



Article

# Automatic Positional Accuracy Assessment of Imagery Segmentation Processes: A Case Study

Juan J. Ruiz-Lendínez <sup>1,\*</sup>, Manuel A. Ureña-Cámara <sup>1</sup> , José L. Mesa-Mingorance <sup>1</sup>   
and Francisco J. Quesada-Real <sup>2</sup>

<sup>1</sup> Department of Cartographic, Geodetic Engineering and Photogrammetry, University of Jaén, 23071 Jaén, Spain; maurena@ujaen.es (M.A.U.-C.); jlmesa@ujaen.es (J.L.M.-M.)

<sup>2</sup> Department of Automatic, Electronics, Architecture and Computer Networks Engineering, University of Cádiz, 11519 Cádiz, Spain; franciscojose.quesada@uca.es

\* Correspondence: lendinez@ujaen.es; Tel.: +34-953212470

**Abstract:** There are many studies related to Imagery Segmentation (IS) in the field of Geographic Information (GI). However, none of them address the assessment of IS results from a positional perspective. In a field in which the positional aspect is critical, it seems reasonable to think that the quality associated with this aspect must be controlled. This paper presents an automatic positional accuracy assessment (PAA) method for assessing this quality component of the regions obtained by means of the application of a textural segmentation algorithm to a Very High Resolution (VHR) aerial image. This method is based on the comparison between the ideal segmentation and the computed segmentation by counting their differences. Therefore, it has the same conceptual principles as the automatic procedures used in the evaluation of the GI's positional accuracy. As in any PAA method, there are two key aspects related to the sample that were addressed: (i) its size—specifically, its influence on the uncertainty of the estimated accuracy values—and (ii) its categorization. Although the results obtained must be taken with caution, they made it clear that automatic PAA procedures, which are mainly applied to carry out the positional quality assessment of cartography, are valid for assessing the positional accuracy reached using other types of processes. Such is the case of the IS process presented in this study.

**Keywords:** automation; positional accuracy assessment; textural imagery segmentation; discrepancy methods



**Citation:** Ruiz-Lendínez, J.J.; Ureña-Cámara, M.A.; Mesa-Mingorance, J.L.; Quesada-Real, F.J. Automatic Positional Accuracy Assessment of Imagery Segmentation Processes: A Case Study. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 430. <https://doi.org/10.3390/ijgi10070430>

Academic Editors: Wolfgang Kainz and Amr Abd-Elrahman

Received: 7 May 2021  
Accepted: 22 June 2021  
Published: 23 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Imagery Segmentation (IS) is a longstanding problem in the computer vision field. According to Marfil et al. [1], there are three formal aspects which must be addressed regarding IS: its definition, its categorization and its main applications. IS could be defined as the process of dividing a certain image into regions—often referred to as regions of interest—which are homogeneous according to certain criteria [1]. These criteria can be very varied and are usually related to certain statistical properties of imagery, such as intensity value, tone, texture, etc. In addition, these regions must have a consistent meaning—meaningful structures—which helps one to simplify image processing tasks [2]. Therefore, it could be stated that the primary aim of IS is to reduce the information contained in an image in order to facilitate its processing and subsequent analysis. However, the scope of IS goes far beyond this basic purpose. According to Kesselman and Dickinson [3], IS can be used as an effective tool for identifying univocal correspondences between salient image features (pixels statistical properties, edges, etc.) and salient model features (objects). In this way, it might be possible to recognize and extract a novel example from a known set of object classes.

IS techniques are very diverse and can be categorized on the basis of three main criteria: method, format of representation and used statistical property. First, IS techniques can be categorized according to the segmentation method used. These methods, in turn,

can be classified into the following major categories [2]: (i) Threshold based segmentation, where pixels are allocated to categories according to the range of values in which a pixel lies [4]. (ii) Edge based segmentation, which is a method based on the detection of edges in an image. These edges represent object contours and are used to identify them [5]. (iii) Region based segmentation, where the pixels that are related to a certain object are aggregated for segmentation using similarity criteria [6]. (iv) Morphological methods, which verify the image with a small template—called a structuring element—which is applied to all possible locations of the input image and generates the same size output [7]. (v) Clustering techniques, where segmentation depends purely on the characteristics of the image and is performed through feature clustering. In summary, it must be noted that every segmentation method is adapted to specific requirements; therefore, it must be the user who decides the algorithm most suited to his own requirements [8].

A second criterion for categorizing IS techniques is the representation of the input image. Thus, instead of IS techniques based on a single representation—which characterizes some of the methods described above—there are other techniques based on hierarchical structures which are widely used in IS [9]. Such is the case of pyramids, or multiresolution pyramids. A pyramid segmentation algorithm is based on the idea of generating a stack of interrelated images with progressively decreased resolution. According to this idea, an image is described using levels [1]. Thus, the base of the pyramid represents the highest resolution level of an input image. The rest of levels are generated using a low-pass filter which reduces the spatial sample rate. This type of representation has important advantages such as the detection and representation of global features at lower-resolution levels thereby ignoring irrelevant details [10,11], the reduction of noise by deleting certain features in lower-resolution levels of the initial image, and the transformation of overall features to local features [12]. Finally, we must note that the use of pyramids significantly reduces computational cost—according to the principle divide and rule—and facilitates IS tasks [10,11].

Although the above categorization criteria can be further divided into other subgroups such as regular and irregular pyramid structures—in the case of the representation of the input image—, we consider it important to establish a third criterion for categorizing IS techniques: the statistical properties of the image—mainly tone and texture—used to segment it. Usually, the IS techniques based on the pixel tone segment the original image by applying clustering methods to its greyscale histogram. In this way, the resulting regions meet the main IS requirements; they are different from a visual perspective and have very similar statistical properties. However, real imagery regions do not share this homogeneous statistical behavior [13]. In agreement with this, the use of the pixel tone is not recommended. In view of this fact, many approaches based on alternative statistical properties have been developed, including texture. Textural IS is based on the use of local texture descriptors. Among these descriptors we can highlight the Local Binary Pattern (LBP). The LBP is a very efficient texture operator due to its discriminative power and, for this reason, it has become the most widely employed operator in textural IS [13–16].

After defining and categorizing IS, the third formal aspect—described by [1]—to address is its applications. Normally, IS is represented by its use for pattern recognition which are applied to multiple areas such as: medicine, biometrics, bioinformatics, facial recognition, etc. With regard to the specific field of Geographic Information (GI), IS has been primarily applied to the recognition and subsequent extraction of spatial objects from remote sensing images—both aerial and satellite imagery. In this sense, we must highlight the studies related to (i) automatic extraction of building features from VHR satellite imagery [17–25], (ii) automatic road network extraction from VHR aerial imagery [26–29] and (iii) vector-imagery conflation procedures based on the use of road intersections—previously extracted—as check points in order to develop matching procedures [14,16,30]. More recently, the applications of IS to GI have been aimed at obtaining the land information required to achieve sustainable urban development, avoiding problems caused by the unorganized urban sprawl [13,31,32]. Such is the case of the approaches of Stryjakiewicz

et al. [32] and Ruiz-Lendínez [13]. Both studies are focused on the municipal district of Poznań (Poland) and assess the degree of abandonment of agricultural farmland. Nevertheless, while in the first case we use classification techniques applied to VHR satellite imagery—Sentinel—in the second one a framework for automatically locating and mapping abandoned farmland using IS techniques based on the textural segmentation of VHR aerial imagery is developed.

However, despite the existence of all these applications in the specific field of GI, there is a fourth formal aspect associated with IS which has still not been addressed in any of them: the assessment of IS results from a positional perspective. Thus, in a field in which positional aspect is critical, it seems reasonable to think that the quality associated with this aspect must be controlled. At this point, the following research question comes up: what does good segmentation mean? There is no simple answer because, among other things, it is difficult to define what a “good” segmentation is [10]. If we adjust for the IS formal definition, a good segmentation would be one which provides regions homogeneous in strict accordance with a given criterion. However, there are many sources of noise that can affect the resulting positional quality in an IS process, making it, therefore, necessary to carry out a measure of this parameter.

Traditionally, the assessment of the segmentation results—and therefore the analysis of the efficiency of a certain IS algorithm—has been based on the decision of a human expert. Some authors call this using qualitative methods [1]. These methods depend on human intuition and, although they may become very useful for evaluating some characteristics of IS algorithms, they do not seem to be the best option when assessing the positional accuracy of the segmentation results in processes involving GI. The main reason is that different human experts may assign different importance levels to different sets of geographic objects. That is why new empirical methods are required. These procedures use test images to assess the IS algorithms. Specifically, they apply a certain IS algorithm to a test image and assess the accuracy of segmentation results from a numerical perspective using measures such as functions and indices. Among them, it is necessary to highlight the Jaccard index [33], the shift variance (SV) [34], the F function [35] and the Q function [36]. Empirical methods can be classified into two types: goodness and discrepancy methods. Goodness methods define the IS process in terms of how well it segmented objects of interest. For this, they use goodness parameters which measure some desirable properties of segmented images. Against them, discrepancy methods assess the IS process comparing the segmentation obtained with a certain IS algorithm with the ideal segmentation by counting their differences. Overall, the assessment of the segmentation results provided by most of the empirical measures—mainly indices—shows a high efficiency, providing results extremely effective when using metrics very straightforward—such in the case of the Jaccard index [33]. However, despite this fact there are many aspects related with GI in which—as will be discussed in our study—the functioning of these metrics is not optimal. These aspects are closely linked to the nature of GI itself and the geographical objects that are on it. We are referring for example to those aspects related with the spatial definition of buildings, city blocks—in urban environments—and agricultural plots—in rural environments. These geographical objects are defined by the shape of their boundaries which means that in order to assess their geometry—after being segmented and extracted by an IS algorithm—it is necessary to assess the geometry of their boundaries. In this sense, our methodology shows notable advantages over other metrics—such as indices—in IS processes involving GI.

In this context, the present study has as a starting point our previous work [13] and its main objective is to develop an automatic positional accuracy assessment (PAA) method for assessing—from an empirical perspective—the positional quality of the regions obtained by means of the application of an IS procedure. Specifically, our case of study was focused on segmented regions from the application of a textural segmentation algorithm to a VHR aerial image. Our segmentation accuracy measure is based on the procedure originally proposed by Ruiz-Lendínez et al. [37,38] and which is outlined in the following

sections. However, although this measure has required an adaptation of methods and procedures previously developed, we consider that it is much more than a simple procedural adaptation of existing methods because it represents the adoption and application of standards and measures, traditionally—and successfully—used in the field of cartographic production, to the specific field of IS. This, together with the advantages over other metrics, represents the true novelty of the present study.

In addition, the efficacy of this measure was also tested by means of its application to the segmented regions obtained from the VHR satellite imagery for the same geographic area. On this basis, this study demonstrates not only that this measure is applicable to any segmentation process, but it may also be helpful in order to select the segmentation method most suited to certain requirements.

## 2. Methodology and Data

### 2.1. Research Outline

In our previous study [13] we implemented an IS algorithm for extracting regions that belong to abandoned farmland employing a non-parametric approach to the texture characterization of a VHR aerial image—originally developed in [39]. Although it is described in detail in [13], here we briefly describe its main steps:

1. Texture characterisation. Texture is locally sampled by the joint distribution of two properties associated with pixels: (i) Local Binary Pattern (LBP) and (ii) a contrast measure (C). LBP is a simple yet very efficient local texture operator capable of characterizing small texture regions [39]. Figure 1a shows the procedure for computing LBP values. After comparing each central pixel with its eight neighbors, the  $3 \times 3$  neighborhood generated must be thresholded by the value of the centre pixel. Thus, the neighbors having a smaller value than the central pixel have bit 0, and the other neighbors having a value equal to or greater than that of the central pixel have bit 1. Then, these binary values of the pixels in the thresholded neighborhood must be multiplied by the weights given to the corresponding pixels. Finally, the values of the eight pixels are added to obtain a number for this neighborhood. With regard to the employed C, its compute is also addressed in Figure 1a.

