MEASURING GROWTH AND CHANGE IN METROPOLITAN FORM.

Progress report on urban form and land use measures.

THE WORLD BANK

City Form Lab

October 10, 2012.

CONTACT

SUTD City Form Lab
20 Dover Drive
138682 Singapore

+65-63036600
cityformlab@mit.edu

MEASURING GROWTH AND CHANGE IN EAST-ASIAN CITIES.

Progress report on urban form and land use measures.

Authors:

Andres Sevtsuk
PI, City Form Lab
Singapore University of Technology and Design

Reza Amindarbari
Researcher, City Form Lab
Singapore University of Technology and Design

Collaborators:

Onur Ekmekci
SUTD City Form Lab

Raul Kalvo
SUTD City Form Lab

Farre Nixon
SUTD City Form Lab

John Fernandez
Massachusetts Institute of Technology

Xin Zhang
Massachusetts Institute of Technology

SECTIONS:

1. INTRODUCTION

INTRODUCTION

This report presents a technical overview of urban form and land-use measures that can be used to describe large-scale expansion and change in the built environments of metropolitan areas. The metrics that we discuss have been studied in conjunction with the World Bank's work on quantifying urban expansion and transformation in the East-Asia and Pacific regions, but similar measures can be adopted for tracking spatial change in metropolitan areas around the world. Quantitative measures of urban form are meant to be useful for the World Bank staff as decision aids for evaluating large-scale urban development trends and detecting spatial effects from previously implemented policies as basic background assessments for large-scale infrastructure investments that the bank engages in. The metrics are also meant to provide a cohesive methodology that allows stakeholders from the metropolitan and regional governments to keep track of the spatial changes taking place in their respective environments, to compare these with other cities in the region, and to explore the relationship between urban policy and the observed development patterns.

Urban expansion is key to explaining the rapid economic growth, accelerating resource consumption, and social transformations that are taking places in Asian cities today. The United Nations projects that between 1.6 and 2.1 billion people will be added to cities around the world by 2030 (UN-HABITAT 2006; UNFPA 2007). China alone is projected to accommodate an additional 318 million inhabitants in cities by 2030 (UNFPA 2007). In conjunction, China's GDP is expected to pass that of the US in 2050 ("Dating Game" 2011). China has already become the largest consumer of concrete, steel and coal – key ingredients that fuel city growth. But the spatial configurations that Chinese cities are taking has also led the nation to become the largest buyer of new automobiles and the most rapidly increasing energy consumer among the OECD countries. A great deal of social, economic and environmental changes taking place in cities are physically manifested in their spatial patterns or the patterns of their metropolitan regions. Having commonly understood metrics for describing these built environments and capturing their change over time is therefore critical for gaining broader insights into the social and economic transformations taking place therein. Moreover, existing spatial development patterns can also reveal critical constraints to growth and potential development opportunities, an efficient operation of land markets, and an effective and equitable spatial resource distribution.

Our work builds upon numerous previous research initiatives that have studied important qualities of metropolitan form (Geddes 1915; Mumford 1961; Webber 1963, 1964; Lynch 1991). The literature on metrics or methods of capturing these qualities is scarcer. Still, most of the measures we present are derived from previous applied studies. We also propose a few new metrics that have not been widely operationalized in comparative studies of metropolitan growth patterns in the past. Our aim has been to evaluate these measures and to propose improvements that would allow them to capture more nuanced and more useful aspects of metropolitan form. Shlomo Angel's study on urban expansion with the World Bank has operationalized variants of a number of these measures: the buildable perimeter, the contiguity index, and the compactness index among others (Angel 2005; Parent, Civco et al. 2009). Alan Betraud's work on comparative population density gradients has implemented a novel dispersion measure and a categorization method to describe the regulatory systems in various countries (Bertraud and Malpezzi 2003). We also borrow from Wheeler's and Burchfield's research on the determinants of urban form in the United States (Wheeler 2008; Burchfield, Overman et al. 2006) and Hall's and Pain's work on the polycentricity and polycentric city regions in Europe (Hall and Pain 2006). We connect some of these measures with more recent GIS-based spatial analysis and network analysis methods (Sevtsuk 2012). Finally, we also refer to some of the more recent research on capturing scaling and repetition in urban metropolitan patterns explored either through fractal analysis or scaling laws (Batty 2006, 2008; Bettencourt, Lobo et al. 2007; Changizi and Destafano 2009).

SCOPE OF THE WORK

The measures discussed below focus primarily on capturing the spatial characteristics of urban form and land-use distribution – qualities of metropolitan areas that can be detected from rather basic remote sensing imagery and that can be readily operationalized in a large number of cities. Measures of metropolitan form are most useful, however, if they can be coupled with other types of data about each city – their economic indicators, social and demographics indicators, transportation surveys, and environmental performance metrics. We do not discuss these latter kinds of data at the metropolitan scale in depth, but emphasize in the latter part of the report how form and land-use characteristics might relate to some of these factors.

We do not attempt to develop an extensive list of all available metrics, but focus on a few, which capture unique and complimentary qualities of metropolitan form. These metrics aim to describe the primary properties of city form that allow us to differentiate development patterns within and between cities. In doing so, it is important to distinguish spatial features that are revealing and consequential from those that are superficial and uninformative. It is less informative, for instance, to capture the details of a unique city block than the overall density of the urban area. In some ways, the challenge of distinguishing the forms of metropolitan areas is analogous to the challenge of distinguishing the physical features of people – both need a concise set of metrics to capture their unique characteristics. Unlike families of plants or animals, cities do not yet have vastly different evolutionary groupings. Their differences are more nuanced and harder to classify. The metrics that describe them therefore need not focus as much on classifying distinct groupings as they do on distinguishing critical differences in individual forms.

INPUT DATA TYPES

In structuring the metrics that describe the physical patterns of urban form, we need to keep in mind that the available data can vary significantly between cities. In a number of cities, the only available data source may be a satellite image with a resolution of 5-10m per pixel. In others, the resolution may be sharper, or the color spectrum wider. In others yet, the input data may come as Geographic Information Systems (GIS) shape files. The indices that follow have kept each of these possibilities in mind. Across all data types we assume, however, that the inputs describe aerial units – they are two-dimensional units that have an area property. Further, we must assume that these areal units are large enough to capture a set of roads and buildings, that is, smaller or larger pieces of the urban development pattern. Keeping this in mind, the input data can come in both raster and vector formats and all the metrics can be computed with both types of data.

Much of the socio-economic data on cities does not depend on the geometry of input data, but rather describes the whole city or part of a city abstractly. The population of a city, its job pool, how much energy it uses, its GDP and other functional indicators therefore do not necessarily come as either raster or polygon data, but as tabular information that can be linked by a common ID to geometric objects.

STRUCTURE OF THE REPORT

The report is structured as follows: We first present a series of indices that characterize metropolitan form, first discussing their previous applications and then illustrating the adjustments and new specifications we propose. The second part focuses on the question of how a sequence of the metrics captured at different snapshots in time can be used to describe growth and change in each metropolitan region. Third, we briefly discuss how the urban form and land use metrics might relate to the potential determinants of metropolitan form – the social, economic and environmental indicators of each respective city. We believe that the formal patterns of cities are important to study only if we have an understanding or a set of hypotheses about the forces that might explain them. Finally, we end by touching on some policy implications and future work directions that result from this work.

2. URBAN FORM METRICS

INTRODUCTION

Metrics of metropolitan form can describe a number of important qualities of a city. Polycentricity, for instance, can describe the degree to which a city's employment is concentrated in its sub-centers – how many significant job centers there are, what their size distribution is and what share of all employment is located in them. Knowing this can help us understand the demands on transportation infrastructure, predict its commuting flows, and analyze the impact of potential spatial investments or policy changes.

Metropolitan form metrics depict quantitative relationships found in geographic data describing a city. All metropolitan form metrics thus first require a collection of geographic data as inputs. Figure 1 lists some useful types of geographic data that are commonly used to describe metropolitan areas. The figure also illustrates how these data can be combined to form metrics.

We can broadly distinguish between two types of geographic input data – geometric and non-geometric. Geometric data refer to geographic information that measures the spatial properties of an urbanized area – its area or perimeter, the number of discontinuous built-up areas within its administrative borders, etc. This data is compiled from geometric measurements of the underlying built environment and can typically be estimated from satellite photos, geometric shape objects in Computer Aided Design (CAD) software or GIS. Non-geometric data, on the other hand, refers to additional information about the environment under study that are not obtained from the quantification of its geometry – a city's energy bill, total vehicle miles traveled or regulatory climate. These data typically come from multiple sources that vary from city to city and cannot be estimated purely spatially by examining its satellite image or two-dimensional pattern. Some non-geometric data, such as a city's population size, or economic output, are closely monitored and therefore readily available in most cities around the world. These basic indicators constitute the more critical non-geometric inputs for the city's metrics. Others, such as the Gini coefficient, vehicle ownership or land prices, are typically not used as inputs for metropolitan form metrics, but can instead be utilized as either dependent or control variables in statistical estimations of the spatial metrics. These types of non-geometric indicators of a city can be correlated or even causally related to the city's geometry. We call these data the potential 'determinants' of form and discuss them further in Section 4 of this report.

Metropolitan form metrics, too, can be conceptually organized into two categories. First, and certainly the most popular, we have metrics that can be directly derived from input data through simple arithmetic, without any need for further spatial analysis. Such metrics include Density and Coverage, both can be found as simple ratios from input data like Administrative Area, Built Area, and Total Population. Second, we have a series of interesting metropolitan pattern metrics, whose calculation requires some additional spatial analysis and computation. In this group one finds indices such as Polycentricity, Compactness, Expandability, etc. We have made a careful effort to keep these pattern metrics as distinct from each other as possible, so that a combination of them can be jointly used as predictors in a regression model without cancelling each other out.

A specific sub-group of these metrics is wholly focused on the shape properties of a city's outline – its Roughness, Depth, or Perimeter Index for example. As the reader may already anticipate, there can be a great degree of correlation between some of these shape metrics. The following individual specifications of metrics will primarily focus on the pattern metrics and simple, directly derived metrics. The shape metrics have already been described in depth by Parent, Civco and Angel (2009).

What unites both these pattern and shape metrics is that each of their calculation requires additional spatial analysis in GIS or other spatial computing environments.

GEOGRAPHIC DATA

GEOMETRIC

Total Built Area
Total Built Premieter
Administrative Area(s)
Total Built Volume
Total Built Floor Area
Total Buildable Area (Convex Hull - Unbuildable)
Total nr. of Continuous Built Areas
Total nr. of Sizable Built Areas (Area > +1 std)
Available Expansion Area (within a 30% area increase buffer)
Convex Hull Area
Total Area per Land-Use Type (buildings, roads, commercial, etc.) (within Convex Hull)

NON-GEOMETRIC

DEMOGRAPHIC
Total Population
Avg. HH Size
Population by Group (Ethnic, Race, Age, etc.)

ECONOMIC
Total Jobs, Firms, Establishments, etc.
GDP/Capita
Car Ownership
Avg. HH Income
Observed Land Prices
Gini Coefficient
Cost of Living
'Big Mac' Index

ENERGY
Avg. HH Enegry Consumption
Material Energy Expenditure
Vehicle Miles Traveled

POLITICAL
Socialist [0-1]
Regulatory Climate [0-n]
Political Stability [0-n]
Corruption [0-n]

CLIMATIC
Temperatures
Sunlight
Flora/Fauna
Aquafers

METROPOLITAN FORM METRICS

DIRECTLY DERIVED

Size
Density
Coverage

COMPUTED WITH EXTRA SPATIAL ANALYSIS

PATTERN
Polycentricity
Compactness
Expandability
Contiguity
Land-Use Mix

SHAPE
Roughness
Depth
Spin Index
Cohesion Index
Girth Index
Perimeter Index
Detour Index
Traversal Index

Figure 1. Input data and metropolitan form metrics.

Image source: http://joshuatwood.com/wp-content/uploads/2012/04/los-angeles-skyline.jpg

SIZE

The land covered by urban use, or its subcategories (e.g. residential, industrial etc.), is probably the primary aspect of a metropolitan area that should be mapped and measured, as major trends of change could be apprehended directly by looking at the snapshots of urban land-uses and their growth and change over time.

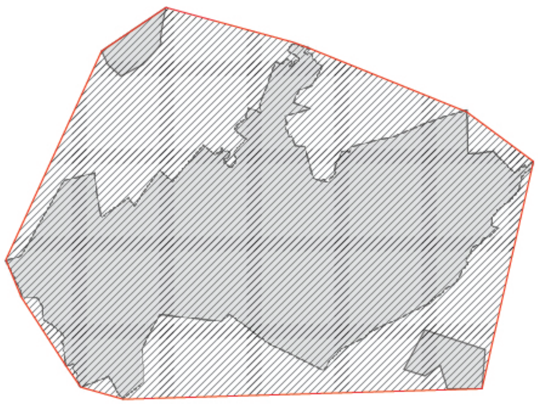


Figure 2. The Convex Hull is defined as the smallest flat polygon that has no concavity in its perimeter and that fully contains all individual polygons of the urban extent of a city.

SIZE

The land covered by urban use, or its subcategories (e.g. residential, industrial etc.), is probably the primary aspect of a metropolitan area that should be mapped and measured, as major trends of change could be apprehended directly by looking at the snapshots of urban land-uses and their growth and change over time. The size of a metropolitan area also provides a crucial basis for other derived spatial metrics, such as density.

