher 400m and 1200m max values.

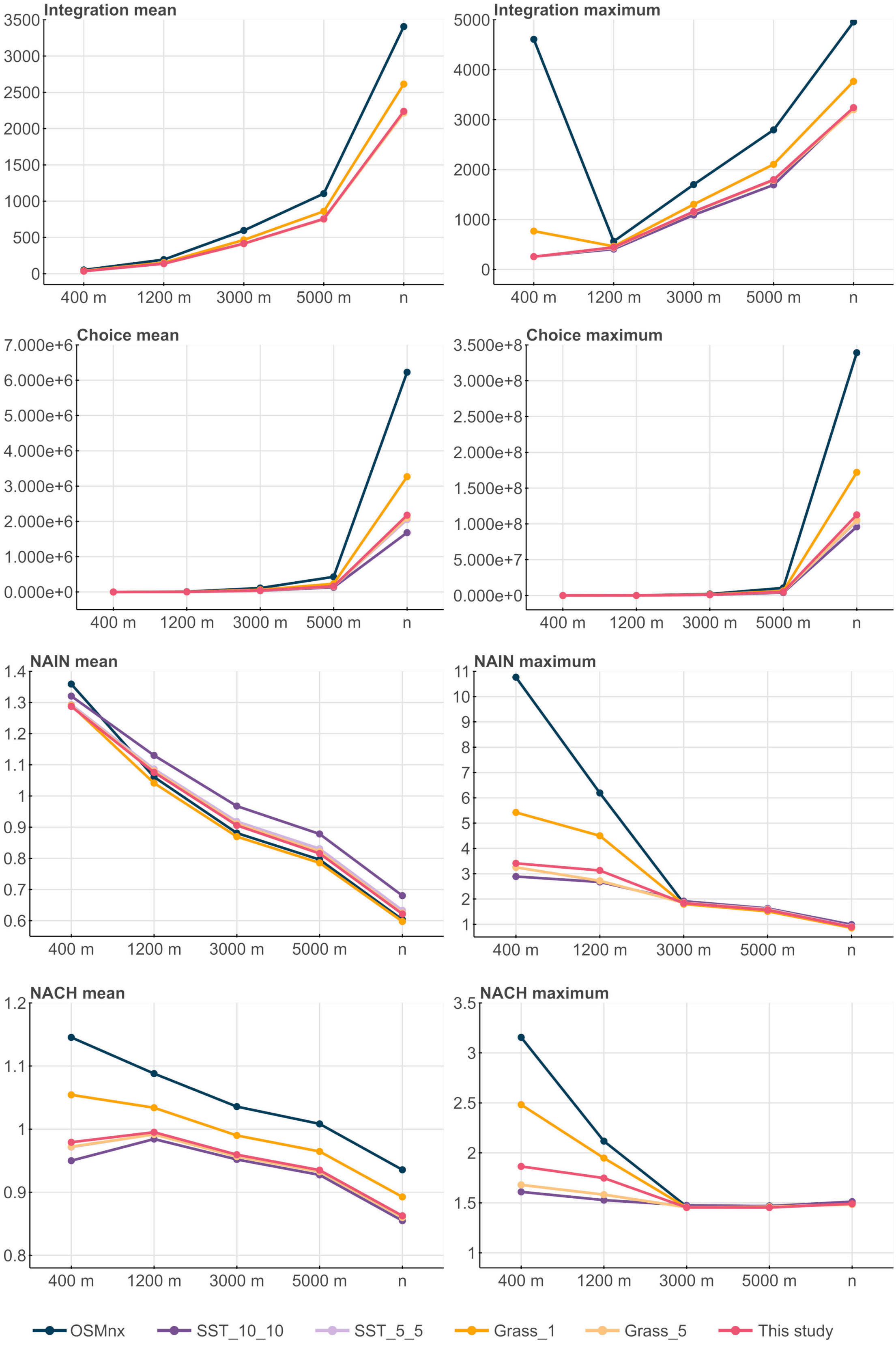

Figure 6: Centralities values mean (left-hand) and maximum (right-hand) values for studied models.