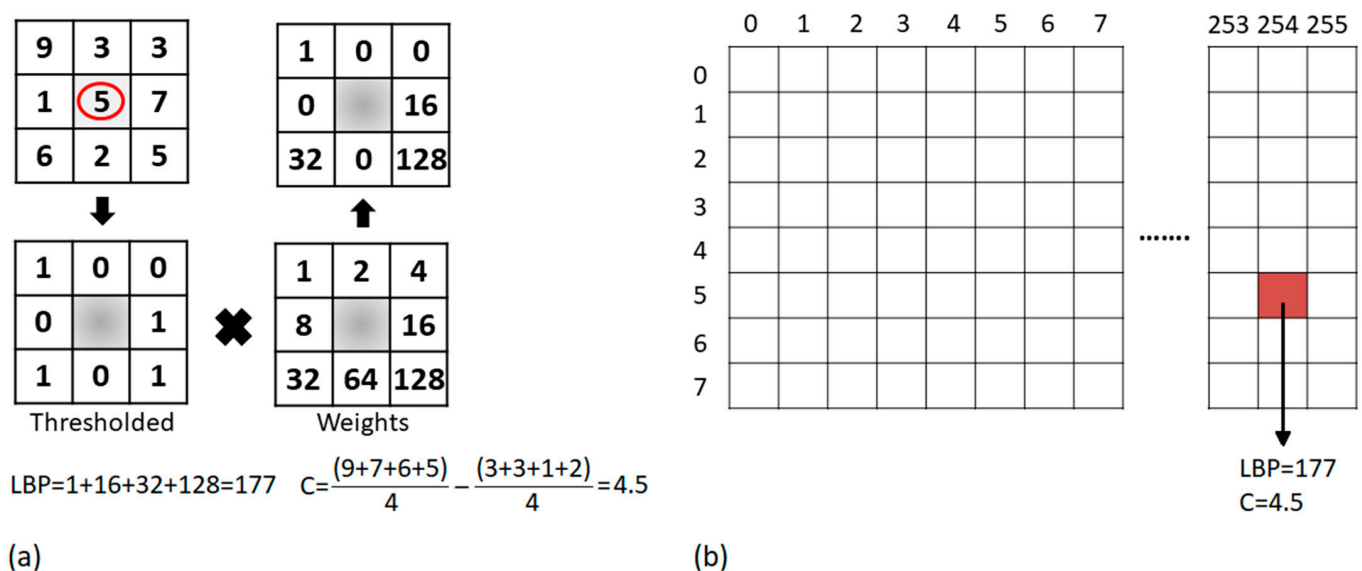


Figure 1. (a) Computation of LBP and C; (b) bidimensional array LBP/C.

Texture is not an image property that may be associated to a single pixel. That is why the image was decomposed into a grid where each of its cells included a fixed number of pixels. In addition, to extract the abandoned farmland regions from the image, the system was trained on a small area of it to learn the texture parameters of LBP/C belonging to

this type of land use. The LBP/C distribution for each cell of the grid was approximated by a discrete two-dimensional histogram—an array—of  $256 \times b$  pixels, where  $b$  is the number of bins for  $C$  (Figure 1b). This number of bins is chosen as a trade-off between the discriminative power and the stability of the texture description [39]. In order to identify two textures as equal or different, LBP/C distributions between two histograms ( $A$  and  $B$ ) were compared using the following likelihood-ratio  $G$  statistic:

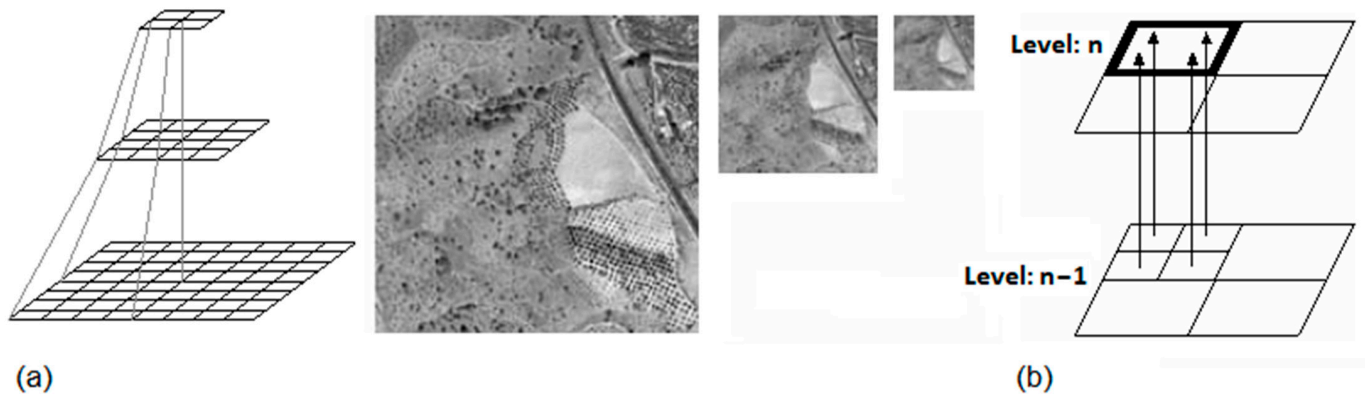
$$G = 2 \left( \begin{array}{l} \left[ \sum_{A,B} \sum_{i=1}^N f_i \log f_i \right] - \left[ \sum_{A,B} \left( \sum_{i=1}^N f_i \right) \log \left( \sum_{i=1}^N f_i \right) \right] - \\ \left[ \sum_{i=1}^N \left( \sum_{A,B} f_i \right) \log \left( \sum_{A,B} f_i \right) \right] + \left[ \left( \sum_{A,B} \sum_{i=1}^N f_i \right) \log \left( \sum_{A,B} \sum_{i=1}^N f_i \right) \right] \end{array} \right) \quad (1)$$

where  $N$  is the number of bins and  $f_i$  is the frequency at bin  $i$ .

2. Hierarchical structure generation. A pyramidal representation was adapted to texture segmentation (Figure 2). A pyramidal architecture is completely defined if we specify how a new level is built (Figure 2a) and how a parent is linked to its children (Figure 2b). In our case, the LBP/C distribution of the grid cells formed the base of this pyramidal structure, and each level  $l$  of the pyramid was a reduced map with one-fourth of the cells of the level immediately below. Each pyramid cell, denoted by  $(x, y, l)$ , had the following parameters associated with it:
  - Homogeneity.  $H(x, y, l)$  ranged from 1 (if the four cells immediately underneath had the same texture) to 0. The setting of  $H$  was based on a uniformity test. Thus, the four cells had the same texture if a measure of relative dissimilarity within that region was lower than a certain threshold  $U$ ,  $(G_{\max}/G_{\min} < U)$ .  $U$  must be set in such a way as to ensure the detection and differentiation of textures. For this reason, it is advisable to choose a small value close to one for this threshold.
  - Texture. If the cell was homogeneous,  $T(x, y, l)$  was equal to the sum of the LBP/C distributions of the four cells immediately underneath. Otherwise, it was set to a fixed value  $T_{NH}$ .
  - Parent link.  $(X, Y)_{(x,y,l)}$ . If  $H(x, y, l)$  was equal to 1, the values of the parent links of the four cells immediately underneath were set to  $(x, y)$ . Otherwise, these four parent links were set to null.
  - Centroid.  $C(x, y, l)$ . The centre of mass of the base region associated with  $(x, y, l)$ .
  - Histogram. Each parent link stored a two-dimensional histogram which characterised the texture of the image region represented by this node. After completing the hierarchical structure generation (step number 2), all the cells belonging to this structure with a homogeneity value equal to 1 and that had no parent were linked to homogeneous regions at the base, defining initial image segmentation.
3. Growth of homogeneous cells. After completing the pyramidal representation, all the cells that presented a homogeneity value equal to 1 and had no parent were linked to homogeneous regions at the base. The growth of these regions was carried out by means of a basic process: The algorithm linked cells whose parent link values were null. Thus, a cell  $(x,y,l)$  was linked to the parent of neighbours  $(xp, yp, l+1)$  when two cells had the same texture.
4. Fusion of homogeneous cells. The neighbouring cells,  $(x1, y1, l)$  and  $(x2, y2, l)$ , were merged if the following four conditions were true:
  - $(X, Y)_{(x1,y1,l)} = \text{null}$ . Therefore, the cell had no parent.
  - $(X, Y)_{(x2,y2,l)} = \text{null}$ . Therefore, the cell had no parent.
  - The cells had a homogeneous texture.  $H_{(x1,y1,l)} = 1$  and  $H_{(x2,y2,l)} = 1$ .
  - The cells had the same texture.



5. Pixel-wise. In order to soften the segmented image, in a post-processing step, the resolution of all blocks in the texture region boundaries was increased until those boundaries were one pixel wide.
6. Abandoned farmland extraction. To identify and separate this type of zones from the rest, we trained the system to learn its texture parameters (LBP, C). Thus, a set of thresholds for these texture parameters was generated.

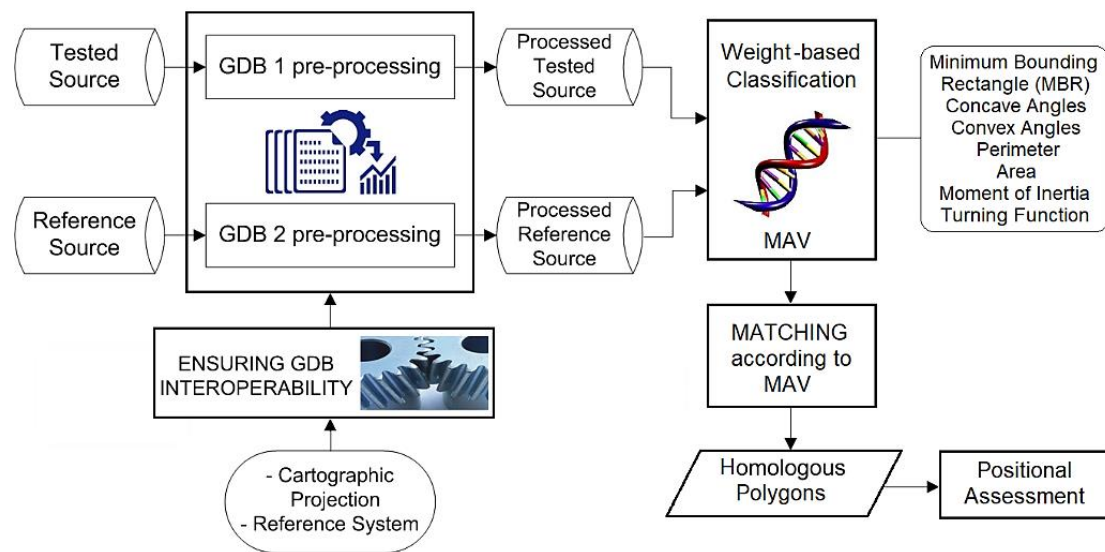


**Figure 2.** Hierarchical structure. (a) Set of pyramidal structures from texture; (b) parent link.

According to the above description and taking into account the three levels of categorization described in the previous section—method, format of representation and statistical property used—our algorithm could be classified as a region-based textural segmentation algorithm which uses a pyramidal structure for describing the contents of the image using multiple representations. This algorithm, which had already been successfully tested in many conflation processes related with the improvement of the alignment between vector data and imagery, allowed us to quantify and locate the abandoned farmland in a heterogeneous urban landscape—Municipality of Poznań—, identifying those areas where investment and speculation processes were causing the abandonment of the best agricultural soils. In that previous study [13], a proposal for assessing the effectiveness of our segmentation procedure was already outlined. However, many issues remained outstanding. In the present study, efforts were made to further improve our approach, addressing in detail some of those issues.