Distinguishing the extent of a city – the border between urban and non-urban land – or the extent of a particular land use is, however, not a trivial task. In many cities, the urban land-cover starts from very compact centers that cover sizable areas and gradually dissolve into peripheral agricultural lands without leaving behind a clear border. This is an important question for analysts who work directly with raw satellite imagery. If the input data comes as pre-categorized Landsat remote-sensing-based land use maps, then the question has already been addressed by the company that composed the maps and the conventions used should be made available along with the data.

CURRENT MEASURES OF LAND USE EXTENT

The World Bank's Platform for Urban Management and Analysis (PUMA) currently keeps tracks of the following metrics:

- The area of urban extent.
- The area of urban extent in each administrative boundary.
- The area of total urban extent in all administrative boundaries that constitute the urban use.
- The area of each land use (either urban or non-urban).
- The area of each land use in each administrative boundary.
- The area of each land use in all administrative boundaries that constitute the urban use.

IMPROVING URBAN EXTENT MEASURES

In addition to the metrics outlined in the PUMA list above, we propose two additional size metrics to be included in the platform:

1. The area of the Convex Hull around all developed polygons. The Convex Hull is defined as the smallest flat polygon that has no concavity in its perimeter and that fully contains all individual polygons of the urban extent of a city (Figure 2).

The Convex Hull can be computed in GIS (Toolbox > Cartography > Masking > Feature Outline Masks) and in other software platforms. It offers a very useful method of defining the joint built and un-built areas of a city in a consistent methodology across cases.

2. The unbuildable area within the Convex Hull. We suggest that the definition of "unbuildable"

should capture areas that are not buildable only due to natural obstacles, such as water-bodies, steep slopes or other natural limitations (Figure 3). The definition of "unbuildable" should not capture currently un-built areas that stem from policy choices (e.g. Central Park in Manhattan). Such a definition will allow for a more reliable and consistent assessment of Density (see Section 2.3).

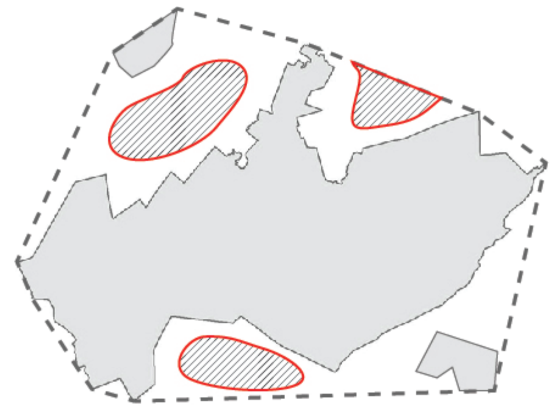
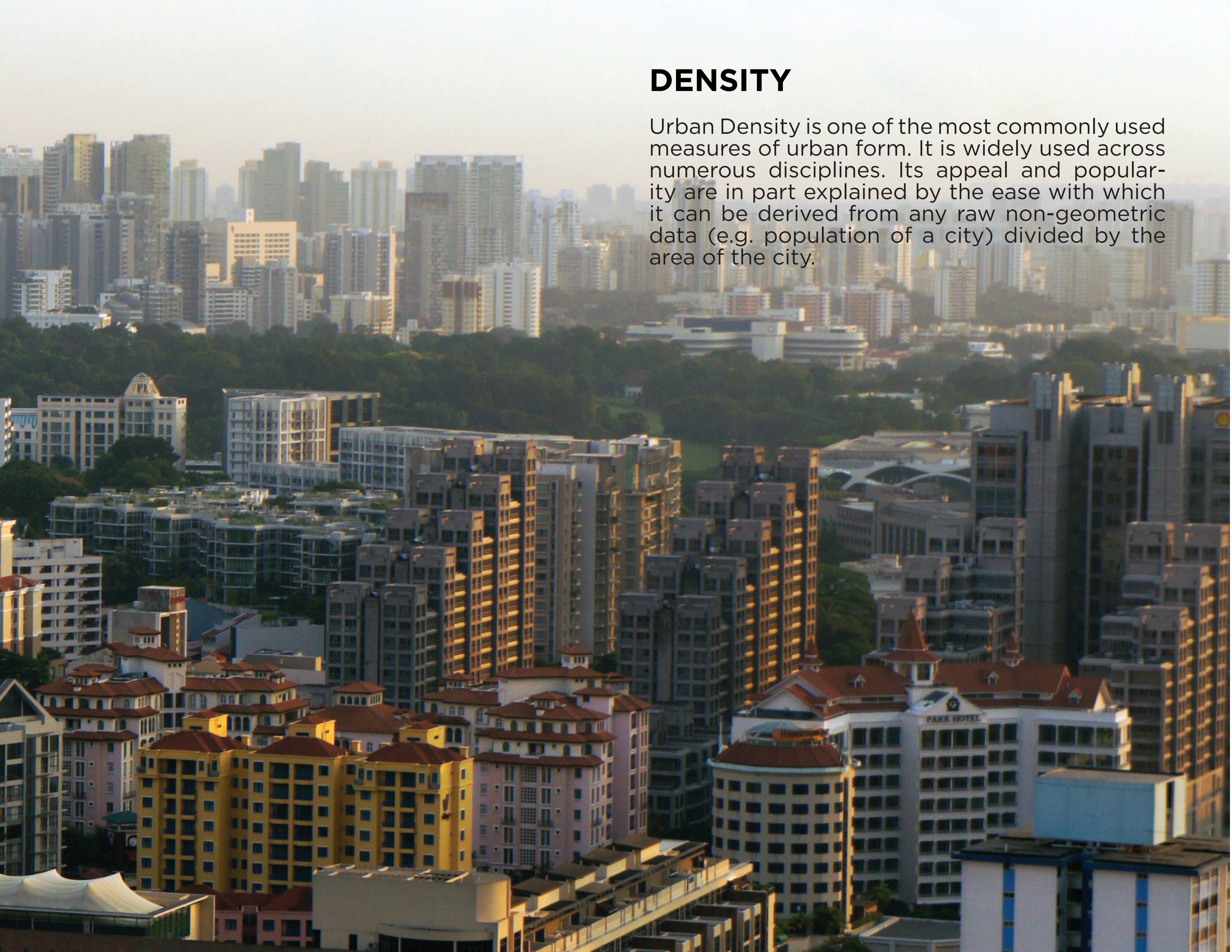


Figure 3. The unbuildable area with the Convex Hull.


Photo by Raul Kalvo

DENSITY

Urban Density is one of the most commonly used measures of urban form. It is widely used across numerous disciplines. Its appeal and popularity are in part explained by the ease with which it can be derived from any raw non-geometric data (e.g. population of a city) divided by the area of the city.



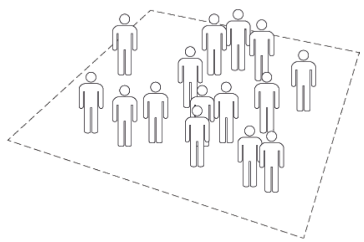


Figure 4. Population density captures the number of inhabitants per unit area of land.

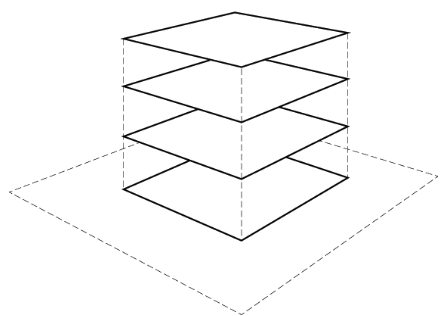


Figure 5. FAR captures gross built floor area per unit area of land.

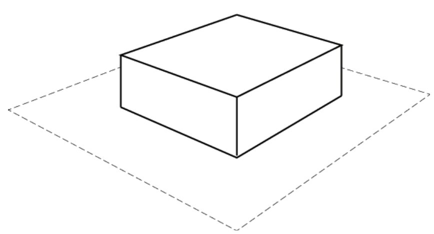


Figure 6. Unit density captures the number of dwellings per unit area of land.

DENSITY

Urban density is one of the most commonly used measures of urban form. It is widely used across numerous disciplines. Its appeal and popularity are in part explained by the ease with which it can be derived from any raw non-geometric data (e.g. population of a city) divided by the area of the city.

Although widely used, there are multiple ways of measuring density, which differ in important ways. Population density – the number of residents per unit area of land – can be similar in Mumbai and Hong Kong, but in terms of built form, the cities are far from alike. Hong Kong is characterized by high-rise towers whereas Mumbai is largely low-rise but extremely dense in ground coverage and its average number of residents per dwelling unit is significantly higher. Likewise, cities or districts that have similar built form can have notably different population densities. The Marais district in central Paris, and the area around the Arc St. Denis in the 10th arrondissement both have a similar built density (approximately FAR 2.2), but the latter has twice the residential density of the former – 20,000 residents per square kilometer in the Marais and 40,000 residents per square kilometer in the 10th arrondissement. This difference is explained by more crowding in the apartments around Arc St. Denis compared to the relatively sparsely populated living quarters in the wealthier Marais. A meaningful comparative interpretation of urban density metrics thus depends on "... what is included and what is excluded to make density figures truly comparable" (Forsyth, 2003).

Density measurements are also affected by the areal units at which the measurements are captured. This issue, which has become known as the Modifiable Areal Unit Problem or MAUP in the literature (Openshaw 1984), is defined as "a problem arising from the imposition of artificial units of spatial reporting on continuous geographical phenomenon resulting in the generation of artificial spatial patterns" (Heywood 1998). A comparison of density figures can be misleading if the areal units that are compared are not analogous – the density of a block should not be compared to the density of a neighborhood,[1] because the type and amount of open space affecting the metrics differs across the two scales. The area of a neighborhood includes road surfaces, sidewalks, and probably open spaces: it is therefore more likely to yield lower outcomes than density measures performed at a block scale.

EXISTING SPECIFICATIONS

Density measures are usually given in the form of a ratio[2], where the numerator shows the amount of resources and denominator the land area under consideration. The list of possible density specifications can be long as the numerator can include a wide range of resources, such as resi-

1 For proper comparison, Tunney Lee and his colleagues have categorized the cases of their Density Atlas into block, neighborhood and district scales.

2 Density may be measured in different forms; e.g. the average set-back of buildings, or the average back-to-back distance of buildings in a block, which are not relevant to the scale of our work.

dential population, the number of dwelling units, or the gross built floor area. Likewise, the base land area shown in the denominator can be calculated in various ways – by excluding or including certain land-use types.

Common specifications of density measures include:

- Residential density: the number of residents per unit area of land (Figure 4).

- Floor area ratio (FAR): the total built floor area per area of base land (Figure 5).

- Unit density: the number of dwellings per unit area of land (Figure 6).

- Employment density: the number of jobs per unit area of land.

A simultaneous examination of residential density, unit density, floor area ratio, and ground coverage can be used to capture typological differences in urban form (see the Density Atlas developed by Tunney Lee et al. at MIT).[3] Mapping employment density, on the other hand, helps us understand the economic structure and probably commuting patterns in a city.

Gross and Net Densities

As discussed above, the types of land-uses that are included in the base area can change the meaning of a density measure and yield either gross or net densities. The base area in gross density includes land-use types that are not directly relevant to the quantity in the numerator. Gross residential density, for instance, includes roads, parking lots, parks, etc. These competing uses are typically excluded from estimations of net residential density. Some scholars (e.g. Bertaud 2004) have used a scalar cut-off threshold to determine whether an open space should be included in a net density measure our not – they include an open (green) space in the base area calculation if it is smaller than a given size threshold.

Reverse Density

Reverse density measures invert the calculation ratio and show how much land is available per person, per dwelling unit or per job. Whereas normal density measures draw our attention to how crowded an area is, reverse density measures emphasize the opposite – the amount of space available to an average person.

Intra-urban Density Analysis

Intra-urban density is often measured based on the area of a census tract. Weighted density measures illustrate how densities are distributed between different census tracts (or other areal units) at an intra-urban level, where tracts are weighted based on their share of the total population.

An important hazard in using census tracts for intra-urban density analysis lies in the fact that census tracts are often drawn at inconsistent scales. Since census tracts typically aim to capture a comparable number of residents, they tend to be larger in low-density areas and smaller in high-density areas of a city.

3 http://densityatlas.org

Image source: http://thecityfix.com/files/2011/09/dharavi.jpg

COVERAGE

Coverage illustrates how large a share of the total urban extent, or a sub-area of the city, is covered by a given land use type. It is a form of Density measure, whose numerator captures only the ground area of different uses. It is most commonly estimated for building footprints – the percentage of the urban area that is covered by buildings.

COVERAGE

Coverage illustrates how large a share of the total urban extent, or a sub-area of the city, is covered by a given land use type. It is a form of density measure, whose numerator captures only the ground area of different uses. Most commonly estimated for building footprints – the percentage of the urban area that is covered by buildings, coverage can also be estimated for other land-use types – residential or commercial land, green spaces, etc.

Two cities with similar population density and FAR can have vastly different urban forms. Parts of Singapore and Manila, for instance, both have high population densities and district-wide FAR, but the former is largely made up of high-rise buildings and the latter of high-density low-rise urban fabric. How both types of configurations can yield a similar population density and FAR is largely explained by their differences in building coverage. In Singapore, building coverage is generally low, and ample green space permeates between tall residential slabs. In low-rise settlements of Manila, buildings and circulation paths between them cover almost all ground.

