Article

Homothetic Behavior of Betweenness Centralities: A Multiscale Alternative Approach to Relate Cities and Large Regional Structures

Diego Altafini and **Valerio Cutini** *

DESTeC—Dipartimento di Ingegneria dell'Energia, dei Sistemi, del Territorio e delle Costruzioni, Università di Pisa, Lungarno Pacinotti 43, 56126 Pisa, Italy; diego.altafini@phd.unipi.it
* Correspondence: valerio.cutini@ing.unipi.it

check for updates

Abstract: Regional configuration can reveal important aspects about city sustainability, as local-regional interactions shape the evolution and inner geography of urban settlements. However, modelling these large-scale structures remains a challenge, due to their sheer size as physical objects. Despite recent improvements in processing power and computing methods, extensive time periods are still required for ordinary microprocessors to model network centralities in road-graphs with high element counts, connectivity and topological depth. Generalization is often the chosen option to mitigate time-constraints of regional network complexity. Nevertheless, this can impact visual representation and model precision, especially when multiscale comparisons are desired. Tests using Normalized Angular Choice (NACH), a Space Syntax mathematical derivative of *Betweenness Centrality*, found recursive visual similitudes in centrality spatial distribution when modelling distinct scaled map sections of the same large regional network structure. Therefore, a sort of homothetic behavior is identified, since statistical analyses demonstrate that centrality values and distributions remain rather consistent throughout scales, even when considering edge effects. This paper summarizes these results and considers homotheties as an alternative to extensive network generalization. Hence, data maps can be constructed sooner and more accurately as "pieces of a puzzle", since each individual lesser scale graph possesses a faster processing time.

Keywords: Betweenness Centralities; homothetic behaviour; space syntax; urban sustainability; urban-regional relations

1. Introduction

Large regional network structures analysis and representation are still regarded as novelties in the field of urban morphology. Due to methodological and practical challenges, the construction of network-based models capable of accurate analysis, with a substantial level of detail at these scales, was hindered until the 2010s. Prior to this period, the rapid expansion in the availability of processing and representation instruments, such as the diffusion of open-source GIS suites, and their integration into in-vogue network analysis methods, allowed the researchers to better comprehend the configurational relations within the road-circulation networks of small scale urban settlements [1], and also evaluate the sustainability of their inner designs. In this sense, these first limited-scale studies laid the theoretical and methodological foundations for urban network spatial analysis, required to further assess greater territorial scales.

Despite this, even after significant advancements that brought exponential improvements on computing power, the ordinary microprocessor is still unable to cope with the complexity of modelling network-based datasets that comprise large extents of territory in acceptable time-lapses. Urban and

regional demands concerning the implementation of these analyses led to an often-extensive use of algorithms driven to the simplification of graphs [2] in an effort to generalize the structures used to depict road-circulation networks and deal with the processing of time-lapses issues. Other more thorough methods to enhance the generalization degree of spatial networks, based on the removal of road-elements deemed less important, were also prevalent [3]. The intent of such methods' is to create graph structures that represent only the regional road-infrastructures central vertices, forming the so-called skeletal road-center line graphs [4], which ignore access roads in highways and certain small urban streets. These extensive generalization principles epitomise the mainstream practices regarding dataset generalization for spatial analyses, having being used in the most recent studies about the spatial configuration of urban, regional [4,5] and supraregional [6] road-circulation networks, achieving robust results.

Those studies share common traits: their configurational analyses are founded on Space Syntax' Angular Analysis [7,8], and all them explore a sole dimension of scale. This latter aspect is, in fact, the reason why robust results were achieved. Angular Analysis can construct models with accurate and detailed depictions of road-circulation networks configurational properties, such as centrality hierarchies [i], but their instrumentation is not well adapted to model regional spaces, since their focus is on representing small scale networks, such as urban spaces [8,9]. In this regard, Angular Analysis [7] requires ulterior dataset adjustments in the form of the aforementioned generalization processes [ii] that allow efficient modelling, within reasonable time-lapses. However, generalization is accompanied by a *trade-off* regarding network incompleteness, which is rather irrelevant for single scale analysis, but can be quite significant when the objective is to assess accurate relations across different territorial scales. This factor happens to hinder one of the most distinguished features of Angular Analysis: the precision in the cartographic representation of network-based systems' real morphologies, while depicting with exactitude their configurational properties.

Although these *time-scale-precision* issues are nontrivial to address, research in the fields of theoretical networks, mathematics and physics may point to possible solutions, as it suggests that recursive regularities in the patterns across different road-circulation networks might exist. If proven true across scales, this could shift paradigms for road-networks cartographical representation. Those studies [10,11] indicate that different urban road-infrastructures tend to have a largely equal underlying network structure, possessing similarities amid their geometric features, while other evidences imply that structural invariances are present on scale-free urban networks, such as those representing *Betweenness Centrality* of urban road-circulation networks [12,13]. These are some indicators that lead to the hypothesis that these network centralities might exhibit visual similitudes (homotheties), hence, some sort of homothetic behavior across different scales.

Michel Chasles [14,15] conceived the term homothety (*homothesis*) in the course of his treaty about the properties of Euclidean geometry, while describing the relationship between spaces with similar form and orientation. In mathematics, a homothetic behavior is characterized by a transformation of an affine space by a factor λ [iii] and results in an invariance of this space form or configuration, albeit its overall scale changes [16]. In this sense, if two objects or parts of those objects have distinct sizes, but conserve the same appearance, they can be considered homothetic. Examples of these properties are the Russian *matryoshkas* dolls and the inflorescence buds of the *Romanesco* broccoli—which is also an example of a natural fractal. Fractals themselves can be considered as an example of homothetic behavior, since they exhibit self-similarity across different scales, a logic that is commonly associated with map scaling in cartography [17] and to scaling laws in urban settlements configuration [18].

In networks, the occurrence of homothetic behaviors would imply that a section of the network, when modelled independently, ought to retain a certain regularity in their distribution of centrality hierarchies (visual similitude) when compared to a larger section, independently modelled as well, that contains it. Hence, the smaller network maintains its overall proportions (configuration, hierarchies and values) across scales. This visual similitude was perceived while apposing several Normalized Angular Choice (NACH) models, a Space Syntax' derivative from mathematical betweenness, in a

GIS-based environment. These models comprised different territorial scales (municipalities and provinces) from Italy's Tuscany Region (Toscana), which prompted further tests using these territorial expanse road-infrastructures as datasets to assess the extent of occurrences and the formal nature of these homothetic phenomena.

This paper summarizes the results of the correlation tests and proceeds to consider the homothetic behavior found in *Betweenness Centrality* as an alternative to extensive generalization in graphs that represent large scale road-circulation networks. If the recursive homothetic behavior is verified across scales, data maps could be constructed sooner and more accurately as "pieces of a puzzle", since each individual lesser scale graph possess a faster processing time. Such theoretical and methodological improvements can aid in the representation of interactions between cities and the regional space in which they are embedded, allowing urban planners to produce assessments that consider these larger territorial extents' configurations in the evaluation for sustainable cities.

2. Datasets and Methods

Datasets employed in the network analyses were derived from a common database: the Tuscan Region Road Graph (*Grafo Viario della Toscana*) [19,20]; a Road-Centre Line (RCL) map representing the whole regional road-circulation network. Its layout follows a typical network structure, composed by arches, continuous polylines referred to as road-elements (*Elemento Stradale*), and node structures, mid-points that set linkages between one or more road-elements, dubbed road-junctions (*Giunzione Stradale*). This source graph is further sectioned using the administrative limits framework [21] to depict the territorial expanses of the region (*regione*), provinces (*province*) and municipalities (*comuni*). From each individual administrative limit area, buffers (1 km for provinces; 300 m for municipalities) are used to preserve most road-circulation network continuities during the rend [iv]. Any road-elements that remained disconnected were removed in the graph revamp. A total of 286 datasets were conceived, corresponding to road-graphs for all Tuscany Region mainland continuous road-infrastructures being scaled in: Regional (1), Provincial (10) and Municipal (275) [v].

Prior to the administrative limits rendering, the regional road-circulation network is submitted to a generalization process using QGIS' GRASS GIS integrated Douglas-Peucker simplification algorithm (v.generalize) [22]. However, a small tolerance (0.1) is applied during the generalization with the intent to reduce the excessive numbers of polylines' vertices in roundabouts and highway accesses, while preserving their road geometries [vi] (Figure 1). Despite being a small tolerance, this has significant effects, described in the results section, regarding reduction of road-elements' overall vertices number, when RCL graphs are converted into Angular-Segment maps using Depthmap X 0.5 [23]. This preprocessing step is a requirement for assessing *Betweenness Centrality* hierarchies through Space Syntax' Angular Analysis [7].
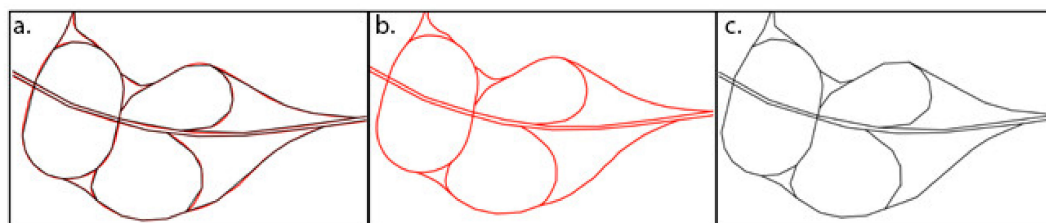

Figure 1. Comparison between (**a**) non-generalized (**b**) and generalized (**c**) road-circulation networks from Tuscan Region Road Graph.