## 2.2. Automatic Discrepancy Method for the PAA of Imagery Segmentation

There is a large number of methods to assess IS segmentation processes. All these methods are focused on different scenarios and can be classified according to different criteria. One of the most accepted classifications is proposed in [1]. In accordance with it, the IS assessment method developed in the present study could be classified as a discrepancy method. Discrepancy methods can be categorized as quantitative and empirical methods, that is to say, procedures based on numerical data which measure explicitly the quality of segmentation results. These types of methods are the most suitable ones for carrying out a formal analysis of GI positional aspects because they are based on the comparison between the ideal segmentation and the computed segmentation by counting their differences [1]. This comparison process constitutes the cornerstone of the functioning of the automatic method developed for assessing the positional accuracy of GI in [37,38]. Although all the details of this methodological approach are extensively addressed in the above-mentioned studies, it is important to recall some of its main characteristics in order to help the reader to understand how this works. As shown in Figure 3, in these studies, the PAA of GI is computed on the basis of the spatial discrepancies automatically obtained from the locations of homologous graphic entities—polygons—stored in two different sources of data—named tested or assessed data and reference data—that must be interoperable from a geometric perspective. In addition, it is important to note that the second dataset must have a higher accuracy than the first one.

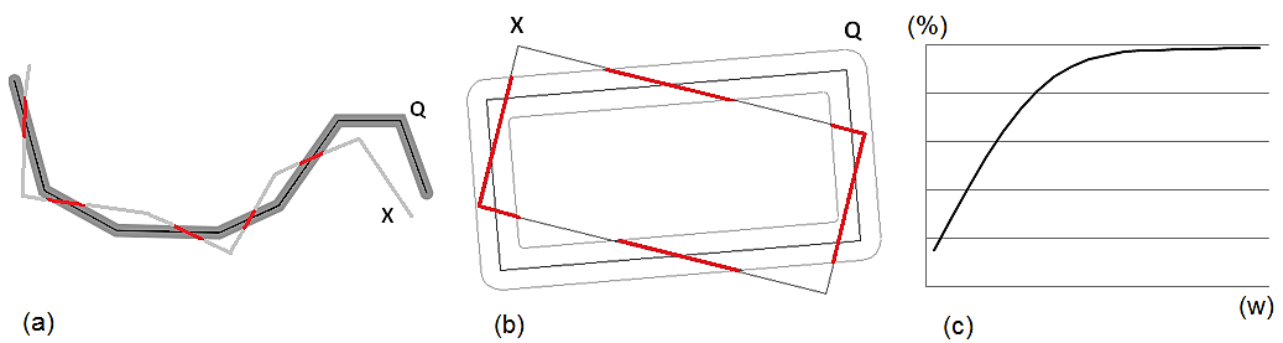


**Figure 3.** Flowchart of the methodological approach developed in [37,38].

Under this perspective of the PAA processes, matching mechanisms between both data sources becomes a key aspect. To that end, in [33,34] the authors proposed an innovative matching mechanism that determined homologous building polygons using a weighted combination of their geometric descriptors. The weights used were computed from a supervised training process using a Real-Coded Genetic Algorithm (RCGA) and a training database composed of 805 pairs of polygons manually matched. (The computing process is an iterative process which ends when sufficiently good quality matches between the polygons are produced. The quality is assessed by a function called the Fitness Function). The geometric descriptors employed quantified the absolute location of the polygons by means of their Minimum Bounding Rectangle (MBR), their geometric properties—perimeter, area—and their shape—moment of inertia and the area of the region below its turning function. In addition, the use of a RCGA allowed to perform the categorization of the matching quality from a quantification of the similarity between polygons by means of a Match Accuracy Value (MAV) ranging from zero to one. Thus, if we consider two matched polygons ( $P_A, P_B$ ), this value is obtained as a linear combination of the  $n$  polygon descriptors—or attributes—( $At(P_A), At(P_B)$ ) and the  $n$  weights ( $W(At)$ )—computed in the training phase—as follows:

$$MAV(P_A, P_B) = \text{MAX} \sum_{i=1}^n \left( \frac{At_i(P_A)}{At_i(P_B)} \cdot W_{At_i} \right), P_A \in \text{Tested Source}, P_B \in \text{Ref. Source} \quad (2)$$

This MAV also allowed to (i) select only 1:1 corresponding polygon pairs among all the possible correspondences (using a thresholded value of MAV), thus, avoiding the acceptance of both erroneously-matched polygons and unpaired polygons, and (ii) set different similarity levels between 1 and 1 corresponding polygon pairs. Finally, having matched pairs of homologous polygons according to the MAV, the procedure used for measuring the positional discrepancies between both sources of data was the single buffer overlay method (SBOM) [40] (Figure 4).



**Figure 4.** (a) Buffer generation by means of the SBOM. (b) Adaptation to line-closed case and (c) empirical probabilistic distribution function.

Originally developed for linear entities, this metric had to be adapted to the line-closed case—polygons (Figure 4b). Thus, for a given pair of lines (X, Q), a buffer of increasing width is generated around the reference boundary line Q in order to compute the percentage of the tested boundary line X (Figure 4a). In the line-closed case, this percentage represents the percentage of the Polygon Perimeter (PP). After this, one can determine the widths which include several percentages of length of the tested element, and represent this by a probability distribution curve (Figure 4c). Finally, this procedure can be repeated with polygons from a sample or with polygons from a complete Geospatial Database (GDB) obtaining an aggregate distribution function of the uncertainty for several levels of confidence.

After having recalled this methodological approach, in this study we focus on applying and adapting it—hereafter this adaptation will be called the automatic discrepancy method (ADM)—to the specific case of regions (polygons which represent rural plots) obtained by means of the application of a textural segmentation algorithm to a VHR aerial image. In general, and particularly in our case, segmented regions computed from a segmentation algorithm would play the role of the tested or assessed data, while the ideal segmented regions would play the role of reference data.

The main innovation of our ADM lies:

1. In the size of the segmented regions population. This population is very large in comparison to the size of the population obtained both with other assessment methods of segmentation results and traditional PAA methods.
2. In the specific characteristics of automatic PAA procedures. The automation of the process implies both a significant cost reduction and a low computational time compared to traditional methodologies (especially if we consider field work).
3. In the fact that our ADM does not depend on human intuition to decide the positional accuracy of a certain segmentation algorithm—as happens in the case of qualitative methods (see [1]).
4. In the fact that the distribution functions provided by our ADM behave as signatures that unequivocally identify the accuracy when two segmented regions are compared. This constitutes a notable advantage over other metrics, such as indices, in IS processes involving GI.

In compensation, our ADM presents the drawback that having a previous ideal segmentation—reference data—is necessary [13]. In addition, the determination of the reference data is not the only problem, there are other critical aspects on which the success of our ADM depends. Thus, and from a statistical and practical perspective, the most controversial aspect of any PAA method is the number, location and categorization of control elements [41], that is to say, the main aspects related to the sample: its size—specifically, its influence on the uncertainty of the estimated accuracy values—its distribution and its categorization. In the specific case of our ADM, it only makes sense to address the first and last of the aforementioned aspects because the sample distribution depends on the spatial distribution of the phenomenon studied—in our case, for example, the sample distribution



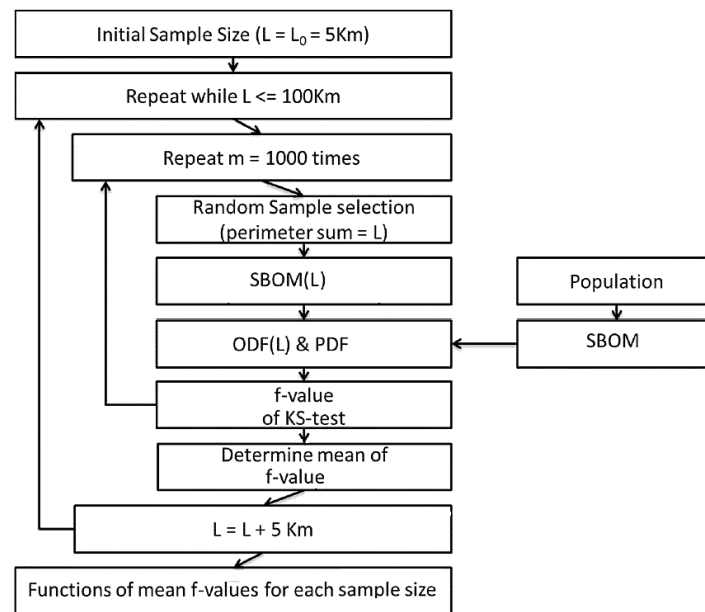
was conditioned to the location of abandoned farmland. Therefore, it is necessary both to establish specific criteria in order to define sample size and to carry out a categorization of the positional accuracy results taking into account characteristic parameters of the segmented regions such as their perimeter length, their number of vertexes, etc.

### 2.2.1. Size of the Sample

In any study where PAA methods are employed it is important to establish specific criteria as well as precise recommendations in order to determine sample size because this parameter might influence the uncertainty of the estimated values. In this sense, the first thing we must perform is to define what sample size means in our case. Thus, in a study whose main goal is to develop a PAA method for assessing the positional accuracy of segmented regions, when we talk about sample size, we are referring to the minimum number of regions (polygons) necessary to make the accuracy results statistically meaningful. As addressed below, this number is determined by means of the length of these regions' perimeter. That is to say, the minimum number of regions employed in order to assess the positional accuracy of a certain IS process is determined by the sum of their perimeters. For example, if the sample size required is 10,000 m and the average perimeter length per region (polygon) is 500 m, 20 regions would be required. With this in mind, the next step must be to adopt a standard which allows us to define an adequate sample size.

Unlike the point-based PAA methods, there are no specific standards concerning adequate sample size related to the use of the line-based PAA methods—as in the case of the SBOM adaptation used by our ADM. Thus, most authors who use them provide the sample size employed in their studies without giving any explanation regarding how that parameter was determined. Such is the case of [40,42,43]. In order to fill this gap, Ariza et al. [41] studied the influence of sample size in terms of uncertainty when automatically estimating the planimetric positional accuracy of urban GDBs by means of these types of methods (buffer methods such as the SBOM).

As shown in Figure 5, this approach was based on a simulation process which consisted of the simulation of samples—randomly extracted from an initial population of homologous urban polygons—of different size  $L$ . The first step was to apply the SBOM to the initial population, thus, computing the population distribution function (PDF). This PDF, which was expressed as a probabilistic distribution function as shown in Figure 2c, was subsequently compared with the distribution function derived from each size  $L$ , functions that were called Observed Distribution Functions (ODFs). In the case proposed in [41],  $L$  ranged from 5 to 100 km with  $\Delta L = 5$  km and the parameter  $L$  used for determining the different sample sizes was the length of the polygons' perimeters, measured on the polygons from the reference source. In addition, for each sample size  $L$ , the simulation was iterated  $m = 1000$  times reproducing in each of them the positional accuracy estimation procedure (SBOM). Having obtained these distribution functions (the PDF and the ODFs), the next step was to compare the similarity between them in order to compute the variability of the estimated planimetric accuracy of the tested GDB. Thus, each of the ODFs was compared with the PDF. This similarity was addressed by means of statistical tests of significance like the Kolmogorov–Smirnov test [44,45]. In this test, the null hypothesis ( $H_0$ ) established that the empirical distribution is similar to the theoretical one. In other words, the null hypothesis established that the observed frequency distribution is consistent with the theoretical distribution (and therefore a good fit). In contrast, the alternative hypothesis ( $H_1$ ) established that the observed frequency distribution is not consistent with the theoretical distribution (poor fit). The interpretation of the KS test may be performed by means of two statistical indicators: an  $f$ -value and a  $p$ -value.



**Figure 5.** Flowchart of the approach developed in [41].

Specifically, the  $f$ -value ranges in the interval  $(0, 1)$  and represents the maximum distance  $D$  between two distribution functions.  $D$  is computed as follows:

$$D = \sup |F_n(x_i) - F_0(x_i)| \quad (3)$$

where  $x_i$  is the value observed in the sample (whose values were previously ordered from lowest to highest),  $F_n(x_i)$  is an estimator of the probability of observing values less than or equal to  $x_i$ , and  $F_0(x_i)$  is the probability of observing lower values or equal to  $x_i$  when  $(H_0)$  is true. With regard to its interpretation, large  $f$ -values imply large discrepancies between the curves representing the distribution functions, while small  $f$ -values imply small discrepancies between the curves that represent the distribution functions.

On the other hand, the  $p$ -value is a probabilistic measure which expresses the level of confidence on the  $f$ -value. It is computed as follows:

$$p - \text{value} = P(D > D_{\text{obs}} / H_0 \text{ is true}) \quad (4)$$

Now, if the  $p$ -value is large, it means that—assuming that  $(H_0)$  is true—the observed value of the  $D$  statistic was expected. Therefore, there is no reason to reject this hypothesis.