EXISTING SPECIFICATIONS

The coverage C_n of land-use type n (LU_n), within an area A, is defined as follows:

$$C_n = \frac{A[LU_n]}{A}$$

The extent of the area A, at which coverage is estimated, can vary depending on our research question and the land-use type in the numerator. For building footprints, the reference area is often the corresponding area of a city block. If the reference area is enlarged to a whole census tract or district, then the resulting coverage metric will expectedly decrease as roads, parks and other open spaces are added to the denominator. Depending on the question, however, coverage can be estimated for an entire urban extent.

IMPROVING THE BUILT-UP COVERAGE MEASURE

Estimating coverage for the urban extent of an entire metropolitan area requires a consistent approach to defining the area. One approach would be to use the administrative area as the denominator of the index. But the size of administrative boundaries in different cities can range widely and may occasionally constitute an area many times as large as the current urban extent. We thus propose to estimate coverage not in the administrative areas, but within the convex hull[1] of all developed polygons (Figure 7) minus the unbuildable areas within the convex hull (see SIZE metric explanation).

Estimating coverage within the convex hull minus unbuildable areas offers a consistent way of measuring how much land is urbanized in each city. The exact shape of the hull and the extent of the

1 The Convex Hull is defined as the smallest flat polygon that has no concavity in its perimeter and that fully contains all individual polygons of the urban extent of a city.

unbuildable area is relative to each city, but the estimation remains consistent and comparable since remaining vacant land within the hull can be said to be vacant by choice.

If the area of the convex hull around the developed polygons is A_{CVX}, the unbuildable area of the convex hull is A_{UNB}, and the total area of all built-up polygons is A_{BLT}, then the built-up coverage C_{BLT} is estimated as follows:

$$C_{BLT} = \frac{A_{BLT}}{(A_{CVX} - A_{UNB})}$$

Alternatively, coverage at the metro scale can also be estimated for particular land uses within the observed built-up area A_{BLT}. Remote sensing data offer an opportunity to measure built coverage consistently and comparably across cities, especially as building footprint data is unavailable for many cities.

The ratio between a given land use area A_{LU} and the total built-up area A_{BLT} captures what share the particular land use – buildings, roads, parking lots, etc. – constitutes of the entire built-up urban area:

$$C_{LU} = \frac{A_{LU}}{A_{BLT}}$$

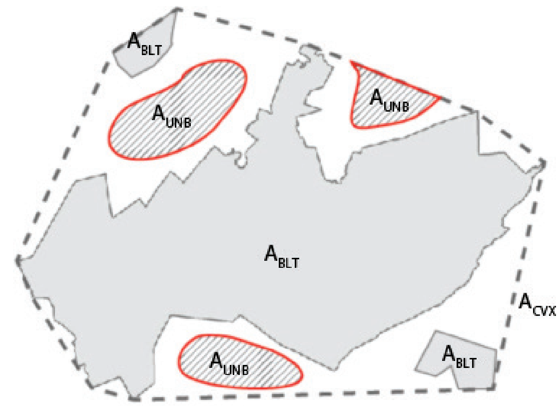


Figure 7 . Convex hull of the urban extent and the unbuildable area with the convex hull.

Image source: http://fmjdata.com/wf/News/25C40E65D1A20BDD6607FEBCB7544FA4/Olympics_aerial_view_credit_London_2-12_fmj_jan12.jpg

POLYCENTRICITY

Every city has its centers, some larger, some smaller. For centuries, planners, economist and geographers have debated why and how centers arise. Most scholars agree that urban centers emerge from transportation cost savings rendered by proximity, from economies of scale in production and service activities, and from spatial externalities between space users.

POLYCENTRICITY

Every city has its centers, some larger some smaller. For centuries, planners, economist and geographers have debated why and how centers arise. Most scholars agree that urban centers emerge from transportation cost savings rendered by proximity, from economies of scale in production and service activities, and from spatial externalities between different space users in a city.

Despite the continued dominance of the monocentric urban model (Alonso 1964, Muth 1969 and Edwin 1967), its primary assumption that jobs are clustered in a single dominant location is inconsistent with the spatial distribution of employment in many contemporary cities. A large body of literature has shown how increases in traffic congestion, land prices and rents, commuting costs and air pollution can jointly shift economies of scale to diseconomies of scale and lead to the emergence of multiple sub-centers. Empirical evidence suggests that even the most monocentric cities, such as Las Vegas or Baltimore, have smaller sub-centers outside the central business district (CBD).

Given the fundamental economic role of a city's centers, changes in their structure and rank order can be telling about the growth and change of the city as a whole. Impacts of job clustering on transportation, energy consumption, or economic inequalities cannot be explained without a reliable metric that captures the characteristic of these clustered landscapes – their polycentricity.

Quantifying polycentricity is challenging since defining a center is far from obvious and the number of centers is not necessarily the only yardstick for polycentricity. There is no universal definition for urban centers; centers can have blurry boundaries and polycentricity can be relative to the size of a city as well as the resolution of the lens with which we examine an urban landscape.

EXISTING SPECIFICATIONS

Polycentricity has been described at both the intra-urban level (polycentric cities), and inter-urban level (polycentric urban regions). Although our focus is on quantifying the geographical aspect of intra-urban polycentricity, previous literature on polycentric urban regions provides inspiration to our work. There are two main types of polycentricity measures at the regional scale that capture the interaction between centers. The Entropy Index (e.g. Limtanakool et al. 2009), for instance, examines how uniformly flows are distributed among centers. It is defined as:

$$EI = -\sum_{i=1}^{L} \frac{(z_i)\ln(z_i)}{\ln(L)}$$

where *EI* is the entropy index, Z_i the ratio of the number of trips from *i* to the total number of trips within the region, and *L* is the number of cities. *EI* is constrained between 0 and 1. It is 0 if all trips involve only one city, and 1 if trips are equally distributed among all cities.

Another group of metrics captures the symmetry of incoming and outgoing trips at each center, or the symmetry of trips between a pair of cities. Limtanakool et al. (2009) have defined Node Symmetry Index (NSI) as follows:

$$NSI_i = \frac{I_i - O_i}{I_i + O_i}$$

, where I_i is the number of trips to city *i* and O_i the number trips originating from *i*.

Given the complexity of trips at the intra-urban level[1] and the difficulty of obtaining similar trip data as the regional level, these inter-urban polycentricity metrics are not ideally suited for use at the intra-urban scale. However, the concept of spatial interaction between centers has inspired the polycentricity metric we propose below, which uses a weight to describe centers (the percentage of resources in a center) analogous to the number of trips to them. In our metric, we use the number of centers, the weights of centers, and their relative size distribution to capture *structure*, *strength*, and *symmetry*.

A key challenge in quantifying polycentricity is the identification of centers. At the inter-urban level, centers are simply cities, but at the city-scale it can be challenging to distinguish centers from non-centers. An area can be a center locally, but not large enough to qualify as a center at the scale of the city or in comparison to other cities. According to Masip (2011) there are two types of methods for identifying urban centers: methods based on analysis of density and those based on analysis of functional relations. The latter, as we discussed above, are typically applied at the regional scale. At the city scale, however, Masip further categorizes density analysis methods into four groups:

1 It can be challenging to distinguish between trips that end at a center and trips through a center to another destination.

a) A sub-center is defined as the second peak beyond CBD, the census tract that has higher employment density than its neighboring tracts. An example of these, McDonald's method (1987), fails to recognize, however, the fact that several contiguous tracts can jointly form a local peak.

b) The second group uses a lower cut-off criterion to identify sub-centers. Some have defined sub-centers by an absolute lower cut-off. Giuliano & Small (1991) defined sub-centers as census tracts with a density of more than 10 employees per acre and at least 10,000 jobs. Others used relative cut-off criteria. García-López (2007, 2008) and Muñiz & García-López (2009) defined sub-centers as "zones with a density higher than the metropolitan average and at least 1% of metropolitan employment."

c) The third group of methods uses the standard monocentric model to identify sub-centers. For instance, McDonald & Prather (1994) defined sub-centers as areas with positive residuals that are significant at a 95% confidence level: i.e. areas with densities significantly higher than expected density in the mono-centric model, based on the distance from the CBD. The main disadvantage of this method is that it presumes the existence and location of a CBD.

d) The last set of methods for identifying sub-centers in Masip's (2011) classification is conceptually similar to the previous method (c), but uses a weighted regression for smoothing the natural logarithm of employment density,

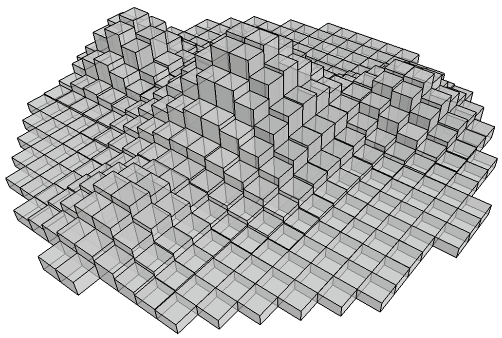


Figure 8. A raster dataset of density measures in a hypothetical polycentric city.

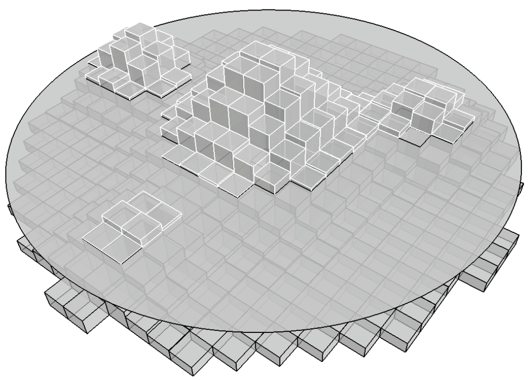


Figure 9a. Job density at urban centers has to be higher than the mean density in the entire city.

and allows for local variation in the flattening rate of the gradient.

Most intra-city indices in the literature measure polycentricity as the number of centers that fit the proposed definitions. Central shortcomings of these indices are that 1) the overall share of jobs that are in centers combined does not affect the outcome; and 2) the relative size balance between centers, too, does not affect the outcome. A city where only 10% of jobs are located in centers (the rest are dispersed) would thus be considered as polycentric as a city where 90% of the jobs are in centers. Likewise, a city with one large center and three small centers would be considered as polycentric as a city with four equally large centers. We have tried to address these shortcomings in the proposed metric for polycentricity below.

PROPOSED POLYCENTRICITY METRIC

Determining the polycentricity of a city requires two stages of analysis. First, we need to determine the total number of centers and the size of each of the centers. Second, given this information, we need to determine how polycentric the city is. Figure 8 illustrates a raster dataset of density measures in a hypothetical polycentric city.

Detecting Centers

In order to first distinguish urban centers in the data, we implement the following three steps:

1) Job density at urban centers has to be higher than the mean density in the entire city (Figure 9a). This threshold is relative to each city and not universally defined. We compute the average job density across all polygons and only retain those whose values exceed the average.

2) If several adjacent polygons pass the previous average cut-off threshold, then these neighboring polygons are grouped together to form joint centers (Figure 9b).

3) A center must contain *n* % or more of all jobs in a city (Figure 10). While many places in a city can have a high number of jobs clustered on a small area of land, an area should not be considered a center if it does not contain an adequate percentage of total jobs in the city. Note that *n* should not be fixed for all city sizes; the minimum cluster size threshold should vary along with the size of the city. We have experimentally determined that "n=10/√population" yields suitable values that adjust intuitively to city size. In a city of 100,000 jobs, it yields a cutoff at 3.16% (3,162 jobs), in a city of 1,000,000 jobs *n* becomes 1% (10,000 jobs) and in a city of 10,000,000 jobs *n* is 0.32% (31,622 jobs).

The third step thus counts the total number of jobs found in each cluster and eliminates the clusters that do not pass the appropriate minimum percentage test, adjusted to city size. The procedure of detecting centers is concluded by assigning an ID to each center that is found, determining the geographic coordinates of the centroid of each center, and summing the total number of jobs in each center.

Center ID	Total Jobs	X-coord.	Y-coord.
1	552,895	1.3667° N	103.7500° E
2	289,637	2.8667° N	101.6100° E
n	*n*	*n*	*n*
Total:	1,689,254		

ESTIMATING POLYCENTRICITY

Having defined the centers, we propose a polycentricity metric that simultaneously satisfies the following criteria:

1) A city with more centers is more polycentric than a city with fewer centers.

2) A city is more polycentric if a greater share of its total job pool is located in its centers.

3) Polycentricity, by definition, implies a shift from the dominance of one center to a state where several centers exist in a balanced competition. We therefore assume that the less any one center dominates, and the more equally balanced the sizes of different subcenters are, the more polycentric a city is.

4) If a new center emerges, while the total number of jobs, the total number of existing centers, and the sizes of existing centers remain constant, then polycentricity of the city increases.

With these criteria in mind, we arrive at an intuitive measure of polycentricity that is based on the number of centers, the sizes of centers, and the size balance between the centers:

$$PC = HI x N x R_c$$

where *PC* is the polycentricity index, *HI* the homogeneity index, *N* the number of centers, and R_c the ratio of the total amount of jobs found in all centers to the total amount of jobs in the city. *HI* measures the degree to which the sizes of centers are homogenous. We define *HI* by using Limtanakool's (2009) Entropy Index *EI* (see Existing Specifications above):

$$EI = -\sum_{i=1}^{L} \frac{(z_i)\ln(z_i)}{\ln(L)}$$

, where *Zi* is the ratio of the number of jobs at center *i* to the total number of jobs in the city, and *L* the number of centers in the city. *EI* is constrained between 0 and 1; it is 0 if all jobs are in a single center and 1 if jobs are equally distributed among all centers.