Developed by Turner [7,24–26], the Angular Analysis consists of a set of methods conceived for assessing spatial network properties and predicting movement potentials on urban road-circulation networks constructed as dual graphs, such as RCL datasets [vii]. These methods require the conversion of RCL graphs into angular segment maps, in which the network *j-graph* is weighted according to the angle (in radians) amid each connected pair of vertices (in dual graphs, the road-elements),

attributing to them a correspondent angular coefficient. System depth is calculated from the shortest angular path among all origin-destination pairs of vertices in the network, therefore, this angular coefficient functions as a weighted topological step that allows for the measurement of *Betweenness Centrality* hierarchies through a simple count. Whenever there are angle variations amid a road pair—including any t-intersections or crossings between two streets—the RCL original continuous polyline will be segmented in two (angular segmentation), and the attribution of a corresponding angular value to each individual road will be ensured. Continuities will persist when no interruption or direction change happens.

Betweenness Centrality analyses are drawn from Space Syntax' Normalized Angular Choice (NACH), an algorithm derived from Angular Analysis [27]. Angular Choice corresponds to mathematical betweenness, since it counts (sums) the number of times each individual road segment *i* is traversed when travelling, through the overall shortest path, towards all potential destination segments, from all possible origin-destination pairs within the network. In this sense, this measure represents the segment through-movement potential or the probability of its use as a system *preferential route* [viii]. Without normalization, *Betweenness Centrality* absolute values are conditioned to the network size—thus, to its depth. The quantity of betweenness acquired by a segment *i* will be related to its integration/segregation and relative position within the network, with segregated designs adding more total and average betweenness to the system than integrated ones. Therefore, Angular Choice measures will highlight only the segments with the highest absolute *Betweenness Centrality* values—the main *preferential route*—while disregarding the rest of road-circulation network *preferential routes* hierarchies due to the high amplitude between top and bottom betweenness counts, resulting in an insufficient representation of through-movement dynamics. In this regard, Angular Choice is unsuitable to comparisons between different systems or scales. The normalization method devised by Hillier, Yang and Turner [27] mitigates this purported "*paradox of choice*" by weighting the calculated centrality values (Angular Choice—ACH) by the corresponding Angular Total Depth (ATD) of a segment *i*, as in the following equation:

$$\text{NACH} = \frac{\log(\text{ACH} + 1)}{\log(\text{ATD} + 3)} \tag{1}$$

With these refinements, centrality values are adjusted according to the Angular Total Depth of its corresponding segment *i*, therefore, distributing the network depth component. In this sense, the greater the network segregation, the more reduced Angular Choice absolute values will be, by being divided by a higher Angular Total Depth value. Normalization will then render visible the lesser *preferential routes*, attributing to them a relative positional weight within the network hierarchy. The inclusion of a logarithm in the equation standardizes the NACH relation into a power law, restricted between 0 and 1.5+ [ix], which deems possible the comparison of road-circulation networks with different depths, hence, non-identical sizes and scales, as their absolute *Betweenness Centrality* values are then brought to an equivalence.

NACH measures were modelled in DepthMap X 0.5 [23] for the 286 individual datasets, then loaded in GIS [22] and further graduated to render spatial networks visual similitudes noticeable. From these models, tests for *Betweenness Centrality* homothetic behavior were elaborated using the following scalar comparisons: regional with provincial (10 cases); regional with municipality (275 cases); and provincial with municipality (275 cases), for a total of 560 individual analyses. To be considered homothetic it was expected that the systems had, beyond perceived visual similitudes in road-circulation network centralities hierarchies' displacement, a minimal degree of correspondence regarding their centrality values. As such, hierarchies ought to repeat in the same position, with similar values, over all the established scales. A perfect homothetic behavior would be deemed true if r-square values were all equal to 1, however, this is unachievable for Angular Analysis' spatial models, as road-circulation networks are susceptible to the natural sensitivity of *Betweenness Centrality* values to discontinuities generated by artificial boundaries, manifested as *edge effects* [28]. To enact comparable thresholds for homothetic behavior definition, tolerance values counting for edge effects occurrences need to be

set. In this case, a minimum r-squared correlation value of 0.8 (Paretian Correlation) was established, meaning that for a system to exhibit significant homotheties, in addition to an observed spatial visual similitude, its numerical values for *Betweenness Centrality* correlation ought to be equal or over the 0.8 r-squared tolerance.

3. Discussion and Results

Prior to an in-depth discussion of *Betweenness Centrality* homothetic behavior in road-circulation networks, several considerations ought to be made regarding the issues that lead to the observation of this phenomenon: the configurational model computation times.

Even though Angular Analysis results in accurate representations for regional scale vehicular movement potentials [4,5,24] and has at its disposal a plethora of datasets, the mandatory angular segmentation preprocessing can be in some cases a hindrance, as it has evident negative impacts on RCL graph-based models computation times. Considering that these graphs use polylines to depict road-circulation networks, which tend to have numerous intermediate vertices to represent exact direction changes in curvatures, substantial numbers of individual road-segments (derived from road-elements vertices angular partition) are generated after preprocessing [x]. In large scale networks, those preprocessing effects are exponential, and time-related drawbacks are certain to emerge due to the recursive iterations required to calculate *Betweenness Centrality* for all origin-destination pairs. In this case, the road-circulation network generalization was not a choice, but an essential step.

Such issues can be further exemplified when non-generalized and generalized Tuscany road-graphs are preprocessed and their computation time-lapses analyzed [xi]. Angular segmentation of the non-generalized regional graph resulted in an individual road-elements (road-segments) increment of over 1000% (from 380,136 to 4,429,530) (Table 1), and in an impractical analysis prospect, given that the processing time-lapse estimation for this model surpassed 7.5 months (Table 2). In this context, even a small tolerance should have provided substantial reductions on road-segment quantities after preprocessing, as most polylines' mid-points could be eliminated without significant changes to the road morphologies (see Figure 1). For instance, while GIS-based Douglas-Peucker (DP) minimal possible tolerance (0.1) for generalization did not impacted total road-element quantities prior to preprocessing (380,136), overall reduction in post-processed road-segments was notable, as their quantities were diminished by 72,34% (4,429,530 to 1,225,140) (Table 1).

Table 1. Comparison between non-generalized and generalized Tuscan Road Graph number of road-elements (road-segments) prior and after angular segmentation pre-processing.

	Number of Road-Elements Prior Pre-Processing	Number of Road-Elements after Pre-Processing	Δ%
Non-Generalized	380,136	4,429,530	1065.25%
Generalized (DP 0.1)	380,136	1,225,140	222.29%
Δ%	0.00%	−72.34%	-

Table 2. Average approximated computation time-lapses for regional, province and municipality *Betweenness Centrality* models for non-generalized and generalized road-circulation networks [xii].

	Non-Generalized (4,429,530 Elements)	Generalized (DP 0.1) (1,225,140 Elements)
Region	≈7.5 Months	≈2.5 Months
Provinces (Average)	≈23–72 h	≈7–18 h
Municipalities (Average)	≈2–8 min	≈0.5–3 min

Estimated time-lapse to regional road-network graph analysis completion was set to around 2.5 months after this generalization process, which was a noticeable improvement when compared to the previous timestamp of 7.5 months. Substantial reductions in processing times were also attained for province and municipality road-graph [xiii] models, as time-lapses after generalization

were brought down to less than a day and from seconds to a couple of minutes depending on size, which allowed expeditious modelling of these datasets. (Table 2). Despite decreases in processing time for municipalities, due to the elevated number of cases (275) and batch process execution impracticality [xiv], it is still less time-consuming to model provincial expanses (10 cases).

Although overall computation times fall after the generalization process, a model completion time-lapse of 2.5 months is still quite high. An option to further mitigate this issue would be an increase in the regional graph degree of generalization, through the Douglas-Peucker simplification algorithm tolerance rise [xv] and the exclusion of less important road-elements. While possible when a single scale assessment is the objective (i.e., Serra and Hillier [4]), since in these cases the skeletal road-graphs can provide a sufficient depiction of systems' configurational properties, this course of action would be detrimental to accuracy when multiscale comparisons ought to be considered. In this sense, model incompleteness tends to be more problematic for those kinds of analyses than lingering edge effects, since partial data representation across scales can bias road-circulation network core centrality hierarchy distribution logics, given that the systems' will be essentially different.