In addition, the graphical representation of these results allows us to determine a sample size that assures, with a given level of probability, that the maximum discrepancy between the sample (the ODFs) and the population (the PDF) is a certain  $f$ -value.

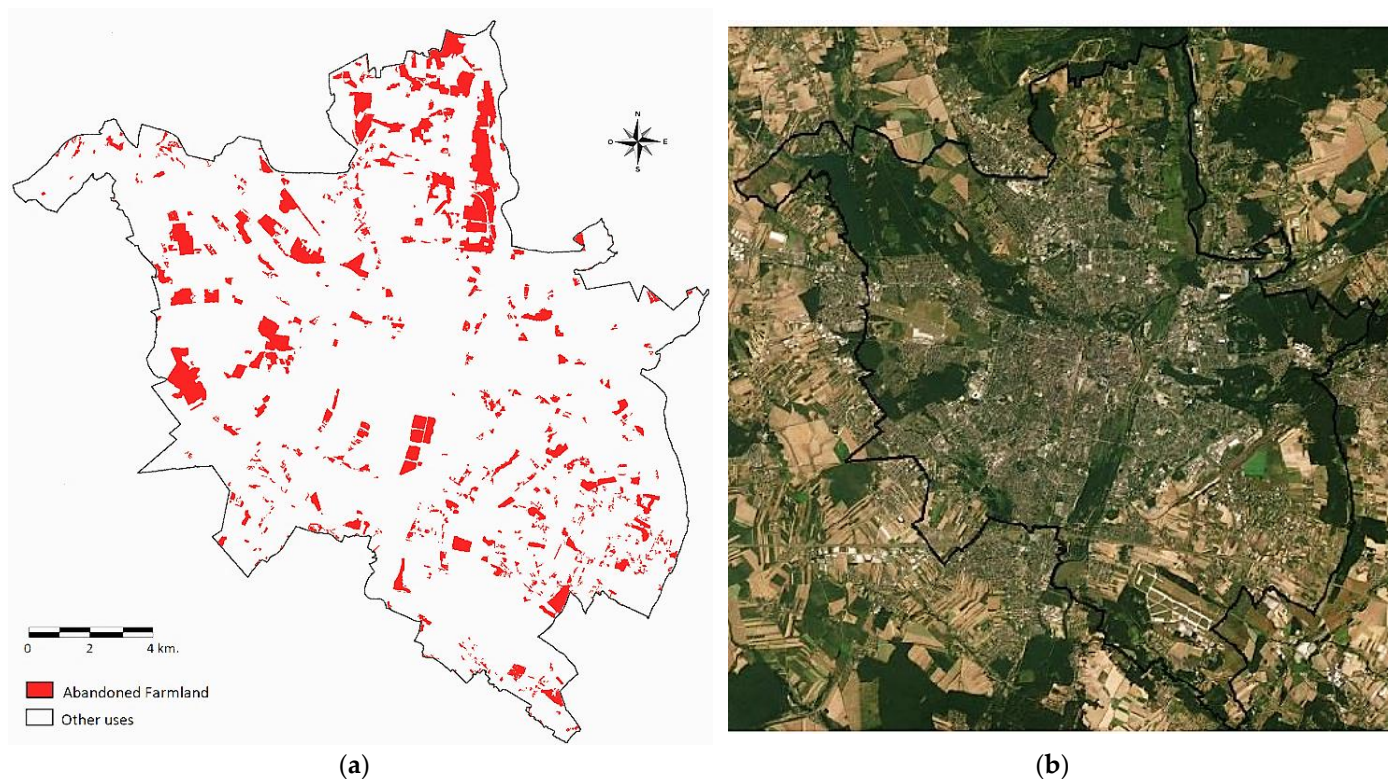
Finally, and as concluded in [41], the pattern of behaviour shown by the results obtained from their study could be extrapolated to other cases, scales and products adapting the simulation process presented here to the characteristics of those products. Therefore, and following this last recommendation, we have applied the simulation process described above in order to give some guidance on the influence of sample size on the results of our ADM when assessing the positional quality achieved in the IS process. The adaptation of the simulation process has consisted mainly in (i) the use of the segmented regions' perimeter (measured on the elements belonging to the reference source) as the parameter  $L$  used for determining the different sample sizes (it was set up according to the following values: from 0.5 km to 20 km with  $\Delta L = 1$  km) and (ii) the change in the number of iterations. In our case  $m = 500$ .

### 2.2.2. Categorization of the Sample

In order to achieve a better understanding of our ADM, it was necessary to carry out an analysis of the positional accuracy results obtained according to a categorization of the sample taking into account parameters such as the perimeter of the segmented regions and the number of vertexes. As in the sample size case, these parameters were measured on the elements belonging to the reference source. Taking into account that the SBOM results are expressed as distribution functions, this categorization must be carried out by means of the comparison of such functions; that is to say, the PDF and the ODFs obtained by classifying the samples according to the mentioned parameters. To that end, we again used the Kolmogorov–Smirnov test [44,45]. In this case, the  $p$ -value was also employed. As mentioned above,  $p$ -values represent a probabilistic measure of  $f$ -values and, therefore, both parameters are closely related. Thus, the closer the  $p$ -value is to the unit, the greater the level of confidence on the corresponding  $f$ -value will be. Therefore, if we apply the Kolmogorov–Smirnov test to two distribution functions with a great similarity between them, the  $f$ -value and  $p$ -value obtained for the pair may be 0.1 and 0.95. [37]. Finally, and in order to carry out the corresponding statistical calculations, the procedure described by Gibbons and Chakraborti [46] was employed, and the  $p$ -value was approximated numerically using the method outlined by Press et al. [47].

### 2.3. Tested and Reference Data

As mentioned above, and although our ADM was also tested by means of its application to other segmentation option, this study has as starting data—tested data—the results obtained from our previous work [13]. Figure 6 shows both of these results (map of abandoned farmland belonging to the municipal district of Poznań (Poland)) and the VHR image from which they were obtained by means of a textural IS procedure.



**Figure 6.** (a) Map of abandoned land as derived from textural segmentation of the aerial image. (b) High-resolution aerial image. Sources: [13,48].

It must be noted that all the aspects related to this procedure are addressed in detail in [13], so they will be not discussed here. There are, however, some aspects referring to the characteristics and specifications of the image which are important to remember for

the development of the present study, for example its origin and resolution. In this sense, we used a high-resolution aerial image provided by World\_Imagery (MapServer) [48]. World\_Imagery provides several products, including satellite and aerial imagery (with resolutions ranging from 0.5 m/pixel to 1 m/pixel). In our case, the resolution—or ground sampling distance (GSD)—of the image selected was 1 m/pixel.

With regard to the ideal segmentation data, this was composed of cadastral data provided by the Poviát Geodetic and Cartographic Documentation Center in Poznań [49]. This center carries out tasks in the field of geodesy and cartography such as creating, maintaining and sharing databases of land and building records (real estate cadastre) or the registering of prices and real estate values. In addition, it is responsible for creating, maintaining and sharing cartography at scales of 1:500, 1:1000, 1:2000 and 1:5000 on the basis of the data contained in the abovementioned databases (cadastral maps and basic maps). Therefore, the cadastral data used as reference data were represented by a sample of cadastral plots (test plots) whose geometry was known with high accuracy (see [13]).

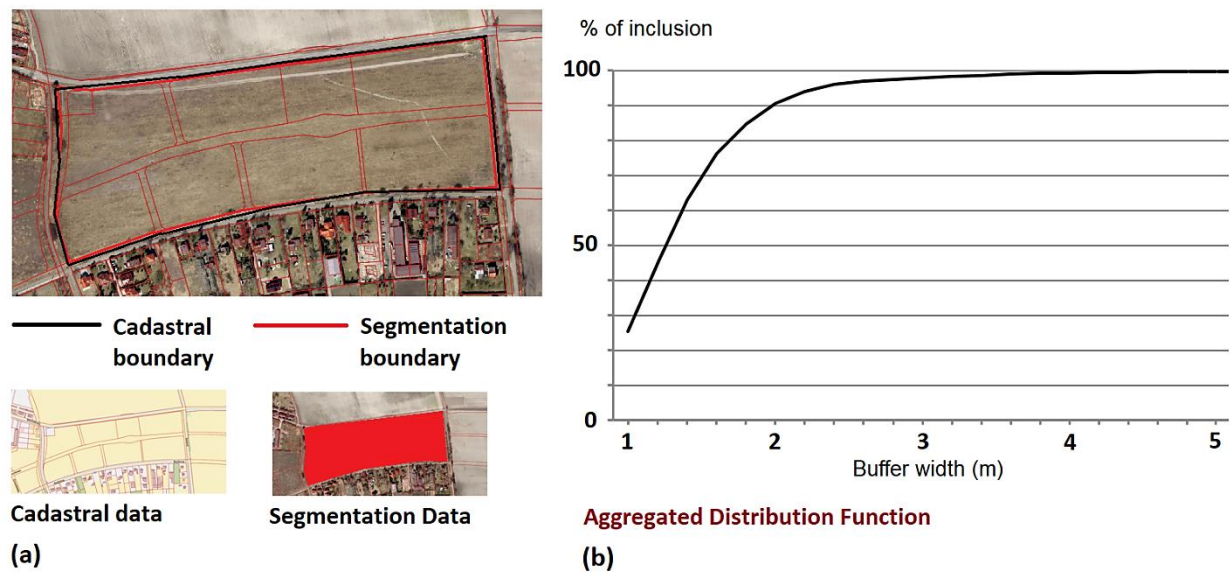
### 3. Results

#### 3.1. PAA of Segmentation Results

Using the population of homologous plots automatically paired by means of the matching procedure outlined in Section 2.2 and the weighted combination of their geometric descriptors described on it, we computed the PDF. Once again, it is necessary to emphasize that this procedure is detailed in [33,34] and that the weights used were the same as those obtained from the supervised training process of the RCGA addressed on it. The reason for this was that in both cases the typology of used polygonal shapes (buildings in the case of [33,34] and agricultural plots in the case of the present study) presents very similar characteristics with regard to certain geometric properties (particularly as regards to, for example, approximate length of the perimeter and area) used by the RCGA in order to carry out the classification.

The computed PDF assesses the positional accuracy of the segmentation plots obtained with our IS algorithm and represents the efficiency—from a positional perspective—with which this algorithm is able to segment an image. Figure 7b presents the resulting PDF by applying the SBOM on our two datasets, determining the percentage of the boundary lines from the regions provided by our IS algorithm that is within the buffer generated on the plot boundary lines from the cadastral data (Figure 7a). With regard to the parameters used and the values achieved, we used buffers with widths from 1 to 5 m, and the final reached uncertainty was 2.4 m for a 95% level of confidence. There are no references in the literature related to specific criteria for the selection of these values (buffer widths). It can be stated that their selection depends on the resolutions of both tested and reference data. Thus, according to the resolution of the image selected—1 m/pixel—and the accuracy with which the geometry of cadastral plots was determined, the values chosen were more than enough to efficiently represent the accuracy of the tested data by means of the corresponding aggregated distribution function. In this regard, we must also note that the choice of values different to those used in this case would not change the final result. In this sense, what really matters is to always employ the same values when two distribution functions are compared. Something similar occurs with the reached uncertainty. Logically, this value represents the accuracy of the IS process and is directly linked to the accuracy levels with which both tested and reference data were produced. Therefore, the only way to improve the accuracy of the obtained uncertainty is to carry out a more precise setting of the parameters that control the functioning of the IS algorithm by means of (i) the improvement of the training process whose main goal is to learn the texture parameters LBP, C belonging to abandoned farmland regions, and (ii) the improvement of the setting of U (Section 2.1) in order to ensure the correct detection and differentiation of textures. Finally, Table 1 summarizes the data employed for each source.





**Figure 7.** (a) Single buffer overlay method (SBOM) adapted to the line-closed case: the black line represents the cadastral boundary (ideal segmentation or source of greater accuracy), and the red line represents the segmentation boundary (tested source). (b) Population distribution function (PDF) obtained.