All overall centrality distributions (Figures 7 and 8) accentuate the city's main structural routes highlighted earlier in Figure 2. At a closer look, there are some differences between models.

SST_10_10 highlights fewer routes forming some of the expansion routes spanning from the old city centre, and highlights less streets connecting and crossing Epitácio Pessoa Avenue (towards east) and Bancários Avenue (towards south) both for integration (Figure 7) and choice values (Figure 8). For these avenues and other structural routes there are instances for the SST_10_10 model where parallel lines are joined at some points and not on others, which might explain loss in linearity with less centrality continuity for some avenues and their crossing streets.

Figure 7: RCL models exhibiting ASA integration, João Pessoa Clip 2.

Choice distribution highlights the main city structural routes in all models (Figure 8). However, for SST_5_5 and, especially, SST_10_10 highest values concentrate more, in less segments. For SST_5_5 highest values locate on Bancários Avenue. On the other hand, a more continuous system of routes bordering the city Atlantic Forest at its northeast and continuing north reaching Epitácio Pessoa Avenue is highlighted by the original OSMnx model, very close to GRASS_5 model distribution and, especially, This Study model.

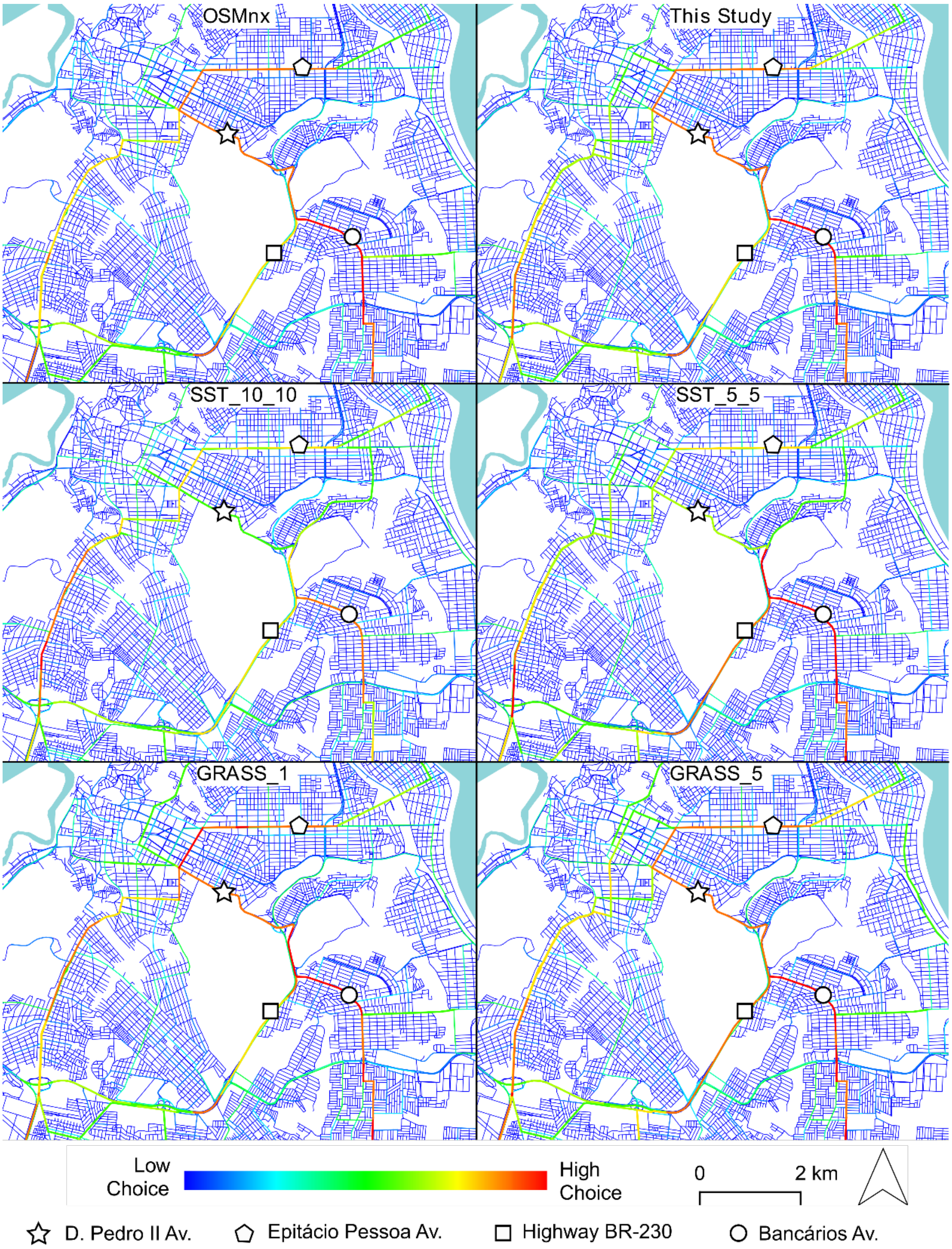

Figure 8: RCL models exhibiting ASA choice, João Pessoa Clip 2

Overall different choice values for SST_10_10 might be impacted from many routes sometimes joining and changing movement hierarchies' readings, as well as for almost straight representations of most roundabouts. Specifically, SST_5_5, that has not shown to join any parallel lines, choice distribution still changes from other models. A specific change between SST_5_5 and other models except SST_10_10 is viewed at BR 230 and D. Pedro II Avenue curved encounter; for these SST models a straighter roundabout representation than those of other models seems to distribute choice values more easily further northeast (route viewed in green), and continuity of routes towards the north of the Atlantic Forest reserve and up to Epitácio Pessoa Avenue. While for other models this circuit north of the Atlantic Forest connecting with Epitácio Pessoa Avenue reaches high to medium-high values viewed in orange or red, for SST model's centrality values are not so high, dropping to yellow (SST_5_5) or even green colours (SST_10_10).