If a new center emerges, while the total number of resources (jobs) and the sizes of existing centers remain constant, *N* and R_c increase. However, *HI* may change in any direction or stay constant depending on the relative size of the new center compared to the size of the pre-existing centers.

The value of our polycentricity metric *PC* thus depends simultaneously on the three factors we set out to integrate: the number of centers, the sizes of the centers, and the relative size distribution between centers that are found in a city.

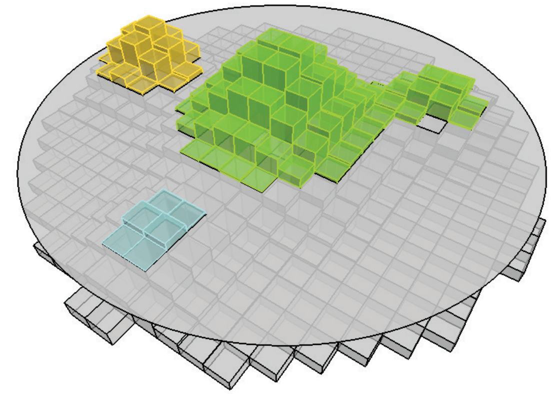


Figure 9b. If several adjacent polygons pass the previous average cut-off threshold, these neighboring polygons are grouped together to form joint centers.

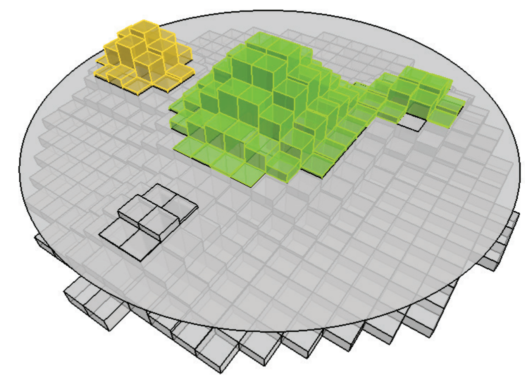


Figure 10. A center must contain *n* % or more of all jobs in a city, where *n* = 10/√population. In a city of 100,000 jobs, it yields a cut-off at 3.16% (3,162 jobs), in a city of 1,000,000 jobs, *n* becomes 1% (10,000 jobs), and in a city of 10,000,000 jobs, *n* is 0.32% (31,622 jobs).

Image source: http://openwalls.com/image/19004/ho_chi_minh_city_1600x1200.jpg

COMPACTNESS

Compactness, and its inverse quality – Dispersion – measure the degree to which the resources of a city – people, buildings, jobs – are spatially spread out; the closer they are located to each other, the more compact the city is.

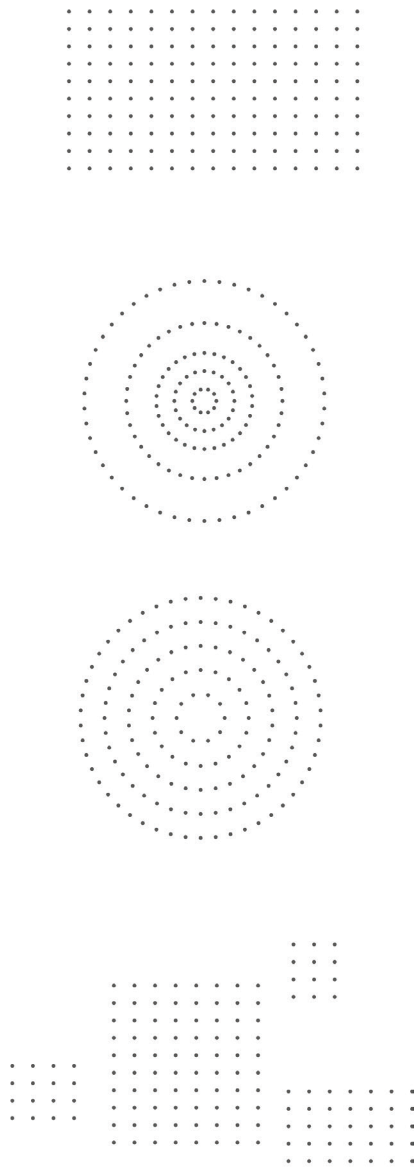


Figure 11. The internal distribution of a city's resources can vary considerably, which remain invisible to a density measure.

COMPACTNESS

Compactness, and its inverse quality – dispersion[1] – measure the degree to which the resources of a city – people, buildings, jobs etc. – are spatially spread out; the closer they are located to each other, the more compact the city is.

Understanding how compactness of metropolitan areas affects the social, economic and environmental performance of cities on the one hand, and detecting the forces that lead to compactness on the other, constitute major research areas in city and regional planning. A clear understanding of how compactness can impact urban energy consumption, economic inequality, car ownership, greenhouse gas emissions and economic efficiency can not be achieved without first capturing the compactness of a city in a quantifiable measure.

EXISTING SPECIFICATIONS

One of the most commonly used measures for capturing the compactness of a city is areal density – the quantity of resources per unit area of land (see section on the DENSITY metric). Areal density assumes, however, that resources are uniformly distributed throughout the study area – it does not distinguish internally homogenous or heterogeneous distributions, contiguous or discontiguous developments, nor the effects that the shape of development can have on its dispersion (Figure 11).

Another way of capturing the compactness of a city's development is achieved by measuring the availability of open space around each piece of developed land. Burchfield, Overnman and their colleagues (2006), for instance, have defined dispersion as the average percentage of open space in the immediate square kilometer around each residential development. As openness, too, is basically a density measures it comes with the same shortcomings in capturing Compactness mentioned above.

If all else equal, spatial distributions are most compact if their aggregate collection forms a circular shape. Circular distributions, as widely witnessed in nature, have the smallest perimeter-to-area ratio of any two-dimensional geometric shape. Capitalizing on this property of circular shapes, a number of researchers have described the observed dispersion of urban resources (people, development, buildings) in comparison to a perfectly circular distribution of the same amount of resources.

Bertaud and Malpezzi (2003) define dispersion as the ratio of average distance from the centroids of population tracts to the CBD to the average distance of the same population from the centroid of a hypothetical circular city of the same size[2]. In their measurements of 48 cities they assume the CBD to be at the geometrical centroid of all population tracts.

1 Dispersion has also been used to characterize urban sprawl. The concept of sprawl, however, involves a number of factors beyond spatial form and remains poorly defined in literature.

2 The average distance of a uniformly distributed population from the center of a disk is equal to two third of the radius of the disk.

The main shortcoming of this measure is that it only captures the relationship of spatial resources (e.g. people) to the city center and overlooks their spatial location with respect to each other. This may be problematic in cities where the CBD is lacking or unimportant (e.g. Randstadt in Holland) or where the CBD is not at the geometric centroid of the tracts (e.g. Singapore). The model is more appropriate for a monocentric city, but ignores the relationships between multiple smaller centers within the city. Second, while this measure weighs distances to the center by population, it is not affected by the absolute changes of density in the city – if the relative size balance in tracts is kept constant, but the population in each tract is increased at the same rate, the dispersion index would remain constant[3].

Angel et al. (2005) have defined compactness (the opposite of dispersion) in a similar vein, but instead of comparing the observed development to a hypothetical circle, they compare it to the actual availability of developable land in that city. Whereas Bertrauds's index relies on distance measurements in defining dispersion, Angel's measure relies on area measurements. Angel et al. account for geographical constrains in the observed area, arguing that "compactness should be restricted to buildable areas, in the sense that a city located on a coast, on a mesa cut up by steep gorges, or in a valley surrounded by steep cliffs can be very compact even if it does not resemble a full disk" (p. 68, Ibid). Their compactness measure is thus defined as the ratio of the observed built-up area to the observed buildable area within "the circle of minimum radius encompassing the consolidated built-up area of the city." The circle, which is used only for the purposes of restricting the geographical extent that is compared, can be defined either as "the minimum radius encompassing the consolidated built-up area of the city" – called the outer circle – or restricted to the main built up area of the city – which we might call the inner circle (Fig. 12).

A key challenge to the index is that if the outer circle is used, a small and consolidated built-up area located far from the main built-up area can significantly impact the minimal radius of the circle. If the calculation of the index is restricted to only the largest continuous built-up area — as performed by Angel's team — then the index may return unreasonably high compactness values in cities with satellite towns that have significant developments outside of the main built-up area (e.g. Paris, Singapore, Seoul). A second important challenge is that the relationship between the built-up area and the buildable administrative area can vary widely between cities. In some, the administrative boundary is cast far and wide, leaving ample room for growth. In others, the administrative boundary remains unchanged for decades allowing the city to hit and leapfrog over its edges.

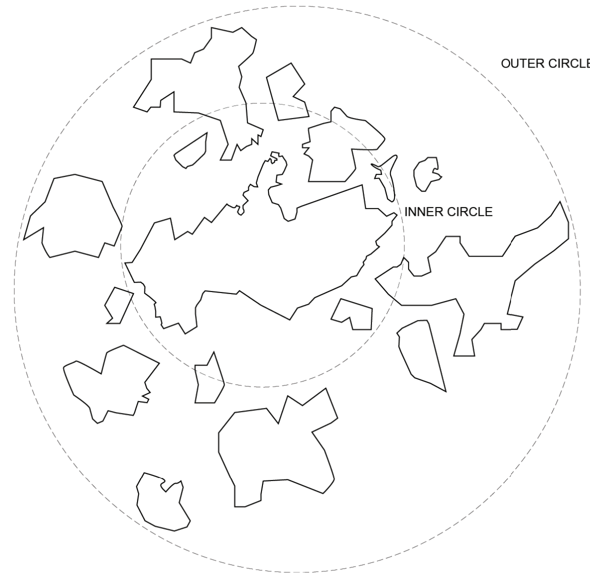


Figure 12. The circle of minimum radius encompassing the consolidated built-up area of the city. Source: Angel et al. (2005).

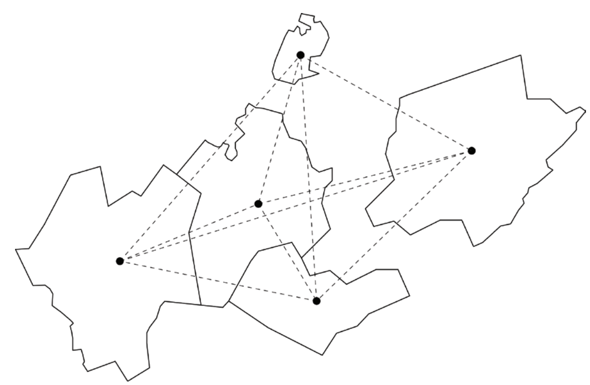


Figure 13. The proposed compactness metric is based on distance measurements between the centroids of built-up areas.

3 However, since the distance from census tracts to the CBD is weighted by population, then keeping the total population constant, but changing the arrangements of density between census tracts does affect the outcome.

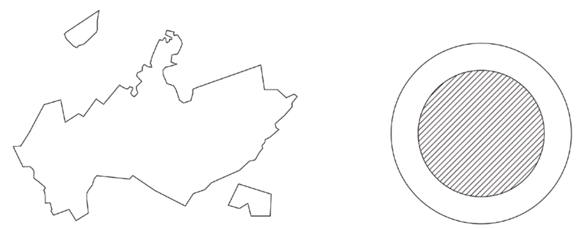


Figure 14. Normalization by a reference case. The area of the larger circle – reference case defined by Bertaud et al – is equal to the total built-up area. The area of our proposed reference disk, the smaller disk, is determined by the total population, and a reference density, rho, which is treated as a constant.

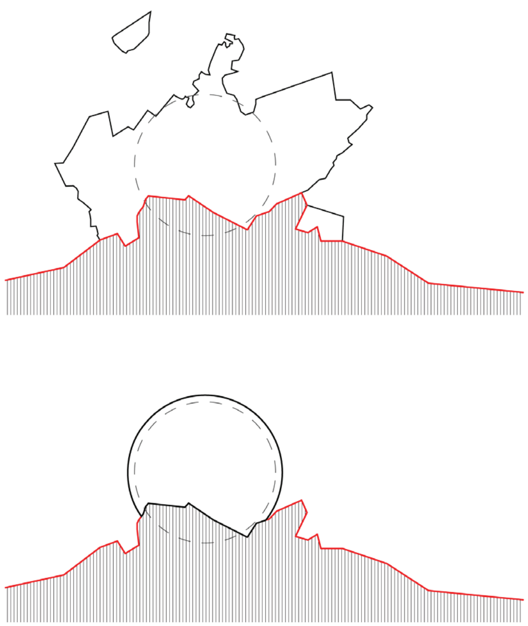


Figure 15. Normalization by a reference case, accounting for geographic constraints. The dashed circle is the reference case without accounting for the geographic constraints, and its area is equal to the area of the reference case polygon that accounts for the geographic constraints.

