Delineating Functional Urban Areas Using a Multi-Step Analysis of Artificial Light-at-Night Data

Nataliya Rybnikova 1,2,3,*, Boris A. Portnov 2, Igal Charney 3 and Sviatoslav Rybnikov 4,5

Abstract: A functional urban area (FUA) is a geographic entity that consists of a densely inhabited city and a less densely populated commuting zone, both highly integrated through labor markets. The delineation of FUAs is important for comparative urban studies and it is commonly performed using census data and data on commuting flows. However, at the national scale, censuses and commuting surveys are performed at low frequency, and, on the global scale, consistent and comparable data are difficult to obtain overall. In this paper, we suggest and test a novel approach based on artificial light at night (ALAN) satellite data to delineate FUAs. As ALAN is emitted by illumination of thoroughfare roads, frequented by commuters, and by buildings surrounding roads, ALAN data can be used, as we hypothesize, for the identification of FUAs. However, as individual FUAs differ by their ALAN emissions, different ALAN thresholds are needed to delineate different FUAs, even those in the same country. To determine such differential thresholds, we use a multi-step approach. First, we analyze the ALAN flux distribution and determine the most frequent ALAN value observed in each FUA. Next, we adjust this value for the FUA’s compactness, and run regressions, in which the estimated ALAN threshold is the dependent variable. In these models, we use several readily available, or easy-to-calculate, characteristics of FUA cores, such as latitude, proximity to the nearest major city, population density, and population density gradient, as predictors. At the next step, we use the estimated models to define optimal ALAN thresholds for individual FUAs, and then compare the boundaries of FUAs, estimated by modelling, with commuting-based delineations. To measure the degree of correspondence between the commuting-based and model-predicted FUAs’ boundaries, we use the Jaccard index, which compares the size of the intersection with the size of the union of each pair of delineations. We apply the proposed approach to two European countries—France and Spain—which host 82 and 72 FUAs, respectively. As our analysis shows, ALAN thresholds, estimated by modelling, fit FUAs’ commuting boundaries with an accuracy of up to 75–100%, being, on the average, higher for large and densely-populated FUAs, than for small, low-density ones. We validate the estimated models by applying them to another European country—Austria—which demonstrates the prediction accuracy of 47–57%, depending on the model type used.

Keywords: functional urban areas (FUAs); boundaries; multiple regression modelling; artificial light-at-night (ALAN); optimal threshold

1. Introduction

More than 50% of the world’s population currently resides in urban areas, and this share is expected to increase to 70% by 2050 [1]. Due to a significant concentration of production factors, urban areas produce approximately 80% of the global GDP [2]. This makes spatial dynamics of urban areas to be important for policy-makers and researchers...
alike. Decision-makers can devise informed development policies, while in the research community, this information can be used to monitor the process of urban growth and the forces behind it [3–5], to assess the impact of urbanization on agriculture and natural landscapes [6], on biodiversity [7], on land surface temperature [8], and other socio-economic and physical phenomena.

Urban growth is characterized by two distinctive components—physical growth and functional change. The former group of attributes reflect changes in impervious surfaces and built-up characteristics, such as building density, building volumes [9,10], as well as population size and density of individual urban settlements [11–13]. Concurrently, functional attributes of urban growth reflect factor mobility, associated with various economic activities, such as commuting, commerce, industrial production and services [14]. Such exchanges are especially intense between urban cores, where a large share of production factors is concentrated, and their surrounding areas. Functionally-integrated clusters, representing geographic entities that consist of a densely inhabited city and a less densely populated commuting zone, both highly integrated through labor markets, are commonly referred to as functional urban areas or FUAs [15]. A FUA is conceptually different from an urban agglomeration, which is commonly defined as a major city surrounded by an adjacent hinterland [16]. The major difference between the two is commuting, which is crucial for delineating FUAs, but is not a prime consideration for the definition of urban agglomerations.

According to the mainstream approach adopted by the European Union (EU) and the Organization for Economic Co-operation and Development (OECD), the boundaries of FUAs are defined in three consecutive steps. First, urban cores are identified as continguities of high-density grid cells with population density of at least 1500 residents per km$^2$ and the total population in the contiguous cells of at least 50,000 residents. Second, local administrative units (LAUs) with at least 50% of their residents living inside the urban core are identified. At the final step, the commuting zone, comprising LAUs, which have at least 15% of their residents employed in the core city, is determined. Together with the central city, these administrative units are assumed to form a single FUA [17].

However, commuting data, needed to perform such delineations, are laborious to collect and are infrequent and sporadic even in developed countries [18]. In addition, different countries and regions report commuting data with different frequencies, and sometime collect them using different definitions and methodologies [18]. As a result, comparable cross-country estimates of FUA boundaries cannot always be obtained.

As artificial light-at-night (ALAN) data are freely available globally and provide a seamless global coverage, the idea of using them for the identification of human activities was investigated in several studies (see inter alia [19–21]). In previous studies, ALAN data were used in health geography [22–27], for the analysis of economic performance of countries and regions [28–31], and in population density research [20,32–36]. The use of such data in the studies of light pollution and its ecological effects is also common [24,37–40].

In recent years there have been attempts to use ALAN data for the identification of urban areas [20,21,41–47]. In one such study, Imhoff et al. [41] examined frequency-based ALAN thresholds for three large metropolitan areas in the U.S.—Miami, Chicago and Sacramento. After the authors analyzed the frequencies of differently lit pixels in the ALAN images, they determined that pixels present with 85%, 89% and 94% frequencies, occupy the areas of approximately same size, such as those reported in the Census for the corresponding metropolitan entities.

In another study, Sutton et al. [20] investigated 2000 cities across the globe, and compared their actual boundaries with those produced by three different frequency-based ALAN thresholds—40%, 80% and 90%. As the study revealed, pixels in the ALAN image, observed with a frequency of 80% or more, correspond to the actual municipal boundaries best, reaching a correlation level of about 68%.

In a separate study, Henderson et al. [21] examined frequency- and intensity-based ALAN thresholds that match the boundaries of San Francisco, Beijing and Lhasa. As the authors of this study have found, the optimal ALAN frequency-based thresholds that produce
the total lit area comparable in size to the Landsat data-derived urban delineations, reach 88% for Lhasa, 97% for Beijing, and 92% for San Francisco, with the corresponding ALAN flux being equal to 19, 30 and 51 digital numbers (DN), respectively. However, the spatial correspondence between metropolitan boundaries, determined using ALAN thresholds, and actual metropolitan delineations was found to be relatively low, not exceeding 8–44%.

It should also be noted that the aforementioned studies focus on the identification of built-up urban contiguities, while, to the best of our knowledge, only one study by Bosker et al. [18] analyzed functional urban delineations based on commuting flows. The authors of this analysis compared varying percentiles of ALAN intensities, reported by the VIIRS/DNB satellite’s sensor for 2015, with commuting delineations in Malaysia. As this study revealed, the best fit of ~40% is observed when 7% commuting frequency delineations are compared with delineations based on the 25th percentile of ALAN intensities.

A possible reason for such a low fit of less than 40% is that FUAs even in the same country differ by the amount of ALAN they emit. As a result, different ALAN thresholds must be used for the delineation of FUA boundaries in different parts of the urban system. In Figure 1, we illustrate this point using two FUAs in France, as an example. As evidenced by this figure, the ALAN threshold of 0.71 nW/cm²/sr fits reasonably well the boundaries of the Paris FUA, but the same threshold fits rather poorly the much smaller Chateauroux FUA, ALAN flux at which boundary does not exceed 0.15 nW/cm²/sr.

Considering that ALAN emissions from different FUAs vary substantially, it is thus important to establish varying ALAN thresholds, which would fit individual FUAs. This task can be performed for each FUA separately. However, in order to be practical, the approach needs to be sufficiently general, to enable its application to different FUAs, both for countries and regions with well-established commuting data and for other locations with unavailable or sparsely available commuting information. In this paper, we develop such an approach and test it against actual FUA delineations.

Figure 1. Cont.
Figure 1. Commuting-based boundaries (black lines) of the Paris (a) and Chateauroux (b) FUAs vs. the ALAN contours (blue lines), representing the 0.71 nW/cm$^2$/sr threshold level.

2. Materials and Methods

2.1. Study Phases

The proposed approach is implemented in several steps, as detailed in Figure 2. The data sources and analysis stages are described in the subsections below.

![Flowchart of study stages](image)

Figure 2. Flowchart of study stages.
2.2. Data Sources

Data for the present study were drawn from the following four main sources:

(1) The ALAN raster maps of France and Spain (see Figure 3), used in the study for model training and validation, and ALAN raster for Austria, used for additional validation of the models’ performance, were clipped from 2015 radiance-calibrated ALAN image downloaded from the VIIRS/SNPP website [48]. The ALAN data used in the study are free of background noise, solar and lunar contamination, and also free from data degraded by cloud cover, and features unrelated to electric lighting (e.g., fires, flares, volcanoes) [49]. In addition, the data underwent an outlier removal procedure, applied to abnormally high radiance pixels that occur infrequently over a year [49]. The image in question is the closest temporal match for other data sources used in the analysis, specifically for the FUA delineations, available for 2011 only (Figure 4). Although ALAN images are available today from the VIIRS-SNPP website on a monthly basis, and, since 2018, as daily composites [50], we opted to use an annual composite image, so as to minimize disturbances resulting from ALAN seasonal fluctuations and weather conditions, such as, e.g., cloud cover, which are often present in monthly and daily composites [50]. The subject image is of a ~500 × 500 m spatial resolution and reports the summarized intensity of nighttime light in nW/cm²/sr for different wavelengths in the 500–900 nm diapason [46]. In the image, ALAN levels vary from 0 to 4187 nW/cm²/sr for France, and from 0 to 550 nW/cm²/sr for Spain (see Figure 3 and Table 1).

(2) Boundaries of FUAs and their cores (see Figure 4) were obtained as shapefiles from the OECD website [51]. These shapefiles are generated using GeoStat grids, based on 2011 commuting data reported in national censuses [51].

(3) The latitudes of the FUA cores’ centroids and distances to the closest major city, used to explain the variance of the optimal ALAN thresholds, were calculated using the above FUA cores’ shapefiles by applying ArcGIS–10.x software tools.

(4) Population density of the FUA cores, and population densities of their 5–15–25 km buffers, also used as explanatory variables for the estimation of the optimal ALAN thresholds, were calculated using 1 × 1 km population grids obtained from the LandScan database for 2011 [52].

Figure 3. Cont.
Figure 3. ALAN maps for continental France (a) and Spain (b). Note: Areas located outside the national borders are marked in blue.