The identification of a recurrent homothetic behavior in road-circulation networks, however, may provide new dimensions to these time-lapse/completeness issues of regional and multiscale analyses in terms of *Betweenness Centrality* computing and representation. As a defining characteristic, homotheties occur when two objects—or parts of those objects—maintain their proportions and retain a visual similitude, whenever scaled by a given factor. For networks, this homothetic property would imply that a section of the system, when modelled independently, should retain a certain regularity in the distribution of its centrality hierarchies' (a visual similitude) whenever compared to a larger section of the network, independently modelled as well, that contains it. Hence, the system would maintain its overall relative proportions (configuration, hierarchies and value). While certain conditions (i.e., congestion, temporary road obstructions, increases in travel-time due traffic) could alter overall real-time traffic flow allocation for each road-element, and therefore change systems' actual *preferential routes*, this does not invalidate the NACH model as a predictive instrument for the network general movement demands in normal conditions. In this regard, the existence of a homothetic behavior in road-circulation networks would indicate that no significant changes in the movement demand patterns—changes in *preferential routes* hierarchy—happen across different scales.

Evidences of homothetic behavior were found when visual similitudes were noticed during the assemblage of Tuscany's provincial and municipal *Betweenness Centrality* models of their road-circulation network in a GIS-based mapset (Figures 2 and 3).

It was then observed that, apart from road-elements located in those systems outer borders [xvi], centrality hierarchies shared an analogous distribution logic in both scales, being located in the same groups of road-elements. In this instance, municipal *preferential route* hierarchies are substantially similar to those patterns visualized in the respective section of the provincial model, despite a distinct overall configurational logic in the system remainder.

Visual similitudes repeated themselves, in greater or lesser extent, over all Tuscany province-municipality comparison cases, and prompted further tests to assess if centrality hierarchy values remained likewise consistent throughout those scales, or if the similitudes were restricted to the visual representation. These tests evaluate the centralities values distribution correspondence degree amid selected scales, being accomplished through correlations between the NACH models. In this case, the r-squared values represent the similarity degree amid provinces and municipalities centrality hierarchy, both numerical and positional, the latter being verified in the correlation graphics through point displacement. Values which surpass the established threshold of 0.8 signal that both network models have significant numerical likeness and therefore can be considered homothetic if visual similitudes are also present—as 0.2 differences are attributed to edge effect [28] distortions [xvii]. Results are aggregated by individual provinces (province) and indicate the average correlation values (r-squared) between all Tuscan municipality road-circulation network models and their respective section in the province road-circulation model (Table 3).
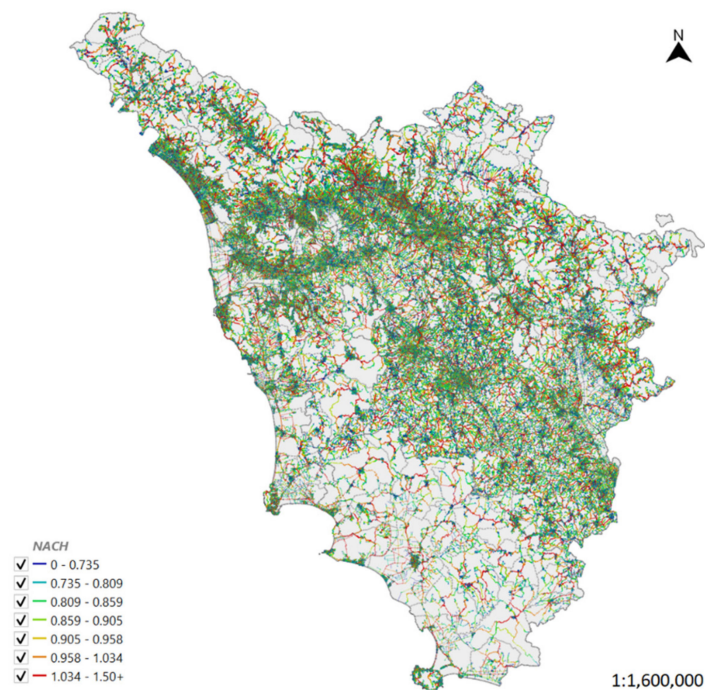
Figure 2. Tuscan Region collective of Municipalities *Betweenness Centrality* (NACH) models.
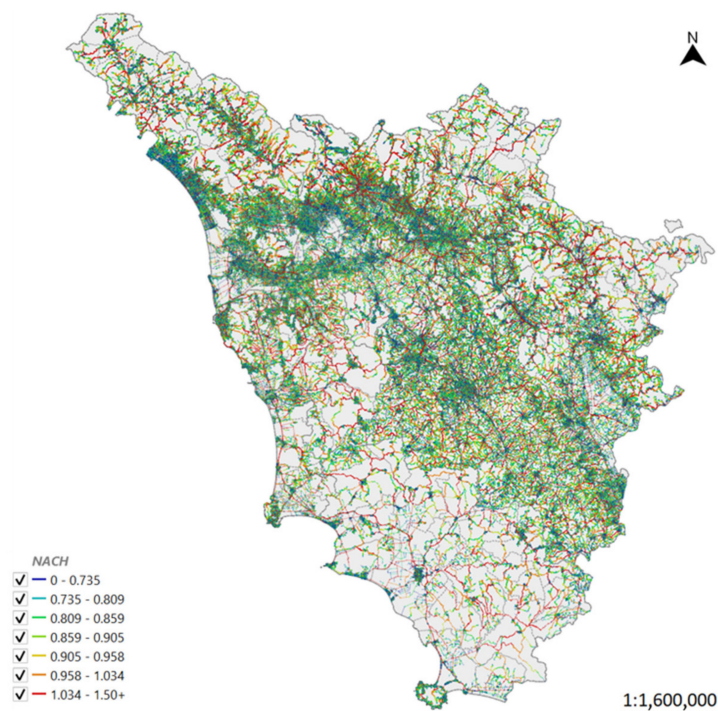

Figure 3. Tuscan Region collective of provinces *Betweenness Centrality* (NACH) models.

Average values for municipality models' correlations with their respective provincial counterparts remain well above the established r-squared threshold (>0.8) and denote that models' centrality hierarchies have an overall similar *Betweenness Centrality* distribution across the scales. While this alone implies strong homothetic behaviors amid these networks, further aspects can be assessed through the correlation analysis graphic examples for Grosseto and Prato municipalities (Figure 4) and observed in their respective provincial and municipal network systems visual similitudes (Figures 5 and 6).

Table 3. *Betweenness Centrality* (NACH) average correlation values (r-squared) and statistics between Province and Municipality road-circulation network models.

| Province × Municipalities—Correlation Data	Analyses	Averages	St. Dev	St. Error
Provincia di Arezzo	38	0.8672	0.0443	0.0072
Provincia di Firenze	42	0.8904	0.0441	0.0068
Provincia di Grosseto	27	0.8514	0.0641	0.0123
Provincia di Livorno	11	0.9003	0.0384	0.0116
Provincia di Lucca	37	0.8685	0.1833	0.0301
Provincia di Massa-Carrara	18	0.8905	0.0591	0.0139
Provincia di Pisa	38	0.8948	0.0405	0.0066
Provincia di Pistoia	22	0.9145	0.0253	0.0054
Provincia di Prato	7	0.9156	0.0316	0.0119
Provincia di Siena	35	0.9046	0.0392	0.0066
Total	275	-	-	-
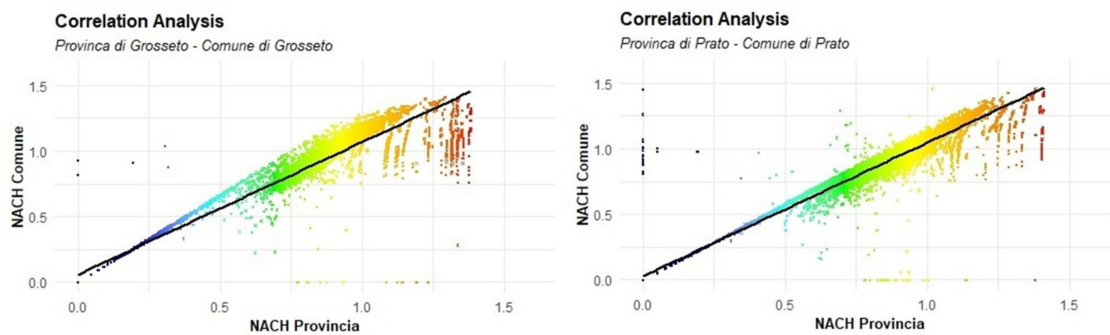
Figure 4. *Betweenness Centrality* values correlation graphics between Municipal and Provincial NACH models for Grosseto and Prato.
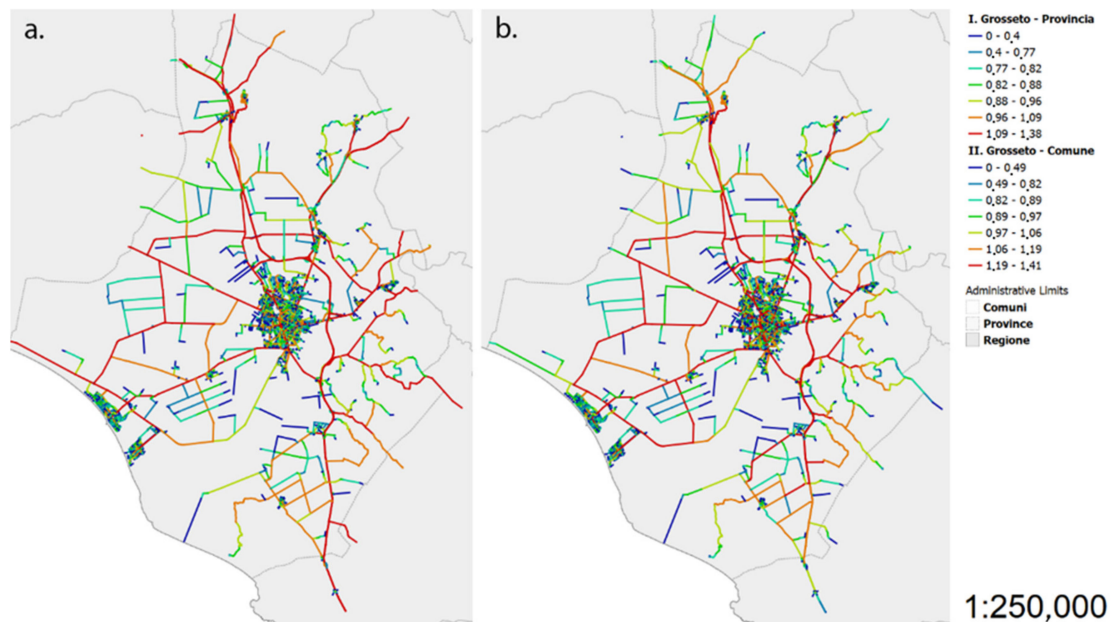

