CHAPTER 3

Understanding and Measuring Urban Layouts

INTRODUCTION

Volume 1: Areas and Densities focused on the physical extents of urban areas on our planet today, their key attributes, and their change over time. Its main thrust was to alert readers—be they policy makers, public officials, academics, civic groups, or interested citizens—to the *quantity* of land converted to urban use and its relation to urban population growth, as well as to key attributes of the resulting physical extent of cities—density, fragmentation, and compactness—and their change over time. As the maps and metrics in Volume 1 clearly illustrate, the majority of cities expand outwards at a faster rate than the population they accommodate. While higher rates of land consumption per capita are largely accounted for by economic growth, by the availability of inexpensive transport, and by the plentitude of land for urban expansion, they may still be a cause for concern, calling for public intervention in urban real estate markets. Slowing down the existing rates of urban expansion would require effective strategies to facilitate the densification of existing urban extents, both by removing regulatory barriers and by addressing local community resistance to densification in its various forms. It may entail, among other things, allowing and promoting smaller dwelling units, smaller plots, higher plot coverage, taller buildings, the transformation of more land to residential use, and the infill of vacant open spaces, both public and private. It may also entail facilitating higher-density development in the expansion areas of cities, permitting, among other things, the construction of multi-family dwellings and small-lot townhouses, and the designation of more lands for residential use.

However, regardless of whether the rapid rate of urban expansion requires public intervention to slow it down or not, there is a second and separate concern that needs to be addressed: None of the attributes described and measured in Volume 1 informs us about the physical layouts of urban areas or about their change over time. It may well be that cities are expanding in an orderly manner—ensuring that they are as productive, as inclusive, and as sustainable as can be—and if they indeed are, then we need not be unduly concerned about the quality of their expansion. But it may also be that cities are expanding in a disorderly manner that is not productive, not inclusive, and not sustainable. In this case, the *quality* of their physical expansion should be of great concern, regardless of its quantitative dimensions.

Cities become more productive when *all* workers have access to *all* jobs; they become more inclusive when they provide decent and affordable housing for all, with residential amenities and good access to these jobs; and they are more sustainable when they provide more of this access with good public transport while preserving public open spaces and areas of high environmental risk from urban development. Cities expand in an orderly manner when they plan, prepare, and secure adequate lands for arterial roads and for streets—as well as their public open spaces—to organize their urban peripheries *before development occurs* in ways that make them more productive, more inclusive, and more sustainable. Whether they do so or not, and whether they are doing it better or worse than before, raises a set of questions that, until now, could not be properly answered: How well laid out are recently built urban peripheries, how are layouts changing over time, and why?

Volume 2: Blocks and Roads begins to provide answers to these questions by the mapping and measurement of urban layouts in the global sample of 200 cities using freely available, high-resolution, *Bing* satellite imagery. More specifically, it addresses three questions:

1. How well laid out are the expansion areas (areas converted to urban use between ~1990 and ~2014) in the global sample of 200 cities?

2. How well laid out are areas converted to urban use before ~1990—the pre-1990 areas—compared to expansion areas in the global sample of 200 cities? and

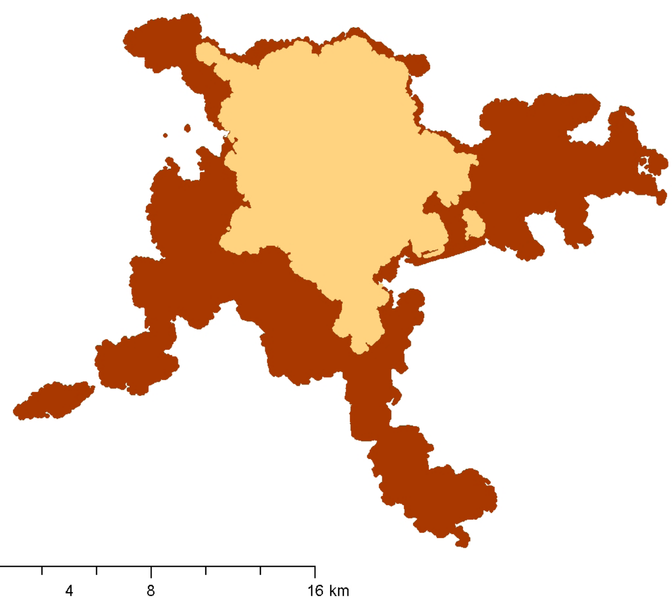
3. How well laid out are the areas converted to urban use in five different time periods —Period

1: before ~1900; Period 2: ~1900 – ~1930; Period 3: ~1930 – ~1960; Period 4: ~1960 - ~1990; and Period 5: ~1990 - ~2014— in a representative sub-sample of 30 cities?