Figure 4. Cont.
Figure 4. FUAs and their cores in continental France (a) and Spain (b).

Table 1. Descriptive statistics of the research variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude of the FUA core’s centroid (dd)</td>
<td>42.757</td>
<td>51.001</td>
<td>47.137</td>
<td>2.366</td>
</tr>
<tr>
<td>Population density of the FUA core (persons per km$^2$)</td>
<td>89.529</td>
<td>2586.130</td>
<td>470.764</td>
<td>382.596</td>
</tr>
<tr>
<td>Population density decline gradient$^a$</td>
<td>1.110</td>
<td>15.409</td>
<td>3.986</td>
<td>2.224</td>
</tr>
<tr>
<td>Distance to the nearest major city (dd)$^a$</td>
<td>0.000</td>
<td>5.900</td>
<td>1.781</td>
<td>1.359</td>
</tr>
<tr>
<td>Average ALAN level (nW/cm$^2$/sr)</td>
<td>0.649</td>
<td>20.411</td>
<td>3.640</td>
<td>3.614</td>
</tr>
</tbody>
</table>

Spain (72 FUAs)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude of the FUA core centroid (dd)</td>
<td>36.110</td>
<td>43.561</td>
<td>40.093</td>
<td>2.381</td>
</tr>
<tr>
<td>Population density of the FUA core (persons per km$^2$)</td>
<td>20.554</td>
<td>3485.360</td>
<td>831.839</td>
<td>794.144</td>
</tr>
<tr>
<td>Population density decline gradient$^b$</td>
<td>0.926</td>
<td>16.232</td>
<td>4.863</td>
<td>3.724</td>
</tr>
<tr>
<td>Distance to the nearest major city (dd)</td>
<td>0.000</td>
<td>5.120</td>
<td>2.202</td>
<td>1.542</td>
</tr>
<tr>
<td>ALAN averaged level (nW/cm$^2$/sr)</td>
<td>0.745</td>
<td>23.129</td>
<td>6.442</td>
<td>4.583</td>
</tr>
</tbody>
</table>

Notes: $^a$ Calculated as straight line distance between a FUA core’s centroid and centroid of the closest FUA with 1.5M+ residents; $^b$ Calculated as the ratio between the population density of the FUA core and that of the core’s buffer with a 5 km width for small FUAs (less than 100,000 residents), a 15 km buffer for medium-size FUAs (100,000–250,000 residents), and a 25 km buffer for large FUAs (over 250,000 residents).

2.3. Initial Determination of the ALAN Thresholds

For the sake of simplicity, let’s assume that the nighttime light source of highest intensity is located at the center of a FUA, and light intensities drop monotonically and uniformly towards the FUA’s periphery (see Figure 5).
Such an assumption might be fully plausible for compact and monocentric urban areas (Figure 6). Under these conditions, the territorial footprint of the FUA’s ALAN emissions follows a perfect circle, and the most frequently observed (i.e., modal) ALAN values are found at the FUA’s outer boundary (Figure 5a). These modal values are also the dimmest ones, and, as such, they effectively define the FUA’s outer boundary (Figure 5b).

If the above assumptions are upheld, the analysis of the frequency distribution of the observed ALAN values can help to identify the ALAN level, which coincides best with the FUA’s boundary. In particular, the researcher needs to choose the modal ALAN value, for which ALAN intensity is expected to be close to zero (Figure 5b).

2.4. Correction for Compactness

The above assumption of monotonic and concentric distribution of ALAN emissions (Figure 5) is upheld only if the boundaries of FUAs that are circularly shaped. However, if a FUA’s shape is not circular, using the modal ALAN value as a delineation threshold would underestimate the actual area of the FUA. Figure A1 in Appendix A, which reports different FUAs’ footprints, helps to illustrate this point. As this figure shows, the more distant the shape of a FUA from a perfect circle, the brighter ALAN values emerge as the most frequent. For such non-circular FUAs, it is thus necessary to correct for compactness, so as to account for a FUA’s shape deviation from a perfect circle.
To perform such a correction, we first estimate the FUA’s compactness \((c)\), calculating it as the ratio between the area of a FUA and the area of its bounding circle [33,54]:

\[
c_{\text{FUA}} = \frac{S_{\text{FUA}}}{S_{BC}}
\]

where \(S_{\text{FUA}} = \text{area of a FUA} \); \(S_{BC} = \text{area of the bounding circle} \), calculated using the Minimum Bounding Geometry tool in the ArcGIS software.

Next, to represent FUAs, which deviate from circular shapes, we model them as ellipses of the same compactness:

\[
c_{\text{El}} = \frac{S_{\text{El}}}{S_{BC}} = \frac{\pi ab}{\pi a^2} = \frac{b}{a} = c_{\text{FUA}}
\]

where \(S_{\text{El}} = \text{area of an ellipse with semi-axes } a \) and \(b \) \((a > b)\).

At the next step, to correct the initially estimated ALAN threshold (see Section 2.3) for a FUA’s compactness, we calculate the radius of the circle, \(r\), which has the maximal intersection with ellipse, \(C_{\text{El}}\). As shown in Box A1 in Appendix A, this radius is equal to:

\[
r = \sqrt{ab}
\]

Lastly, we estimate the percentile of the ALAN value distribution, \(p^*\), corrected for compactness (see Box A1 in the Appendix A for the justification):

\[
p^* = \frac{2}{\pi} \arcsin \left( \frac{1 - c}{1 + c} \right)
\]

According to (4), for compact shapes, which are close to a circle, i.e., for which \(c \to 1\), the optimal ALAN threshold percentile \((p^*)\) tends to the dimmest ALAN value \((p^* \to 0)\), while for prolonged shapes with \(c \to 0\), \(p^* \to 100\), that is, the optimal ALAN threshold will tend to the highest ALAN percentile (see Figure 7).

![Figure 7](image-url)  

**Figure 7.** Relationship between a FUA’s compactness \((c)\) and the optimal ALAN percentile \((p^*)\). Note: Shapes deviating from a perfect circle are assumed to be elliptical; see text for explanations.

### 2.5. Regression Modelling

After the optimal ALAN threshold is identified for each FUA by determining the modal ALAN value (see Section 2.3), and corrected for compactness (Section 2.4), we link the estimated threshold values to several explanatory variables, characterizing the FUA cores, so as to determine these variables’ load on the optimal ALAN threshold value. To model these relationships, the following generic regression equation is used:

\[
ALAN_i = b_0 + b_1 \times Lat_i + b_2 \times D_i + b_3 \times PD_i + b_4 \times PDD_i + \varepsilon_i
\]
where $ALAN_i$ is the optimal ALAN threshold for FUA $i$ (nW/cm$^2$/sr); $Lat_i$ is latitude of the FUA core’s centroid (decimal degrees, dd); $D_i$ is distance to the nearest major city, calculated between a given FUA core’s centroid and the centroid of the nearest FUA with more than 1.5M residents (dd); $PD_i$ is population density of the FUA core (persons per km$^2$); $PDD$ is population density decline gradient, calculated as the ratio between the FUA core’s population density and population density in the FUA core’s buffer with a 5 km width for small FUAs (under 100,000 residents), a 15 km width for mid-sized FUAs (100,000–250,000 residents), and a 25 km width for large FUAs (over 250,000 residents); $b_0..b_4$—regression coefficients, and $\varepsilon$ is a random error term.

The predictors used in the model are expected to contribute to the ALAN threshold’s variance due to varying reasons. In particular, population density is known to be closely associated with ALAN flux (see inter alia [33,35,55]). Concurrently, population density gradient might capture changes in the pattern of population density around the FUA core. Concurrently, distance to the nearest major city is likely to show how local development patterns are modulated by proximity to major urban concentrations [55]. In addition, as population concentrations in high latitudes often require more artificial illumination, especially during long winters [40], FUA’s latitude is also included into the model as a potential predictor.

In the analysis, we tested different functional forms of the models, and determined that the logarithmic transformation of the $PD$ and $PDD$ variables provides the best results, by improving the regression fit substantially ($p < 0.05$). The initial analysis was performed in the IBM SPSSv.25 software using its multiple regression module. To ensure the normality distribution of the dependent variable, $ALAN_i$, we applied Box-Cox transformation procedure, to redefine the ALAN thresholds [56].

In addition to ordinary least square regressions (OLS), we also tested “random forest” regressions. Such regressions imply building an ensemble of “decision trees”, each of which “voting” for a certain level of the dependent variable, with subsequent averaging of the estimates across all the decision trees [57]. In the present analysis, we implemented a standard realization of the “random forest” regression (the TreeBagger module) in the MATLAB v.R2020x software [58]. During the estimation procedure, two parameters were a matter of choice—the number of independent variables used for the individual decision tree construction and the number of decision trees that comprise the forest. To ensure the comparability of the results, we used all independent variables, covered by the analysis, for the decision trees’ construction, and defined number of trees to be equal to 100, which is usually considered to be a reasonable number for reaching a generalization error convergence (see for example [57,59]). Each decision tree was built for 80% of randomly selected observations.

### 2.6. Adjustment for Contiguity

When the analysis is performed, any given ALAN threshold level might identify several clusters of identically lit pixels, some of which might be related to a given FUA, while other pixels might be located elsewhere. Therefore, to identify the ALAN pixels relevant to a given FUA, the following analytical procedure was implemented. First, for each FUA, we identified pixels that overlap the FUA’s core area, considering the core boundary information as an initial input (see Section 2.2: Data Sources). Next, for each pixel selected thereby, we analyzed all the pixels in its surroundings. If the ALAN value of a neighboring pixel was lower or equal to that of the pixel under analysis but greater than the ALAN threshold identified for the FUA (see Sections 2.3 and 2.4), the pixel in question was considered to be a part of the FUA analyzed. We have continued this procedure as long as all the pixels, which satisfy the above criteria, maintained a spatial contiguity. Then, for each FUA, we selected local administrative areas (LAUs), most of which area (that is, >50%) is occupied by the pixels identified thereby. These LAUs were considered to be a part of a given FUA (the MATLAB code for contiguity adjustment can be obtained from the authors upon request).
2.7. Initial Validation

To assess the performance of the estimated models (see Section 2.5), we analyzed the degree of correspondence between the empirically determined (see Section 2.4) and model-predicted ALAN thresholds adjusted for contiguity (see Section 2.6). To this end, the model estimated for France was used to predict the ALAN thresholds for individual FUAs in Spain and vice versa. In order to assess the extent to which the empirically determined and model-predicted ALAN thresholds coincide, we used different metrics, including Pearson correlation coefficients, standard error of the estimates (SEE), and weighted mean squared errors (WMSE).