PROPOSED MEAURE OF COMPACTNESS

As mentioned earlier, a compactness index should capture the degree to which the resources of a city (e.g. people, buildings, jobs, etc.) are spread out. Put alternatively, the metric should capture how closely different parts of the city are accessible to each other (Figure 13). There is an analogous measure in transportation research, called "Gravity" (Hansen 1959). The gravity index of a location is proportional to the total amount of resources available to that location and inversely proportional to the travel cost of reaching them:

$$G_i = \sum_{j \in G-\{i\}} \frac{W[j]}{e^{\beta \cdot d[i,j]}}$$

, where G_i is the gravity index for location *i*, *W[j]* the size or attractiveness of the destination *j*, and *d[i,j]* the distance between locations *i* and *j*, and beta is the exponent that controls the effect of distance decay between *i* and *j*. Distance *d[i,j]* can be measured from the centroid of polygon *i* to the centroid of polygon *j*.

Computing the gravity index for each built-up polygon in the metropolitan area and taking the mean result across all polygons can thus capture how compactly these resources are situated with respect to each other. If weighted by the size of the resources, then the spatial relationships between larger destinations have a proportionately stronger effect on the index than smaller destinations.

$$G = \frac{\sum_{i \in G} W[i].G_i}{\sum_{i \in G} W[i]}$$

However, the weighted average gravity is obviously impacted by the total amount of the resources in the city. It would not make sense to get a higher average gravity in a larger city that has more people than in a smaller city that has less people, even if the latter is more compactly distributed. The mean gravity measure above, therefore cannot be used to compare compactness across cities of different size unless it is reasonably normalized. We propose three reasonable ways to proceed with the normalization:

Normalization by a Reference Case

The reference city is defined as a circular city with its number of resources (population, jobs, or built-up pixels) similar to the actual city, where the resources are uniformly distributed. Different from the cylindrical city defined by Alain Bertaud et al. the area of our reference city is not necessarily equal to the total area of the real case study city. As discussed above, using the city's own in the normalization will cancel density from the picture. Therefore, we propose to introduce a fixed density rho to the reference city — for example rho=1000 unit per square kilometer. As rho is a constant, the area can be derived from the given population in every city. Note that the same rho value should be used in all cities. If the weighted average gravity in the reference case is G_0, then the compactness index, *C*, would be:

$$C = \frac{G}{G_0}$$

Normalization by a Reference Case, accounting for Geographic Constraints

As discussed by Angel and his colleagues, we are often interested in knowing how compact a city is within its own geographic constraints. This can be achieved by super-imposing the geographic constraints on the reference case. The reference city, in this case, is a polygon with a uniform density rho, derived by subtracting the unbuildable land from the circle *C*. The reference circle shares the same geometric center as the observed city, such that the remaining area is equal to the area of the reference case calculated by the method in (a) (Figure 14).

Calculating the weighted average gravity for the reference case in which we account for geographic constraints is not feasible with a formula since the reference polygons have variable regions of unbuildable geometry. In order to be consistent in both types of reference cases introduced here, we suggest calculating the weighted average Gravity by applying a grid of 30x30 meters on the reference case and assigning an equal weight (*w*) to each pixel (Figure 16). The *w* of each pixel is the total amount of resources (population) divided by the number of cells in the reference case. The gravity for each pixel is then measured between the centroids of all pixels using the same Gravity index as above.

Normalization by Total Resources.

The mean gravity index across all observed polygons can also be normalized by the total amount of resources (i.e. population) in the city. This can be defined as follows:

$$G_{i-norm} = e.\frac{G_i}{W_{total}}$$

, where W_{total} is the total amount of resources in the city (e.g. population). The compactness index *C* is then given as the weighted average of the normalized gravity indices of all polygons *i*:

$$C = \frac{\sum_{i \in G} W[i].G_{i-norm}}{\sum_{i \in G} W[i]}$$

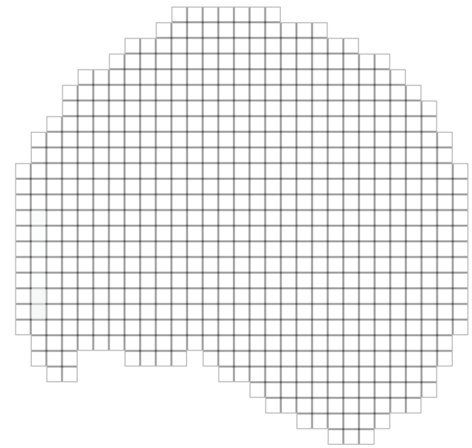


Figure 16. The 30x30m grid on top of the reference case.

Image source: http://low-tax-asia.com/wp-content/uploads/2012/05/Hong_Kong_China_02.jpg

EXPANDIBILITY

The Expandability metric aims to capture constraints to a city's growth by quantifying the availability of buildable land beyond the urban extent, within a defined peripheral area. Quantifying Expandability is fundamental to explaining sprawl, segregation, density and land prices.

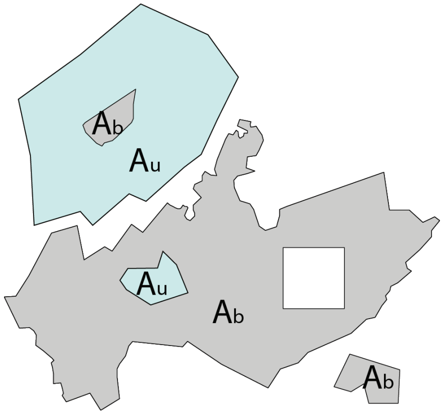


Figure 17. Existing built-up area A_b and unbuildable area A_u.

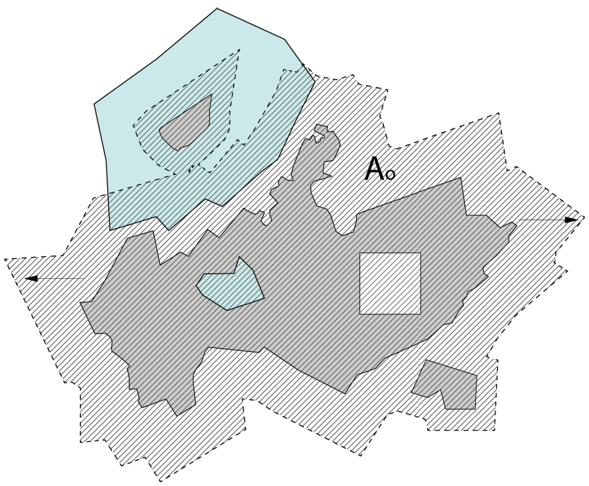


Figure 18. The Idealzied expansion, A_o; The area of A_o is twice the existing built-up area A_b.

EXPANDABILTY

A key determinant of a city's growth is the availability of buildable land in its vicinity. The availability of space to grow not only affects the rate of possible growth but also its character. Cities that are constrained by geographic features, such as water bodies or steeply sloped land, grow very differently from those with no barriers around them. The former, for instance, leave no room for leapfrog development and set serious physical limits on sprawl; the latter allow for spatially spread-out and fragmented growth. The Expandability metric aims to capture these constraints by quantifying the availability of buildable land beyond the urban extent within the non-urban realm. Quantifying expandability is fundamental to explaining sprawl, segregation, density and land prices.

EXISTING SPECIFICATIONS

The key challenge in quantifying the expandability of a city is to define a reasonable zone for the measurement of the buildable area around existing urban clusters, which could be analyzed consistently for expansion across cities.

City authorities are interested in knowing how much developable land is available in their administrative area, so they naturally look at the available land within the administrative boundaries of their city. Actual growth may occur, however, well beyond the existing administrative boundaries. The size of administrative boundaries in different cities can range widely and may occasionally constitute an area many times as large as the current urban extent. Studies of land supply and land demand management (e.g. Hopkins and Knaap 2000) have used the urban growth boundary as a limit for land supply. Growth boundaries, however, are inadequate for our purpose since few cities have legally implemented urban growth boundaries. Furthermore, the definition and regulation of growth boundaries varies widely across cities.

Albert Saiz's (2010) *Geographic Determinants of Housing Supply* is one of the rare studies on measuring developable lands that has disregarded administrative boundaries. Saiz's search area constitutes a 50-kilometer radius from the centroid of the city. But as Wendell Cox (2011) has rightly pointed out, an invariant search radius makes results incomparable in cities of different size. While in larger cities a 50-kilometer radius may barely cover the built-up area, for smaller towns it may contain several times their existing urban extent.

PROPOSED EXPANDABILITY METRIC

In order to specify a measure that captures the availability of buildable land in a consistent way across cities of different size, we propose an expandability metric that satisfies the following conditions:

a) The search radius should be measured from the edges of urban extent rather than its centroid.

b) The search radius should be relative to city size.

Let us first represent the existing built-up area as A_b, and the unbuildable area that contains natural obstacles like mountains or lakes as A_u (see Figure 17). Second, we can generate an idealized expansion of the existing built area by offsetting the boundaries of all polygons in A_b so far that the total built-up area of each polygon doubles[1], and then merging the polygons together into an overall doubled area that we call A_o (Figure 18). The offset radius at which we precisely double the area of A_b cannot be mathematically pre-determined, but it can be found in a simple automated iteration of offsets that check the expanded area against the original area until the right radius is found. While the area increase should be modeled consistently across cities, the offset radius will vary from city to city.

In the third step, we subtract all the unbuildable areas (A_u) from the offset area A_o to find the area A_e that is actually available for expansion beyond mountains and water (Figure 19). Finally, we compute our expandability index as the ratio between the expansion area A_e and the existing built area A_b (Figure 20):

$$Expandability = A_e / A_b$$

The index tells us how many times the current built-up area can expand within a 100% buffer from its current edges. The Expandability Index is constrained between 0 and 1; if all the land within the specified offset buffer is buildable, the index is 1, and if none of the land in the specified offset buffer is buildable, the index is 0.

When we consider areas that are unbuildable, we can categorize such land into two groups:

1) Land that is unbuildable due to natural obstacles, such as mountains or water.

2) Land that is unbuildable due to human policy choices, such as parks, protected areas or urban growth boundaries.

The two types of unbuildable land can have a different effect on a city's growth. Those areas that are put aside as unbuildable due to conscious human regulation could have a higher risk of being re-zoned than natural obstacles. But even natural obstacles are not set in stone – Singapore has expanded its shorelines by 20% in 40 years, and a number of cities have historically leveled mountains to make way for urbanization. These risks can be integrated into our Expandability index by considering the financial and technological capacity of the city and allowing a small fraction of the obstacles to be overturned each year.

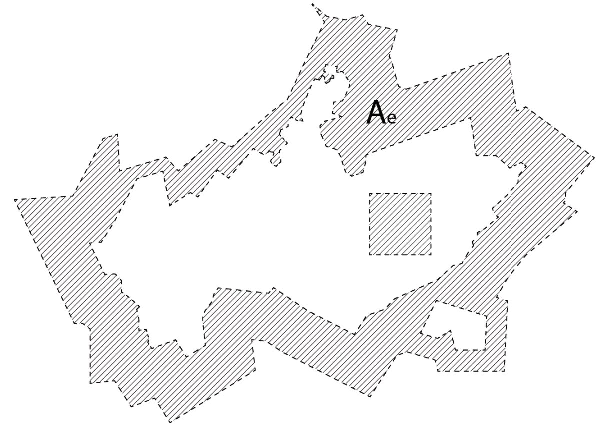


Figure 19. The expansion area A_e is found by subtracting all the unbuildable areas A_u from the oidealized ffset area A_o.

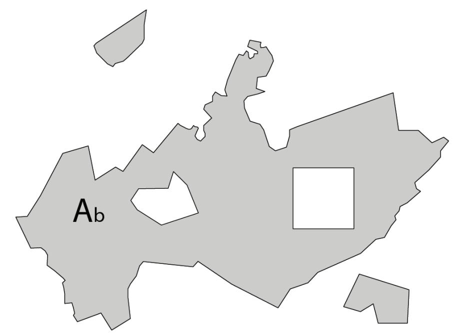


Figure 20. The final expandability metric is computed as the ratio between the expansion area A_e and the existing built area A_b.

1 The particular choice of 100% expansion roughly matches the average 20-year growth that was observed across 66 cities in Angel's study (2005). The radius can be adjusted as needed by the analyst to reflect typical annual growth, 5-year growth or other growth of existing built area.

Image source: http://www.vad1.com/photo/stock/a84-30-5.jpg

DISCONTIGUITY

In order to gain a deeper quantitative description of the Discontiguity of metropolitan form, we need to look at the rank order and relative size difference between discontinuous urban clusters.

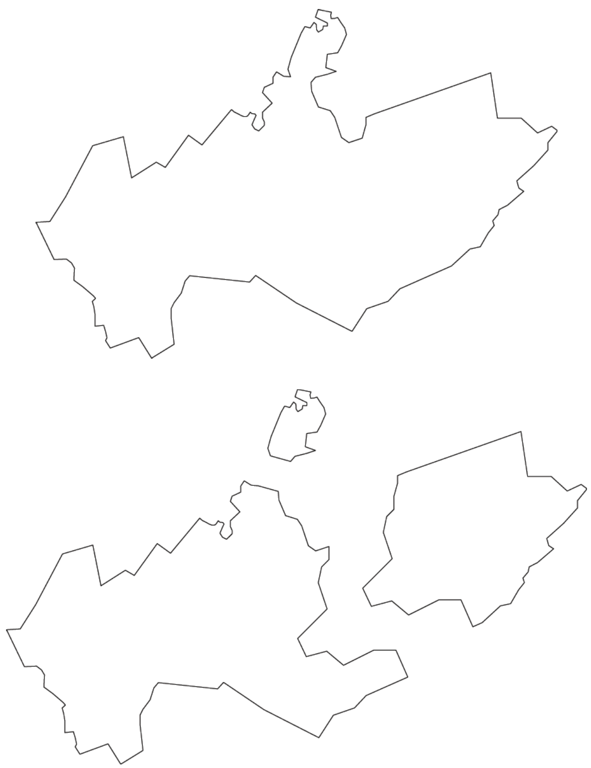


Figure 21. Continuous urban extent (top) and discontinuous urban extent (bottom)

DISCONTIGUITY

While the number of urban clusters and their size provide a general description of the discontiguity of an metropolitan area (see Figure 1), they do not tell us much about the structure of the city's fragmentation. In order to gain a deeper quantitative description of the discontiguity of metropolitan form, we need to look at the rank order and relative size difference between discontinuous urban clusters.