**Table 1.** Data employed for each source.

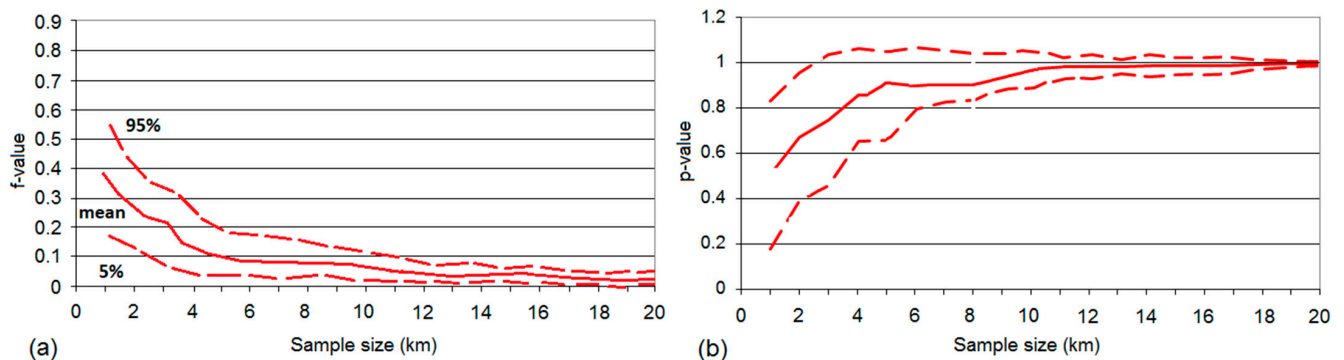
Polygons Matched	Number of Regions/Number of Polygons		Total Number of Vertexes		Total Perimeter (m)	
	Segmented Image	Cadastral Data	Segmented Image	Cadastral Data	Segmented Image	Cadastral Data
376	391	376	-	4021	-	187,785

### 3.2. Size of the Sample

As mentioned in Section 2.2.1, and in order to determine the influence of sample size on the variability of the results of our ADM when assessing the positional quality achieved in the segmentation process, we replicated the experiment developed in [37] in accordance with the procedure outlined above. In addition, let us remember that (i) this variability was computed by comparing the similarities between each ODF—obtained from each sample of length  $L$ —and the PDF described in the above section (ii)  $L$  was set up according to the following values: from 0.5 km to 20 km with  $\Delta L = 1$  km, and (iii) the number of iterations ( $m$ ) carried out was 500—reproducing in each of them the positional accuracy estimation procedure (SBOM). With regard to the results, we must note that they are presented by means of two different graphical representations which show the frequency distance  $f$ -value (Figure 8a) and its associated  $p$ -value (Figure 8b) obtained for each sample size. In addition, as in the abovementioned study, along with the curve which represents the mean value of these parameters the curves which correspond to the 5% and 95% percentiles were also represented.

With regard to the shape of the curves obtained—mean  $f$ -values, mean  $p$ -values, and their associated 5% and 95% percentile curves—it was as one would expect from this type of estimation process:  $f$ -values decrease and  $p$ -values increase when sample size  $L$  increases. In addition, the variability between  $f$ -values of the 5% and 95% percentiles decreases when  $L$  increases.



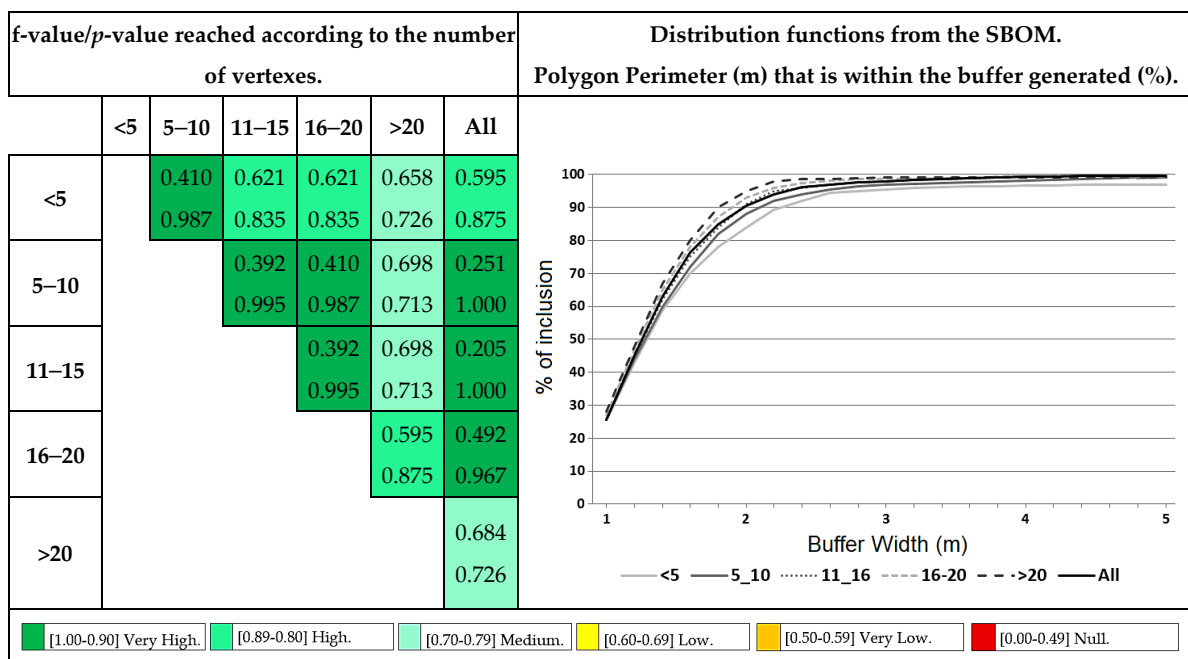


**Figure 8.** (a) Statistical  $f$ -value parameter for the Kolmogorov–Smirnov test as a function of sample size (km) for the SBOM. (b) Statistical  $p$ -value parameter for the Kolmogorov–Smirnov test as a function of sample size (km) for the SBOM.

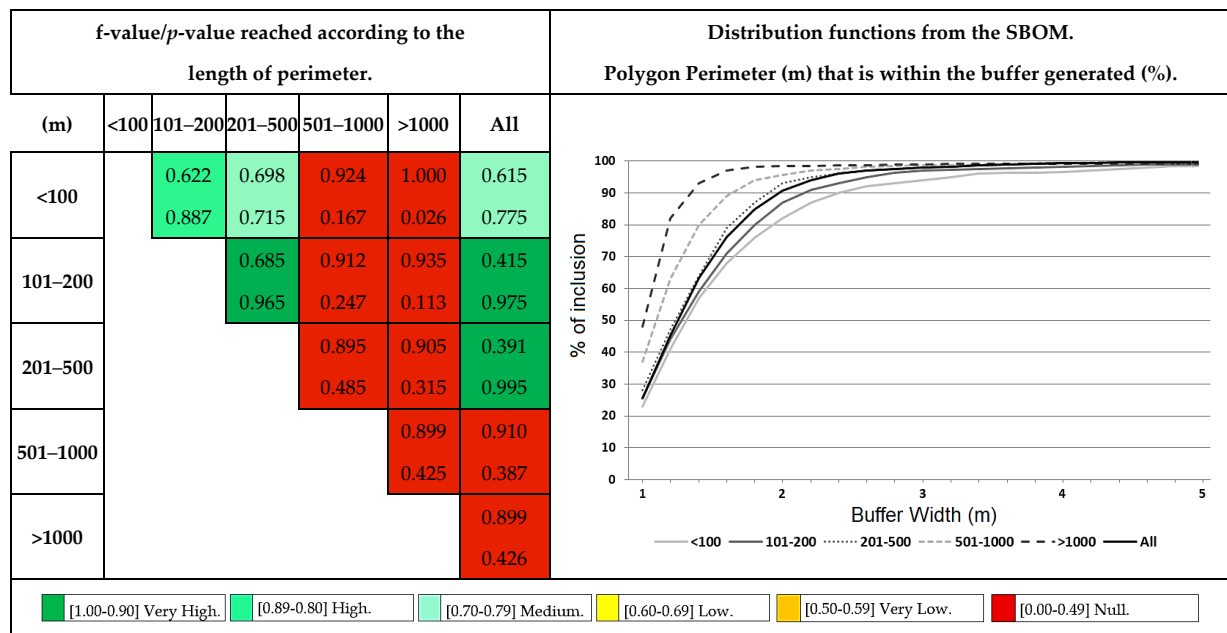
### 3.3. Categorization of the Sample

As stated above, this categorization was carried out by means of the comparison between the PDF and the several ODFs derived from the SBOM, and obtained by classifying the polygons—belonging to reference data or ideal segmentation—taking into account two parameters: the number of vertexes and the perimeter length. With regard to the first, the polygons (segmented regions) were classified in the following categories: <5, 5–10, 11–15, 16–20 and >20 vertexes. In the case of the perimeter length, the polygons were classified in the following categories: <100, 101–200, 201–500, 501–1000 and >1000 m.

The results are shown in Figures 9 and 10. Specifically, Figure 9 presents the similarity between the PDF and ODFs derived from the SBOM by means of both the  $f$ -value and  $p$ -value (computed from the Kolmogorov–Smirnov test) according to the number of vertexes. On the other hand, Figure 10 presents the similarity between the PDF and ODFs derived from the SBOM by means of both the  $f$ -value and  $p$ -value (computed from the Kolmogorov–Smirnov test) according to the perimeter length. In addition, and in order to facilitate the interpretation of the results, each of the cells of the results figures were colored according to the  $p$ -value reached.



**Figure 9.** Similarity between distribution functions derived from the SBOM.  $f$ -value and  $p$ -value (Kolmogorov–Smirnov test) according to the number of vertexes.



**Figure 10.** Similarity between distribution functions derived from the SBOM. f-value and *p*-value (Kolmogorov–Smirnov test) according to the length of perimeter.

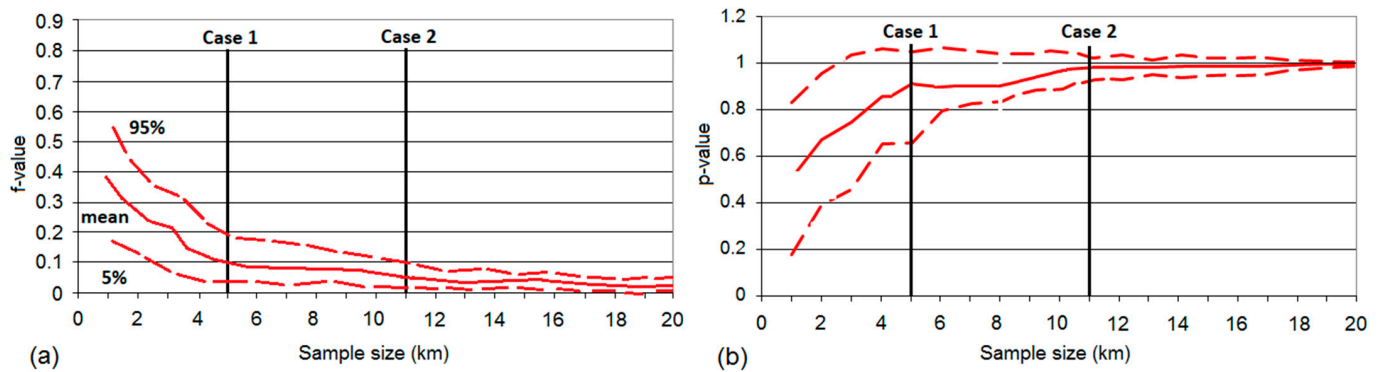
#### 4. Discussion

This paper presents an automatic PAA method for assessing the positional quality of the regions obtained by means of the application of a textural segmentation algorithm to a VHR aerial image. Based on numerical data, this method (named ADM) measures explicitly the quality of segmentation results starting from the idea proposed in [33,34]. Its greatest strength is that it does not depend on human intuition to decide the positional accuracy of a certain segmentation algorithm—as happens in the case of qualitative methods (see [1]). In addition, the automation of the process implies both a significant cost reduction and a low computational time compared to traditional methodologies (especially if we consider fieldwork).