Next, we compared the FUAs’ delineations, either empirically determined using the modal ALAN values (see Section 2.4) and adjusted for contiguity (see Section 2.6), or model-predicted (see Section 2.5), with commuting-based FUAs’ delineations (see Section 2.2). To perform such a comparison, we used the Jaccard Index (JI), which estimates the share of intersection within the union of the two sets relative to these shapes’ union [60]:

\[
JI(FUA_C, FUA_T) = \frac{|FUA_C \cap FUA_T|}{|FUA_C \cup FUA_T|}
\]

where \( FUA_C \) = the set of local autonomous units (LAUs) forming a FUA defined by commuting, and \( FUA_T \) = set of LAUs within either an empirically determined or model-predicted FUA boundary. The value of the index in question ranges from zero, when no intersection between the two sets is present, to one, when the two sets completely coincide and their intersection is equal to their union [60].

2.8. Second-Step Validation

For an additional validation, we applied the models estimated for France and Spain to FUAs in another European country—Austria (Figure 8).

![Figure 8. FUAs in Austria used for the models' validation.](image)

Although Austria differs from the two other countries under analysis in terms of size, urbanization level, topography, and FUAs’ location, it was chosen for an additional model validation, to demonstrate that the estimated models perform reasonably well even in this specific case. As all FUAs in this country are located apart from each other (see Figure 8), this country is considered particularly suitable for the intended validation.

The validation procedure was carried out in the following four steps. First, we determined the optimal ALAN thresholds for each FUA empirically (see Sections 2.3 and 2.4). Second, we used the ALAN-threshold identification models, estimated for France and Spain (see Section 2.5), to predict optimal ALAN thresholds for FUAs in Austria, using relevant input variables (see Sections 2.2 and 2.5), and, then, adjusted these estimates for contiguity (see Section 2.6). Considering that FUAs in Austria are located in close proximity to international borders, the input information was not limited to the areas inside Austria only. For instance, population density-decline gradient and distance to the closest
major city were calculated regardless of the state borders. Third, we assessed the correspondence between the empirically determined and models-predicted ALAN thresholds using Pearson correlation coefficients, SEE, and WMSE. Finally, we compared delineations, based on model-predicted ALAN thresholds, with commuting-based delineations, while expanding the study area by a 50-km buffer around the Austrian border, to cover the parts of FUAs located outside Austria and potentially extending into neighboring countries. As in the previous stage of the analysis (Section 2.7), the comparison of the shapes was performed using $JI$.

3. Results

3.1. Optimal ALAN Thresholds

The descriptive statistics of the ALAN thresholds, estimated by the multi-step approach described in Section 2.1, are reported in Table 2, separately for France and Spain, both as ALAN percentiles and actual ALAN levels in nW/cm$^2$/sr. As evidenced by this table, the optimal ALAN thresholds identified for individual FUAs appear to vary widely, ranging from 0.15 to 9.91 nW/cm$^2$/sr for France, and from 0.13 to 8.23 nW/cm$^2$/sr for Spain.

Table 2. Descriptive statistics of the identified ALAN thresholds.

<table>
<thead>
<tr>
<th>Country/Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>France (Number of FUAs = 82)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• ALAN percentile (0–100)</td>
<td>13.610</td>
<td>45.290</td>
<td>25.623</td>
<td>6.533</td>
</tr>
<tr>
<td>• ALAN threshold (nW/cm$^2$/sr)</td>
<td>0.150</td>
<td>9.910</td>
<td>0.664</td>
<td>1.218</td>
</tr>
<tr>
<td>Spain (Number of FUAs = 72)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• ALAN percentile (0–100)</td>
<td>12.660</td>
<td>46.460</td>
<td>29.355</td>
<td>7.279</td>
</tr>
<tr>
<td>• ALAN threshold (nW/cm$^2$/sr)</td>
<td>0.130</td>
<td>8.230</td>
<td>1.026</td>
<td>1.518</td>
</tr>
</tbody>
</table>

The most frequent (i.e., modal) ALAN values for all FUAs under analysis are reported in Figure 9, separately for France and Spain. In both countries, the modal ALAN values are not identical to the dimmest ones, thus pointing out that the “circularity” assumption (see Section 2.3) is violated. The bottom sub-figures report ALAN thresholds, corrected for compactness using the approach described in Section 2.4. As it can be seen from the comparison of the upper and bottom diagrams, modal ALAN thresholds, corrected for compactness are closer to the dimmest ALAN values than before the correction (especially for France), albeit differences in distributions are still valid.

Figure 9. Cont.
Figure 9. The modal and dimmest ALAN values estimated for individual FUAs before correcting for compactness ((a) = France; (b) = Spain) and after correcting for compactness ((c) = France; (d) = Spain). Notes: The column numbering (axis X) refers to FUA numbers listed in Table A1 of the Appendix. FUAs are sorted in an ascending order according to their modal ALAN values (upper diagrams) or according to compactness-based ALAN thresholds (bottom diagrams).

3.2. Explaining the Variance of the Observed ALAN Thresholds

In Table 3, we report the results of OLS analysis, linking individually determined optimal ALAN thresholds with geographic and socio-economic attributes of the FUAs’ core areas. As evidenced by Table 3, the predictors used in the analysis help to explain ~74% of the ALAN threshold variance ($R^2 = 0.739–0.740$). Characteristically, in both models, significant predictors are nearly identical and exhibit the same signs: population density (+); population density gradient (−); latitude (+), and distance to the nearest major city (−) ($p < 0.01$).

Table 3. Factors affecting ALAN threshold values estimated for individual FUAs (Method—OLS; Dependent variable—ALAN optimal threshold level, Box-Cox transformed with $\alpha = -0.55$).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Model 1 (France)</th>
<th></th>
<th>Model 2 (Spain)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Beta</td>
<td>t</td>
<td>B</td>
</tr>
<tr>
<td>(Constant)</td>
<td>-11.192</td>
<td>-</td>
<td>-8.466 *</td>
<td>-11.846</td>
</tr>
<tr>
<td>Latitude (dd)</td>
<td>0.106</td>
<td>0.246</td>
<td>4.209 *</td>
<td>0.140</td>
</tr>
<tr>
<td>Population density of the FUA core, persons per km$^2$ (ln)</td>
<td>1.151</td>
<td>0.751</td>
<td>11.143 *</td>
<td>1.158</td>
</tr>
<tr>
<td>Population density gradient (ln)</td>
<td>-1.369</td>
<td>-0.685</td>
<td>-10.349 *</td>
<td>-1.209</td>
</tr>
<tr>
<td>Distance to the nearest major city (dd)</td>
<td>-0.137</td>
<td>-0.183</td>
<td>-3.035 *</td>
<td>-0.147</td>
</tr>
<tr>
<td>N of obs.</td>
<td>82</td>
<td></td>
<td></td>
<td>72</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.739</td>
<td></td>
<td></td>
<td>0.740</td>
</tr>
<tr>
<td>$r$</td>
<td>0.866</td>
<td></td>
<td></td>
<td>0.812</td>
</tr>
<tr>
<td>SEE</td>
<td>0.533</td>
<td></td>
<td></td>
<td>0.629</td>
</tr>
<tr>
<td>WMSE</td>
<td>4.521</td>
<td></td>
<td></td>
<td>2.718</td>
</tr>
<tr>
<td>$F$</td>
<td>54.43 *</td>
<td></td>
<td></td>
<td>46.304 *</td>
</tr>
</tbody>
</table>

Notes: $^a$ unstandardized regression coefficients; $^b$ standardized regression coefficients; $^c$ t-statistic and its significance level; SEE = standard error of the estimates; WMSE = weighted mean squared error; $F = F$-statistics; $^* 0.01$ significance level.

As random forest regressions do not provide explicit estimates of the explanatory variables’ coefficients, we do not report these models here, but should remark that these estimates in terms of correlation with the actual ALAN threshold levels are similar to the
OLS estimates reported in Table 3 ($r = 0.856$ for France and $r = 0.883$ for Spain, as opposed to $r = 0.866$ for France and $r = 0.812$ for Spain in the OLS models), while in terms of SEE they are poorer ($SEE = 0.913$ for France and $SEE = 0.817$ for Spain in comparison to $SEE = 0.533$ for France and $SEE = 0.629$ for Spain; see Table 3).

However, in terms of WMS$_E$, random forest regressions are much better ($WMS_E = 0.945$ for France and $WMS_E = 0.545$ for Spain in comparison to $WMS_E = 4.521$ for France and $WMS_E = 2.718$ for Spain in the OLS models; see Table 3). Considering this result, we use the ALAN threshold estimates, produced by the random forest regressions, in the following analysis.

### 3.3. Model Cross-Validation

In Figure 10, we report the correspondence between the empirically determined and model-predicted ALAN thresholds. For this analysis, the model estimated for France (see Table 3) is applied to FUAs in Spain and vice versa. As evidenced by this figure, the estimates are fairly congruent, with $r > 0.819$.

![Figure 10](image1.png)

**Figure 10.** Models cross-validation results for France (a) and Spain (b).

### 3.4. Model-Estimated vs. Commuting-Based FUAs' Delineations

Figure 11 shows several most successful examples of FUAs’ delineations, generated by the proposed approach. Concurrently, in Figure 12, we report actual FUA delineations and model estimates for all FUAs in continental France and Spain. In addition, in Table 4, we report the degree of correspondence between the model-estimated and commuting-based delineations, assessed using $JI$.

![Figure 11](image2.png)

**Figure 11.** Cont.
Figure 11. Examples of FUAs featuring compactness-based boundaries (blue lines), model-based boundaries (green lines) and commuting-based boundaries (black lines): Paris (a) and Madrid (b) (see text for explanations).

Figure 12. Commuting-based (a) vs. model-estimated (b) delineations of FUAs in France and Spain.

Table 4. Values of the Jaccard index (JI) calculated for FUAs of different size and population density.

<table>
<thead>
<tr>
<th>FUA Type</th>
<th>N. of Obs.</th>
<th>Delineations Derived from Compactness-Based ALAN Thresholds</th>
<th>Delineations Derived from Model-Based ALAN Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>All FUAs under analysis</td>
<td>154</td>
<td>0.342</td>
<td>0.158</td>
</tr>
<tr>
<td>FUAs in:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• France</td>
<td>82</td>
<td>0.304</td>
<td>0.116</td>
</tr>
<tr>
<td>• Spain</td>
<td>72</td>
<td>0.385</td>
<td>0.186</td>
</tr>
<tr>
<td>FUAs by class:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• 1&amp;2 (Smallest)</td>
<td>93</td>
<td>0.327</td>
<td>0.154</td>
</tr>
<tr>
<td>• 3 (Medium)</td>
<td>55</td>
<td>0.351</td>
<td>0.154</td>
</tr>
<tr>
<td>• 4 (Largest)</td>
<td>6</td>
<td>0.499</td>
<td>0.187</td>
</tr>
</tbody>
</table>
Table 4. Cont.

