geographies

MDPI

Article

Measuring Data Quality from Building Registers: A Case Study in Italy

Gianluigi Salvucci [1], Donato Scarpitta [2], Marco Maialetti [3], Kostas Rontos [4], Stefano Bigiotti [5], Adele Sateriano [6,*] and Alessandro Muolo [7]

1 Italian National Institute of Statistics (ISTAT), Via Cesare Balbo 16, I-00184 Rome, Italy; salvucci@istat.it
2 Independent Researcher, Santa Marina, I-84067 Salerno, Italy; scarpittadonato@libero.it
3 Independent Researcher, I-00195 Rome, Italy; avvmarco.maialetti@gmail.com
4 Department of Sociology, University of the Aegean, University Hill, EL-81100 Mitilini, Greece; k.rontos@soc.aegean.gr
5 Department of Agricultural and Forest Sciences (DAFNE), Tuscia University, Via S. Camillo de Lellis, I-01100 Viterbo, Italy; stefano.bigiotti@unitus.it
6 Mediterranean Sustainable Development Foundation (MEDES), Sicignano degli Alburni, I-84029 Salerno, Italy
7 Department of Methods and Models for Economics, Territory and Finance (MEMOTEF), Sapienza University of Rome, Via del Castro Laurenziano 9, I-00161 Rome, Italy; muolo.1916725@studenti.uniroma1.it
* Correspondence: adele.sateriano.pul@gmail.com

Abstract: Geographic data quality is a complex issue requiring continuous operational improvements. Considering this to be one of the most topical research and technical issues in official statistics and environmental monitoring, this study re-connects the operational dimension of 'geographic data quality' with the broader issue of monitoring the quality of official statistics. By estimating the accuracy of public (spatially explicit) data, this study illustrates an operational framework with an exploratory exercise in estimating the geographic data quality characteristic of a specific information source within the official statistical system (i.e., building registry) in a given European country, namely, Italy. The results of this exercise provide a paradigmatic example of profound innovation in the activities of statistical services in Europe and specifically the Italian National Statistical Institute (Istat), transitioning from independent (and poorly connected) field surveys to an integrated system of registries. Since several studies are based on spatially explicit survey units, it is essential to estimate the quality of geographical data, especially those derived from information sources where space is topical information, such as (local, regional, or national) building registers. Thanks to the results of an empirical exercise applied to Italian building registers, the present article will discuss the issue of data accuracy, considered the main issue related to monitoring geographic data quality, from an official statistics' perspective. Statistical indicators will be proposed for the assessment of systematic and random errors of spatially explicit measures, possibly enabling a quali-quantitative improvement in the semantic content of building registers that address the inherent requirements of official statistics. Such indicators have some positive implications for the entire system of official statistics in Italy and, for generalization, within the European Statistical System.

Keywords: official statistics; census; sampling survey; public database; precision

check for updates

Citation: Salvucci, G.; Scarpitta, D.; Maialetti, M.; Rontos, K.; Bigiotti, S.; Sateriano, A.; Muolo, A. Measuring Data Quality from Building Registers: A Case Study in Italy. *Geographies* **2024**, *4*, 596–611. https://doi.org/10.3390/geographies4030032

Academic Editor: Eliseo Clementini

Received: 12 August 2024
Revised: 26 August 2024
Accepted: 3 September 2024
Published: 9 September 2024

Copyright: © 2024 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

Present (and future) official statistics, especially on the old continent, are primarily (and will be increasingly) based on the information content of publicly available databases that are not specifically collected in the framework of censuses or sampling surveys [1]. Most of them—developed with the general aim of assessing public action and government policies' impacts—especially those with a spatially explicit dimension, may become accurate and point sources of official statistics [2]. Additionally, such information sources

are becoming increasingly relevant in any research field, from environmental monitoring to territorial studies, from regional science to spatial demography, and from urban planning to rural sociology [3]. Applied economics, geography, ecology, and agronomy are likely the academic disciplines benefiting more intensively from such methodological and applied improvements [4]. From this perspective, building registers and (non-statistical) data sources on human settlements may represent a coherent example of public (often freely available) archives that may provide, after accurate verification and testing, relevant official statistics [5]. They routinely display standard requirements and characteristics (e.g., semantic coherence, users' reliability, easily updating over time, geographical precision, perfectly documented errors associated with any measure, and documented estimates in the case of missing data) typical of such data sources [6]. Additionally, building registers are eminently spatially explicit sources of information, so the verification of their routine use as official statistics is a challenging task for spatial analysis and geographical science at large [7].