THE SELECTION OF AREAS FOR ANALYSIS

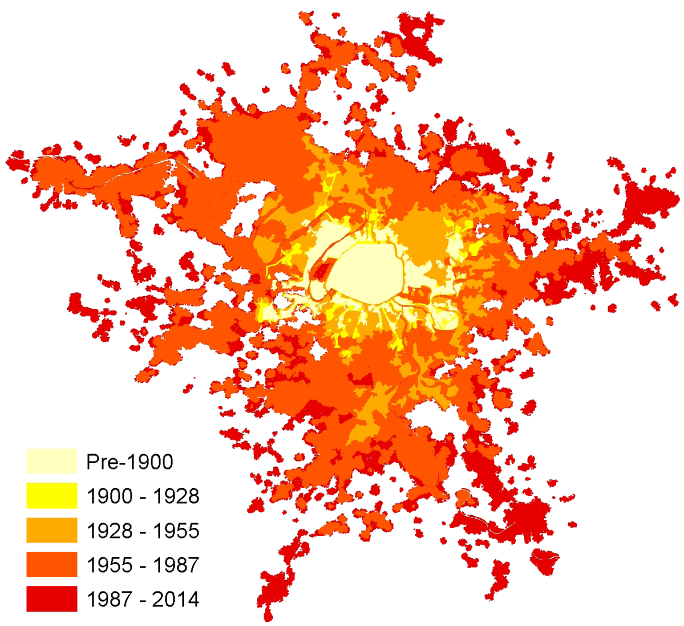
The answers to the first two questions require maps identifying the pre-1990 areas and the expansion areas of all the 200 cities in the global sample. These maps can be drawn from the urban extent maps for each city in the global sample compiled in Volume 1 of this Atlas. The map of the pre-1990 area of Addis Ababa, Ethiopia, for example, is simply the map of its urban extent in ~1990. For purposes of analysis, we combined the areas converted to urban use between ~1990 and ~2000 and between ~2000 and ~2014 into one area, referring to it as the expansion area. The map of the expansion area of the city is then simply the map of its urban extent in ~2014, with its pre-1990 area excluded (see figure 3.1). That said, in the maps of the 200 cities presented in the main section of this volume of the Atlas, the two areas are shown as two distinct expansion areas.

FIGURE 3.1:
The Pre-1990 Area (ochre) and the Expansion Area (red) of Addis Ababa, Ethiopia.



Answers to the third question posed in the previous section require maps that show the history of the expansion of the 30 cities in the global representative sample of cities over the past two centuries. These maps were created and presented in the first edition of the *Atlas of Urban Expansion* (S. Angel et al. 2011, 260-319), and they are not reproduced in full in this edition of the Atlas. Instead, summary maps for all 30 cities, showing the areas converted to urban use in each of the five periods listed in the previous section are given in the pages pertaining to these cities. An example of a summary map for Paris, France, showing the areas developed in five consecutive periods is given in figure 3.2.

FIGURE 3.2:
Area converted to urban use in the five different time periods in Paris, France.

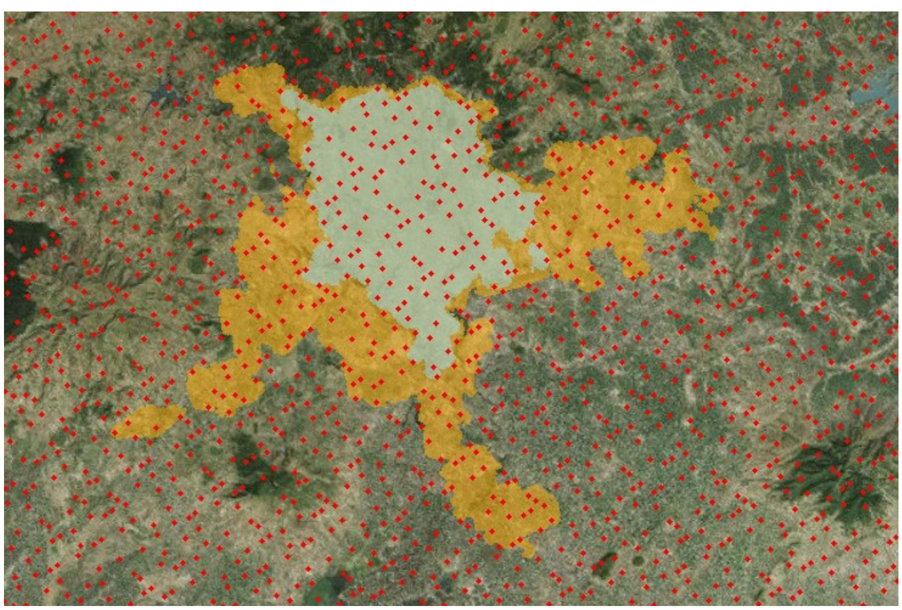


We can answer a number of important questions regarding the quality of urban layouts by patiently digitizing and analyzing high-resolution satellite imagery. *Bing* imagery, for example, is now freely available worldwide (similar *Google Earth* imagery is still proprietary) and can be analyzed using identical methods in each city in the sample, thus ensuring that results are consistent and comparable. That said, such studies are very labor intensive. In small cities, digitizing key features of urban layouts can be carried out in almost the entire urban extent of the city—including both its pre-1990 area and its expansion area. In larger cities, some layout features, like the presence of arterial roads, can be digitized and analyzed for the city as a whole, but more detailed features—like residential types, the share of the land in roads, block sizes, or plot sizes—cannot be. They can be more thoroughly investigated by sampling a limited number of 10-hectare locales at random in each of the areas of interest in the city, calculating the relevant metrics from these sampled locales, and generalizing the results for these areas of interest as a whole. In the largest cities, even the presence of arterial roads may need to be determined by sampling as well, in our case by sampling randomly selected one-kilometer squares throughout their urban extents. In broad terms, this is the procedure followed in generating the maps and metrics for this volume of the Atlas.

Most of the analysis focused on digitizing and analyzing randomly selected 10-hectare locales in the 200 cities in the global sample. All in all, a total of 20,795 locales were digitized, approximately 100 locales per city, on average. In addition, a total of 5,638 additional locales were analyzed in the sub-sample of the 30 cities used to study changes in urban layouts over a longer period: ~1900 - ~2014.