EXISTING SPECIFICATION

Efforts to quantify the contiguity of urban form or its reverse quality – discontiguity – have been rare. The most popular description of contiguity, developed by Angel and his colleagues (2005), describes contiguity as the ratio between the main (largest) built-up area of the city and the sum total built-up area of the city. The more built-up area is concentrated into the single largest cluster, the more contiguous the city is. This measure is easy to compute and is useful as long as the main built-up area constitutes a large portion of the total built extent of the city. But the metric is not well suited to distinguish forms of discontiguity when a city is made of multiple larger or smaller built-up clusters, with a large portion of the total built-up area located outside of the biggest cluster. The metric does not account for rank-size relationships between individual discontinuous areas beyond the largest cluster.

PROPOSED DISCONTIGUITY MEASURE

Similar to Angel et al. (2005), we assume that the fewer the total number of discontinuous developments, the more contiguous a metropolitan area is (Figure 21).

Although our proposed description is a measure of discontiguity rather than contiguity, it is in essence similar to the measure developed by Angel and his colleagues (2005). While based on the relative size of urbanized clusters, the metric additionally accounts for the areas of all clusters that are smaller than the largest cluster (Figure 22). We define discontiguity as follows:

$$DC = \sum_{n=1}^{N} \left(\frac{\sum_{i=n+1}^{N} A_i}{A_n}\right)\left(\frac{\sum_{i=n}^{N} A_i}{A_{total}}\right)$$

where *DC* is the discontiguity of the built-up area, *N* the number of urbanized clusters, A_n the area of cluster *n*, and A_{total} the joint area of the urban extent. Note that $A_n \geq A_{n+1}$, so that the denominator in the first part of the index always compares other areas to the largest continuous area.

The key improvement of the proposed index is that it accounts for the size relationships between all the clusters in the city by calculating the same ratio – the area of each cluster that is smaller than the largest cluster to the area of the largest cluster – and summing up these ratios, weighted by their share of the total area.

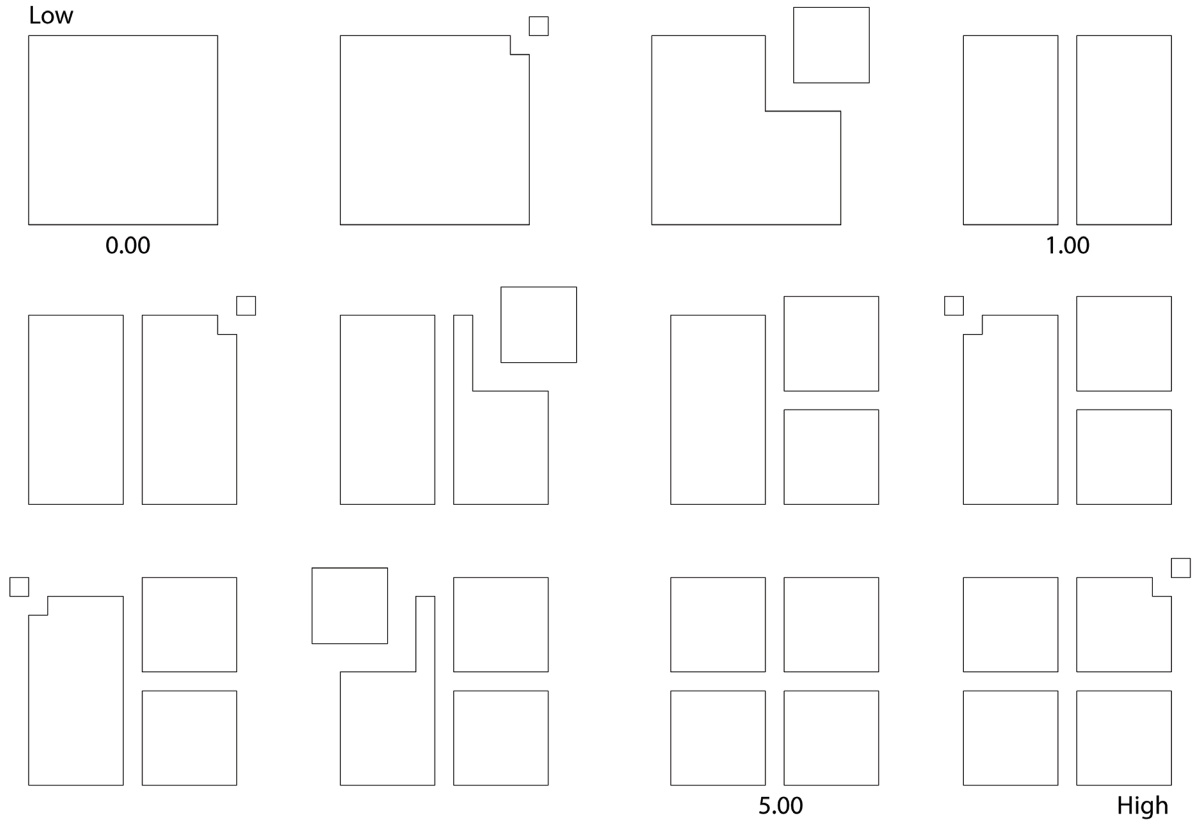


Figure 22. Interpretation of the discontiguity metric. The top left configuration has the lowest discontinuity results, the bottom right one the highest. The numbers indicate the actual computed results for selected configurations.

Image source: http://www.gaudetfamily.org/Photos/BangkokLG.jpg

LAND-USE MIX

Land-use Mix constitutes an important characteristic of the built environment whose impacts on traffic congestion, energy consumption, real estate values and crime are extensively discussed in planning literature. While mixed-use developments are widely promoted by planners, supporting quantitative evidence of their effects is underdeveloped and even contradictory.

LAND-USE MIX

Land-use Mix constitutes another important characteristic of the built environment whose impacts on traffic congestion, transportation energy consumption, real estate values and crime rates are extensively discussed in planning literature. While mixed-use developments are widely promoted by planners, supporting quantitative evidence of their effects is underdeveloped and even contradictory. This is partly attributable to a lack of intuitive and commonly accepted metrics that can capture how mixed or segregated the land-uses of an urban area are. Perhaps more importantly, there has been little discussion on what levels and configurations of mixing are actually desirable.

EXISTING MEASURES OF LAND-USE MIX

There are two popular types of metrics that capture land-use mixing. The first focuses on the number of different uses that are found in a given area, allowing comparative areas to be ranked according to the number of land-use types (1-n) they accommodate. The second focuses on the relative balance between uses; it tells us how heterogeneous or homogenous the land-use pattern of an area is based on how equally the area is occupied by different uses.

The former may not be meaningful if different uses occupy notably different amounts of land. This shortcoming is addressed in the latter, but heterogeneity indices, too, have important shortcomings. Most of such indices weigh all uses equally and assume that an equal distribution of each type of use is the benchmark to compare an observed pattern against. Cities do not have an equal share of all land-uses – a much larger share of land is typically used for residential purposes than commercial purposes. Industrial and transportation lands often top commercial land, as the former tend to accommodate rather land-intensive activities. Second, not all types of land-uses tend to mix with each other at equal likelihoods (Hess et al. 2001). A mixture between commerce and housing is far more likely to be seen in most cities than a mixture between industry and housing. Military land and agricultural land tend to stay apart from other land uses for logistical, security, and economic reasons. Instead of an equal weighting and mixing benchmark, a land-use mix metric could use a statistically "expected weighting" that is based on the observed citywide balance of land-use types, and an "expected mixing" ratio that is based on desirable examples.

PROPOSED MEASURE OF LAND-USE MIX

A meaningful implementation of a land-use mix metric requires that the metric be estimated at the scale of small intra-urban subdivisions, not the whole city together.[1] Land-use mixing can be estimated with most accuracy if the input data is given as pixels or raster cells, where each pixel contains information about the land uses it accommodates.

For each raster cell, we propose a land-use mix

1 At the whole city level, land uses always appear mixed, even if they are completely segregated at a finer scale.

metric that illustrates how closely the distribution of observed uses in a given neighborhood around that cell correspond to an expected distribution. There is no consensus on how large the evaluation neighborhood ought to be, but we propose to use a standard of one square kilometer (1 km^2) that is large enough to detect use mixing and convenient to derive in many cities (Burchfield 2006).

Let us define the square kilometer neighborhood of a given cell as *i* and a particular land use of interest *n* as LU_n. Since there is often more than one land use in the square kilometer around the cell, we can give a weight *w* to each of these land uses, based on how much area of the neighborhood they occupy. $w[LU_{n:i}]$ thus denotes the weight of land use *n* within the square kilometer neighborhood of cell *i*. We can now specify $S_{n:i}$ as the share of LU_n among all land uses of interest with the area *i*:

$$S_{n:i} = \frac{w[LU_{n:i}]}{\sum_{n=1}^{N} w[LU_{n:i}]}$$

$S_{n:i}$ is given as a ratio between the weight of land use *n* within *i* and all other land uses of interest (including *n* itself) within *i*. $S_{n:i}$ thus ranges between 0 and 1.

Knowing $S_{n:i}$ – how big a share of *i* is covered by land use *n* among all uses of interest – allows us to estimate how closely this observed coverage matches an expected coverage of that same land use *n* within the area of *i*. For that we need to first specify the expected share of the same land use in *i* as $ES_{n:i}$. $ES_{n:i}$ can be determined in a number of ways depending on the intentions of the analyst; we propose to base its specification on the following criteria:

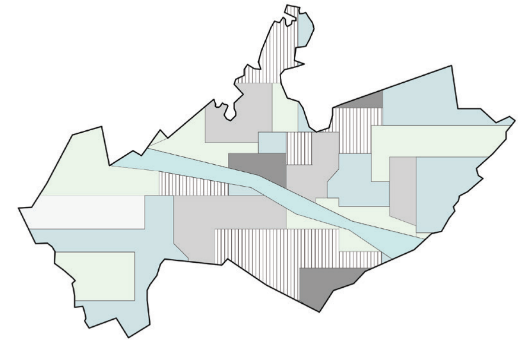
1) The expected distribution $ES_{n:i}$ of land use *n* in area *i* should depend on the city-wide total balance of all land-uses of interest. Given the presence of land use *n* in the whole area of the city, we can determine the statistical likelihood of its presence in any smaller sub-area of the city.

2) The expected distribution should also depend on the likelihood of co-location between any pair of land uses. A mixture between commerce and housing is more likely in most cities than industry and housing. The "expected mixing" ratio should, however, not reflect the city-wide average, but instead a desirable scenario based on exemplary locations.

Having set the benchmark for each area *i*, we can now estimate a "matching index" ratio $M_{n:i}$ that shows how closely the observed coverage of land use n matches the expected distribution of that same land use *n* within the area of *i*. The "matching index" ratio $M_{n:i}$ is determined by first finding the difference between the observed share of land use *n* in *i* from the expected share of the same land use *n* in *i*, and then subtracting the absolute value of this difference from one:

$$M_{n:i} = 1 - |S_{n:i} - ES_{n:i}|$$

The absolute value alone would tell us how much the observed value deviates from the expected value, and subtracting this deviation from one tells us how closely, in terms of percentage, each observed land use matches its expected target.



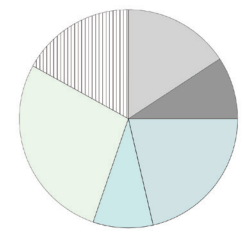


Figure 23. The expected land-use mix should be based on the city-wide distribution. Different cities include different types and balances of land-uses.