4 DISCUSSION

Although some earlier work argued that RCL model preparation for space syntax ASA should approximate RCL to axial maps (Kolovou *et al.*, 2017; Krenz, 2018), there does not seem to be a strong reason for this. ASA already incorporates angular changes and thus, as proposed by Turner (2007), minimises effects of cartographic differences between representations. Moreover, it can be difficult to change RCL model default representation without individual interpretation problems. If a specific significant route width is set from which two parallel lines would be joined or merged (as the dual line removal mentioned by Krenz, 2018) this can impact on route continuity reading, as happens especially for SST_10_10 model. As route width changes sometimes in the same route, this might explain why some routes in SST_10_10 João Pessoa model have some stretches and points joining two routes' sides discontinuously and breaking each side linear continuity. This also mixes possible centrality differences between two sides of these avenues. This is the case, for instance, for stretches of Epitácio Pessoa, Ruy Carneiro and Bancários avenues, as well as for other routes in João Pessoa city centre. OSM RCL does not represent separate lines for all streets with different traffic directions. Separate lines appear where there are physical curbs separating street traffic directions. These might indeed work as separate spatial and movement entities with different network connections and, indeed, flows. If roundabouts (even those without building in the middle) were to be erased one might easily find some issues noted for parallel lines joining. As roundabouts come in many different shapes and sizes, to erase all would also be to equate all. When feasible to be processed by ASA, OSM finer grain representation can be taken advantage of rather than changed, as modelled separations indeed appear where there are potential movement barriers, or at least conflict. This aligns with Turners reasoning "if space syntax is to be used as a model of movement, then it ought to incorporate ideas from traffic modelling to make a fully coherent model of the built environment." (Turner, 2007, p. 540). The main issue for RCL for ASA is the overall number of