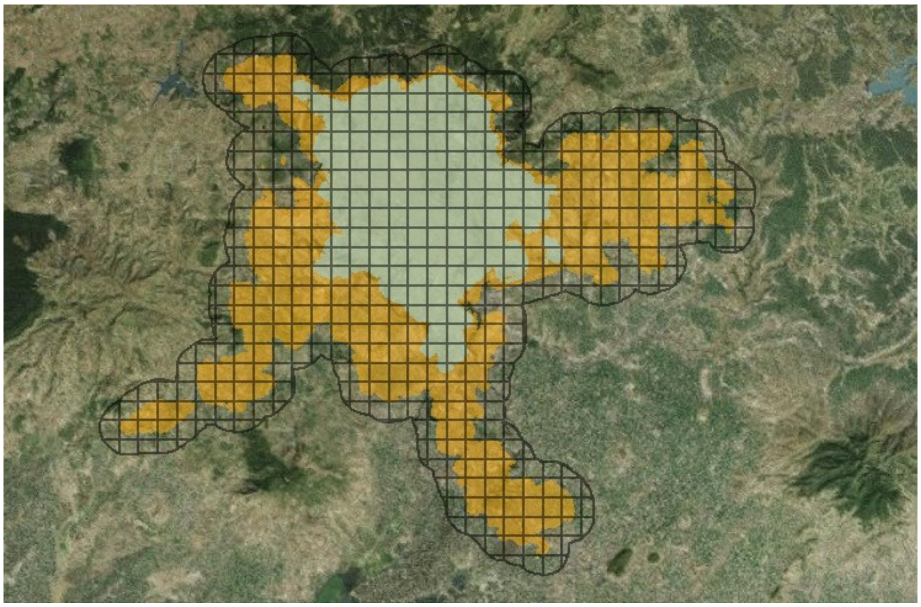
The locations of these locales in a given city were determined by combining a quasi-random series of numbers known as a Halton Sequence with the XY (latitude and longitude) origins of a bounding box that encompassed the city as a whole. This procedure generates a set of points in two-dimensional space that appear to be randomly distributed but cover the space more evenly than a set of points generated at random. A particular Halton Sequence, using the same initial XY origin to generate point coordinates, always generates the same set of points in the same order. We used one tenth of a degree of longitude and latitude as XY origins to generate points for every city in the global sample. The set of points generated for the study area of Addis Ababa, Ethiopia, is shown in figure 3.3. Subsequently, in every area of interest, say the expansion area of Addis Ababa, we initially selected 40 points for analysis in the order they were identified by the Halton sequence.

FIGURE 3.3:
Quasi-random placement of potential 10-hectare locales for the analysis of urban layouts in the study area of Addis Ababa, Ethiopia, using a Halton sequence.



As noted earlier, in each city in the sample we digitized the arterial road network in its entire urban extent in order to determine the share of the relevant area within that extent that was within walking distance of an arterial road, as well as to estimate the share of the built-up area in arterial roads. To do so in an orderly fashion, we placed a one-kilometer grid over the entire urban extent of the city and then identified and digitized arterial roads in each of the grid squares. Since arterial roads within walking distance of a built-up area may include roads outside the urban extent, we also included areas within one-kilometer of the edge of the urban extent in our analysis. The one-kilometer grid for the 2014 urban extent of Addis Ababa, Ethiopia, is shown in figure 3.4.

FIGURE 3.4:
The one-kilometer grid used to identify arterial roads in the urban extent of Addis Ababa, Ethiopia, including a one-kilometer-wide buffer around it.



In the largest cities in the global sample, identifying, digitizing, and analyzing arterial roads in their entire urban extent proved unnecessary. Instead, we selected a set of one-kilometer squares at random, using our set of Halton points described earlier. We then identified the arterial roads within each square, as well as in an area within one-kilometer of its edge, so as to be able to determine the share of the area within the square that was within walking distance of an arterial road, a road that could well be outside that square. The resulting set of randomly placed 3-by-3-kilometer areas (their rounded edges are the result of being one kilometer away from the corners of the one-kilometer squares) used to identify, digitize, and analyze arterial roads on the periphery of Tokyo, Japan, is illustrated in figure 3.5.

To summarize: Measuring the attributes of urban layouts requires a focus on high-resolution satellite imagery which, in turn, requires a more careful selection of representative areas for analysis within the urban extents of cities in the global sample. In order to study the changes over time in the attributes of urban layouts, we divided the urban extents of all cities in ~2014 into two: pre-1990 and post-1990 areas. To study changes in these attributes over a longer time period, we also differentiated the urban extents of 30 cities into five periods, spanning the twentieth century and the first fifteen years of the present century. In the absence of sampling, the study of urban layouts in the global sample of cities would be a daunting task. We rendered it doable by sampling 10-hectare locales within the urban extents of cities.

The actual locales randomly selected for digitizing and analysis of the change in urban layouts over time in Paris, France, are shown in figure 3.6. In a number of the largest cities in the global sample, we also sampled a number of one-kilometer squares throughout their urban extents to map enough of their arterial road networks to calculate the various metrics associated with them.

FIGURE 3.5:
Randomly selected 3-by-3-kilometer areas used to identify, digitize, and analyze arterial roads (in yellow) on the urban periphery of Tokyo, Japan.