Figure 5. Comparison of *Betweenness Centrality* hierarchy distribution between the Grosseto Province (Provincia) (**a**) section and the Grosseto Municipality (Comune) (**b**) road-circulation network models (NACH).

Figure 6. Comparison of *Betweenness Centrality* hierarchy distribution between the Prato Province (Provincia) (**a**) section and the Prato Municipality (Comune) (**b**) road-circulation network models (NACH).

While both graphics have similar comportment regarding centralities, as most of their values tend to converge in a linear distribution, which highlights the network homotheties existence, higher centrality hierarchy values (red and orange) exhibit somewhat divergent displacements, phenomena that are repeated in greater or lesser degree for all the remaining correlation graphics. This perturbation is attributed to edge effects, since values shift toward provincial NACH models, which not incur in these effects when sectioned at this scale due to network continuities. Another factor is that perturbations are more pronounced on Grosseto (Figure 4), which, when compared to Prato (Figure 5), has a prevalence of long and sparsely connected road-segments near the borders, endorsing that these kinds of configurations suffer greater *Betweenness Centrality* value decrements derived from edge effects, and explains differences on the homothety degree across the networks.

Following regional model completion, tests for *Betweenness Centrality* homothetic behavior were expanded in order to assess correlation between centrality values' distribution amid municipalities and the whole regional road-circulation network. Furthermore, these tests were compared to results obtained for the province-municipalities correlations and their analyses reviewed to investigate if the homothetic behavior is recurrent throughout scales. As before, the results are organized based on the province in which each municipality is placed. However, for these cases, they indicate the average correlation values (r-squared) among all Tuscany municipalities' road-circulation network models and their respective section in the regional road-circulation model (Table 4).

Even though region-municipality correlation averages for r-squared suffer an overall decrease in each province, when compared to their respective pairs from province-municipality correlation averages (Table 3), their values are still within the established threshold (>0.8). These results suggest that the observed homothetic behaviors continue to happen across scales, and with relative regularity, since most values stand substantially close to province-municipality numbers, which indicates that most centrality hierarchy cores tend to maintain placement correspondence. In this aspect, despite value differences, the centrality distribution logics remain quite similar to those verified in the assessment of province-municipality numbers, as correlation graphics conserve their tendencies of linear convergence for lower centrality values and maintain predominant divergent comportments for the higher centrality values. This can also be addressed through the correlation graphic examples for Grosseto and Prato municipalities (Figure 7).

Table 4. *Betweenness Centrality* (NACH) average correlation values (r-squared) and statistics between Region and Municipality road-circulation network models.

| Region × Municipalities—Correlation Data	Analyses	Averages	St. Dev	St. Error
Provincia di Arezzo	38	0.8423	0.0453	0.0074
Provincia di Firenze	42	0.8685	0.0479	0.0074
Provincia di Grosseto	27	0.8167	0.0750	0.0144
Provincia di Livorno	11	0.8513	0.0470	0.0142
Provincia di Lucca	37	0.8577	0.1616	0.0266
Provincia di Massa-Carrara	18	0.8583	0.0605	0.0143
Provincia di Pisa	38	0.8660	0.0539	0.0087
Provincia di Pistoia	22	0.8905	0.0274	0.0058
Provincia di Prato	7	0.8476	0.0514	0.0194
Provincia di Siena	35	0.8862	0.0446	0.0075
Total	275	-	-	-
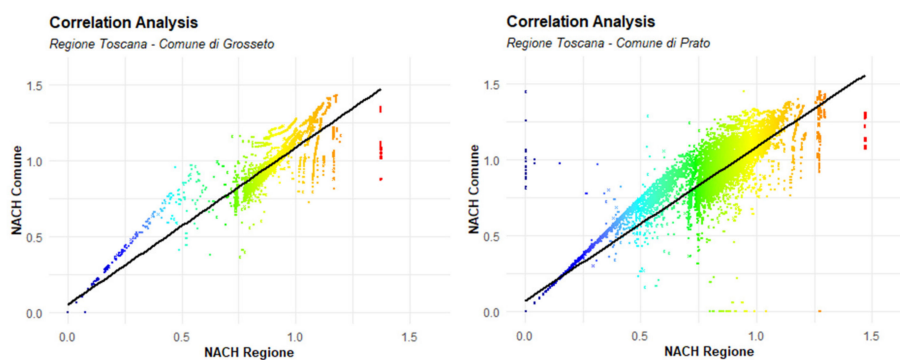
Figure 7. *Betweenness Centrality* value correlation graphics between Municipal and Regional NACH models for Grosseto and Prato.

Investigation into the value decrements in regional-municipality average correlations amid NACH models reveals that these can be attributed to general increases in centrality values for the regional road-circulation network. Those derive from system continuities (network wholeness), as value increments tend to happen, above all, in *preferential routes* with medium and high hierarchies. Considering that edge effects are stronger on high value road-elements near the borders, the greater differences between the regional and municipalities scale models' degrees of continuity tend to be the causes of greater value/hierarchy distortions. This can be well observed in the municipalities from Livorno province (Figure 8), which have an important regional *preferential route* located close to their provincial boundaries and in-between its municipalities' borders.

Results for region-provinces correlations demonstrate a remarkable correspondence between these NACH models, as the r-squared values for the 10 cases not only surpassed the established significance threshold (>0.8), but are also all set above the 0.9 mark, with averages around 0.96 (Table 5), which indicates noteworthy overall correspondence between the models.

Compared to the previous analyses for region-municipality correlations, the homothetic behavior verified among provinces models and their regional model counterparts is far more noticeable. These regularities are exemplified in the centrality hierarchy distribution of Lucca province—highest correlation—and Prato province—lowest correlation—(Figures 9 and 10), both of which have substantial visual similitudes with their regional model section. Distortions can, however, be visualized and attributed to edge effects, which, while minimal, are the main causes of region-province model NACH value differences. Configurational characteristics are responsible for the overall increases in correlation as, although in greater quantities, due to provinces' network size, the road-elements near the borders tend to be shorter when compared to their inner continuities, and are also considerably shorter in

comparison to the road-segments at the municipality scale, hence these "more compact" systems exhibit a higher degree of homothety.

These correlation tests revealed that apposing provincial scale networks provided a much closer depiction of the regional road-circulation network *preferential routes* than the collective of municipality models. Hence, this indicates that even though the homothetic behavior exists and is recurrent across scales, using closer scales provides better results as it diminishes the only perceived source of distortion in these networks: the edge effects. While it is possible to confirm the main hypothesis with these results —that the recurrent homothetic behavior can be used as an alternative to road-circulation network representation of large scale systems in order to diminish processing times—further examinations of the possible outlier cases are needed to ascertain where this behavior happens inconsistently.