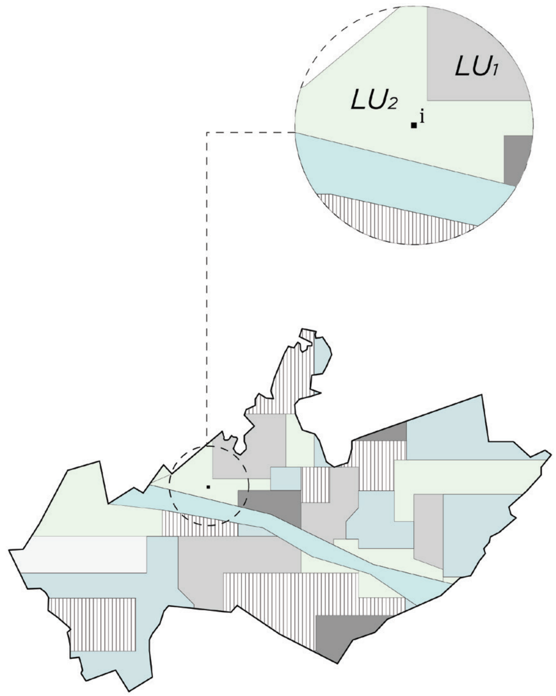


Figure 24. Land-use Mix around location i. LU_1 and LU_2 are the target land uses. $w[LU_{1:i}]$, the share of LU_1 within the immediate one square-kilometer around i, is 0.15 Likewise, $w[LU_{2:i}]$ is 0.45. S_i, the total share of LU_1 and LU_2 is 0.6 (0.15+0.45). $S_{1:i}$ is 0.15/0.6, (0.25) and $S_{2:i}$ is 0.75. If SD_1, the expected share of LU_1 is 0.3 and SD_2 is 0.7, then both $M_{1:i}$ and $M_{2:i}$ will be 0.95 and finally land-use mix index of i, MX_i, *is 0.6x0.95x0.95= 0.54*

If residential use, for instance, constitutes 20% of the observed land cover ($S_{n:i}$) in the square kilometer around *i*, and the expected cover ($ES_{n:i}$) is 60%, then the matching index $M_{n:i}$ for this particular land use around location *i* is: 1 - |0.2 – 0.6| = 0.6.t

Finally, we need one more ingredient before finding the total land use mix index around *i* – the total observed share of all land uses of interest around *i*, which we call S_i. Using S_i allows us to focus our analysis on only a selected set of land uses and ignoring others without compromising the validity of the index. If we are only interested in the mixing between commercial and residential land, then we specify S_i to only include these two uses. It is found as follows:

$$S_i = \frac{\sum_{n=1}^{N} w[LU_{n:i}]}{w_i}$$

, where S_i is the observed share of all land uses of interest around *i*, the sum in the numerator denotes the total observed area of the land uses of interest around *i*, and w_i is the absolute total of all land uses around *i*, including not only the ones that interest us, but all land uses. If the weights are measured in land area (e.g. in square kilometers), which is commonly the case, and our chosen neighborhood area around *i* is 1 km2, then w_i is always 1. However, we indicate this total as w_i instead of simply one for scenarios in which the weights might be measured in other units, such as total floor areas of buildings, etc.

The final land use mix index MX_i around location *i* is given by multiplying the observed share of all land uses of interest around *i* (S_i) with the product of all individual matching indices $M_{n:i}$ around *i*:

$$MX_i = S_i \cdot \left(\prod_{n=1}^{N} M_{n:i} \right)$$

This land use metric tells us how closely the distribution of all land uses of interest around location *i* correspond to their expected distribution (Figure 23). MX_i always ranges between 0 and 1. MX_i is at its maximum value when the land uses in the immediate square kilometer around *i* perfectly match the expected distribution. MX_i is zero when none of the expected uses are found in the area of *i*.

In order to obtain a combined index for the entire city, we can simply take the average of all individual MX_i indices:

$$MX = \frac{\sum_{i=1}^{N} MX_i}{N}$$

The combined city-wide land-use mix metric *MX* tells us how closely the average distribution of land uses across all analysis areas in the city corresponds to their expected distribution.

The key advantage of this proposed land-use mix metric lies in its flexibility in working with different combinations of uses, as well as the fact that it can be calibrated for different expected distributions. It can be used to evaluate land use mixing for only a narrow set of uses or all uses found in a city. A similar metric can also be used for evaluating other types of spatial mixing or segregation, such as the spatial mixing of different demographic, income, or racial groups.

SUMMARY TABLE

Metric Name	Description	Example Specifications	Basic Data Required	Optional Data Required	Computational Requirements	Related Indicators
SIZE	Estimates the area of a land-use category.	Size of urban extent (built-up area)	Total built-up area	Convex hull polygon around the built-up area; Area of unbuildable lands within the convex hull polygons; Area of land-use categories (observed/remotely sensed).	Simple arithmetic.	Regulatory climate, car ownership, transit ridership, land prices, crime rate, transport energy consumption, corruption, vehicle miles traveled, GDP, avg. HH income, Gini coefficient, cost of living, water resources, temperature, total population, water resources, total nbr. of jobs/firms/establishments,
COVERAGE	Estimates the ground cover of a land-use type within a total area.	Coverage of built-up area within the convex hull.	Total built-up area, convex hull polygon around built-up area, area of unbuildable lands within the convex hull.	Area of buildings footprints (aggregated), Area of any land-use category (observed/remotely sensed).	Simple arithmetic.	Land price, total nbr. of jobs/establishments, crime rate, transport energy consumption, vehicle miles traveled, regulatory climate, temperature, water resources.
DISCONTIGUITY	Quantifies the degree to which a city is fragmented into discontinous built-up areas. The metric jointly increases by the number of dicsontinuous developments and the size inbalance between the developments.		Areas of all individual built-up polygons.		Iterative arithmetic.	Climate, water resources, crime rate, share of aging population, regulatory climate, car ownership, transit ridership, land prices, transport energy consumption, vehicle miles traveled.
COMPACTNESS	Indicates the average spatial accessibility between separate built up areas – the higher these accessibilities, the more compact a city is.	Compactness of population, compactness of built-up pixels, compactness of building footprints.	Geometric built-up polygons with population data.	Census tracts with data on population/jobs/establishments/etc. Remotely sensed built-up pixels, building footprints, building volumes.	Calculating X,Y coordinates of the centroid (of pixels/polygons), distance measurement, iterative arithmetic.	GDP, regulatory climate, car ownership, avg. HH income, transit ridership, land prices, transport energy consumption, vehicle miles traveled, crime rate, temperature, total nbr. of jobs/firms/establishments

Metric Name	Description	Example Specifications	Basic Data Required	Optional Data Required	Computational Requirements	Related Indicators
POLYCENTRICITY	Estimates the degree to which a city's employment (or other activity) is concentrated in centers. Polycentricity depends simultaneously on three factors: a) the number of centers; b) the size-balance between centers and c) the share of total employment that is located in centers.	Employment polycentricity, polycentricity of the built-up area, population polycentricity.	Geometric census tracts with employment data.	Raster employment density map, census tracts with population density, remotely sensed built-up pixels.	Boolean functions (merge), simple arithmetic, iterative kernel density measurement (for optional data).	Transport energy consumption, GDP, regulatory climate, car ownership, avg. HH income, Gini coefficient, total nbr. jobs/firms/establishments.
EXPANDABILITY	Illustrates how much space is available for development beyond the city's current borders in a given distance threshold.		Geometric built-up areas, unbuildable polygons (water bodies, steep slopes).		Boolean functions (subtract), area calculation, offset, iterative arithmetic.	Land price, water resources, cost of living, Gini coefficient, total nbr. Of jobs/firms/establishments.
LAND-USE MIX	Captures the degree to which the observed distribution of land uses corresponds to an expected distribution.	Mix of residential, transportation, commercial and green land-use categories in a city.	Geometric polygons of land-use categories.	High-resolution land-use categories.	Clip, area calculation, iterative arithmetic.	Regulatory climate, share of aging population, car ownership, transit ridership, land prices, crime rate, transport energy consumption, corruption, vehicle miles traveled.

3. GROWTH

GROWTH

The metrics defined in the previous section offer snapshots of built environments at particular points in time. The metrics alone, however, do not tell us much about growth or change that takes place in cities over time. Cross-sectional measures of metropolitan form offer the basis for understanding trends in growth and change of urban development patterns, but to detect the changes, the metrics need to be captured at multiple time points and additional tools are required. In this section, we present a brief overview of mapping, visualizing, and measuring tools that can be potentially used to capture changes in the urban forms and land-use metrics over time.

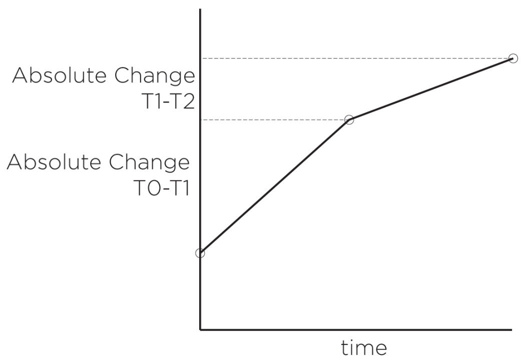


Figure 25. Trend analysis models a dependent variable on the y-axis and time on the x-axis.

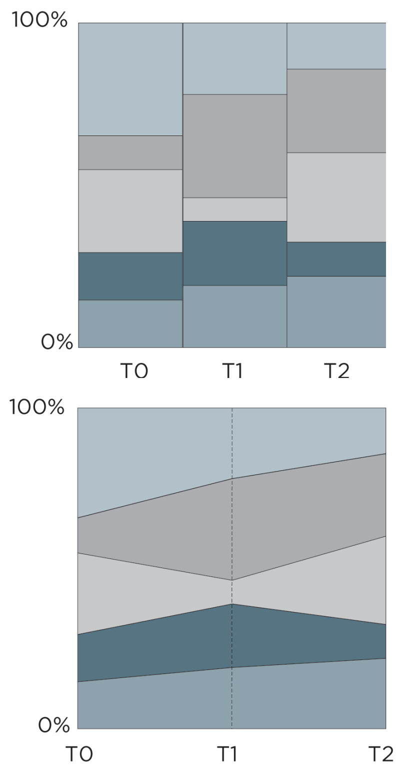


Figure 26. Shift-share analysis illustrates proportional changes in quantitites over time. The total quantity adds up to 100% at any given point in time.

TREND ANALYSIS

The most popular way of capturing change is to measure the differences in the absolute value of a metric at different points in time. Thus a comparison of two measurements over a given time period can tell us whether a city is growing denser or sparser, bigger or smaller, more polycentric or monocentric. In order to construct a trend, at least three data points over time are needed (e.g. 1990, 2000, 2010). Plotting the trend-line on a graph, with time on the x-axis, can help us visualize the rate of change of over time.

SHIFT-SHARE ANALYSIS

Another tool for visualizing change is offered in a "shift-share" graph. A shift-share graph depicts how big a share of a total amount a subset category consumes over time. It can be used, for instance, to depict the balance of a city's land uses over time, where the total always adds up to 100%, but the share of each land-use may vary across time. A shift-share graph allows us to both track individual share trends as well as the combined share balance across all categories over time.

MAP OVERLAY

A superimposition of maps from different times is not only an intuitive way of visualizing growth, but also a powerful method for capturing changes in unexpected qualities of spatial environments that may not be directly available in the urban form metrics.

Through simple Boolean functions, for example, leapfrog developments can be distinguished from extensions that are contiguous to existing built-up areas. Finer scale building data can be used to distinguishing infill from extension and so on.

Map overlays also allow us to capture transformations of land-use patterns over time. An old industrial site, whose location has gradually become more central, may be recycled as new residential land, or an informal settlement may transform into a consolidated middle-class area. This process of land recycling can be understood by merely looking at the change in the overlaid land-use map.

REGRESSION MODELS OF CHANGE

A trend diagram visualizes the bivariate relationship between a metric and time, but does not control for other potentially important variables that can impact a metric of interest. A longitudinal regression model, or growth model, is a regression model in which one of the independent variables is time. The model can be used to capture the relationship between an urban form metric and time, while keeping other important covariates (Willett, Singer 2003).

MULTILEVEL MODELS

Multilevel models, also known as hierarchical linear models, or nested models, are statistical models of parameters that vary at more than one level. In the context of studying metropolitan form, multilevel models can be used to study relationships between the question variable and independent variables at multiple scales – what portion of variation in the dependent variable is explained by variations in the independent variables within cities or between cities. The choice of levels could distinguish variables that are captured for a city as a whole – total urban energy expenditure – or in intra-urban tracts – what is the economic output in each census tract. Multilevel models offer a powerful way to move urban expansion analysis beyond the comparison of multiple cities to a two-level comparison of differences between cities and within cities (Snijders and Bosker 1999).

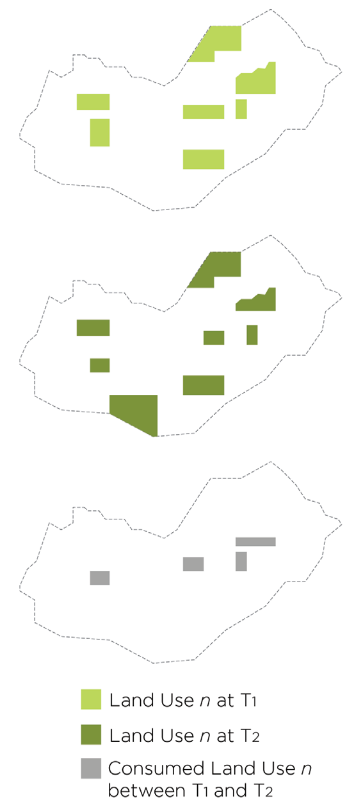


Figure 27. Map overlay analysis of land-use growth.

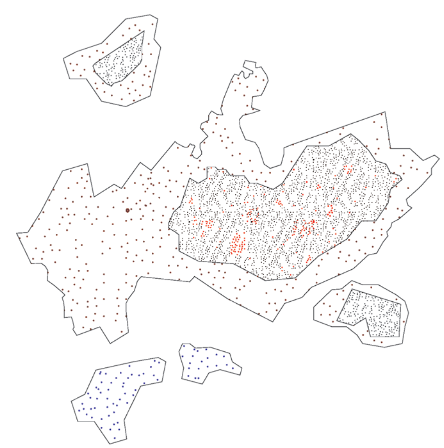


Figure 28. Map overlay analysis of fine-grain land-use change.