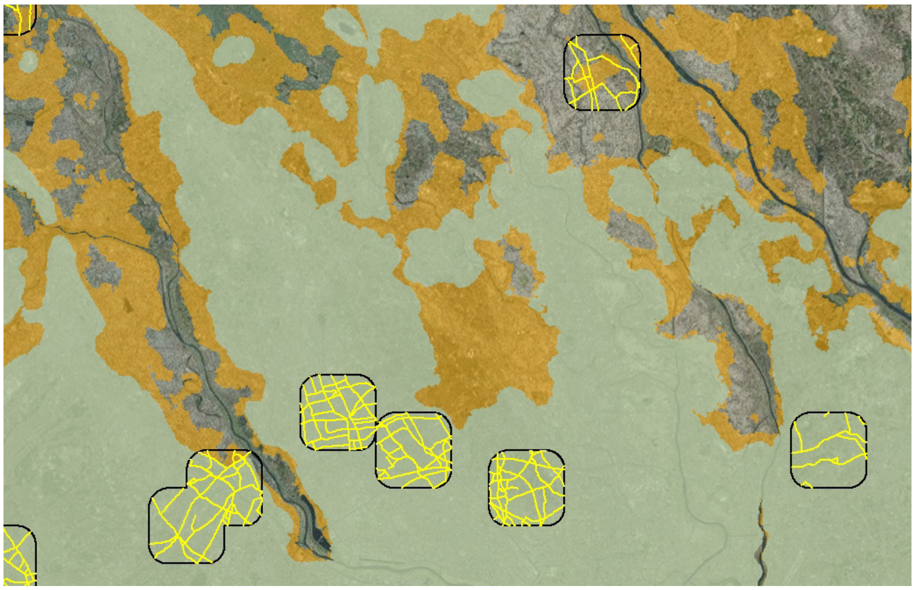
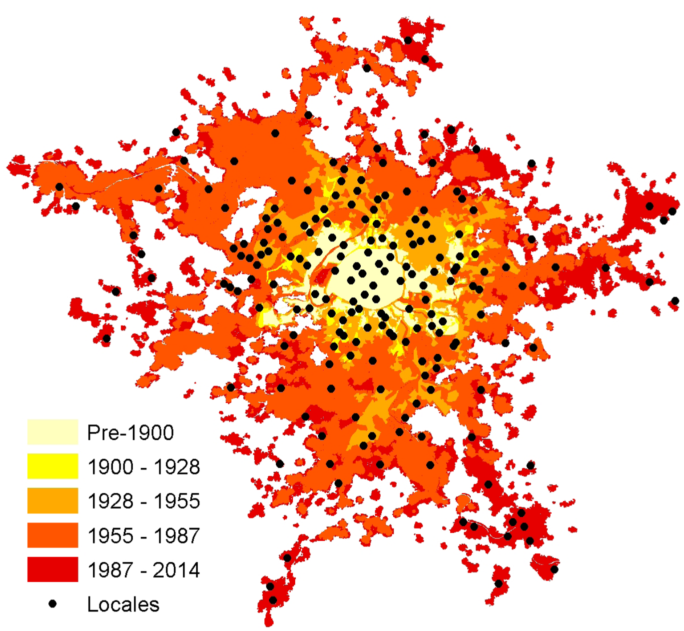


FIGURE 3.6:
Actual locales selected in a quasi-random process from the total number of available locales in the study area of Paris, France, to record the changes in the attributes of urban layouts over time.



MEASURING KEY ATTRIBUTES OF URBAN LAYOUTS

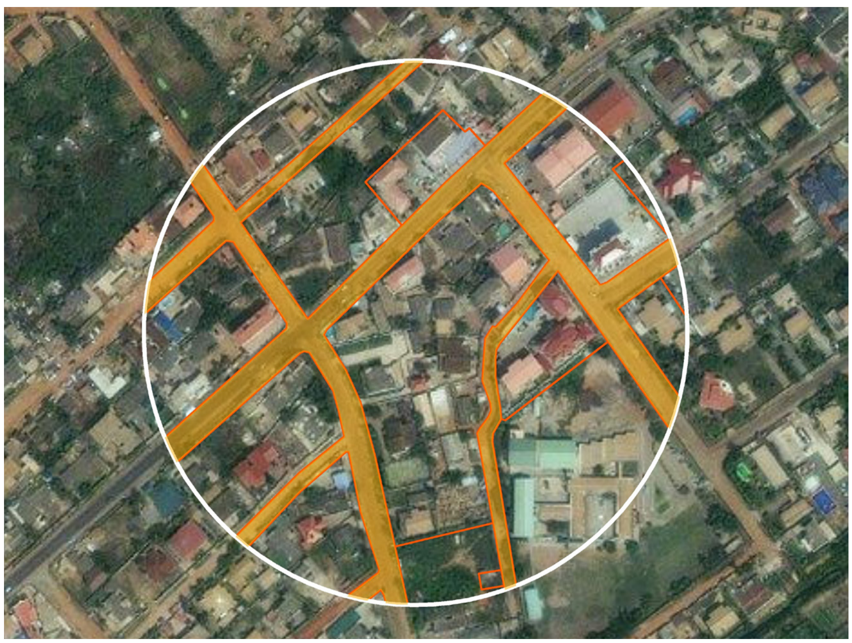
Within each 10-hectare locale, manual digitization techniques were used to identify, map, and measure the physical characteristics of its urban fabric. The primary focus was on the quality and orderliness of their block and road layouts, the quality of their visible infrastructure, the size of blocks and residential plots, and the density of street intersections. Orderliness or disorderliness that can be assessed from satellite imagery largely comes down to the way in which public space is used to organize the urban fabric (through road and block layouts), the level of infrastructure that is provided in a given area (indicative of formality or informality), and the form of dwellings, both through identification of plot boundaries and through a visual assessment of building types. Much of this work falls on the image analyst. With that in mind, detailed rules were developed to assist the analyst in classifying the imagery. We summarize these rules here so that the method we used to arrive at the maps and metrics in the following Atlas pages can be better understood and easily replicated.