Using the population of homologous plots automatically paired by means of the matching procedure detailed in [34] and the SBOM as the tool for measuring discrepancies, we computed the PDF which assesses the positional accuracy of the segmented plots obtained with our IS algorithm with regard to the source of higher accuracy (cadastral plots). The PDF indicated a final uncertainty of 2.4 m for a 95% level of confidence. In addition, this result was validated by means of the contrast between the original image and the segmented image using the area of the segmented regions as a reference parameter. As with any PAA method, and in order to minimize the uncertainty of the estimated accuracy values, we focused our study on two key aspects: (i) the analysis of how sample size influences the variability of the results reached by our ADM, and (ii) the categorization of the sample.

With regard to the first aspect, Figure 8 constitutes a key result of this study because the curves shown in it can be used for determining the sample size for our ADM in two different ways [37]: (i) in order to define a sample size that assures a certain value of mean discrepancy *f* between the sample (the ODF) and the population (the PDF), and (ii) in order to define a sample size that assures, with a probability of 95%, that the maximum discrepancy between the sample (the ODF) and the population (the PDF) is *f*. Figure 11a illustrates—through an example—how the *f* and *p* curves can be used to determine the size of the sample according to the two above mentioned premises. In the first case (case 1), if we wish for the mean discrepancy *f* between the sample and the population to be 10%, then we have to determine the point where the line corresponding to level 0.1 (*f*) crosses the continuous line. After this, we must check the value—belonging to the abscissa axis—

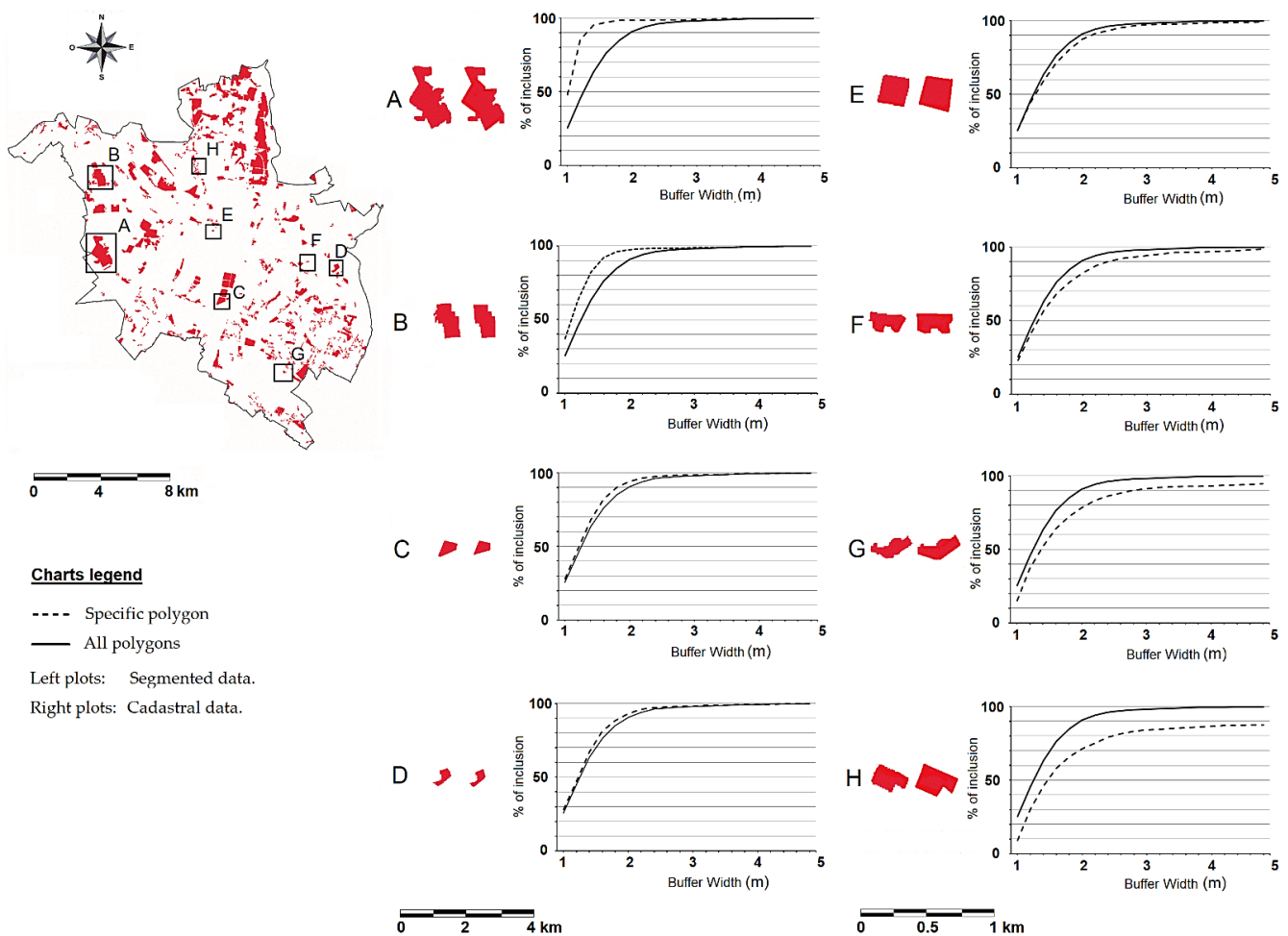
which corresponds to that point. In the second case (case 2), following the same procedure as above, we have to determine the point where the line corresponding to level 0.1 ( $f$ ) crosses the line corresponding to the 95% percentile line. Finally, we must check the value - belonging to the abscissa axis - which corresponds to that point. The results are, respectively,  $L = 5$  km (case 1) and  $L = 11$  km (case 2). In addition, the  $p$ -values corresponding to both cases are shown in Figure 11b.



**Figure 11.** (a) Illustration of the use of the  $f$ -value graphic for determining a mean discrepancy between the sample and the population (case 1) or a maximum discrepancy (case 2), and (b)  $p$ -value obtained for case 1.

In any case, it must be noted that the curves presented in Figures 8 and 11 show tendencies of a general situation of estimating parameters by sampling. Therefore, they must be used with care, to obtain an approximate idea of the sample size required in order to minimize the uncertainty of the accuracy estimated values, so they should not be interpreted as a standard. Finally, and although the curves were derived from two specific datasets, the process can be extended to other cases. To that end, it would suffice to apply a simulation procedure similar to that described above.

With regard to the characterization of the sample, all the ODFs derived from the SBOM and obtained by classifying the polygons according to the number of vertexes present a similar behavior to the PDF. Thus, the similarity between most distribution functions was classified as high or very high. In any case, the two classes whose behavior is further away from the population are those composed of polygons with a number of vertexes higher than 20 and lower than 5. The first case represents the polygons—or segmented regions—for which the highest segmentation accuracy levels were reached. The second one, on the contrary, represents the regions for which the lowest segmentation accuracy levels were reached. The homogeneity observed in the previous case does not occur when classifying the polygons according to the length of perimeter. In this case there are two classes that stand out from the rest: those composed of polygons with the highest length of perimeter (501–1000 m and >1000 m). The positional accuracy for the polygons belonging to these two classes is very far above the rest. The opposite occurs for the polygons with the lowest length of perimeter (included in the following classes: <100 m and 101–200 m). Figure 12 shows the distribution functions corresponding to eight different polygons (plots) classified according to their length of perimeter. As shown, best levels of accuracy correspond to polygons with the highest length of perimeter (plots from A to D). On the other hand, worse levels of accuracy correspond to polygons with the lowest length of perimeter (plots from E to H).



**Figure 12.** Distribution functions corresponding to eight different polygons (plots) classified according to their length of perimeter.

In view of the above, it could be stated that polygons with a larger value for the perimeter length present a greater positional accuracy regardless of the number of vertices they have. This means that a high number of vertices tends to mean a great length of perimeter; nevertheless, the opposite is not always true: a reduced number of vertices need not mean a smaller value for the perimeter length. This last point can explain the differences between both distribution functions—Figures 9 and 10.

Finally, the trend showed in Figure 12 was confirmed by carrying out a correlation analysis between the positional accuracy (for a 95% level of confidence) and the length of perimeter for a random sample of the polygons (plots) assessed. Figure 13 shows the result from both a graphic and a numerical perspective (coefficient).

#### 4.1. ADM Applied to Textural Segmentation vs. ADM Applied to Sentinel Classification

The efficacy of our ADM was also tested by means of its application to the classification results obtained by Strykiewicz et al. [32]. The reason for employing this study was two-fold: (i) because it was applied to the same geographic area (the municipal district of Poznań), and (ii) because its results were checked from a thematic perspective by means of an external validation based on a visual inspection procedure through field visits to a set of test plots. In the aforementioned study, we provided a better understanding of how seasonal Sentinel data may be used to quantify abandoned agricultural land. However, its results were not tested from a positional perspective. Figure 14 shows a comparison between the PAA reached by means of the application of our ADM to both methods (textural segmentation and Sentinel classification).

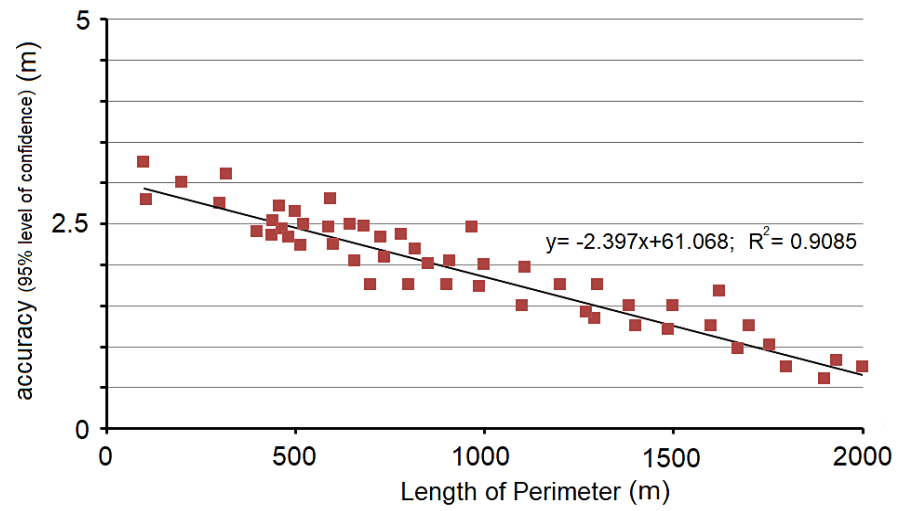


Figure 13. Correlation analysis between the positional accuracy (for a 95% level of confidence) and the length of perimeter for a random sample of the polygons (plots).