<table>
<thead>
<tr>
<th>FUA Type</th>
<th>N. of Obs.</th>
<th>Delineations Derived from Compactness-Based ALAN Thresholds</th>
<th>Delineations Derived from Model-Based ALAN Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Population density in the FUA core, people per km$^2$ (ln)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• ≤5</td>
<td>18</td>
<td>0.235</td>
<td>0.097</td>
</tr>
<tr>
<td>• &gt;5</td>
<td>136</td>
<td>0.356</td>
<td>0.159</td>
</tr>
<tr>
<td>• &gt;6.5</td>
<td>48</td>
<td>0.440</td>
<td>0.177</td>
</tr>
<tr>
<td>• &gt;7.5</td>
<td>12</td>
<td>0.551</td>
<td>0.190</td>
</tr>
<tr>
<td>Population density in the core’s buffer zone, people per km$^2$ (ln)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• ≤5</td>
<td>101</td>
<td>0.313</td>
<td>0.137</td>
</tr>
<tr>
<td>• &gt;5</td>
<td>53</td>
<td>0.398</td>
<td>0.180</td>
</tr>
<tr>
<td>• &gt;6</td>
<td>11</td>
<td>0.525</td>
<td>0.220</td>
</tr>
<tr>
<td>• &gt;6.3</td>
<td>4</td>
<td>0.638</td>
<td>0.277</td>
</tr>
<tr>
<td>Ratio between population density in the core and the core’s buffer zone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• ≤2</td>
<td>25</td>
<td>0.303</td>
<td>0.151</td>
</tr>
<tr>
<td>• &gt;2</td>
<td>129</td>
<td>0.350</td>
<td>0.159</td>
</tr>
<tr>
<td>• &gt;5</td>
<td>43</td>
<td>0.363</td>
<td>0.153</td>
</tr>
<tr>
<td>• &gt;10</td>
<td>11</td>
<td>0.467</td>
<td>0.184</td>
</tr>
</tbody>
</table>

As can be seen in Table 4, the calculated $I_J$ values range between 0.30 and 0.64, being higher for large FUAs ($I_J = 0.499–0.507$) than for small FUAs ($I_J = 0.33–0.34$). For densely populated FUAs, the match between the commuting-based and ALAN-based delineations is especially high, reaching 0.557–0.638, or 56–64% (see Table 4). [The $I_J$ values for all the French and Spanish FUAs are reported in Figure A2 in Appendix A].

3.5. Second-Step Validation

In Table 5, we report ALAN threshold values for FUAs in Austria, calculated using the ‘French’ and ‘Spanish’ models (Table 3), and compared to individually fitted ALAN thresholds. As evidenced by this table, the ALAN thresholds, estimated using the French and Spanish models, correspond to the individually fitted ALAN thresholds quite well, with $r > 0.77$ and $SEE < 0.82$. Yet, in terms of WMSE, the French model performs poorer in comparison to the Spanish model (WMSE = 0.711 vs. WMSE = 10.102, respectively). In Figure 13, we report FUAs’ delineations obtained by averaging the estimates obtained using the French and Spanish models (see Table 3).

Table 5. Individually fitted vs. model-estimated ALAN thresholds for FUAs in Austria.

<table>
<thead>
<tr>
<th>FUA</th>
<th>Individually Fitted</th>
<th>Estimated Using the “French” Model</th>
<th>Estimated Using the “Spanish” Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vienna</td>
<td>0.34</td>
<td>0.65</td>
<td>1.56</td>
</tr>
<tr>
<td>Graz</td>
<td>0.23</td>
<td>0.48</td>
<td>1.05</td>
</tr>
<tr>
<td>Linz</td>
<td>0.24</td>
<td>0.49</td>
<td>1.04</td>
</tr>
<tr>
<td>Salzburg</td>
<td>0.27</td>
<td>0.49</td>
<td>1.08</td>
</tr>
</tbody>
</table>
Table 5. Cont.

<table>
<thead>
<tr>
<th>FUA</th>
<th>ALAN Threshold, nW/cm²/sr</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individually Fitted</td>
</tr>
<tr>
<td>Innsbruck</td>
<td>0.26</td>
</tr>
<tr>
<td>Klagenfurt</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Performance indicators (in relation to individually fitted ALAN threshold)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>-</td>
<td>0.783</td>
</tr>
<tr>
<td>SEE</td>
<td>-</td>
<td>0.216</td>
</tr>
<tr>
<td>WMSE</td>
<td>-</td>
<td>0.711</td>
</tr>
</tbody>
</table>

Figure 13. Commuting-based vs. models-estimated delineations of FUAs in Austria (see text for explanations).

4. Discussion and Conclusions

The delineation of geographic boundaries of FUAs is important for comparative urban studies. However, using commuting data for this task is not always feasible due to difficulties in data collection. In the present study, we suggested and tested an approach, based on the analysis of ALAN data. As ALAN is emitted from roads, frequented by commuters, and by buildings surrounding roads, ALAN emissions can be used, as we hypothesize, for the identification of FUAs.

We verify this hypothesis using data on commuting-based delineations available for France and Spain, applying a multi-step approach. First, we fit the ALAN threshold for each individual FUA, using the modal value of the ALAN frequency distribution. Next, we explain this threshold by a multiple regression analysis, using several characteristics of the FUAs’ cores, such as latitude of the core’s centroid, distance to the closest major city, population density, and density decline gradient. Although the boundaries of the FUA core areas used as an initial input are not generated by the analysis per se, such boundaries, if not a priori available, can be identified easily using Global Human Settlement [61] or LandScan [52] as contiguities of densely populated grids. Lastly, we cross-validate the obtained models for three European countries.

As our analysis indicates, the degree of correspondence between the individually fitted and model-predicted ALAN thresholds is relatively high ($r > 0.819$), with Jaccard Index values reaching up to 75% for France and up to 100% for Spain.

Our results are more robust than those obtained by Bosker and colleagues [18] for FUAs in Indonesia, according to which the correspondence between ALAN-based and 15% commuting-based FUA delineations did not exceed 28%. We explain the improvement, obtained in the present study, by the use of individually-fitted ALAN thresholds, based on the analysis of modal values, corrected for compactness.
To the best of our knowledge, this study is the first that estimates the optimal ALAN thresholds that approximate the boundaries of individual FUAs, using readily available, or easy-to-compute, characteristics of the FUAs’ cores, such as latitude of the core’s centroid, distance to the closest major city, population density and population density decline gradient, combined with ALAN flux data.

The proposed modelling approach might be useful for FUA delineations in countries and regions, for which commuting data are unavailable, as well as for places, in which commuting data are not updated on a regular basis, and for a comparative analysis of countries and regions, which use different commuting-assessment procedures. Using our modeling approach, FUAs’ boundaries can be determined in the following steps. First, the boundaries of FUAs’ cores should be identified. If such boundaries are not readily available, they can be determined as contiguities of high-density grid cells, using input sources, such as Global Human Settlement [61] or LandScan [52] grids. The procedure might follow the algorithm described in Dijkstra et al. (2019): Grid cells with population density of at least 1500 residents per km$^2$ are identified. Afterward, the grid cells identified thereby are grouped into contiguous area with a total population of at least 50,000 residents. For such areas, the development and locational characteristics are identified next, including the latitude of the contiguity’s centroid, distance to the closest major city, population density and population density decline gradient (see Section 2.5). These characteristics of the core areas are then used as predictors in the ALAN-threshold estimation models, reported in Section 2.5 for either France or Spain, or both, to obtain the optimal ALAN threshold estimates for each individual FUA. Finally, a VIIRS-DNB raster is used, to select pixels, corresponding to the estimated ALAN threshold, and to identify LAUs associated with such pixels’ contiguities, as detailed in Section 2.6.

The present study has several limitations. While for some FUAs, our estimates are quite accurate, reaching the levels of accuracy of 74–100%, for other, typically smaller FUAs, our estimations are less accurate. We assume that the reason might be that commuting-based boundaries rely mainly on work-related commuting, while omitting other human flows, such as travels for leisure, services, and social activities. In contrast, the suggested ALAN approach captures human activities at large. In addition, the ALAN-based approach might omit areas occupied by functions that operate mainly at daytime and emit much less light at night. For smaller FUAs, this source of error might by more pronounced than for large FUAs, where many functions operate around the clock. Another possible reason for a relatively low correspondence between some commuting- and ALAN-based delineations might be due to the fact that many FUAs are not monocentric, or might have a shape which is far from circular or elliptic, which we considered for modelling. For such cases, further studies might be needed to reflect more complex situations, in which FUA is either polycentric, or adjacent to other FUAs and their boundaries overlap or merge.

It should also be noted that in this study, we investigated the performance of the proposed method by applying it to three well-developed countries in Europe—France, Spain, and Austria. Yet, question remains about the models’ applicability to countries outside Europe and to countries in mid-latitudes, and, especially, to less-developed countries. We expect that applying the models to such countries might result in the overestimation of the optimal ALAN thresholds and thus in the underestimation of the commuting extent (the evidence for this conclusion is provided in [21]). Therefore, a follow-up investigation of the applicability of the proposed models to less developed countries and regions might be needed. Additionally, we need to acknowledge that a temporal mismatch between ALAN and actual FUAs’ delineation exists. Whether it might affect the results of the analysis should be clarified in future studies, after newer commuting data become available.

Author Contributions: Conceptualization, B.A.P. and I.C.; methodology, N.R., S.R. and B.A.P.; software, S.R. and N.R.; formal analysis, N.R.; data curation, N.R.; writing—original draft preparation, N.R.; writing—review and editing, B.A.P., I.C. and N.R.; overall study supervision, B.A.P. All authors have read and agreed to the published version of the manuscript.
**Funding:** This research was funded by the Council for Higher Education of Israel (postdoctoral scholarship of N.R.).