Blocks and Roads

Classification of satellite imagery is fundamentally an exercise in pattern recognition. As with all pattern recognition, the first task in identifying the elements of a locale involves making a primary distinction between these elements. In our case, that distinction is between block space and road space. Road space consists of all land that is currently or potentially used by either pedestrians or vehicles to travel from one place to another. We seek to identify the *right-of-way* of streets and roads, containing both the area that is currently in use and any lands that are clearly reserved for future use. All of these areas constitute road space. Block space, in turn, consists of all other uses, including open space and off-street parking areas. Road space and block space add up to the entire area of a 10-hectare locale. In other words, all space that is not road space is block space, and all block space is assigned a land use. The division of a high-resolution satellite image of a locale in Accra, Ghana, into street space and block space is illustrated in figure 3.7.

Block space is subdivided into units identified as *blocks*. Individual blocks are areas that are continuously bounded by roads or vacant open spaces (for instance, a block at the edge of a built-up area that borders on farmland). Any given block might contain several different land uses (say, apartment buildings on one end, single-family homes in the middle, and a school at the far end). Blocks and block space can be further subdivided into *plots*, individual parcels of land that would likely be identified as separate properties in a cadaster. Any given block is composed of either one large plot or a series of smaller plots. Much like the identification of rights-of-way, plot boundaries are identified through surface indicators, pattern recognition, and comparisons with nearby areas. The concept of the plot is very important in differentiating residential categories. A suburban plot in a formal residential area might contain several structures—a house, a garage, and a toolshed, for example. We were not interested in measuring the dimensions of these structures in this Atlas. Instead, our goal was to measure the use of the underlying land so as to get a sense of the shares of land in different uses. When land uses are determined, it is the land use of the *plot* as a whole that is determined and measured, not the land occupied by a specific building. The same principle holds true when assessing patterns to determine land use in a larger area: The key is to focus on the pattern of plot boundaries and not on building footprints.

FIGURE 3.7:
The division of a high-resolution satellite image of a 10-hectare locale in Accra, Ghana, into road space (light brown) and block space (orange borders)



Land Use Categories

Each city in the global sample has specific residential and non-residential typologies, along with unique characteristics of form and layout that deserve recognition and study in their own right. However, in order to study land use on a global scale, the land use categories must be simple enough and broad enough to be identified in any city in the world, encompassing (to the maximum extent possible) the whole range of land use types found in cities. Following a review of numerous land use classifications, we narrowed our classification to seven land uses that could be reliably identified in high-resolution satellite imagery, with a focus on four types of residential land use: (1) open space; (2) non-residential areas; (3) atomistic settlements; (4) informal land subdivisions; (5) formal land subdivisions; (6) housing projects; and (7) road space.

1. *Open Space* includes open countryside, forests, cultivated lands, parks, vacant lands that have not been subdivided, cleared land, and water bodies: seas, rivers, lakes, and canals.

2. *Non-Residential Areas* include all built-up areas, both public and private, that are not in residential use.

3. *Atomistic Settlements* are areas with irregular layouts that were clearly not subdivided or laid out before residential construction took place. This category includes squatter settlements that grew incrementally without an overall plan, homes built on irregular parcels of land, or homes built on rural plots that were not regularly subdivided before their conversion to urban use.

4. *Informal Land Subdivisions* are areas that have been subdivided for urban use, but that lack visible evidence of conformity to land subdivision regulations such as regular plot dimensions, paved roads, streetlights, or sidewalks. That said, structures in these informal land subdivisions, although different in size and form, are typically laid out along straight or almost-straight roads, with regular intersections and standardized widths. Blocks are also regular or semi-regular in size and shape, when topography permits.

5. *Formal Land Subdivisions* are similar in layout to informal layouts, but exhibit a higher level of regularity, a higher level of provision of infrastructure, and better connections to existing roads. All roads must be paved for an area to be classified as a formal land subdivision. Sidewalks and streetlights are often visible as well.

6. *Housing Projects* range from large apartment tower projects to suburban tract housing. Housing projects share one feature: their structures must be essentially homogenous. These are projects in which all structures are built by a single developer using variations on the same plan.

7. *Road Space* includes the rights-of-way of lanes, streets and roads, both paved and unpaved, containing both the area that is currently in use and any lands that are clearly reserved for future use.

The four types of residential land use are illustrated with examples in figure 3.8. These types were chosen to reflect stages in the evolution of the housing sector, from a state of weaker planning skills and traditions, less regimented property-right and regulatory regimes, low availability of capital, and an absence of housing finance, to a state of stronger planning and regulatory regimes and a broader availability of capital. The housing sector is at its most basic in atomistic settlements, where the organization of the settlements is insufficient even to ensure consistent plot size or road width and where dwellings are located haphazardly and constructed over time. The housing sector is at its most complex when it is able to support large, planned housing projects, whether private or public, with access to capital, constructed from start to finish over a short period of time. The characterizations of these seven land use categories were used by analysts to determine the land uses within blocks in the 10-hectare locales, taking into account that a single block surrounded by roads or open spaces on all sides may contain more than one of six land uses.