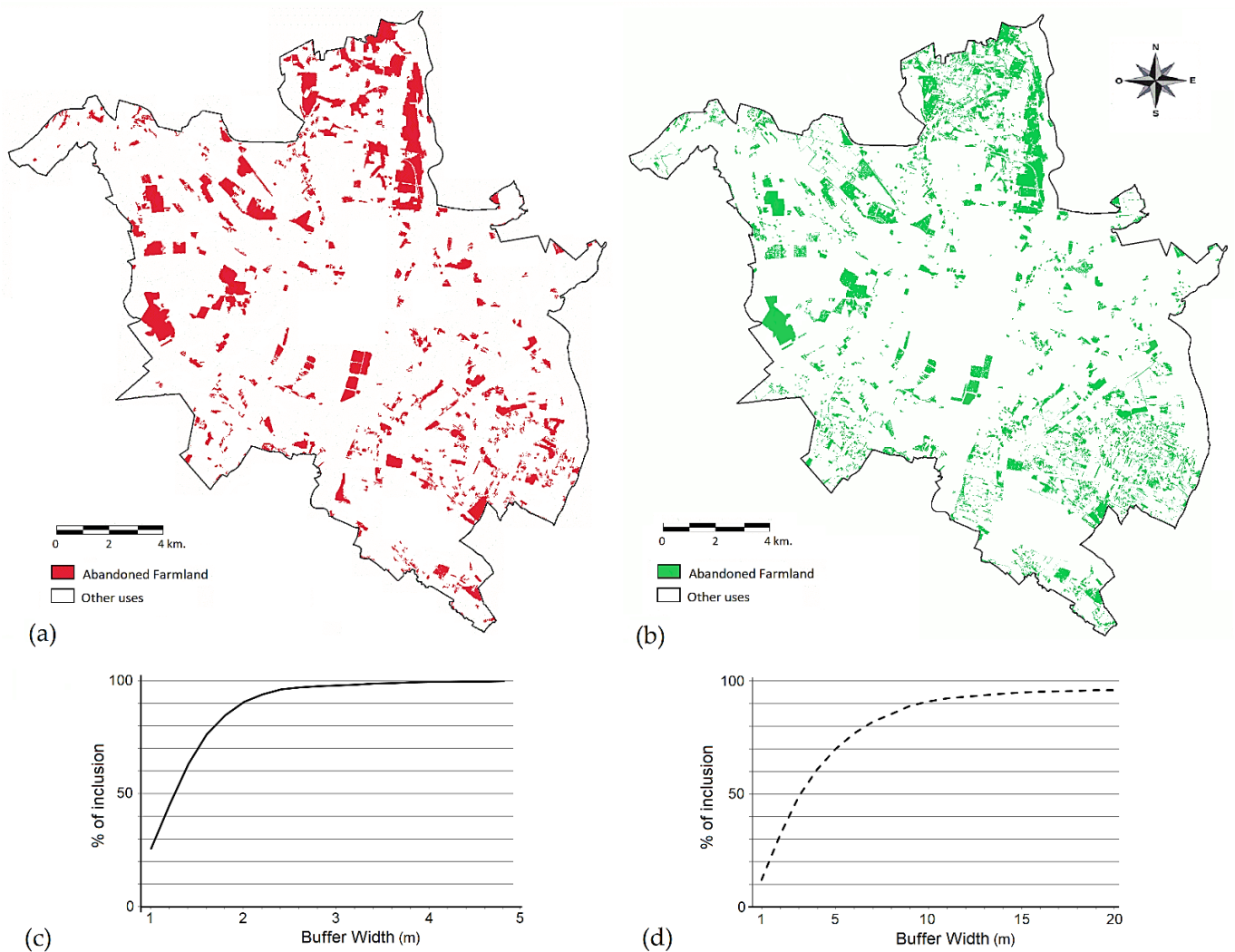


Figure 14. Distribution functions corresponding to the application of our ADM to: (a,c) textural segmentation procedure; (b,d) Sentinel classification.



In the first case (Figure 14a,c), the reached uncertainty was 2.4 m for a 95% level of confidence, while for the latter (Figure 14b,d) the reached uncertainty was 15.1 m for the same level of confidence. Logically, both results were conditioned by the imagery resolution. In the case of Sentinel, 10 m; in the case of the aerial image provided by World Imagery (MapServer) [48], 1 m. Nevertheless, they demonstrate that our ADM can be applied to different types of IS processes. In addition, and on this basis, our measure may be helpful in order to select the segmentation method most suited to certain requirements. For example, in the two cases above mentioned we might conclude that, although both are efficient in order to locate abandoned farmland, textural segmentation is more appropriate when the aim is to provide information about the exact area of abandoned arable land at any given time.

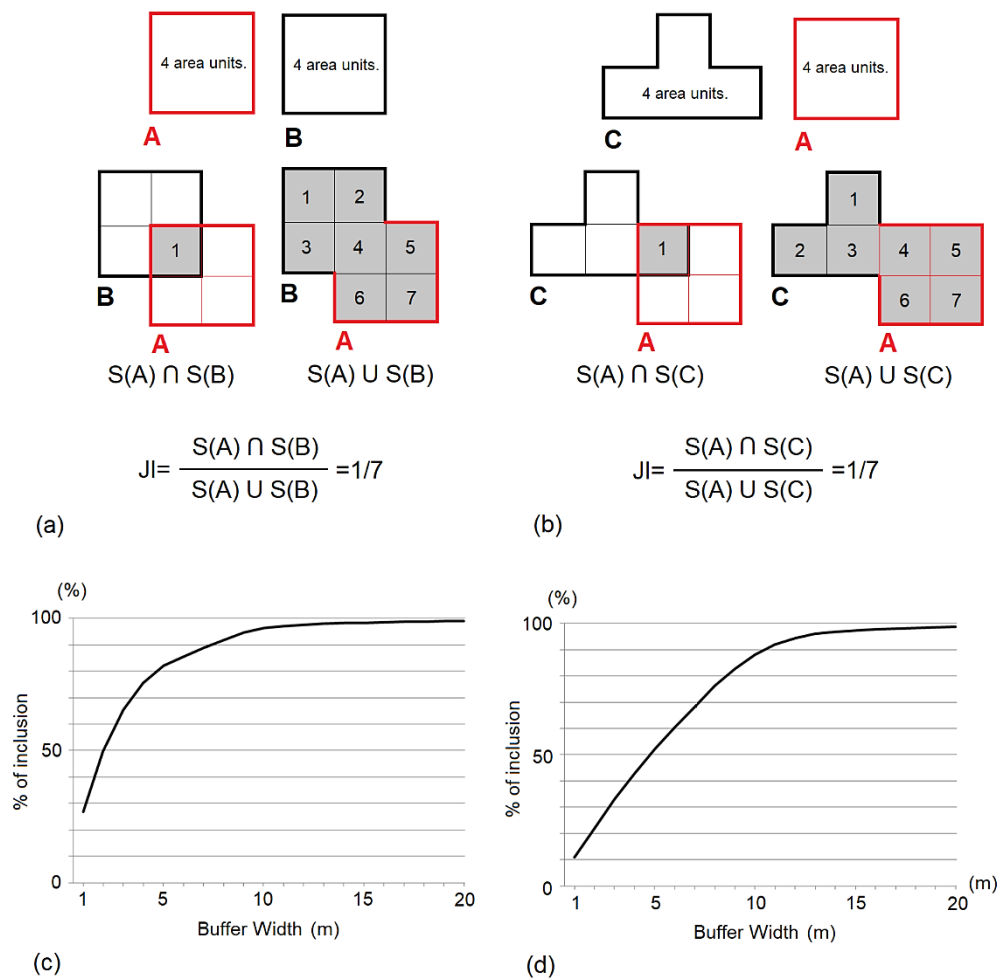
#### 4.2. ADM vs. Segmentation Assessment Indices

As stated in the Introduction Section, the assessment of the segmentation results provided by most of the goodness indices shows a high efficiency. These indices provide results extremely effective in comparison with the extraordinary simplicity of the metrics which they employ. Such in the case of the Jaccard index (JI) [33]. The Jaccard index, also called the Intersection-Over-Union index, is one of the most commonly used metrics in segmentation assessment processes. The JI could be classified as an empirical method that measures the similarity between the ground truth and the segmentation result obtained by means of the application of an IS algorithm. Specifically, it relates the area of overlap between the computed segmentation and the ground truth and the area of union between the computed segmentation and the ground truth according to Equation 5. In this sense, it could be stated that the JI is similar to our ADM because both use an ideal segmentation in order to measure the quality of a certain segmentation process.

$$JI = \frac{\text{Area of overlap}}{\text{Area of Union}} \quad (5)$$

This metric ranges from 0 to 1 (0–100%) with 0 signifying no overlapping and 1 signifying perfect overlapping segmentation. However, the functioning of indices in general, and particularly of the JI seems not to be the most adequate in PAA processes involving GI. As mentioned above, geographical objects are defined by the shape of their boundaries which means that in order to assess their geometry—after being segmented and extracted by an IS algorithm—it is necessary to assess the geometry of their boundaries. Thus, although our IS algorithm is a region-based textural segmentation algorithm, our ADM is a boundary-based metric. Figure 15 illustrates this.

Figure 15a shows two synthetic regions (polygons) segmented and subsequently matched, and that have the same area ( $SA = SB = 4$  area units). Region A belongs to reference data—ground truth or cadastral data in our case—and region B belongs to tested data—region segmented by the application of a certain IS algorithm. On the other hand, Figure 15b shows the same above region (polygon) A belonging, as before, to reference data (ground truth or cadastral data) but that, in this case, was matched with region C—segmented by the application of another IS algorithm. Moreover, as in the previous case, both regions have the same area ( $SA = SB = 4$  area units). If we compute the JI in both cases, the result is exactly the same (1/7) because both in the first case and in the second one the overlapping area is equal to one and the area of union is equal to seven. Therefore, two spatial regions (B and C) with the same area but with a different geometric shape can be assessed with the same quality index when compared with the same geographic truth (region A). Furthermore, Figure 15c,d show the results obtained by applying our ADM to the cases shown in Figure 15a,b. Now, both distribution functions behave as signatures that unequivocally identify the accuracy when the segmented regions are compared—regions A and B in the case of Figure 15c, and regions A and C in the case of Figure 15d.



**Figure 15.** D: (a) assessment of segmentation results by means of the Jaccard index for regions A and B, (b) assessment of segmentation results by means of the Jaccard index for regions A and C, (c) assessment of segmentation results by means of the ADM for regions A and B, and (d) assessment of segmentation results by means of the ADM for regions A and C.

## 5. Conclusions

The majority of the proposals for assessing the effectiveness of IS processes are qualitative approaches which, based on human intuition, determine the accuracy of a certain segmentation algorithm. There are other methods which use test images to assess the IS algorithms. Specifically, they apply a certain IS algorithm to a test image and measure the accuracy of segmentation results from a numerical perspective. The most extended of them are goodness methods. These methods define the IS process in terms of how well it segmented objects of interest. For this, they use goodness parameters which measure some desirable properties of segmented images. However, these types of methods do not seem to be the most efficient when assessing the positional accuracy of the segmentation results in processes involving GI—VHR aerial imagery or satellite imagery. Thus, and in order to carry out a formal analysis of GI positional aspects derived from IS processes, the most suitable methods are discrepancy methods because they share the same conceptual principles as the procedures used to carry out the quality control of cartography.

The present study has made it clear that automatic PAA procedures, which are mainly applied to the quality control of cartography, are valid for assessing the positional accuracy achieved with other types of processes. Such is the case of the IS process presented in this study. Thus, and having as a starting point the regions obtained by means of the application of a textural segmentation algorithm to a VHR aerial image [13], we developed an ADM for assessing—from an empirical perspective—the positional accuracy of these regions. This ADM was based on the automatic procedure originally proposed in [33,34]

and for its development particularly close attention was paid to two aspects: the influence of sample size on the segmentation results and the characterization of the sample.

The results obtained must be taken with caution because, although they are promising, it is not easy to find exact mathematical criteria which allow us to determine whether a segmentation is good or not. Thus, in most cases, the success of IS processes still depends on the nature of the segmentation problem. In any case, the simulation of samples—randomly extracted from an initial population of homologous polygons—of different size  $L$  has allowed us to compute the variability of the estimated planimetric accuracy of the tested segmentation data by comparing the similarities between each of the ODFs and the PDF and, thus, in turn, to determine the sample size for our ADM under different premises. In addition, the classification of the segmented regions (polygons) according to different parameters, such as the number of vertexes and the length of the perimeter, has allowed us to conclude that the size of the polygons is a significant feature to consider when assessing the PAA of the IS processes by means of our ADM. Thus, the similarity studies between distribution functions derived from the SBOM according to both parameters categorize the polygons with the highest number of vertexes and the highest length of perimeter as the polygons with the greatest accuracy.

Finally, our ADM was applied to a specific area for polygons representing rural plots of abandoned land, and to imagery with specific resolutions or GSD. However, it would be interesting to know its behavior in other geographical frameworks, with other typologies of spatial objects, such as rural plots with different land uses—in the case of rural environments—buildings and city blocks—in the case of urban environments—and applied to imagery with a higher resolution. In this regard, it must be noted that our ADM and its procedural approach is transferable to other areas, spatial objects and imagery resolutions. To do so, a precise configuration of the parameters that control the functioning of the IS algorithm will only be required. As mentioned in Section 3.1, these parameters are (i) the training process whose main goal is to learn the texture parameters LBP,  $C$  belonging to the spatial regions to segment, and (ii) the threshold  $U$ . The correct setting of  $U$  will ensure the correct detection and differentiation of textures in the different cases that may arise. Examples of all this can be found in [13,39].