**Data Availability Statement:** Initial data and processing codes are available from N.R. upon request.

**Acknowledgments:** We express our gratitude to three anonymous reviewers for the highly valuable comments.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Appendix A**

**Table A1.** FUAs’ identification (ID number = number in axis X in Figure 9 in the main manuscript).

<table>
<thead>
<tr>
<th>ID Number</th>
<th>FUA Code</th>
<th>FUA Name</th>
<th>ID Number</th>
<th>FUA Code</th>
<th>FUA Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FR067</td>
<td>Quimper</td>
<td>1</td>
<td>ES053</td>
<td>Ciudad Real</td>
</tr>
<tr>
<td>2</td>
<td>FR061</td>
<td>Niort</td>
<td>2</td>
<td>ES058</td>
<td>Avila</td>
</tr>
<tr>
<td>3</td>
<td>FR021</td>
<td>Poitiers</td>
<td>3</td>
<td>ES046</td>
<td>Merida</td>
</tr>
<tr>
<td>4</td>
<td>FR053</td>
<td>La Rochelle</td>
<td>4</td>
<td>ES016</td>
<td>Toledo</td>
</tr>
<tr>
<td>5</td>
<td>FR051</td>
<td>Troyes</td>
<td>5</td>
<td>ES057</td>
<td>Ponferrada</td>
</tr>
<tr>
<td>6</td>
<td>FR086</td>
<td>Evreux</td>
<td>6</td>
<td>ES527</td>
<td>Jaen</td>
</tr>
<tr>
<td>7</td>
<td>FR066</td>
<td>Saint-Brieuc</td>
<td>7</td>
<td>ES050</td>
<td>Manresa</td>
</tr>
<tr>
<td>8</td>
<td>FR077</td>
<td>Roanne</td>
<td>8</td>
<td>ES011</td>
<td>Santiago de Compostela</td>
</tr>
<tr>
<td>9</td>
<td>FR035</td>
<td>Tours</td>
<td>9</td>
<td>ES040</td>
<td>Talavera de la Reina</td>
</tr>
<tr>
<td>10</td>
<td>FR059</td>
<td>Chalon-sur-Saone</td>
<td>10</td>
<td>ES054</td>
<td>Chiclan de la Frontera</td>
</tr>
<tr>
<td>11</td>
<td>FR093</td>
<td>Brive-la-Gaillarde</td>
<td>11</td>
<td>ES545</td>
<td>Lorca</td>
</tr>
<tr>
<td>12</td>
<td>FR104</td>
<td>Chalons-en-Champagne</td>
<td>12</td>
<td>ES528</td>
<td>Lleida</td>
</tr>
<tr>
<td>13</td>
<td>FR076</td>
<td>Belfort</td>
<td>13</td>
<td>ES059</td>
<td>Ourense</td>
</tr>
<tr>
<td>14</td>
<td>FR025</td>
<td>Besancon</td>
<td>14</td>
<td>ES523</td>
<td>Leon</td>
</tr>
<tr>
<td>15</td>
<td>FR073</td>
<td>Tarbes</td>
<td>15</td>
<td>ES031</td>
<td>Lugo</td>
</tr>
<tr>
<td>16</td>
<td>FR505</td>
<td>Charleville-Mezieres</td>
<td>16</td>
<td>ES043</td>
<td>Lugo</td>
</tr>
<tr>
<td>17</td>
<td>FR038</td>
<td>Le Mans</td>
<td>17</td>
<td>ES043</td>
<td>Ferrol</td>
</tr>
<tr>
<td>18</td>
<td>FR037</td>
<td>Brest</td>
<td>18</td>
<td>ES034</td>
<td>Caceres</td>
</tr>
<tr>
<td>19</td>
<td>FR068</td>
<td>Vannes</td>
<td>19</td>
<td>ES515</td>
<td>Burgos</td>
</tr>
<tr>
<td>20</td>
<td>FR096</td>
<td>Albi</td>
<td>20</td>
<td>ES519</td>
<td>Albacete</td>
</tr>
<tr>
<td>21</td>
<td>FR050</td>
<td>Montbeliard</td>
<td>21</td>
<td>ES014</td>
<td>Pamplona</td>
</tr>
<tr>
<td>22</td>
<td>FR074</td>
<td>Compiègne</td>
<td>22</td>
<td>ES041</td>
<td>Palencia</td>
</tr>
<tr>
<td>23</td>
<td>FR022</td>
<td>Clermont-Ferrand</td>
<td>23</td>
<td>ES542</td>
<td>Basin</td>
</tr>
<tr>
<td>24</td>
<td>FR023</td>
<td>Caen</td>
<td>24</td>
<td>ES017</td>
<td>Badajoz</td>
</tr>
<tr>
<td>25</td>
<td>FR506</td>
<td>Colmar</td>
<td>25</td>
<td>ES544</td>
<td>Linares</td>
</tr>
<tr>
<td>26</td>
<td>FR036</td>
<td>Angers</td>
<td>26</td>
<td>ES510</td>
<td>Donostia-San Sebastian</td>
</tr>
<tr>
<td>27</td>
<td>FR019</td>
<td>Orleans</td>
<td>27</td>
<td>ES062</td>
<td>Sanlucar de Barrameda</td>
</tr>
<tr>
<td>28</td>
<td>FR049</td>
<td>Lorient</td>
<td>28</td>
<td>ES033</td>
<td>Girona</td>
</tr>
<tr>
<td>29</td>
<td>FR058</td>
<td>Chambery</td>
<td>29</td>
<td>ES013</td>
<td>Oriedo</td>
</tr>
<tr>
<td>30</td>
<td>FR069</td>
<td>Cherbourg</td>
<td>30</td>
<td>ES009</td>
<td>Valladolid</td>
</tr>
<tr>
<td>31</td>
<td>FR018</td>
<td>Reims</td>
<td>31</td>
<td>ES022</td>
<td>Vigo</td>
</tr>
<tr>
<td>32</td>
<td>FR090</td>
<td>Chateauroux</td>
<td>32</td>
<td>ES054</td>
<td>Benidorm</td>
</tr>
<tr>
<td>33</td>
<td>FR056</td>
<td>Angouleme</td>
<td>33</td>
<td>ES051</td>
<td>Granada</td>
</tr>
<tr>
<td>34</td>
<td>FR063</td>
<td>Beziers</td>
<td>34</td>
<td>ES044</td>
<td>Pontevedra</td>
</tr>
<tr>
<td>35</td>
<td>FR020</td>
<td>Dijon</td>
<td>35</td>
<td>ES537</td>
<td>Alcoy</td>
</tr>
<tr>
<td>36</td>
<td>FR064</td>
<td>Arras</td>
<td>36</td>
<td>ES514</td>
<td>Almeria</td>
</tr>
<tr>
<td>37</td>
<td>FR057</td>
<td>Boulogne-sur-Mer</td>
<td>37</td>
<td>ES552</td>
<td>Igualada</td>
</tr>
<tr>
<td>38</td>
<td>FR016</td>
<td>Nancy</td>
<td>38</td>
<td>ES547</td>
<td>Sagunto</td>
</tr>
<tr>
<td>39</td>
<td>FR014</td>
<td>Amiens</td>
<td>39</td>
<td>ES026</td>
<td>Coruna (A)</td>
</tr>
<tr>
<td>40</td>
<td>FR048</td>
<td>Annecy</td>
<td>40</td>
<td>ES007</td>
<td>Murcia</td>
</tr>
<tr>
<td>41</td>
<td>FR079</td>
<td>Saint-Quentin</td>
<td>41</td>
<td>ES037</td>
<td>Puerto de Santa Maria, El</td>
</tr>
<tr>
<td>42</td>
<td>FR045</td>
<td>Pau</td>
<td>42</td>
<td>ES516</td>
<td>Salamanca</td>
</tr>
<tr>
<td>43</td>
<td>FR006</td>
<td>Strasbourg</td>
<td>43</td>
<td>ES021</td>
<td>Alicante</td>
</tr>
</tbody>
</table>
Table A1. Cont.

<table>
<thead>
<tr>
<th>ID Number</th>
<th>FUA Code</th>
<th>FUA Name</th>
<th>ID Number</th>
<th>FUA Code</th>
<th>FUA Name</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>France</strong></td>
<td><strong>Spain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>FR215</td>
<td>Rouen</td>
<td>44</td>
<td>ES073</td>
<td>Elda</td>
</tr>
<tr>
<td>45</td>
<td>FR082</td>
<td>Beauvais</td>
<td>45</td>
<td>ES048</td>
<td>Guadalajara</td>
</tr>
<tr>
<td>46</td>
<td>FR304</td>
<td>Melun</td>
<td>46</td>
<td>ES046</td>
<td>Gandia</td>
</tr>
<tr>
<td>47</td>
<td>FR011</td>
<td>Saint-Étienne</td>
<td>47</td>
<td>ES020</td>
<td>Cordoba</td>
</tr>
<tr>
<td>48</td>
<td>FR084</td>
<td>Creil</td>
<td>48</td>
<td>ES023</td>
<td>Gijon</td>
</tr>
<tr>
<td>49</td>
<td>FR214</td>
<td>Valence</td>
<td>49</td>
<td>ES533</td>
<td>Marbella</td>
</tr>
<tr>
<td>50</td>
<td>FR046</td>
<td>Bayonne</td>
<td>50</td>
<td>ES532</td>
<td>Algeciras</td>
</tr>
<tr>
<td>51</td>
<td>FR026</td>
<td>Grenoble</td>
<td>51</td>
<td>ES004</td>
<td>Seville</td>
</tr>
<tr>
<td>52</td>
<td>FR065</td>
<td>Bourges</td>
<td>52</td>
<td>ES035</td>
<td>Torrevieja</td>
</tr>
<tr>
<td>53</td>
<td>FR060</td>
<td>Chartres</td>
<td>53</td>
<td>ES018</td>
<td>Logrono</td>
</tr>
<tr>
<td>54</td>
<td>FR099</td>
<td>Frejus</td>
<td>54</td>
<td>ES028</td>
<td>Reus</td>
</tr>
<tr>
<td>55</td>
<td>FR039</td>
<td>Avignon</td>
<td>55</td>
<td>ES522</td>
<td>Cadiz</td>
</tr>
<tr>
<td>56</td>
<td>FR024</td>
<td>Limoges</td>
<td>56</td>
<td>ES005</td>
<td>Saragossa</td>
</tr>
<tr>
<td>57</td>
<td>FR205</td>
<td>Nice</td>
<td>57</td>
<td>ES006</td>
<td>Malaga</td>
</tr>
<tr>
<td>58</td>
<td>FR034</td>
<td>Valenciennes</td>
<td>58</td>
<td>ES508</td>
<td>Jerez de la Frontera</td>
</tr>
<tr>
<td>59</td>
<td>FR008</td>
<td>Nantes</td>
<td>59</td>
<td>ES065</td>
<td>Linea de la Concepcion, La Plata</td>
</tr>
<tr>
<td>60</td>
<td>FR010</td>
<td>Montpellier</td>
<td>60</td>
<td>ES012</td>
<td>Vitoria</td>
</tr>
<tr>
<td>61</td>
<td>FR040</td>
<td>Mulhouse</td>
<td>61</td>
<td>ES505</td>
<td>Elche/Eix</td>
</tr>
<tr>
<td>62</td>
<td>FR047</td>
<td>Annemasse</td>
<td>62</td>
<td>ES521</td>
<td>Huelva</td>
</tr>
<tr>
<td>63</td>
<td>FR007</td>
<td>Bordeaux</td>
<td>63</td>
<td>ES001</td>
<td>Madrid</td>
</tr>
<tr>
<td>64</td>
<td>FR004</td>
<td>Toulouse</td>
<td>64</td>
<td>ES070</td>
<td>Irún</td>
</tr>
<tr>
<td>65</td>
<td>FR043</td>
<td>Perpignan</td>
<td>65</td>
<td>ES520</td>
<td>Castellon de la Plana/Castello de la Plana</td>
</tr>
<tr>
<td>66</td>
<td>FR044</td>
<td>Nîmes</td>
<td>66</td>
<td>ES015</td>
<td>Santander</td>
</tr>
<tr>
<td>67</td>
<td>FR052</td>
<td>Saint-Nazaire</td>
<td>67</td>
<td>ES525</td>
<td>Tarragona</td>
</tr>
<tr>
<td>68</td>
<td>FR017</td>
<td>Metz</td>
<td>68</td>
<td>ES002</td>
<td>Barcelona</td>
</tr>
<tr>
<td>69</td>
<td>FR009</td>
<td>Lille</td>
<td>69</td>
<td>ES003</td>
<td>Valencia</td>
</tr>
<tr>
<td>70</td>
<td>FR003</td>
<td>Lyon</td>
<td>70</td>
<td>ES019</td>
<td>Bilbao</td>
</tr>
<tr>
<td>71</td>
<td>FR012</td>
<td>Le Havre</td>
<td>71</td>
<td>ES506</td>
<td>Cartagena</td>
</tr>
<tr>
<td>72</td>
<td>FR519</td>
<td>Cannes</td>
<td>72</td>
<td>ES039</td>
<td>Avilés</td>
</tr>
<tr>
<td>73</td>
<td>FR207</td>
<td>Lens - Lievin</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>FR209</td>
<td>Douai</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>FR032</td>
<td>Toulon</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>FR062</td>
<td>Calais</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>FR001</td>
<td>Paris</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>FR203</td>
<td>Marseille</td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>FR042</td>
<td>Dunkerque</td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>FR324</td>
<td>Martigues</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>FR208</td>
<td>Henin - Carvin</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>FR013</td>
<td>Renne</td>
<td>82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>France</strong></th>
<th><strong>Spain</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FR324</td>
</tr>
<tr>
<td>2</td>
<td>FR047</td>
</tr>
<tr>
<td>3</td>
<td>FR039</td>
</tr>
<tr>
<td>4</td>
<td>FR040</td>
</tr>
<tr>
<td>5</td>
<td>FR048</td>
</tr>
<tr>
<td>6</td>
<td>FR065</td>
</tr>
<tr>
<td>7</td>
<td>FR082</td>
</tr>
<tr>
<td>8</td>
<td>FR208</td>
</tr>
<tr>
<td>9</td>
<td>FR304</td>
</tr>
<tr>
<td>10</td>
<td>FR505</td>
</tr>
<tr>
<td>11</td>
<td>FR049</td>
</tr>
<tr>
<td>12</td>
<td>FR067</td>
</tr>
<tr>
<td>13</td>
<td>FR066</td>
</tr>
</tbody>
</table>
Table A1. Cont.