FIGURE 3.8:
Four types of residential land use identified in locales, using high-resolution satellite imagery: Atomistic settlements (top left), informal land subdivisions (top right), formal land subdivisions (bottom right), and housing projects (bottom left).

Plots, Blocks, and Intersections

The dimensions of residential plots in formal and informal land subdivisions are of interest because they may tell us, for example, whether large plot sizes in formal subdivisions are leading to high rates of land consumption per capita or whether small plot sizes in informal subdivisions reflect a discrepancy between minimum official plot sizes and those offered in the informal market. It is possible to measure plot sizes in land subdivisions using high-resolution satellite imagery when plots are relatively uniform. In these cases, it is possible to identify the boundaries between plots, to count the plots, and to determine their widths and depths. To measure plot dimensions in residential subdivisions, a block that had an array of plots of uniform size was identified and two lines were drawn along two of its edges. Each line was tagged with the number of plots along it, creating an estimate of typical plot depth and width in that area. This procedure is illustrated in figure 3.9. In this example, the length of the block (160 meters) is divided by 22 and its depth (40 meters) is divided by 2 to yield a typical plot size of 7.3-by-20 meters or 146m^2.

FIGURE 3.9:
Arriving at a typical plot size in an informal subdivision in Guadelajara, Mexico, by measuring overall block length and depth and dividing each dimension by the number of plots along it.

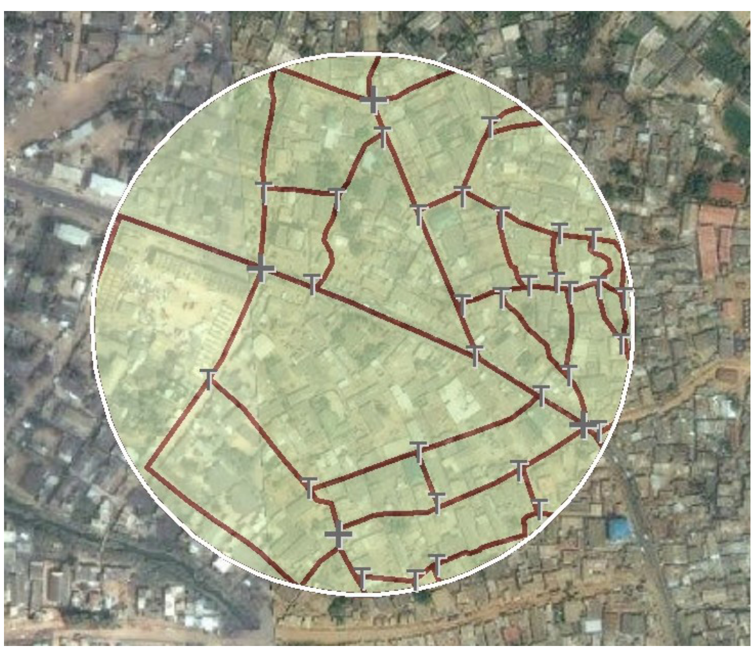
The size of city blocks or, alternatively, the density of 4-way intersections compared to 3-way ones in typical city neighborhoods is of interest because neighborhoods with small blocks and with high 4-way intersection densities facilitate walking and bicycling, reducing the reliance on private automobiles and making the urban environment healthier and more convivial. It is indeed possible to measure the size of blocks and the density of both 3-way and 4-way intersections using high-resolution satellite imagery, and we did indeed measure them in all locales.

To measure block sizes and intersection density, the analysis of locales required the digitization of road *medians* (the lines along the middle of roads). This was done for all blocks in every locale, and included the digitization of medians along the entire perimeter of all blocks within the locale, including those clipped by the circular boundary of the locale. It is important to note that using this procedure implied that the area of blocks was calculated as the entire area enclosed by road medians, including the area of roads. The procedure for identifying and mapping blocks is illustrated in figure 3.10. The density of intersections was calculated by counting the intersections within the locale and dividing their total by the built-up area of the local, excluding areas identified as open space. The procedure for identifying and counting road intersections is illustrated in figure 3.11. In this example, there are 4

4-way intersections, 33 3-way intersections, and a total area of 9.3 hectares (or 0.093 km^2) in built-up areas. The 3-way intersection density in this locale is therefore 354 per km^2 and the 4-way intersection density is 43 per km^2.

FIGURE 3.10:
Identifying all the blocks in a typical locale by digitizing the road medians around them, including blocks that are clipped by the circular boundary of the locale.

FIGURE 3.11:
Identifying all the 3-way and 4-way road intersections in a typical locale by digitizing the road medians within the locale (4-way intersections are marked with a + and 3-way intersections are marked with a T).



Arterial Roads

Arterial roads in cities are of interest because they are essential for integrating urban labor markets—providing access, by all transport modes, from all residences to all workplaces in the city—and the more integrated their labor markets, the more productive they are. The road network in every country typically forms a three-tier hierarchy of primary, secondary, and tertiary roads. Central or state governments usually plan, acquire land, finance, construct, and maintain the primary intercity road network that connects the country together. Municipalities typically plan, acquire land, finance, construct, and maintain the secondary or arterial road network within their jurisdictions. In many cases, private developers of residential neighborhoods or of commercial, office, and industrial projects typically plan, acquire land, finance, and construct the tertiary roads that serve buildings within their projects. In many other cases, municipalities plan and build the tertiary road network as well. The network of arterial roads is a classic public good (i.e., users cannot be effectively excluded from using it). Since it is a public good, there is no market mechanism that can ensure that arterial roads are in adequate supply in appropriate locations. In other words, a shortage of arterial roads may be a form of market failure. This means that it is up to public authorities to supply arterial roads in adequate quantities, in the right locations throughout the city, preferably before development takes place. Whether or not this happens in practice can only be determined by observation and measurement.