There are many possible directions to continue this research. Thus, in the future and taking into account that, as mentioned above, the success of IS processes still depends on the nature of the segmentation problem, we plan to use other approaches based on artificial intelligent, such as methods based on convolutional neural networks, with training series related to a given specific problem. On the other hand, we plan to apply our ADM to different segmentation cases, including new imagery platforms and IS algorithms, studying the signification of other polygon features than those described above.

**Author Contributions:** Conceptualization, Juan J. Ruiz-Lendínez; Methodology, Juan J. Ruiz-Lendínez and Manuel A. Ureña-Cámara; Software, José L. Mesa-Mingorance and Francisco J. Quesada-Real; Investigation, Juan J. Ruiz-Lendínez and Manuel A. Ureña-Cámara; Writing—original draft preparation, Juan J. Ruiz-Lendínez; Writing—review and editing, Juan J. Ruiz-Lendínez and José L. Mesa-Mingorance; Supervision, Juan J. Ruiz-Lendínez, Manuel A. Ureña-Cámara, José L. Mesa-Mingorance and Francisco J. Quesada-Real; Project administration, Juan J. Ruiz-Lendínez; Funding acquisition, Juan J. Ruiz-Lendínez. All authors have read and agreed to the published version of the manuscript.

**Funding:** Financial support from grant 358 PID2019-106195RB-100 (Spanish Ministerio de Ciencia, Innovación y Universidades).

**Institutional Review Board Statement:** The study did not require ethical approval.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data sharing is not applicable to this article.

**Acknowledgments:** We would like to acknowledge the staff of the Institute of Socio-Economic Geography and Spatial Management of the Adam Mickiewicz University of Poznań for their cooperation and technical support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Marfil, R.; Molina-Tanco, L.; Bandera, A.; Rodríguez, J.; Sandoval, F. Pyramid segmentation algorithms revisited. *Pattern Recognit.* **2006**, *39*, 1430–1451. [[CrossRef](#)]
2. Manjula, K. Role of Image Segmentation in Digital Image Processing for Information Processing. *Int. J. Comput. Sci. Trends Technol.* **2015**, *3*, 312–318.
3. Keselman, Y.; Dickinson, S. Generic model abstraction from examples. *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **2001**, *1*, 856–863.
4. Singh, P.; Chadha, R. A Novel Approach to Image Segmentation. *Int. J. Adv. Res. Comput. Sci. Soft. Eng.* **2013**, *2*, 24–27.
5. Gurcan, R.; Erer, I.; Kent, S. An edge detection method using 2-D autoregressive lattice prediction filters for remotely sensed images. In Proceedings of the Geoscience and Remote Sensing Symposium, IGARSS'04, Anchorage, AK, USA, 20–24 September 2004.
6. Yogamangalam, R.; Karthikeyan, B. Segmentation techniques comparison in image processing. *Int. J. Eng. Technol.* **2013**, *5*, 307–313.
7. Serra, J.; Salembier, P. Connected operators and pyramids. In Proceedings of the SPIE: Image Algebra and Morphological Image Processing IV, San Diego, CA, USA, 12–13 July 1993.
8. Boiangiu, C.; Dinu, O.; Popescu, C.; Constantin, N.; Petrescu, C. Voting-Based Document Image Skew Detection. *Appl. Sci.* **2020**, *10*, 2236. [[CrossRef](#)]
9. Jolion, J.; Montanvert, A. The adaptive pyramid, a framework for 2D image analysis. *Comput. Vis. Image Underst.* **1992**, *55*, 339–348. [[CrossRef](#)]
10. Kropatsch, W.; Haxhimusa, Y. Grouping and segmentation in a hierarchy of graphs. In *Computational Imaging II*; SPIE Press, Digital Library: Bellingham, WA, USA, 2004; pp. 193–204.
11. Meer, P. Stochastic image pyramids. *Comput. Vis. Graph. Image Process.* **1989**, *45*, 269–294. [[CrossRef](#)]
12. Bister, M.; Cornelis, J.; Rosenfeld, A. A critical view of pyramid segmentation algorithms. *Pattern Recognit. Lett.* **1990**, *11*, 605–617. [[CrossRef](#)]
13. Ruiz-Lendínez, J. Abandoned Farmland Location in Areas Affected by Rapid Urbanization Using Textural Characterization of High Resolution Aerial Imagery. *Int. J. Geo-Inf.* **2020**, *9*, 191. [[CrossRef](#)]
14. Ruiz-Lendínez, J.J.; Rubio-Campos, T.J.; Ureña-Cámara, M.A. Automatic extraction of road intersections from images based on texture characterization. *Surv. Rev.* **2011**, *43*, 212–225. [[CrossRef](#)]
15. Malhotra, A.; Sankaran, A.; Mittal, A.; Vatsa, M.; Singh, R. Fingerphoto Authentication Using Smartphone Camera Captured Under Varying Environmental Conditions. In *Using Computer Vision, Pattern Recognition and Machine Learning Methods for Biometrics*; Academic Press: Oxford, UK, 2017; pp. 119–144.
16. Ruiz-Lendínez, J.J.; Mackiewicz, B.; Motek, P.; Strykiewicz, T. Method for an automatic alignment of imagery and vector data applied to cadastral information in Poland. *Surv. Rev.* **2019**, *51*, 123–134. [[CrossRef](#)]
17. Shi, J.; Malik, J. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence. IEEE Trans. Geosc. Remote Sens.* **2000**, *22*, 888–905.
18. Benediktsson, J.; Pesaresi, M.; Arnason, K. Classification and Feature Extraction for Remote Sensing Images from Urban Areas Based on Morphological Transformations. *IEEE Trans. Geosc. Remote Sens.* **2003**, *41*, 1940–1949. [[CrossRef](#)]
19. Tseng, Y.; Wang, S. Semi-automated building extraction based on CSG model-image fitting. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 171–180. [[CrossRef](#)]
20. Jaynes, C.; Riseman, E.; Hanson, A. Recognition and reconstruction of buildings from multiple aerial images. *Comput. Vis. Image Underst.* **2003**, *90*, 68–98. [[CrossRef](#)]
21. Chen, Y.; Wang, J. Image Categorization by Learning and Reasoning with Regions. *J. Mach. Learn. Res.* **2004**, *5*, 913–939.
22. Ünsalan, C.; Boyer, K. A system to detect houses and residential street networks in multispectral satellite images. *Comput. Vis. Image Underst.* **2005**, *98*, 423–461. [[CrossRef](#)]
23. Yee, R. *Architectural Drawing a Visual Compendium of Types and Methods*; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2013.
24. Ok, A. A new approach for the extraction of aboveground circular structures from Near-Nadir VHR satellite imagery. *Geosciences and Remote Sensing. IEEE Trans. Geosc. Remote Sens.* **2014**, *52*, 3125–3140. [[CrossRef](#)]
25. Ok, A.; Senaras, C.; Yuksel, B. Automated Detection of Arbitrarily Shaped Buildings in Complex Environments from Monocular VHR Optical Satellite Imagery. *IEEE Trans. Geosc. Remote Sens.* **2013**, *51*, 1701–1717. [[CrossRef](#)]
26. Gruen, A.; Li, H. Road extraction from aerial and satellite images by dynamic programming. *ISPRS J. Photogramm. Remote Sens.* **1995**, *50*, 11–20. [[CrossRef](#)]
27. Mayer, H.; Steger, C. Scale-Space Events and Their Link to Abstraction for Road Extraction. *ISPRS J. Photogramm. Remote Sens.* **1998**, *53*, 62–75. [[CrossRef](#)]
28. Bückner, J. Model based road extraction for the registration and interpretation of remote sensing data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **1998**, *32*, 85–90.
29. Poonam, S.; Pandey, A. Automatic urban road extraction using airborne laser scanning-altimetry and high-resolution satellite data. *J. Indian Soc. Remote Sens.* **2009**, *37*, 223–231.

30. Chen, C.; Knoblock, C.; Shahabi, C. Automatically conflating road vector data with orthoimagery. *Geoinformatica* **2006**, *10*, 495–530. [[CrossRef](#)]
31. Gradinaru, S.; Kienast, F.; Psomas, A. Using multi-seasonal Landsat imagery for rapid identification of abandoned land in areas affected by urban sprawl. *Ecol. Indic.* **2019**, *96*, 79–86. [[CrossRef](#)]
32. Strykiewicz, T.; Królewicz, S.; Ruiz-Lendínez, J.J.; Mickiewicz, B.; Motek, P. Abandoned agricultural land quantification in urban areas using high resolution satellite imagery. In Proceedings of the RSA Central and Eastern Europe Conference: Metropolises and Peripheries of CEE Countries: New Challenges for EU, National and Regional Policies, Lublin, Poland, 11–13 September 2019.
33. Shi, R.; Ngan, K.; Li, S. Jaccard index compensation for object segmentation evaluation. In Proceedings of the International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014.
34. Prewer, D.; Kitchen, L. Soft image segmentation by weighted linked pyramid. *Pattern Recognit. Lett.* **2001**, *22*, 123–132. [[CrossRef](#)]
35. Liu, J.; Yang, Y. Multi-resolution color image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1994**, *16*, 689–700.
36. Borsotti, M.; Campadelli, P.; Schettini, R. Quantitative evaluation of color image segmentation results. *Pattern Recognit. Lett.* **1998**, *19*, 741–747. [[CrossRef](#)]
37. Ruiz-Lendínez, J.J.; Ariza-López, F.J.; Ureña-Cámara, M.A. Automatic positional accuracy assessment of geospatial databases using line-based methods. *Surv. Rev.* **2013**, *45*, 332–342. [[CrossRef](#)]
38. Ruiz-Lendínez, J.J.; Ureña-Cámara, M.A.; Ariza-López, F.J. A Polygon and Point-Based Approach to Matching Geospatial Features. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 399. [[CrossRef](#)]
39. Ojala, T.; Pietikäinen, M. Unsupervised texture segmentation using feature distributions. *Pattern Recognit.* **1999**, *32*, 477–486. [[CrossRef](#)]
40. Goodchild, M.; Hunter, G. A simple positional accuracy measure for linear features. *Int. J. Geogr. Inf. Sci.* **1997**, *11*, 206–299. [[CrossRef](#)]
41. Ariza-López, F.J.; Ruiz-Lendínez, J.J.; Ureña-Cámara, M.A. Influence of Sample Size on Automatic Positional Accuracy Assessment Methods for Urban Areas. *Int. J. Geo-Inf.* **2018**, *7*, 200. [[CrossRef](#)]
42. Van Niel, T.; McVicar, T.R. Experimental evaluation of positional accuracy estimates from linear network using point and line based testing methods. *Int. J. Geogr. Inf. Sci.* **2002**, *16*, 455–473. [[CrossRef](#)]
43. Mozas-Calvache, A.T. Exactitud Posicional de Elementos Lineales en Cartografía. Ph.D. Thesis, University of Jaén, Jaén, Spain, 2008.
44. William, H.; Brian, P.; Saul, A.; William, T. Numerical Recipes. In *C: The Art of Scientific Computing*, 2nd ed.; Cambridge University Press: Cambridge, UK, 1992.
45. Marsaglia, G.; Tsang, W.W.; Wang, J. Evaluating Kolmogorov's distribution. *J. Stat. Softw.* **2003**, *8*, 1–4. [[CrossRef](#)]
46. Gibbons, J.D.; Chakraborti, S. *Nonparametric Statistical Inference*; Marcel Dekker Ltd.: London, UK, 2003.
47. Press, W.A.; Vetterling, W.T.; Teukolsky, S.A.; Flannery, B.P. *Numerical Recipes in Fortran: The Art of Scientific Computing*; Cambridge University Press: Cambridge, UK, 1992.
48. World Imagery. Available online: [http://goto.arcgisonline.com/maps/World\\_Imagery](http://goto.arcgisonline.com/maps/World_Imagery) (accessed on 10 January 2021).
49. Poviát Geodetic and Cartographic Documentation Center in Poznań. Available online: <https://podgik.poznan.pl/> (accessed on 10 January 2021).