<table>
<thead>
<tr>
<th>ID Number</th>
<th>FUA Code</th>
<th>FUA Name</th>
<th>ID Number</th>
<th>FUA Code</th>
<th>FUA Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>FR214</td>
<td>Valence</td>
<td>14</td>
<td>ES062</td>
<td>Sanlucar de Barrameda</td>
</tr>
<tr>
<td>15</td>
<td>FR209</td>
<td>Douai</td>
<td>15</td>
<td>ES540</td>
<td>Chiclana de la Frontera</td>
</tr>
<tr>
<td>16</td>
<td>FR068</td>
<td>Vannes</td>
<td>16</td>
<td>ES070</td>
<td>Irun</td>
</tr>
<tr>
<td>17</td>
<td>FR027</td>
<td>Lens - Lievin</td>
<td>17</td>
<td>ES514</td>
<td>Almeria</td>
</tr>
<tr>
<td>18</td>
<td>FR053</td>
<td>La Rochelle</td>
<td>18</td>
<td>ES046</td>
<td>Gandia</td>
</tr>
<tr>
<td>19</td>
<td>FR506</td>
<td>Colmar</td>
<td>19</td>
<td>ES053</td>
<td>Ciudad Real</td>
</tr>
<tr>
<td>20</td>
<td>FR064</td>
<td>Arras</td>
<td>20</td>
<td>ES057</td>
<td>Ponferrada</td>
</tr>
<tr>
<td>21</td>
<td>FR084</td>
<td>Creil</td>
<td>21</td>
<td>ES528</td>
<td>Lleida</td>
</tr>
<tr>
<td>22</td>
<td>FR050</td>
<td>Montbeliard</td>
<td>22</td>
<td>ES532</td>
<td>Algeciras</td>
</tr>
<tr>
<td>23</td>
<td>FR077</td>
<td>Roanne</td>
<td>23</td>
<td>ES028</td>
<td>Reus</td>
</tr>
<tr>
<td>24</td>
<td>FR056</td>
<td>Angouleme</td>
<td>24</td>
<td>ES039</td>
<td>Aviles</td>
</tr>
<tr>
<td>25</td>
<td>FR079</td>
<td>Saint-Quentin</td>
<td>25</td>
<td>ES054</td>
<td>Benidorm</td>
</tr>
<tr>
<td>26</td>
<td>FR069</td>
<td>Cherbourg</td>
<td>26</td>
<td>ES521</td>
<td>Huelva</td>
</tr>
<tr>
<td>27</td>
<td>FR012</td>
<td>Le Havre</td>
<td>27</td>
<td>ES519</td>
<td>Albacete</td>
</tr>
<tr>
<td>28</td>
<td>FR086</td>
<td>Evreux</td>
<td>28</td>
<td>ES537</td>
<td>Alcoy</td>
</tr>
<tr>
<td>29</td>
<td>FR519</td>
<td>Cannes</td>
<td>29</td>
<td>ES527</td>
<td>Jaen</td>
</tr>
<tr>
<td>30</td>
<td>FR063</td>
<td>Beziers</td>
<td>30</td>
<td>ES033</td>
<td>Girona</td>
</tr>
<tr>
<td>31</td>
<td>FR058</td>
<td>Chambery</td>
<td>31</td>
<td>ES525</td>
<td>Tarragona</td>
</tr>
<tr>
<td>32</td>
<td>FR090</td>
<td>Chateauroux</td>
<td>32</td>
<td>ES011</td>
<td>Santiago de Compostela</td>
</tr>
<tr>
<td>33</td>
<td>FR096</td>
<td>Albi</td>
<td>33</td>
<td>ES040</td>
<td>Talavera de la Reina</td>
</tr>
<tr>
<td>34</td>
<td>FR052</td>
<td>Saint-Nazaire</td>
<td>34</td>
<td>ES031</td>
<td>Lugo</td>
</tr>
<tr>
<td>35</td>
<td>FR057</td>
<td>Boulogne-sur-Mer</td>
<td>35</td>
<td>ES505</td>
<td>Elche/Elx</td>
</tr>
<tr>
<td>36</td>
<td>FR059</td>
<td>Chalon-sur-Saone</td>
<td>36</td>
<td>ES026</td>
<td>Coruna (A)</td>
</tr>
<tr>
<td>37</td>
<td>FR060</td>
<td>Chartres</td>
<td>37</td>
<td>ES522</td>
<td>Cadiz</td>
</tr>
<tr>
<td>38</td>
<td>FR061</td>
<td>Niort</td>
<td>38</td>
<td>ES016</td>
<td>Toledo</td>
</tr>
<tr>
<td>39</td>
<td>FR022</td>
<td>Clermont-Ferrand</td>
<td>39</td>
<td>ES544</td>
<td>Linares</td>
</tr>
<tr>
<td>40</td>
<td>FR010</td>
<td>Montpellier</td>
<td>40</td>
<td>ES018</td>
<td>Logrono</td>
</tr>
<tr>
<td>41</td>
<td>FR020</td>
<td>Dijon</td>
<td>41</td>
<td>ES533</td>
<td>Marbella</td>
</tr>
<tr>
<td>42</td>
<td>FR019</td>
<td>Orleans</td>
<td>42</td>
<td>ES510</td>
<td>Donostia-San Sebastian</td>
</tr>
<tr>
<td>43</td>
<td>FR026</td>
<td>Grenoble</td>
<td>43</td>
<td>ES552</td>
<td>Igualada</td>
</tr>
<tr>
<td>44</td>
<td>FR025</td>
<td>Besancon</td>
<td>44</td>
<td>ES073</td>
<td>Elda</td>
</tr>
<tr>
<td>45</td>
<td>FR076</td>
<td>Belfort</td>
<td>45</td>
<td>ES019</td>
<td>Bilbao</td>
</tr>
<tr>
<td>46</td>
<td>FR023</td>
<td>Caen</td>
<td>46</td>
<td>ES015</td>
<td>Santander</td>
</tr>
<tr>
<td>47</td>
<td>FR038</td>
<td>Le Mans</td>
<td>47</td>
<td>ES003</td>
<td>Valencia</td>
</tr>
<tr>
<td>48</td>
<td>FR045</td>
<td>Pau</td>
<td>48</td>
<td>ES529</td>
<td>Ourense</td>
</tr>
<tr>
<td>49</td>
<td>FR007</td>
<td>Bordeaux</td>
<td>49</td>
<td>ES520</td>
<td>Castellon de la Plana/Castello de la Plana</td>
</tr>
<tr>
<td>50</td>
<td>FR021</td>
<td>Poitiers</td>
<td>50</td>
<td>ES004</td>
<td>Seville</td>
</tr>
<tr>
<td>51</td>
<td>FR034</td>
<td>Valenciennes</td>
<td>51</td>
<td>ES516</td>
<td>Salamanca</td>
</tr>
<tr>
<td>52</td>
<td>FR044</td>
<td>Nimes</td>
<td>52</td>
<td>ES001</td>
<td>Madrid</td>
</tr>
<tr>
<td>53</td>
<td>FR035</td>
<td>Tours</td>
<td>53</td>
<td>ES002</td>
<td>Barcelona</td>
</tr>
<tr>
<td>54</td>
<td>FR008</td>
<td>Nantes</td>
<td>54</td>
<td>ES005</td>
<td>Saragossa</td>
</tr>
<tr>
<td>55</td>
<td>FR006</td>
<td>Strasbourg</td>
<td>55</td>
<td>ES006</td>
<td>Malaga</td>
</tr>
<tr>
<td>56</td>
<td>FR004</td>
<td>Toulouse</td>
<td>56</td>
<td>ES007</td>
<td>Murcia</td>
</tr>
<tr>
<td>57</td>
<td>FR042</td>
<td>Dunkerque</td>
<td>57</td>
<td>ES009</td>
<td>Valladolid</td>
</tr>
<tr>
<td>58</td>
<td>FR017</td>
<td>Metz</td>
<td>58</td>
<td>ES044</td>
<td>Pontevedra</td>
</tr>
<tr>
<td>59</td>
<td>FR003</td>
<td>Lyon</td>
<td>59</td>
<td>ES017</td>
<td>Badajoz</td>
</tr>
<tr>
<td>60</td>
<td>FR104</td>
<td>Chalons-en-Champagne</td>
<td>60</td>
<td>ES020</td>
<td>Cordoba</td>
</tr>
<tr>
<td>61</td>
<td>FR036</td>
<td>Angers</td>
<td>61</td>
<td>ES022</td>
<td>Vigo</td>
</tr>
<tr>
<td>62</td>
<td>FR016</td>
<td>Nancy</td>
<td>62</td>
<td>ES043</td>
<td>Ferrol</td>
</tr>
<tr>
<td>63</td>
<td>FR073</td>
<td>Tarbes</td>
<td>63</td>
<td>ES048</td>
<td>Guadalajara</td>
</tr>
<tr>
<td>64</td>
<td>FR046</td>
<td>Bayonne</td>
<td>64</td>
<td>ES006</td>
<td>Cartagena</td>
</tr>
<tr>
<td>65</td>
<td>FR093</td>
<td>Brive-la-Gaillarde</td>
<td>65</td>
<td>ES001</td>
<td>Granada</td>
</tr>
<tr>
<td>66</td>
<td>FR099</td>
<td>Frejus</td>
<td>66</td>
<td>ES008</td>
<td>Jerez de la Frontera</td>
</tr>
<tr>
<td>67</td>
<td>FR203</td>
<td>Marseille</td>
<td>67</td>
<td>ES023</td>
<td>Leon</td>
</tr>
<tr>
<td>68</td>
<td>FR205</td>
<td>Nice</td>
<td>68</td>
<td>ES542</td>
<td>Basin</td>
</tr>
<tr>
<td>69</td>
<td>FR001</td>
<td>Paris</td>
<td>69</td>
<td>ES538</td>
<td>Avila</td>
</tr>
</tbody>
</table>
Table A1. Cont.