We identified and digitized arterial roads throughout the urban extents of all cities in the sample. As noted earlier, in the largest cities in the sample we opted to sample locations selected at random and to identify and digitize arterial roads only in these sampled locations. The information obtained from digitizing arterial roads was then used to calculate the share of the built-up area within walking distance to arterial roads, the average beeline distance to an arterial road, and the density of arterial roads. All of these measures provide some insight, for the first time, on the availability of arterial roads in cities the world over, as well as on its change over time.

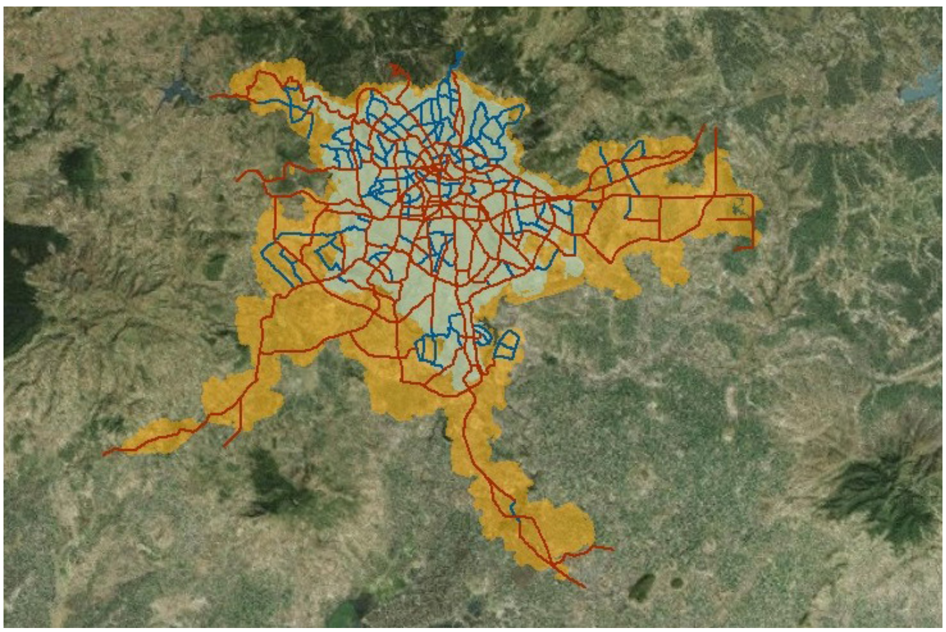
All roads that fall within the urban footprint (or its surrounding one-kilometer buffer) were considered as possible arterial roads. Likely candidate roads were identified in three data sources: Java Open Street Map, Google maps, and Bing maps, where roads are available as map layers. On any of these three road map layers, roads having through-connectivity are distinguished by width and color. Analysts examined each one-kilometer grid square in the urban extent to identify arterial roads. A candidate road was identified as an arterial road when it met two criteria:

1. It connected to other arterial roads, forming part of a network that extends throughout the city; and

2. It connected to the nearby minor roads. Limited access roads (freeways or expressways) were not considered arterial roads, even though they were connected to other arterial roads.

When an analyst identified a road as *arterial*, they differentiated it further into two categories: Wide and Narrow, where wide roads were those having rights-of-way of 18-meters or more. The network of wide and narrow arterial roads in the urban extent of Addis Ababa, Ethiopia in 2014 is shown in figure 3.12. The same procedure was followed in identifying wide and narrow arterial roads in randomly selected 3-by-3-kilometer squares in the largest cities in the sample, as previously shown (figure 3.4).

FIGURE 3.12:
The network of arterial roads in the urban extent of Addis Ababa, Ethiopia in 2014, distinguishing wide arterial roads (brown) from narrow ones (blue).



Urban Layout Metrics

In each city in the global sample of 200 cities, we initially selected at random 40 locales for analysis in its pre-1990 area and 40 locales in its expansion area, a total of 80 locales per city or 16,000 locales for the global sample as a whole. Key layout features of these locales, observed in high-resolution satellite imagery, were then digitized, analyzed, and stored. The digital files associated with locales were processed in ArcGIS using a Python script that calculated the following metrics for each locale:

- **Land Use**
    - Share of land in open space (open space in locale/area of locale);
    - Share of built-up area in non-residential use (non-residential land in locale excluding roads/area of locale);

- Share of the built-up area in residential use (all area in residential use in locale/built-up area of locale);
- Share of built-up area occupied by roads (area in roads/built-up area)
- Share of the residential area not laid out before development (area of atomistic settlements/residential area);
- Share of the residential area in informal land subdivisions (area in informal land subdivisions/residential area);
- Share of the residential area in formal land subdivisions (area in formal land subdivisions/residential area);
- Share of the residential area in housing projects (area in housing projects/residential area);
- Share of the residential areas laid out before development (area in both formal and informal land subdivisions/residential area);
- Share of locale that is gridded [visual assessment of the presence of orthogonal street grids in the locale and their assignment to three categories: not gridded, partially gridded (covering 10-90% of the locale area), and gridded (covering 90% or more of the locale area)].
- Average plot size in informal land subdivisions; and
- Average plot size in formal land subdivisions.