4. DETERMINANTS

DETERMINANTS

Metrics of metropolitan form illustrate certain physical properties of the cities' development patterns. A longitudinal analysis of these metrics over time may tell us how these patterns are changing – is a city growing more compact or sprawling, are its centers consolidating around particular locations or is employment becoming more scattered? An examination of metrics that this report has presented may help us answer questions of *how much*, *where* and *how* the formal patterns of cities are changing. But the analysis of metropolitan form alone does not offer deep insights into *why* we observe these patterns or *what* explains a certain change in the data.

In order to understand why urban spatial patterns take particular forms or follow certain development trajectories, we need to test hypotheses that relate metropolitan form metrics to other non-geometric data, which we call the *determinants* of form. In reality, cause and effect between spatial forms and social processes are very difficult to untangle. It is unlikely, for instance, that growing economic output is clearly the egg and urban expansion around it the chicken. Rather, an existing urban environment with its reputation and history, probably also affects where economic development happens and how effectively economies operate or expand. Businesses are more likely to locate in places with advantageous spatial capital, but spatial capital is also more likely to develop at a higher rate around economies that do well. Where might this chain of circular causality begin? Each environmental intervention is likely to be influenced by pre-existing environmental conditions, but once implemented, the intervention becomes part of the environment for the subsequent interventions.

Though the causal chain of factors that lead to a particular metropolitan form poses a fascinating theoretical paradox, its full resolution might not be so important from a practical point of view. What matters more in the practice of urban planning and management is that we have some understanding of the implications of our own interventions in the built environment.

Given that a city is planning to cut its transportation energy expenditures by half for instance, we might desire to know how such a change would affect the existing development pattern of the city. Or given that a new job center is planned, one might like to assess its potential impact on the existing employment districts or traffic patterns in the city. In situations where interventions are planned in a pre-existing urban context, the chain of causality thus shrinks to a manageable set of variables, which spatial and non-spatial indicators of the built environment can begin to inform.

Overall, the relationship between metropolitan form and different social, economic and environmental factors is poorly understood. Much more work is needed in this area. In Figure 28 we simply present a basic overview of some relationships between geometric form and non-geometric indicators that have been proposed in the literature on the subject. Developing a better understanding of these relationships can help us:

a) Provide a measurable basis for understanding not only what the current urban expansion trends are, but where they came from and where they might be headed.

b) Inform us of what policies and planning interventions might possibly achieve certain effects of development patterns.

c) Provide a basis for evaluating the consequences of already implemented policies and design or planning interventions.

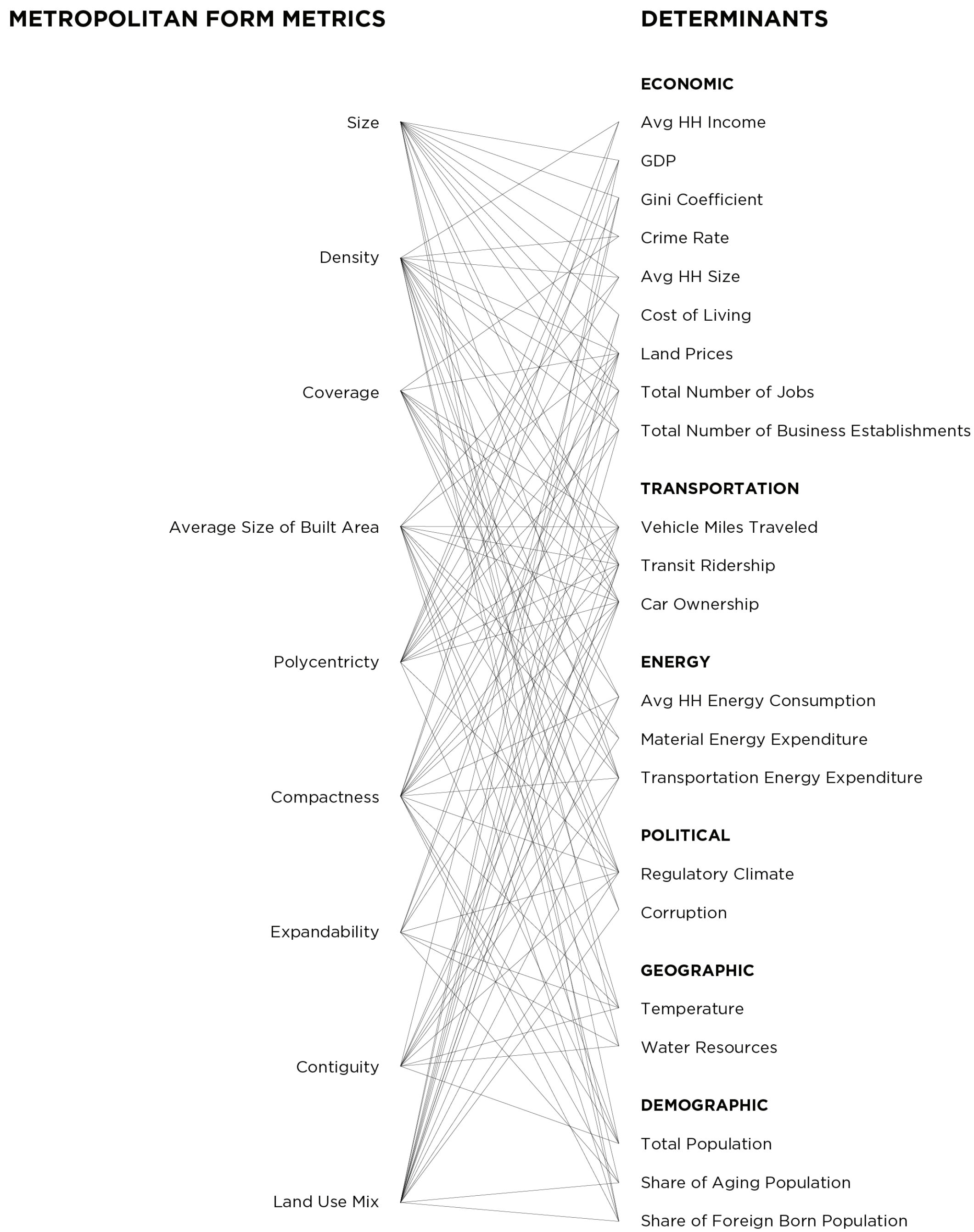


Figure 29. “Determinants” of metropolitan form.

5. POLICY IMPLICATIONS

POLICY IMPLICATIONS

The significant advancements in GIS-based computational spatial analysis tools and methods of collecting geo-referenced data have profoundly changed our understanding of urban processes. It seems that our more dynamic understanding of city growth and change have brought the idea of a static utopia — an ideal urban form— to an end. Rather than seeking a unique global "best answer", which should be imposed on the built environment, future policies, based on a deep understanding of existing trends in the city, should aim to achieve maximum profit through minimal interventions and investments. This requires a dynamic monitoring of urban processes both at inter-urban and intra-urban levels through measurable evidence, which has been the main focus of our work, and independent of the vehicles of policies, can be applicable to: a) directly regulating the built environment — such as zoning, growth boundaries, and building codes— b) incentives or c) public investment in the public realm and infrastructure.

The more measurable evidence for local socio-economic, environmental, and spatial performance of cities becomes available through efforts of spatial analysis, the more localized policies can be. The tradition of evidence based planning is still new and much work remains to be done not only in analysis methods, but also in adaptation institutionalization and up-keep on behalf of stakeholders. Yet, there are also global patterns in the processes of urbanization: phenomena that are similar across numerous cities, despite differing local characteristics. They include phenomena like the rapid rate of urban expansion, the sparsening of development patterns, and spatial segregation of social classes. Studying these global patterns may suggest policies that are similar in their direction, but different in detail.

REFERENCES

Angel S., Sheppard S.C., & Civco, D.L. (2005). *The dynamics of global urban expansion*. Transport and Urban Development Department, The World Bank, Washington, DC.

Alonso, W. (1964). *Location and land use*. Cambridge: Harvard University Press.

Batty, M. (2006). Rank clocks. *Nature*, 444, pp.592–596.

Batty, M. (2008). Size, scale and shape of cities. *Science*, Vol. 319(No. 5864), pp.pp. 7669–771.

Bertaud, A. (2004). The spatial organization of cities: Deliberate outcome or unforeseen consequence? *Working Paper*, Institute of Urban and Regional Development, University of California at Berkeley.

Bertaud, A., & Malpezzi, S. (2003). The spatial distribution of population in 48 world cities: Implications for economies in transition. *Unpublished manuscript*.

Bettencourt, L.M.A. et al. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences*, 104(17), pp.7301–7306.

Burchfield, M., Overman, H. G., Puga, D., & Turner, M. A. (2006). Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics*, 121(2), 587-633.

Burgalassi, D. (2010). Defining and measuring polycentric regions. The case of Tuscany. *Discussion Paper*, n. 101, Dipartimento di Scienze Economiche, Università di Pisa.

Changizi, M. & Destafano, M. (2009). Common scaling laws for city highway systems and the mammalian neocortex. *Complexity*, 15, pp.11–18.

Cox, W. (2011). Constraints on housing supply: Natural and regulatory. *Econ Journal Watch*,8(1),13-27.

Geddes, P. (1915). *Cities in evolution: An introduction to town planning movement and to the study of civics*. London: Williams & Norgate.

Forsyth, A. (2003). Measuring density: Working definitions for residential density and building intensity. *Design Brief*, 8, Design Center for American Urban Landscape, University of Minnesota.

García-López, M. A. (2008). Manufacturas y servicios en la RMB, cambios en la estructura espacial de su empleo. *Revista de Estudios Regionales*, 83,197-224.

García-López, M. A. (2007). estructura espacial del empleo y economías de aglomeración: El caso de la industria de la región metropolitana de Barcelona. *Architecture, City & Environment*, 4, 519-553.

Giuliano, G. & Small, K. A. (1991). Subcenters in Los Angeles region. *Regional Science and Urban Economics*, 21, 163-182.

Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Planning Association*, 25(2), 73–76.

Hall, P.G. & Pain, K., (2006). *The polycentric metropolis : Learning from mega-city regions in Europe*. London, Sterling, VA: Earthscan.

Hess, P. M., Moudon, A. V., & Logsdon M. G. (2001). Measuring land use patterns for transportation research. *Transportation Research Record*, 1 780, 17-24.

Hopkins, L. D., & Knaap G.J. (2000). Portland Oregon:

an inventory approach and its implication for database design. In A. V. Moudon & M. Hubner, eds. *Monitoring land supply with geographic information systems: Theory, practice and parcel-based approaches*, New York: Wiley.

Heywood, D. I. (1998). I*ntroduction to geographic analysis*. New York, Addison Wesley Longman.

Limtanakool, N., Schwanen, T., & Dijst M. (2009). Developments in the Dutch urban systems on the basis of flows. *Regional Studies*, 43(2), 179–196.

Lynch, K.,(1991). The pattern of the metropolis. In T. Banarjee & M. Southworth, eds. *City sense and city design: Writings and projects of Kevin Lynch*. MIT Press, 47–64.

Masip, J. (2011). Polycentrism and emerging sub-centers in the restructuring of metropolitan systems: The case of Barcelona metropolitan region (RMB). *58th Annual North American Meetings of the Regional Science Association International (NARSC-RSAI)*, 1-25.

McDonald, J., & Prather, P. (1994). Suburban employment centers: The case of chicago. *Urban Studies*, 31, 201-218.

McDonald, J. F. (1987). The identification of urban employment subcenters. *Journal of Urban Economics*, 21, 242-258.

Mills E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, 57(2),197-210

Mumford, L. (1961). *The city in hisory*, New York: Harcourt, Brace & World.

Muñiz, I.,& García-López, M. A. (2009). Policentrismo y sectores intensivos en información y conocimiento. *Ciudad y Territorio Estudios Territoriales*, 160.

Muth R F. (1969). *Cities and housing: the spatial pattern of urban residential land use*. Chicago, IL: University of Chicago Press

Openshaw, S. (1984). *The modifiable areal unit problem*. Geo Books.

Parent J., Civco D. and Angel S. (2009). *Shape metrics*. Center for Land Use Education and Research

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125 (3), 1253-1296.

Sevtsuk, A. & Mekonnen, M. (2012). Urban network analysis toolbox. *International Journal of Geomatics and Spatial Analysis*, 22(2), 287–305.

Snijders, T. A. B., Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, California: SAGE

Singer, J. D., Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford: Oxford University Press.

The dating game: We invite you to predict when China will overtake America. (2011). *The Economist Online*.

UNFPA (2007). *State of world population*. UN.

UN-HABITAT (2006). *State of the world's cities*. UN.

Webber, M. (1963). Order in diversity: community without propinquity. In L. Wingo, ed. *Cities and space: The future use of urban land*. Baltimore: Johns Hopkins Press, p. 261.

Webber, M. (1964). The urban place and the nonplace urban realm. In M. Webber, ed. *Explorations into urban structure*. Philadelphia, University of Pennsylvania Press.

Wheeler, S.M. (2008). The evolution of built landscapes in metropolitan regions. *Journal of Planning Education and Research*, 27, 400–416.

City Form Lab
at SUTD, established in collaboration with MIT | **cityform.mit.edu**

20 Dover Drive
138682 Singapore

+65-63036600
cityformlab@mit.edu