<table>
<thead>
<tr>
<th>ID Number</th>
<th>FUA Code</th>
<th>FUA Name</th>
<th>ID Number</th>
<th>FUA Code</th>
<th>FUA Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>FR051</td>
<td>Troyes</td>
<td>70</td>
<td>ES546</td>
<td>Merida</td>
</tr>
<tr>
<td>71</td>
<td>FR009</td>
<td>Lille</td>
<td>71</td>
<td>ES545</td>
<td>Lorca</td>
</tr>
<tr>
<td>72</td>
<td>FR011</td>
<td>Saint-Etienne</td>
<td>72</td>
<td>ES515</td>
<td>Burgos</td>
</tr>
<tr>
<td>73</td>
<td>FR013</td>
<td>Rennes</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>FR014</td>
<td>Amiens</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>FR018</td>
<td>Reims</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>FR024</td>
<td>Limoges</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>FR032</td>
<td>Toulon</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>FR037</td>
<td>Brest</td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>FR062</td>
<td>Calais</td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>FR043</td>
<td>Perpignan</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>FR215</td>
<td>Rouen</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>FR074</td>
<td>Compiègne</td>
<td>82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure A1. Cont.
Figure A1. Light emission distribution from the center of a monocentric FUA, modelled by different geometric shapes (left panel) and distributions of ALAN in corresponding FUAs (right panel).

Figure A2. Jaccard Index for the estimated delineations, derived from the compactness-based (a,b) and model-based (c,d) ALAN thresholds: FUAs in France (a,c) and Spain (b,d). Note: The column numbering refers to FUA numbers listed in Table A2 below. In the graphs, FUAs are sorted in descending order according to their JI values.

Table A2. FUAs' identification (ID number = number on the X-axis in Figure A2 above).

<table>
<thead>
<tr>
<th>ID Number</th>
<th>FUA Code</th>
<th>FUA Name</th>
<th>FUA Code</th>
<th>FUA Name</th>
<th>FUA Code</th>
<th>FUA Name</th>
<th>FUA Code</th>
<th>FUA Name</th>
<th>FUA Code</th>
<th>FUA Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FR006</td>
<td>Strasbourg</td>
<td>ES065</td>
<td>Linea de la</td>
<td>FR006</td>
<td>Strasbourg</td>
<td>ES065</td>
<td>Linea de la</td>
<td>ES065</td>
<td>Linea de la</td>
</tr>
<tr>
<td>2</td>
<td>FR037</td>
<td>Brest</td>
<td>ES015</td>
<td>Santander</td>
<td>FR049</td>
<td>Lorient</td>
<td>ES501</td>
<td>Granada</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>FR047</td>
<td>Annemasse</td>
<td>ES001</td>
<td>Madrid</td>
<td>FR047</td>
<td>Annemasse</td>
<td>ES001</td>
<td>Madrid</td>
<td>ES001</td>
<td>Madrid</td>
</tr>
<tr>
<td>4</td>
<td>FR024</td>
<td>Limoges</td>
<td>ES540</td>
<td>Chiclana de la</td>
<td>FR039</td>
<td>Avignon</td>
<td>ES540</td>
<td>Chiclana de la</td>
<td>ES540</td>
<td>Chiclana de la</td>
</tr>
<tr>
<td>5</td>
<td>FR069</td>
<td>Cherbourg</td>
<td>ES002</td>
<td>Barcelona</td>
<td>FR003</td>
<td>Lyon</td>
<td>ES515</td>
<td>Burgos</td>
<td>ES515</td>
<td>Burgos</td>
</tr>
<tr>
<td>ID</td>
<td>Subfigure (a)</td>
<td>Subfigure (b)</td>
<td>Subfigure (c)</td>
<td>Subfigure (d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>FR039</td>
<td>Avignon</td>
<td>Granada</td>
<td>FR203</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>FR042</td>
<td>Dunkerque</td>
<td>Besancon</td>
<td>FR024</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>FR043</td>
<td>Perpignan</td>
<td>Perpignan</td>
<td>FR043</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>FR001</td>
<td>Paris</td>
<td>Bordeaux</td>
<td>ES538</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>FR023</td>
<td>Caen</td>
<td>Besancon</td>
<td>ES009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>FR062</td>
<td>Calais</td>
<td>Calais</td>
<td>ES506</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>FR052</td>
<td>Saint-Nazaire</td>
<td>Palencia</td>
<td>ES001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>FR067</td>
<td>Quimper</td>
<td>Salamanca</td>
<td>ES068</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>FR022</td>
<td>Clermont-Ferrand</td>
<td>Cadiz</td>
<td>ES063</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>FR505</td>
<td>Mezieres</td>
<td>Valladolid</td>
<td>FR008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>FR025</td>
<td>Nice</td>
<td>Besancon</td>
<td>ES532</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>FR039</td>
<td>Lille</td>
<td>Besancon</td>
<td>ES025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>FR034</td>
<td>Valenciennes</td>
<td>Besancon</td>
<td>ES529</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>FR035</td>
<td>Tarbes</td>
<td>Marbella</td>
<td>ES501</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>FR046</td>
<td>Bayonne</td>
<td>Niort</td>
<td>ES061</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>FR066</td>
<td>Saint-Brieuc</td>
<td>Le Havre</td>
<td>ES012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>FR034</td>
<td>Lorient</td>
<td>Le Havre</td>
<td>ES055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>FR050</td>
<td>Montbeliard</td>
<td>Le Havre</td>
<td>ES061</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>FR057</td>
<td>Boulogne-sur-Mer</td>
<td>Coruna (A)</td>
<td>FR214</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>FR003</td>
<td>Lyon</td>
<td>Aviles</td>
<td>ES045</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>FR008</td>
<td>Nantes</td>
<td>Nice</td>
<td>ES005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>FR032</td>
<td>Toulon</td>
<td>Nantes</td>
<td>ES057</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>FR506</td>
<td>Colmar</td>
<td>Puerto de Santa Maria, El</td>
<td>FR096</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>FR040</td>
<td>Mulhouse</td>
<td>Ciudad Real</td>
<td>FR505</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>FR012</td>
<td>Le Havre</td>
<td>Colmar</td>
<td>ES031</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>FR519</td>
<td>Cannes</td>
<td>Algeciras</td>
<td>ES506</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>FR004</td>
<td>Toulouse</td>
<td>Sanlucar de Barrameda</td>
<td>ES017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>FR084</td>
<td>Creil</td>
<td>Puerto de Santa Maria, El</td>
<td>FR096</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>FR065</td>
<td>Bourges</td>
<td>Ponteferrada</td>
<td>FR022</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>FR203</td>
<td>Marseille</td>
<td>Louise</td>
<td>ES056</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>FR028</td>
<td>Henin - Carvin</td>
<td>Ferrol</td>
<td>FR215</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>FR036</td>
<td>Angers</td>
<td>Gandia</td>
<td>FR004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>FR010</td>
<td>Montpellier</td>
<td>Toulouse</td>
<td>ES014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>FR053</td>
<td>La Rochelle</td>
<td>Puerto de Santa Maria, El</td>
<td>FR065</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>FR074</td>
<td>Compiégne</td>
<td>Mulhouse</td>
<td>ES040</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>FR044</td>
<td>Nîmes</td>
<td>Donostia-San Sebastian</td>
<td>FR519</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>FR017</td>
<td>Metz</td>
<td>Jerez de la Frontera</td>
<td>FR016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>FR082</td>
<td>Beauvais</td>
<td>Burgos</td>
<td>FR073</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>FR038</td>
<td>Martigues</td>
<td>Biarritz</td>
<td>FR012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>FR003</td>
<td>Poitiers</td>
<td>Reims</td>
<td>FR011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>FR077</td>
<td>Sens</td>
<td>Sens</td>
<td>FR104</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>FR056</td>
<td>Angoulême</td>
<td>Jarnac</td>
<td>FR047</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>FR013</td>
<td>Rennes</td>
<td>Guadalajara</td>
<td>FR058</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>FR025</td>
<td>Besancon</td>
<td>Toledo</td>
<td>FR067</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>FR058</td>
<td>Angoulême</td>
<td>Badajoz</td>
<td>FR209</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>FR050</td>
<td>Agen</td>
<td>Évora</td>
<td>FR324</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>FR060</td>
<td>Saint-Etienne</td>
<td>Calais</td>
<td>ES506</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>FR060</td>
<td>Chartres</td>
<td>Elche/Eix</td>
<td>FR069</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>FR034</td>
<td>Chalon-sur-Saone</td>
<td>Le Mans</td>
<td>ES031</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>FR032</td>
<td>Tours</td>
<td>Vitoria</td>
<td>FR079</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>FR035</td>
<td>Erreux</td>
<td>Lens - Lavois</td>
<td>FR207</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>FR079</td>
<td>Saint-Quentin</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>FR014</td>
<td>Amiens</td>
<td>Saint-Brieuc</td>
<td>FR090</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>FR215</td>
<td>Rouen</td>
<td>Doubs</td>
<td>FR020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>FR058</td>
<td>Chambery</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>FR051</td>
<td>Troyes</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>FR061</td>
<td>Niort</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>FR020</td>
<td>Dijon</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>FR104</td>
<td>Chalon-en-Champagne</td>
<td>Virieux</td>
<td>FR086</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>FR090</td>
<td>Chateaurenoux</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table A2. Cont.**
Box A1. Estimation of the compactness-based ALAN threshold (derivation).