• **Roads**

- Share of roads less than 4-meters-wide (length of roads less than 4-meters-wide in locale/length of total road network in locale);
- Share of roads that are 4-to-8-meters-wide (length of roads 4-8-meters-wide in locale/length of total road network in locale);
- Share of roads that are 8-to-12-meters-wide (length of roads 8-12-meters-wide in locale/length of total road network in locale);
- Share of roads that are more than 16-meters-wide (length of roads more than 16-meters-wide in locale/length of total road network in locale); and
- Average road width in locale.

- **Block Layout**
  - Average block size (hectares);
  - The density of 3-way intersections (number per square kilometer of locale area);
  - The density of 4-way intersections (number per square kilometer of locale area);
  - Share of intersections that are 4-way (ratio of 4-way intersections to total number of intersections in locale);
  - The Walkability Ratio (The average ratio of the beeline distance and the street travel distance for 40 pairs of sample points within the locale that are more than 200-meters apart);

In addition to calculating metrics for individual locales, a number of metrics were calculated for the arterial road network identified in each city:

- **Arterial Roads**
  - The average density of all arterial roads (linear kilometers of arterial roads/square kilometers of urban extent);
  - The average density of wide (18m+) arterial roads (linear kilometers of wide arterial roads/square kilometers of urban extent);
  - Average beeline distance to all arterial roads (meters);
  - Average beeline distance to wide arterial roads (meters);
  - Share of the urban extent within walking distance (625m) of all arterial roads; and
  - Share of urban extent within walking distance (625m) of wide arterial roads.

Data for each locale is stored in four files: (1) Locale boundary file; (2) Blocks file; (3) Plot measurement file; and (4) Street medians file. Arterial roads data is stored in two additional files: (5) Arterials master file; and (6) Arterials study area file. All of the data is stored in shapefile format and can be downloaded on a city-by-city basis or in batches at www.atlasofurbanexpansion.org.

The Atlas pages that follow provide average values for the locales in each area of interest in each of the 200 cities in the global sample for many, but not all, of these metrics. Some metrics were chosen over others as more illustrative of the quality of urban layouts in cities at the present time. Tables summarizing these metrics in Excel format are given following the city-focused pages.

Improving the confidence in the metric averages

The metrics that we calculated exhibited a high degree of variation across locales within a city. This intra-city variability poses a challenge for making correct inferences. More specifically, in order to detect statistically significant differences in the mean value of a metric across cities, precise estimates of the mean value of a metric within a city are needed. Although the sample average for a given metric—say, the average share of the built-up area in roads—might differ in two cities, the number of locales in each city might not be large enough to reject the null hypothesis that the two means are equal to each other. We can improve the precision of our estimates by adding locale observations to each city, but additional locales entail additional costs, in terms of both time and money.

Given the time and cost associated with extracting data from each locale, the study leading to the production of this volume of the Atlas operated with a budget allowing for the analysis of approximately 20,000 locales in the 200 cities in the global sample. All in all, some 30 analysts worked for an average of 90 days each to digitize and analyze these locales. We initially allocated 80 locales to each city in the sample, 40 in the pre-1990 area of the city and 40 in its expansion area. Then, rather than equally dividing the remaining 4,000 locales evenly among all cities, these locales were allocated using a rule to improve the overall precision of our subsequent estimates of city averages. This rule was based on the understanding that some cities are more complex than others and feature more variability in key metrics of interest. Adding locales to these cities may therefore be especially useful in improving the precision of our estimates.

We chose to focus on three principle metrics, or 'variables of interest', that are of key importance in assessing the quality of urban layouts: (1) the share of the built-up area in roads; (2) the share of residential land in atomistic settlements; and (3) the share of residential land in informal land subdivisions. Each sampled locale provides values for each one of these three metrics. For each city, given a set of sampled locales, we can calculate the sample average and sample standard deviation of each variable of interest. The method chosen to add locales to particular cities uses the information on the averages and standard deviations for these three metrics to improve the statistical power to detect differences between hypothesized means in the cities in the global sample (For a general discussion of statistical power see Casella and Berger, 2002, pp. 382-383). The procedure we followed involved the following steps:

- Initially, allocate 80 locales to each city;
- Calculate the statistical power associated with one-sided hypothesis tests for each of the variables of interest in all the cities in the sample;
- Create a power index for each city, which is the average statistical power associated with the tests for the three variables of interest;

- Sort cities on the basis of the power index from lowest to highest;
- Select the 20 cities with the lowest rankings on the power index;
- Add 10 new locales to each of these 20 cities, then calculate new metrics and new power indices;
- Rank cities again, using this new information;
- Repeat the process until all 4,000 new locales have been allocated.

It should be noted that in some cities, the expansion area is sufficiently small that it might be completely covered with locales, either before the initial 80 locales are randomly chosen or before the termination of the procedure for adding locales. As soon as it becomes impossible to add another locale that does not overlap with the existing locales, no more locales are added to a given city. As noted earlier, all in all, 20,795 locales were digitized and analyzed, a maximum of 270 locales in Cairo, Egypt and a minimum of 25 locales in Zhijin, China. Unfortunately, the addition of locales at this scale does not yet ensure that the average values reported in the Atlas pages that follow are significantly different from each other.

There are two pages in Volume 2 of the Atlas for each city in the global sample of 200 cities, arranged in alphabetical order in the following pages. They are followed by Atlas pages with maps and metrics for the 30 cities for which we have data on urban layouts that were created from 1800 onwards. These maps and metrics pages are followed by summary tables in Excel format that provide metric values for all attributes shown in the individual city tables.