Standard deviations (Tables 3 and 4) indicate which provinces have wider differences among municipality r-squared values for each scale. These deviations were used to identify eventual discrepancies in homothetic behavior, individuate the outliers, and analyze their particular cases. Overall values, however, tended to be small, which indicates relative concentration of values around the averages and the absence of significant outliers in most cases, thus confirming regularities in behavior. Organizing the data for the municipality correlation degree for both scales in graphics and allocating them to their respective value range aids in the visualization of the instances in which r-squared values are not inside the established correlation threshold (>0.8) (Figure 11).
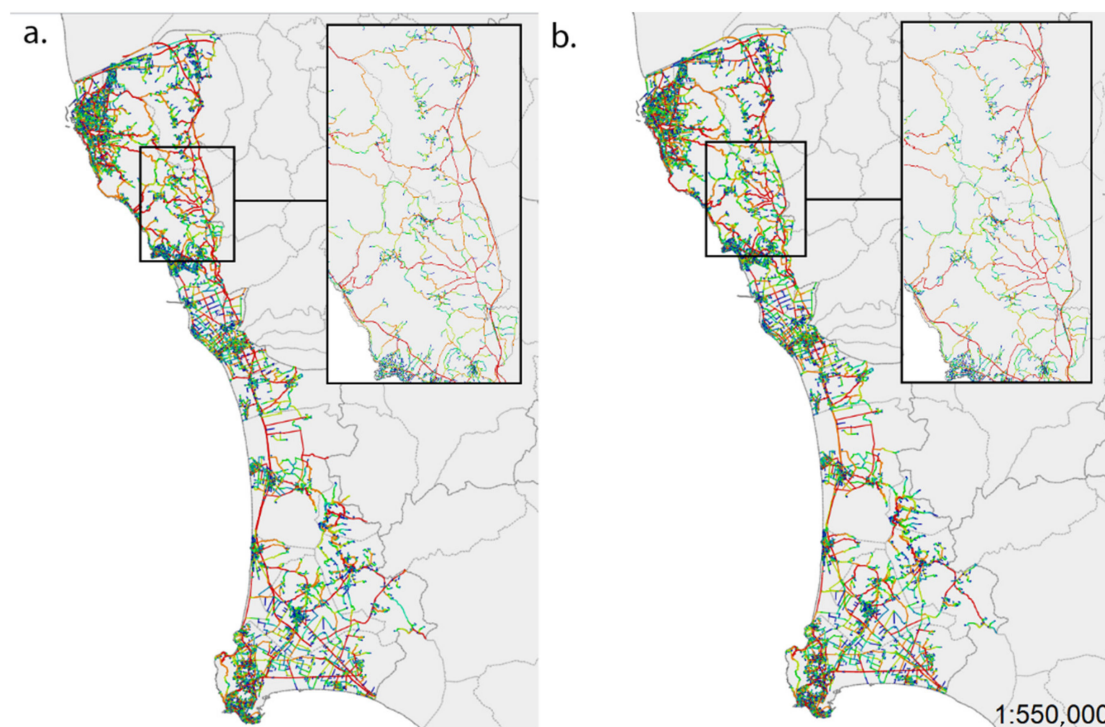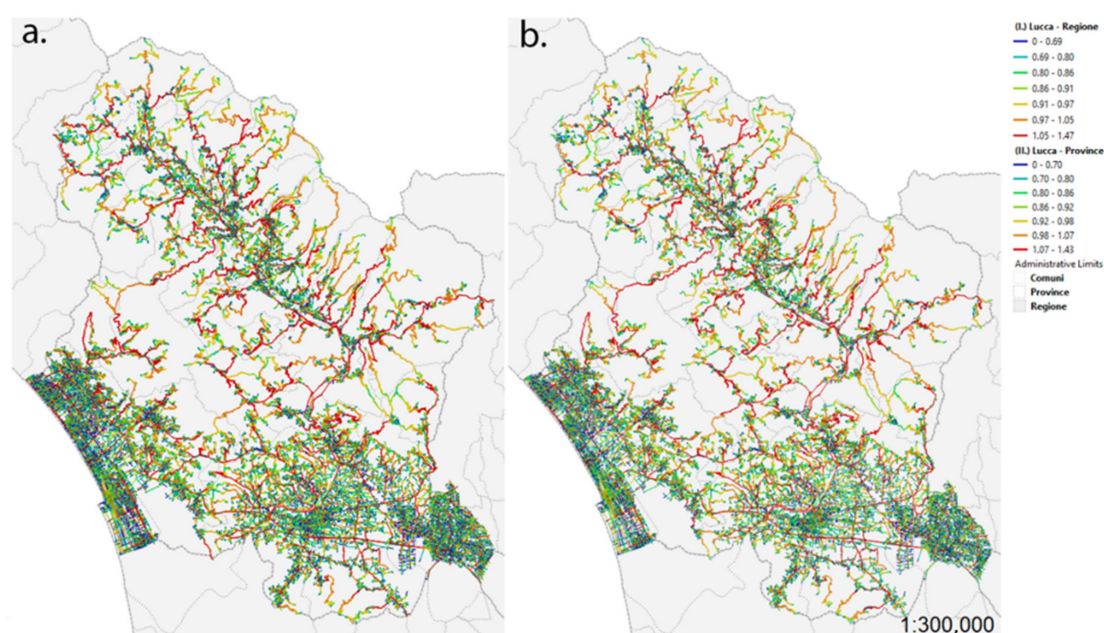

Figure 8. Comparison of *Betweenness Centrality* hierarchy distribution amid the regional section (**a**) and the municipality (**b**) road-circulation network models of the Livorno province.

Lucca province exhibits a higher standard deviation when compared to the remaining provinces, a situation that is repeated for both region-municipality and in province-municipality cases, and in the latter case it is the sole province to have r-squared values under the 0.7 range. These results indicate the presence of outliers and prompted a further analysis of their particular configuration. Tables below indicate the municipality maximum (Table 6) and minimum (Table 7) r-squared values for correlation. No significant outliers were found for the maximum values.

Table 5. *Betweenness Centrality* correlation between Regional and Provincial road-circulation network models (NACH).

Regione × Province—Correlation Data	Values	Statistics	
Provincia di Arezzo	0.9685	Average	0.9598
Provincia di Firenze	0.9772	St. Dev	0.0210
Provincia di Grosseto	0.9583	St. Error	0.0066
Provincia di Livorno	0.9403	Maximum	0.9821
Provincia di Lucca	0.9821	Minimum	0.9107
Provincia di Massa-Carrara	0.9649	Highest Correlation	Provincia di Lucca
Provincia di Pisa	0.9463	Lowest Correlation	Provincia di Prato
Provincia di Pistoia	0.9693		
Provincia di Prato	0.9107		
Provincia di Siena	0.9807		

Figure 9. Comparison of *Betweenness Centrality* hierarchy distribution between the Regional section (**a**) and the Lucca Province (**b**) road-circulation network models.