The figure below illustrates our assumptions: We model actual FUAs’ shapes by ellipses with axes \(a\) and \(b\) \((a > b)\).

\[\text{(1)}\]

\[\text{(2)}\]

\[\text{Box A1. Estimation of the compactness-based ALAN threshold (derivation).}\]

A. Optimal Radius Calculation

To calculate the radius of the circle (\(r^*\)), ensuring maximal intersection with the ellipse we should define and maximize Jaccard index: 

\[J I = \frac{|S c \cap S E|}{|S c \cup S E|} \rightarrow \max,\] where \(S_c\) = area of the circle, and \(S_E\) = area of the ellipse. For this sake, we should calculate and differentiate the following function:

\[JI(r) = \int_0^i y_{\text{ellipse}}(x)dx + \int_i^r y_{\text{circle}}(x)dx\]

where \(y_{\text{circle}} = \sqrt{r^2 - x^2}\) is equation of circle and \(y_{\text{ellipse}} = \frac{b}{a} \sqrt{a^2 - x^2}\) is equation of ellipse.

Limit of integration \(i\) is defined as \(x\) coordinate of intersection \(y_{\text{circle}}\) and \(y_{\text{ellipse}}\):

\[\sqrt{r^2 - x^2} = \frac{b}{a} \sqrt{a^2 - x^2}\] (A2)

Both integrals in Equation (A1) are of the same type, which are calculated in the same way:

\[\int \sqrt{k^2 - x^2} dx = \left\{ \begin{array}{l}
  x = k \sin(y) \\
  dx = k \cos(y) dy
\end{array} \right\} = \int \sqrt{k^2 - k^2 \sin^2(y)} + k \cos(y) dy

= \int k^2 \cos^2(y) dy = \frac{k^2}{2} \int (\cos(2y) + 1) dy = \frac{k^2}{2} \left( \frac{\sin(2y)}{2} + y \right)\]

(A3)

Proceeding from the equations of circle and ellipse, limit of integration \(i\) (formula (A2)), and the integral calculation (formula (A3)), let’s consequentially compute the integrals in \(JI(r)\). Thus, the first integral in nominator will look like the following:

\[\int_0^i y_{\text{ellipse}}(x)dx = \frac{b}{a} \int_0^i \sqrt{rac{r^2 - x^2}{a^2 - r^2}} + \frac{b}{a} \sqrt{rac{r^2 - x^2}{a^2 - r^2}}\]

\[= \frac{ab}{2} \left( \frac{r^2}{a^2 - b^2} + a \sin \left( \frac{x}{2} \right) \right) \bigg|_{0}^{\sqrt{\frac{r^2 - b^2}{a^2 - b^2}}} \]

(A4)

The second one will be equal to

\[\int_0^i y_{\text{circle}}(x)dx = \frac{\pi r}{4} \left( \frac{r^2}{a^2} + a \sin \left( \frac{x}{2} \right) \right) \bigg|_{0}^{\sqrt{\frac{r^2 - b^2}{a^2 - b^2}}} = \frac{ab}{2} \left( \frac{r^2}{a^2 - b^2} + a \sin \left( \frac{x}{2} \right) \right)\]

(A5)
Box A1. Cont.

The nominator of \( J(I) \) will equal to \( (A4) \) + \( (A5) \):
\[
\int_0^i y_{ellipse}(x) dx + \int_i^a y_{circle}(x) dx = \frac{ab}{2} \sin \left( \frac{\sqrt{r^2 - b^2}}{a^2} \right) - \frac{a^2}{2} \sin \left( \frac{\sqrt{r^2 - a^2}}{b^2} \right) + \frac{\pi r^2}{4} \tag{A6}
\]

Actually, the denominator of \( J(I) \), representing the union of \( S_C \) and \( S_E \), is equal to the sum of the quarter of corresponding areas diminished by the intersection of \( S_C \) and \( S_E \), calculated in formula \( (A6) \):
\[
\int_0^i y_{circle}(x) dx + \int_i^a y_{ellipse}(x) dx = \frac{1}{4} (r \tau^2 + \tau ab) - \left( \frac{ab}{2} \sin \left( \frac{\sqrt{r^2 - b^2}}{a^2} \right) - \frac{a^2}{2} \sin \left( \frac{\sqrt{r^2 - a^2}}{b^2} \right) + \frac{\pi r^2}{4} \right) = \tag{A7}
\]
\[
\tau^2 \sin \left( \frac{\pi}{2} \sqrt{\frac{r^2 - b^2}{a^2}} \right) - \frac{ab}{2} \sin \left( \frac{\pi}{2} \sqrt{\frac{r^2 - a^2}{b^2}} \right) + \tau ab = \tag{A8}
\]

Derivative of \( J(I) \) will be equal to
\[
\frac{dJ}{dr} = z' (ab + r^2) - 2\pi r + \pi \tau ab = \tag{A9}
\]

To find the maximum of the function \( J(I) \), let’s put equal to zero its derivative and define \( r \):
\[
-2\tau ab \left( \sin \left( \frac{\pi}{2} \sqrt{\frac{r^2 - b^2}{a^2}} \right) + \sin \left( \frac{\pi}{2} \sqrt{\frac{r^2 - a^2}{b^2}} \right) \right) + \pi \tau ab = 0; \quad \sin \left( \frac{\pi}{2} \sqrt{\frac{r^2 - b^2}{a^2}} \right) + \sin \left( \frac{\pi}{2} \sqrt{\frac{r^2 - a^2}{b^2}} \right) = \frac{\pi}{2}; \quad \sin \left( \frac{\pi}{2} \sqrt{\frac{r^2 - b^2}{a^2}} \right) + \sin \left( \frac{\pi}{2} \sqrt{\frac{r^2 - a^2}{b^2}} \right) = \tag{A10}
\]
\[
\left( \sin (\alpha + \beta) = \sin(\alpha) \sqrt{1 - \sin^2(\beta)} + \sqrt{1 - \sin^2(\alpha) \sin(\beta)} \right) = \frac{\sqrt{r^2 - b^2} \sqrt{a^2 - r^2}}{\sqrt{a^2 - b^2}} = \sin \left( \frac{\pi}{2} \right) = 1 \quad \sqrt{r^2 - b^2} \sqrt{a^2 - r^2} = r(a - b) \quad r^2 - b^2 = r^2(a - b)^2 \quad r^2 = ab \quad r = a \sqrt{\frac{b}{a + b}} \]

B. Optimal Percentile Calculation

Optimal percentile \( p^* \) will equal to the share of area \( (2) \) (see figure above) of the area of ellipse:
\[
p^* = \frac{\int_i^a y_{ellipse}(x) dx - \int_i^a y_{circle}(x) dx}{\frac{\pi r^2}{4}} \tag{A11}
\]

Under defined optimal radius \( r = \sqrt{ab} \), limit of integration \( i \) is equal to:
\[
a \sqrt{\frac{r^2 - b^2}{a^2 - b^2}} = a \sqrt{\frac{b}{a + b}} \tag{A12}
\]
Thus, proceeding from the equations of circle and ellipse, limit of integration \( i \) (formula (A12)), and the integral resolution (formula (A3)), \( p^* \) will equal to:

\[
p^* = \frac{1}{\pi} \left( \left( \frac{b^2}{a^2} + \frac{a^2}{b^2} \right) \left( 1 - \left( \frac{a}{b} \right)^2 \right) + \sin^{-1} \left( \frac{a}{b} \right) \right)
- \left( \frac{b^2}{a^2} + \frac{a^2}{b^2} \right) \left( 1 - \left( \frac{a}{b} \right)^2 \right) + \sin^{-1} \left( \frac{a}{b} \right) \right)
\]

\[
= \frac{1}{\pi} \left( \left( \frac{a^2}{b^2} + \frac{b^2}{a^2} \right) + \sin^{-1} \left( \frac{b}{a} \right) \right)
- \left( \frac{a^2}{b^2} + \frac{b^2}{a^2} \right) + \sin^{-1} \left( \frac{b}{a} \right) \right)
\]

\[
= \left( 2 \pi \right) \left( \frac{a-b}{a+b} \right) \sin^{-1} \left( \frac{a-b}{a+b} \right)
\]

Finally, putting compactness \( c \) of the ellipse with axes \( a \) and \( b \) (\( a > b \)) to be equal to their ratio between the ellipse’s area and the area of the bonding circle \( c = \frac{Area \ of \ Ellipse}{Area \ of \ Bonding \ Circle} = \frac{\pi \alpha \beta}{\pi \alpha^2} = \frac{b}{a} \), optimal percentile \( p^* \) will be equal to

\[
p^* = \frac{1}{\pi} \arcsin \left( \frac{a-b}{a+b} \right) = \frac{1}{\pi} \arcsin \left( \frac{\sqrt{a^2 - b^2}}{a} \right) = \frac{1}{\pi} \arcsin \left( \frac{1-c}{1+c} \right)
\]

### References


11. Ng, E. Policies and technical guidelines for urban planning of high-density cities - air ventilation assessment (AVA) of Hong Kong. Build. Environ. 2009, 44, 1476–1488. [CrossRef]


54. Li, W.; Goodchild, M.F.; Church, R.L. An Efficient Measure of Compactness for 2D Shapes and its Application in Regionalization Problems. *Int. J. Geogr. Inf. Sci.* 2013, 27, 1227–1250. [CrossRef]