nodes and segments, that decreases significantly with preparation steps presented and discussed in this paper without changing street network connections and route representation.
Viability to access accurate open RCL models has been improving rapidly in recent years by OSM RCL model quality and catchment so that some issues addressed earlier such as area catchment and topologic inconsistencies (Dalton, Peponis and Dalton, 2003; Dhanani *et al.*, 2012; Kolovou *et al.*, 2017; Krenz, 2018) were not found to be an issue for this study. Ways to access and prepare automatedly RCL for ASA also have been improving. Boeing's (2017) recent OSMnx library is an easy interface to select OSM street networks and easily define system sizes and limits while doing this, as well as easily set different types of systems of movement (e.g., walkable, drivable). However, the OSM categorization could still improve so that considered drivable routes do not include routes inside private condominiums. Some private condominium routes appear in João Pessoa model, although many were not included.

Different types of system catchment can be used according with the research investigation. Although the catchment set for João Pessoa municipality exemplified different RCL preparations results, a different system catchment could be more functional as amenities and urban infrastructure at neighbouring municipalities are sometimes close and continuous to João Pessoa. The osmnx.graph module (Boeing, 2017) provides a system considering a road network distance catchment from an address. For instance, a 20 km distance from João Pessoa provides the road network spanning up to 20 km from its old city centre covering all João Pessoa and fairly continuous occupations, including parts of Cabedelo (north), Conde (south) Bayeux – where the airport is set - and Santa Rita (West).

Findings indicate that default SST (set at 10 for threshold and 10 for simplify angular changes) and Grass (set at 1 for Douglas generalization) parameters are not ideal for city ASA. While one default setting, SST_10_10, exhibited the highest decrease in number of segments, some corners were cut and routes distorted affecting centrality distribution. On the other hand, the Grass default processing decreased segment numbers very little. While this step still improved the model comparing with the original OSMnx, further thresholds can and should be used satisfactorily, facilitating processing and levelling centrality values.

Our proposal –shared step by step – helped view better adjustments for existing tools. This study exploration via CAD and GIS allowed more forms exploration and to understand better available steps at other interfaces. However, these are not openly accessible via free programs, and, having found an open and easier tool, there is currently no apparent need found for further development in this sense.

From the OSMnx RCL preparation steps and outcoming models, the GRASS vector generalize tool set at 5 for Douglas generalization seems the best fit for use at this moment. All the steps are easily available at openly accessible QGIS and GRASS platform, and results reach a simplified model with a high decrease in segment numbers, while maintaining street network form representation.

5 CONCLUSIONS

Understanding that models' representation impact on results (Marshall et al., 2018) and the need for automated steps for city models analysis (Batty and Rana, 2002), this paper described steps for automated OSM RCL models capture and preparation with good ASA performance. This aims to contribute in facilitating more standardized and rapid city space syntax analysis.

Boeing's (2017) OSMnx library provides different ways of selecting catchment areas and types of movement routes, in this case using drivable routes for city analysis. From the tested models and taken steps, the Grass_5 (with parameters set at 5 for Douglas) calibration seems the best tuned model with easy preparation - performed in free platforms with few steps via QGIS GRASS tool -, while reading centralities also with few (if any) distortions regarding the original OSM representation. These processes enable updated city modelling processing worldwide.
The possible advantages for tuning into an automated and more precise street network representation can go beyond advantages for space syntax ASA centrality results. As these prepared models might have more to say about actual street form crossings, street segment sizes and significant angular changes in legibility, models with the same preparation steps can help compare forms, sizes and connections between cities or urban fractions.

REFERENCES

Altafini, D. and Cutini, V. (2020) 'Homothetic Behavior of Betweenness Centralities: A Multiscale Alternative Approach to Relate Cities and Large Regional Structures', *Sustainability*, 12(19), p. 7925. doi:10.3390/su12197925.

Batty, M. (2021) 'Defining Urban Science', in Shi, W. et al. (eds) *Urban Informatics*. Singapore: Springer (The Urban Book Series), pp. 15–28. doi:10.1007/978-981-15-8983-6_3.