Minimum r-squared values (Table 7), demonstrate that overall lower correlations, especially in the province-municipalities cases, do not deviate much from the 0.8 threshold. When verified, these systems tend to exhibit extensive road-elements near the borders and are more susceptible to edge effects [28], that are affected differently depending on their model scale, as can be observed in the Montemignaio municipality (Arezzo province) example (Figure 12).
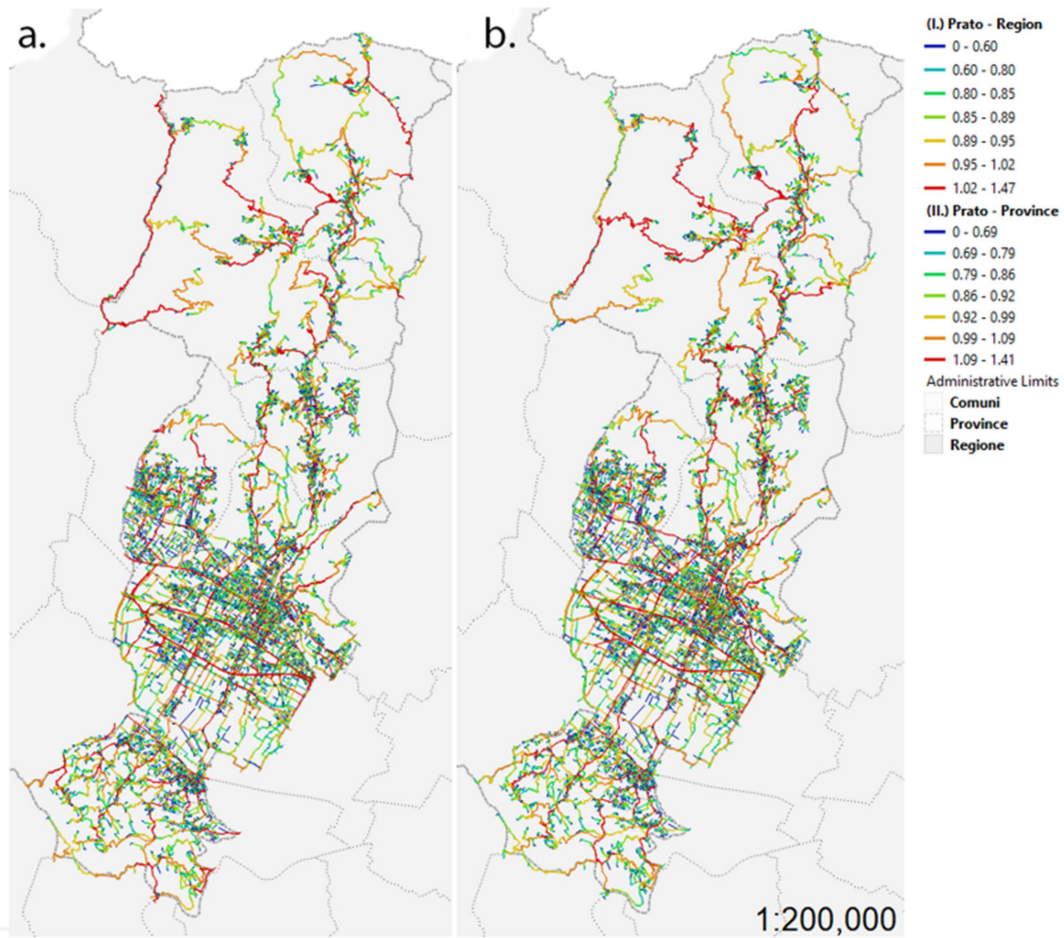
Figure 10. Comparison of *Betweenness Centrality* hierarchy distribution between the Regional section (**a**) and the Prato Province (**b**) road-circulation network models.
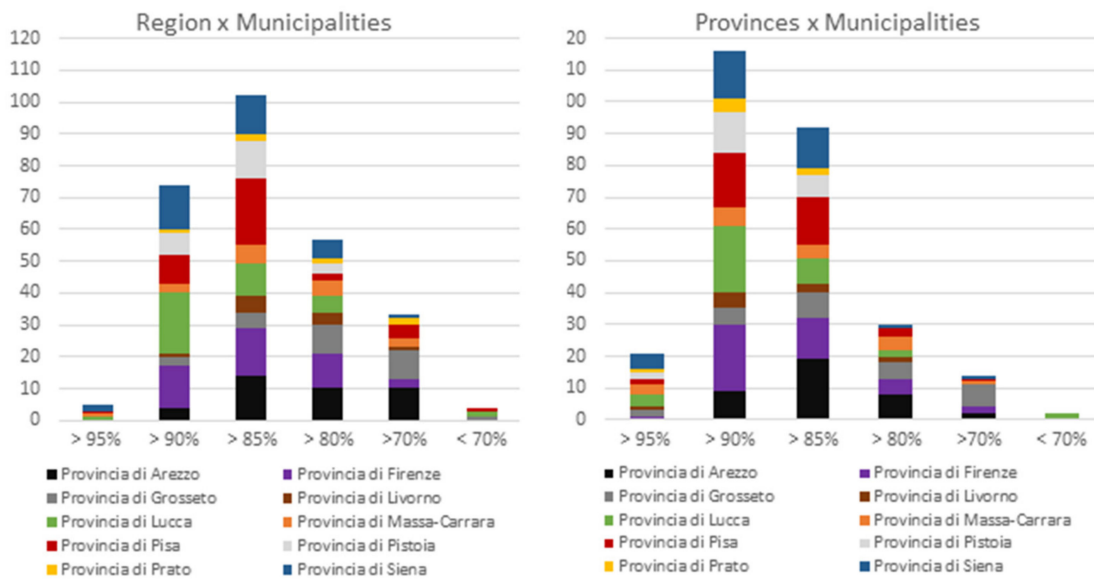

Figure 11. Distribution of municipality r-squared values according to established ranges for region-municipality and province-municipality NACH model correlation.

Table 6. *Betweenness Centrality* correlation between Regional and Provincial road-circulation network models (NACH).

Province	Max.—Region × Municipality	Value	Max.—Province × Municipality	Value	Δ
Arezzo	Arezzo	0.9177	Arezzo	0.9492	0.0315
Firenze	Dicomano	0.9448	Dicomano	0.9528	0.0080
Grosseto	Monte Argentario	0.9433	Monte Argentario	0.9735	0.0300
Livorno	Piombino	0.9407	Piombino	0.9735	0.0328
Lucca	Altopascio	0.9504	Altopascio	0.9618	0.0114
Massa-Carrara	Massa	0.9531	Massa	0.9742	0.0211
Pisa	Cascina	0.9504	Cascina	0.9633	0.0129
Pistoia	Pistoia	0.9466	Pistoia	0.9720	0.0087
Prato	Vaiano	0.9003	Montemurlo	0.9565	0.0068
Siena	Siena	0.9583	Siena	0.9660	0.0077

Table 7. *Betweenness Centrality* correlation between Regional and Provincial road-circulation network models (NACH).

| Province	Min.—Region × Municipality	Value	Min.—Province × Municipality	Value	Δ
Arezzo	Montemignaio	0.7614	Montemignaio	0.7157	−0.0457
Firenze	Sesto Fiorentino	0.7227	Sesto Fiorentino	0.7522	0.0295
Grosseto	Magliano in Toscana	0.6896	Cinigiano	0.7426	0.0530
Livorno	Sassetta	0.7456	Sassetta	0.8408	0.0952
Lucca	Stazzema (Nord)	0.1668	Seravezza (Nord)	0.0320	−0.1348
Massa-Carrara	Zeri	0.7302	Mulazzo	0.7917	0.0615
Pisa	C. di Val di Cecina (Nord)	0.6909	C. di Val di Cecina (Nord)	0.7478	0.0569
Pistoia	Serravalle Pistoiese	0.8440	Serravalle Pistoiese	0.8711	0.0271
Prato	Carmignano	0.7518	Carmignano	0.8702	0.1184
Siena	Radicofani	0.7420	Radicofani	0.7719	0.0299
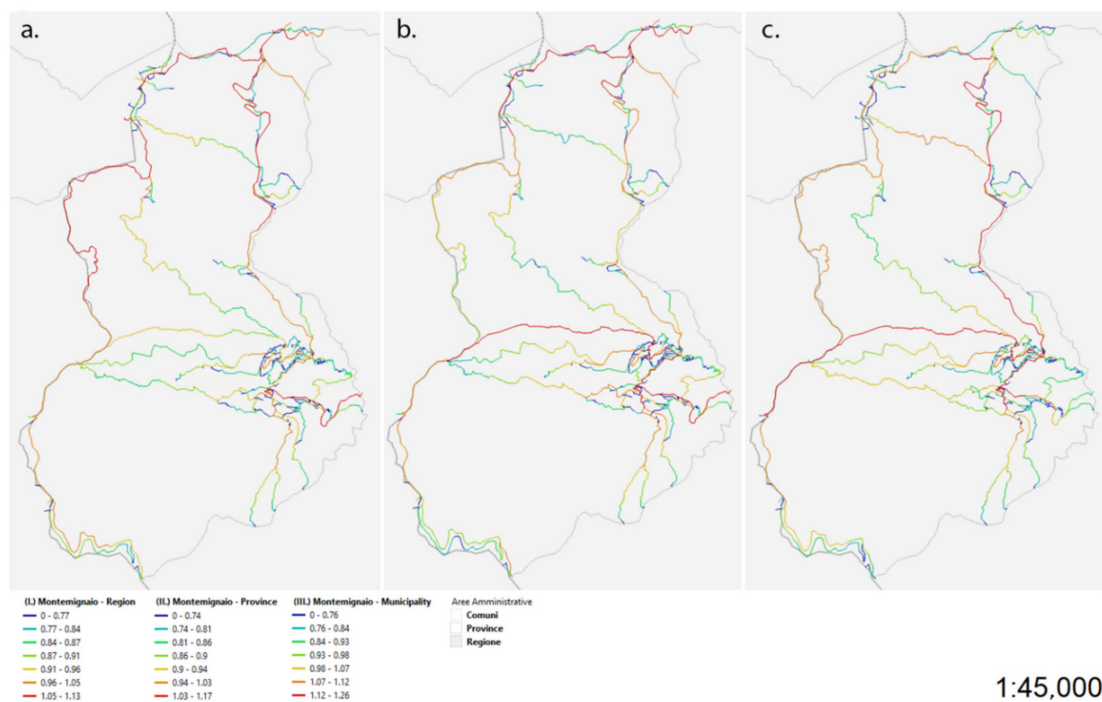
Figure 12. Comparison of *Betweenness Centrality* hierarchy distribution of Montemignaio Municipality for the Regional (**a**) Province (**b**) and Municipality (**c**) road-circulation networks models.