Batty, M. and Rana, S. (2002) 'Reformulating Space Syntax: The Automatic Definition and Generation of Axial Lines and Axial Maps', *CASA Working Papers*, 58.

Dalton, N., Peponis, J. and Dalton, R. (2003) 'To tame a TIGER one has to know its nature: extending weighted angular integration to the description of GIS road-centerline data', in *4th International Space Syntax Symposium*. London: University College London, p. 65.1-65.10.

Dhanani, A. *et al.* (2012) 'From the axial to the walked line: evaluating the utility of commercial and user-generated street network datasets in space syntax analysis', in *Proceedings of the 8th International Space Syntax Analysis*. Santiago de Chile: PUC, p. 8211:1-8211:32. Available at: http://sss8.cl/8211.pdf (Accessed: 14 October 2020).

Freeman, L.C. (1977) 'A Set of Measures of Centrality Based on Betweenness', *Sociometry*, 40(1), p. 35. doi:10.2307/3033543.

Gil, J. *et al.* (2015) 'The space syntax toolkit: Integrating depthmapX and exploratory spatial analysis workflows in QGIS', in. *10th International Space Syntax Symposium*, London: Space Syntax Laboratory, The Bartlett School of Architecture, University College London, p. 148:1-148:12.

Hillier, B. and Hanson, J. (1984) *The social logic of space*. Cambridge, England: Cambridge University Press.

Hillier, B. and Iida, S. (2005) 'Network and psychological effects in urban movement', in *Spatial information theory*. Springer, pp. 475–490. Available at: http://link.springer.com/chapter/10.1007/11556114_30 (Accessed: 23 July 2014).

Hillier, B., Yang, T. and Turner, A. (2012) 'Normalising least angle choice in Depthmap-and how it opens up new perspectives on the global and local analysis of city space', *Journal of Space Syntax*, 3(2), pp. 155–193.

Husain, A. and Vaishya, R.C. (2018) 'Road surface and its center line and boundary lines detection using terrestrial Lidar data', *The Egyptian Journal of Remote Sensing and Space Science*, 21(3), pp. 363–374. doi:10.1016/j.ejrs.2017.12.005.

Kolovou, I. *et al.* (2017) 'ROAD CENTRE LINE SIMPLIFICATION PRINCIPLES FOR ANGULAR SEGMENT ANALYSIS', in *Proceedings of the 11th Space Syntax Symposium*. Lisboa, p. 16.

Krenz, K. (2018) *Network Centralities in Polycentric Urban Regions: Methods for the Measurement of Spatial Metrics*. University College London.

Marshall, S. *et al.* (2018) 'Street Network Studies: from Networks to Models and their Representations', *Networks and Spatial Economics*, 18(3), pp. 735–749. doi:10.1007/s11067-018-9427-9.

Minaei, M. (2020) 'Evolution, density and completeness of OpenStreetMap road networks in developing countries: The case of Iran', *Applied Geography*, 119, p. 102246. doi:10.1016/j.apgeog.2020.102246.

Pezzica, C., Cutini, V. and Souza, C.B.D. (2019) 'Rapid Configurational Analysis using OSM Data: Toward the use of Space Syntax to orient post-disaster decisions', in *Proceedings of the 12th International Space Syntax Symposium*. Beijing: Beijing Jiaotong University, p. 18.

Ratti, C. (2004) 'Space Syntax: Some Inconsistencies', *Environment and Planning B: Planning and Design*, 31(4), pp. 487–499. doi:10.1068/b3019.

Turner, A. (2007) 'From axial to road-centre lines: a new representation for space syntax and a new model of route choice for transport network analysis.', *Environment and Planning B: Planning and Design*, 34(3), pp. 539–555. doi:10.1068/b32067.

Van Nes, A. and Yamu, C. (2021) *Introduction to Space Syntax in Urban Studies*. Cham: Springer.