The case of Lucca province, nonetheless, is quite different from the other outliers, because in both comparison scales, the municipalities of Stazzema (Nord) in the region-municipality case, and Seravezza (Nord) for the province-municipality case, exhibit minimum r-squared values that are significantly lower than their pairs. These are in fact very particular situations, because both of these road stretches

were modelled as discontinuous parts of the main municipal road-circulation network, due to the fragmented territorial division of the municipalities, as can be observed in Figure 13, since Stazzema (Nord) is a continuity of Seravezza and Seravezza (Nord) a continuity of Stazzema.
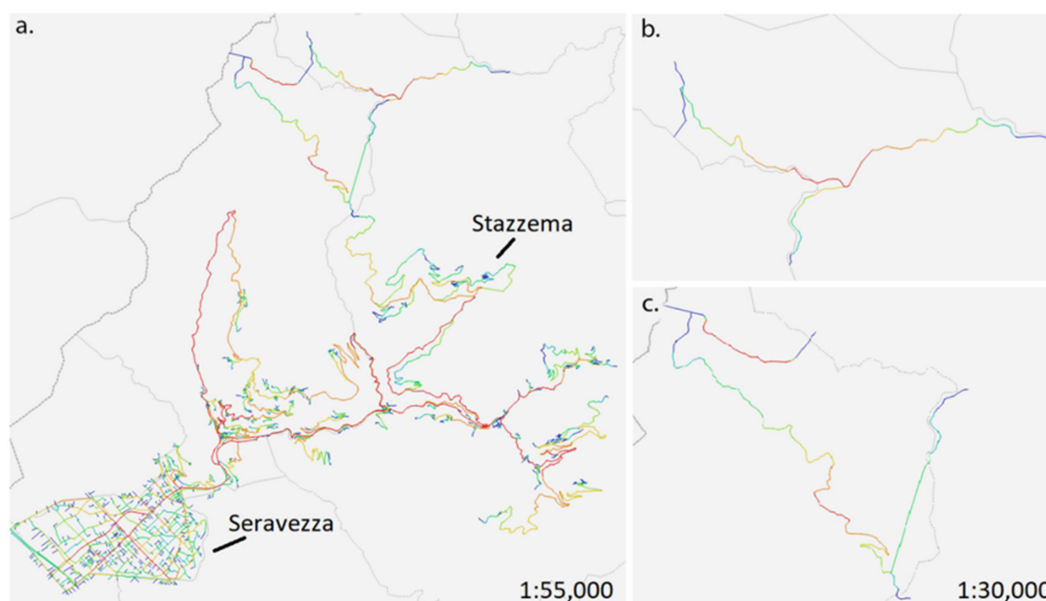

Figure 13. Discontinuities and road-circulation network characterization (**a**) of Stazzema (Nord) (**b**) and Seravezza (Nord) (**c**) *Betweenness Centrality* models.

In these cases, no homothetic behavior (visual similitude) can be perceived in relation to regional or provinces models, as these stretches are sparsely connected and segregated, being placed near system's borders. *Betweenness Centrality* verified for these systems, therefore, behave as local centralities, not being related to the remainder of the system. While this does not invalidate exploring homothetic behaviors as a means to represent large regional road-circulation network configurational properties in shorter spans of time and with reasonable accuracy, it prompts further studies which investigate the reasons for this occasional behavior.

4. Conclusions

Large spatial network analyses and the representation of regional road-infrastructures are scale associated problems. In these instances, scale ought to be interpreted as time, since the models' processing pace—and, therefore, the cartographic data availability for further assessments—often be constrained by network dimension.

While configurational analyses, such as the in-vogue Space Syntax' Angular Analysis, are able to construct models with accurate and detailed depictions of road-circulation network properties—such as centrality hierarchies—their methods are not tailored to model region-wide areas. Being focused on the representation of plazas, cities and other small-scale networks, Space Syntax's developers did not consider these possible time-scale related issues whilst developing its instrumentation. In this sense, Angular Analysis falls under the inevitability of network complexity reduction and requires, as the usual course of action, ulterior adjustments to the available Road-Centre Line (RCL) databases, in the form of generalization processes, to enable efficient spatial modelling. Without a doubt RCL graph generalization is an important and sometimes indispensable step for these network-based analyses, as a reduction in often-excessive polyline vertices numbers will considerably impact computation time-lapses, above all in Space Syntax, due to the angular segmentation preprocessing required by Angular Analysis. However, extensive generalizations of spatial data are accompanied by a *trade-off* regarding network incompleteness, which is significant whenever the assessment of accurate

relations among different territorial scales is the objective. This, in the end, hinders one of the most distinguished features of Angular Analysis—its accurateness in the cartographic representation of configurational properties alongside the real morphologies of urban network-based systems. From this perspective, the assessment of interactions between cities and regions, an important component for sustainability analysis at the urban scale, becomes unidimensional, given the inability of the analysis to consider and evaluate the whole territorial configurational context in which the cities are embedded. Further improvements in the representation of local-regional relations and road-circulation network representation can serve as guidance for tailored territory-oriented policies that assess economic activity displacement and resilience. Understanding regional movement demands and apposing them to spatial data and models that can assess movement logics related to supply—such as travel-time analysis, activities placement and infrastructure characteristics—enable to precisely determine the location of territorially exposed places. Those areas are susceptible to the under-usage of building stock, or abandonment and divestment conditions, urban phenomena observed and faced in many Tuscan municipalities and, as well, throughout other Italian regions.

Exploring road-circulation networks homothetic behavior, a property found in *Betweenness Centrality*-based NACH models, could be an alternative to extensive generalization for large-scale networks, and may provide more accurate scale-comparable representations. This property suggests that visual similitudes (homotheties) are present in *Betweenness Centrality* (NACH) models across scales. Thus, road-circulation networks will possess similar centrality hierarchy distributions and values, forming an actual image of a map section when juxtaposed to a network of greater or lesser scale. The reduction in processing time-lapses that can be attained, due to the use of smaller-sized networks, will be equivalent, or even greater than the gains made through the use of extensive generalization, without the morphological distortions derived from over-simplification.

Correlation tests demonstrated that homothetic behavior occurs with relative regularity across all analyzed territorial scales. In these instances, r-squared correlations indicate that the centralities' distribution in municipalities models tends to be similar, both in position and in value, to their respective sections in provincial and regional road-network models. The same pattern was also observed—and with more evident correlations—for provinces and their respective sections of the regional network. Even though the attained correlations denote robust homothetic behavior, distortions are still present, as no r-squared value is equal to one (perfect homotheties). Investigations about the nature and causes of these homothetic behavior asymmetries amid models revealed that those come, in greater or lesser degree, from edge effects, an implication from the network discontinuities close to systems' borders, that result from the creation of artificial boundaries—in this case, the section made according to the administrative limits. While edge effects somewhat impact model accurateness, these distortions do not invalidate the homothetic behavior logic observed in core centralities, since beyond edge effects area of influence, the immediate border, models remain substantially similar. Hence, consideration of network homotheties as an alternative to large scale network analyses may allow the construction of region-wide *Betweenness Centrality* cartographical representations with smaller sections of maps—like pieces in a puzzle—with comparable smaller processing times and higher relative detail.

Further research is required to better understand the formal or mathematical causes that allow *Betweenness Centrality* to exhibit this homothetic behavior in networks, and to determine whether this property can be used or reproduced for other kinds of centrality measures based on sum, or even for other types of spatial networks. Other research efforts could be directed into the conception of an algorithm capable of compensating for the distortions caused by edge effects on these networks, a problem that, if solved, could aid the exploration of original concepts on fields of theoretical network analysis and applied mathematics. Future studies could also employ the proposed methods—alongside other spatial datasets—as guidelines for the recommendation of territorial policies, which constitute crucial developments for applied research.

Notes:

[i]　Centrality hierarchies can predict movement or flow in an urban settlement, important for assessing aspects of human relationships [9] or even economic activities distribution.

[ii]　Angular Analysis [7] has a preprocessing stage called angular segmentation, that serves to improve the level of derailment of centrality measures, allowing its calculation with respect to each road-element of the network. However, this process also greatly increases the total quantity of road-elements, requiring generalization to enable processing in reasonable time-lapses

[iii]　In Euclidean geometry this factor λ is called the similitude ratio and determines the degree of magnification of a group of vectors [16].

[iv]　It is important to remark that province and municipality sections are not territorially strict, because for network analyses the natural graph continuities ought to be respected over administrative limits to ensure system wholeness. Groups of road-elements were maintained whenever their removal would cause a network gap, even when comprised in a neighboring territorial unit. This adjustment was used mainly where road-elements had small segments in another municipality territory.

[v]　Tuscany has 273 municipalities, counting the Tuscan archipelago. Some mainland municipalities had to be sectioned in two, due to network discontinuity, a consequence of Tuscany's fragmented territorial division.

[vi]　Constructed map differs from the UK Ordnance Survey graphs used in recent studies regarding large scale regional network analyses [4] as this generalization does not reduce it into a skeletal RCL, which converts the roundabouts and interchanges to single line crossings. Hence, the simplification method used results in a more accurate regional road-circulation network representation with better comparability on smaller urban scales.

[vii]　Turner [24,26] proposes using Angular Analysis and RCL graphs to represent movement potentials, since those datasets are readily available throughout countries' GIS suites, and therefore do not require time consuming previous construction as, at the time preferred, Space Syntax' axial maps. These models resulted in more accurate spatial representations of road-circulation networks morphology than axial analysis models.

[viii]　NACH is a probabilistic index of "*movement demand*" for a road-network, thus, it highlights *preferential routes* only considering the overall system configuration, not accounting for real-time conditions i.e. congestion, travel time. Still, tests indicate significant correspondences amid NACH estimations and real-time movement [4].

[ix]　Hillier, Yang and Turner [27] theoretical tests state that *Betweenness Centrality* values can sometimes surpass the 1.5 threshold in specific cases, where there are ample differences between mean and total depth.

[x]　The same thing occurs with orthogonal grids, a common pattern on planned urban settlements, as the t-intersections are split into multiple road-elements. Analysis of orthogonal grid structures was one of the motivations behind angular segmentation preprocessing and angular analysis development [7].

[xi]　Preprocessing RCL graphs to angular segment maps in DepthMapX 0.5 [23] is often instantaneous and takes no more than a couple of seconds even when large road-networks are considered. Therefore, it is simple to verify total road-element numbers after angular segmentation and estimate time-lapse of completion.

[xii]　(Estimated in DepthmapX 0.5 [23], modelling of closeness centralities also included. Processing time-lapse values for an overclocked Intel i7—8700k (4.7 GHz); 16 GB of RAM.

[xiii]　(Province road-graphs, after generalization, ranged from 80,000 to 320.000 elements following preprocessing, a proportional reduction, since preprocessed non-generalized graphs possessed from 140,000 to 540,000 elements. Municipality road-graphs also exhibited the same logic.

[xiv]　It is possible to run multiple models at the same time in DepthMapX 0.5 [23], however, each dataset must be imported and preprocessed individually prior to Angular Analysis, which is per se time-consuming.

[xv]　Higher tolerances (DP 0.5 and DP 1) were tested and, while the post-processing segment number reductions after preprocessing provided a reasonable decrement in processing times-respectively 1.2 months and 0.7 months, the distortions in the road-circulation network morphologies were quite significant, even in greater scales.

[xvi]　*Edge effects* [28] seem to be responsible for these differences in centralities values along the borders, which happen due to the network discontinuity caused by the road-graph section. Under the NACH algorithm, however, high centrality values tend to be more sensitive to these effects, with decrements in value being greater on edge road-elements that are long and sparsely connected.

[xvii]　While 0.8 (the Paretian correlation) is attributed as a minimum value correspondence to verify homotheties, as under this value core centralities tend to have significant changes; ideal values should be in the range between 0.85 and 0.99. Perfect homotheties (1.0) are impossible in road-circulation networks due to lingering edge-effects.

References

1. Jiang, B.; Claramunt, C. Integration of Space Syntax into GIS: New Perspectives for Urban Morphology. *Trans. GIS* **2002**, *6*, 295–309. [CrossRef]
2. Douglas, D.; Peucker, T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Can. Cartogr.* **1973**, *10*, 112–122. [CrossRef]
3. Jiang, B.; Claramunt, C. A Structural Approach to the Model Generalization of an Urban Street Network. *GeoInformatica C* **2004**, *8*, 157–171. [CrossRef]
4. Serra, M.; Hillier, B. Angular and Metric Distance in Road Network Analysis: A nationwide correlation study. *Comput. Environ. Urban Syst.* **2019**, *74*, 194–207. [CrossRef]
5. Hanna, S.; Serras, J.; Varoudis, T. Measuring the Structure of Global Transportation Networks. In Proceedings of the 6th International Space Syntax Symposium, Istanbul, Turkey, 12–15 June 2007; Kubat, A., Ed.; Istanbul Technical University: Istanbul, Turkey, 2007.
6. Braga, A.; Zampieri, F.L.; Fauri, A.; Campos, H.Á.; Ruckert, A. Plural territorial integration and road network configuration in South America. In Proceedings of the 11th International Space Syntax Symposium, Lisbon, Portugal, 3–7 July 2017; Heitor, T., Ed.; Instituto Superior Técnico: Lisbon, Portugal, 2017; pp. 1–17.
7. Turner, A. Angular Analysis. In Proceedings of the 3rd International Symposium on Space Syntax, Atlanta, GA, USA, 7–11 May 2001; Peponis, J., Wineman, J.D., Bafna, S., Eds.; Georgia Institute of Technology: Atlanta, GA, USA, 2001; pp. 1–11.
8. Hillier, B. *Space Is the Machine: A Configurational Theory of Architecture*; Space Syntax Ltd.: London, UK, 2007. [CrossRef]
9. Jiang, B. Ranking spaces for predicting human movement in an urban environment. *Int. J. Geogr. Inf. Sci.* **2009**, *23*, 823–837. [CrossRef]
10. Kalapala, V.; Sanwalani, V.; Clauset, A.; Moore, C. Scale invariance in road networks. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **2006**, *73*, 1–6. [CrossRef] [PubMed]
11. Chan, S.H.Y.; Donner, R.V.; Lämmer, S. Urban road networks–Spatial networks with universal geometric features? A case study on Germany's largest cities. *Eur. Phys. J. B* **2011**, *84*, 563–577. [CrossRef]
12. Barthélemy, M. *Betweenness Centrality* in large complex networks. *Eur. Phys. J. B* **2004**, *38*, 163–168. [CrossRef]
13. Kirkley, A.; Barbosa, H.; Barthelemy, M.; Ghoshal, G. From the *Betweenness Centrality* in street networks to structural invariants in random planar graphs. *Nat. Commun.* **2018**, *9*. [CrossRef] [PubMed]
14. Chasles, M. *Traité de Géométrie Supérieure*; Bachelier: Paris, France, 1852.
15. Chasles, M. *Rapport sur les Progrès de la Géométrie*; Impr. Nationale: Paris, France, 1870.
16. Berger, M. *Geometry I*; Springer: Berlin/Heidelberg, Germany, 1987; Volume 1, p. 432.
17. Jiang, B. Spatial Heterogeneity, Scale, Data Character, and Sustainable Transport in the Big Data Era. *Int. J. Geo-Inf.* **2018**, *7*, 167. [CrossRef]
18. Jiang, B. New Paradigm in Mapping: A Critique on Cartography and GIS. *Cartogr. Int. J. Geogr. Inf. Geovis.* **2019**, *54*, 193–205. [CrossRef]
19. Regione Toscana. Direzione Urbanistica e Politiche Abitative—Sistema Informativo Territo-riale e Ambientale—SITA.: Grafo Stradario e Ferroviario Della Regione Toscana. 2019. Available online: http://www502.regione.toscana.it/geoscopio/cartoteca.html (accessed on 2 February 2020).
20. Regione Toscana. Specifiche Tecniche per la Gestione E L'aggiornamento del Grafo Viario e Della Numerazione Civica. Dipartimento Generale delle. Politiche Territoriali, Ambientali e per la Mobilita—Settore Pianificazione del Sistema Integrato della Mobilità e della Logistica, Firenze. 2019. Available online: http://www502.regione.toscana.it/geoscopio/cartoteca.html (accessed on 22 September 2020).
21. Regione Toscana. Direzione Urbanistica e Politiche Abitative–Sistema Informativo Territo-riale e Ambientale—SITA.: Limiti Amministrativi. 2019. Available online: https://www502.regione.toscana.it/geoscopio/cartoteca.html (accessed on 2 February 2020).
22. QGIS 3.12.: 'Bucaresti'. Available online: https://www.qgis.org/en/site/forusers/download.html (accessed on 6 January 2020).
23. Varoudis, T. DepthmapX 0.5. 2014. Available online: https://varoudis.github.io/depthmapX/Last (accessed on 6 January 2020).
24. Turner, A. From axial to road-centre lines: A new representation for space syntax and a new model of route choice for transport network analysis. *Environ. Plan. B Plan. Des.* **2007**, *34*, 539–555. [CrossRef]

25. Turner, A. *Getting Serious with DepthMap: Segment Analysis and Scripting*; UCL: London, UK, 2008; pp. 18–25.

26. Turner, A. Could a road-centre Line be an axial line in disguise? In Proceedings of the 5th International Space Syntax Symposium, Delft, The Netherlands, 13–17 June 2005; Van Nes, A., Ed.; TU Delft: Delft, The Netherlands, 2005.

27. Hillier, B.; Yang, T.; Turner, A. Normalising least angle choice in Depthmap–and how it opens up new perspectives on the global and local analysis of city space. *J. Space Syntax* **2012**, *3*, 155–193.

28. Gil, J. Examining "edge effects": Sensitivity of spatial network centrality analysis to boundary conditions. In Proceedings of the 10th International Space Syntax Symposium, London, UK, 13–17 July 2015; Karimi, K., Ed.; University College London: London, UK, 2015; pp. 186–189.