The correct positioning of a statistical unit at a sub-municipal level becomes crucial in the context of registry-based censuses, one of the most important applications of public (non-statistical) registers within an official statistics' operation, namely, the total enumeration of population or economic activities [8]. In a traditional census, the enumerator, armed with a sectional itinerary or map, ensures the accurate geocoding of the collected information by knowing the specific portion of the territory [9]. However, when integrating various sources, such as population and cadastral data, challenges arise [10]. Despite the provisions of the registry regulation that facilitate geocoding at the population section level, the cadastral system operates on a different basis [11]. Consequently, aligning diverse information sources may lead to misalignments stemming from the geographic quality of the utilized registries [12]. It is not merely a matter of measuring the positional accuracy of individual statistical units but a matter of ensuring the coherence of information, a basic feature of any official statistics when released to the final stakeholders (e.g., scholars, practitioners, and general users). This means that individuals recorded in the registry as residing at certain addresses must seamlessly align with the corresponding residences through the specific relationships delineated within the registry system [13]. The substantial difference between various technical and mapping approaches is reflected in the possibly diverging statistical releases of vastly different data sources and informative knowledge [14]. While cartographers' goal is to represent the complexity of a given territory in a simplified image, considerations about the quality of geographic data as a possible source of official statistics stem from the challenge of cartographic production and are largely overcome, or surmountable, in a world based on Geographic Information Systems [15]. In other words, there is not a unique source that can populate, alone, an official statistics' building register [16]. Conversely, a complex framework of data sources' integration should be implemented when preparing a complete building register supporting census operations and other official statistics' surveys [17]. Among several other issues, matching different information sources implies an intense activity of technical harmonization and a continuous estimation of the degree of completeness of the newly built-up archive, assumed as a proxy of statistical accuracy, likely the most relevant dimension in any official statistical system, e.g., [18]. A practical implementation of these operations is a novel aspect of official statistics dealing with spatially explicit data.

Considered one of the most topical research and technical issues in official statistics, environmental monitoring, and territorial analysis at large, the present study connects this challenging topic with the broader issue of monitoring the quality of official statistics and estimating the precision of public, spatially explicit data—considering both systematic and, eventually, survey errors [19]. With this perspective in mind, we propose an operational framework with a mensuration exercise for spatially explicit data provided by a specific source of the official statistical system [20], namely, a building registry, in a given European country, namely, Italy. This country serves as a paradigmatic example of a profound innovation in the activities of the National Statistical Institute (hereafter,

Istat) transitioning from field (sparse and basically independent) surveys to an integrated (and statistically coherent) system of registries, whose degree of spatial integration is actually undergoing active improvement [21]. Focusing on the relationship between public (non-statistical) registers and the intrinsic requirements of official statistics in a spatially explicit world [22], the exercise documents an operational framework assessing—through simplified indicators—the intrinsic quality of various (non-statistical) sources of spatial data taken as multiple examples of building registers [23]. The empirical results of this exercise allow a broader discussion on the methodology for verifying the quality, precision, update, readability, and coherence of non-statistical (spatially explicit) data sources [24]. When reaching complete compliance, these criteria may represent the basic requirements for inclusion in the portfolio of official statistics in a given country (e.g., Italy) or a broader geographical range (e.g., European Union). When reaching partial compliance [25], these criteria can, however, guide statisticians in the final decision of whether such information can be routinely used in official procedures (e.g., population censuses).

In summary, the present work will address the main issues related to monitoring geographic data quality, with a focus on building registers in Italy [26]. Having this framework in mind, statistical indicators will be proposed (and preliminarily calculated) when assessing systematic and sampling errors in spatially explicit measures [27], possibly enabling a quali-quantitative improvement in the semantic content of registers, with positive implications for the entire system of official statistics in Italy and, for generalization, all over Europe. Aimed at filling this research gap—of primary interest for the informed development of any national statistical system facing the challenge of big (spatial) data—the present paper is organized as follows. Section 2 introduces the empirical methodology developed and applied in the present study. Section 3 summarizes the main results, considering important issues for official statistics (e.g., accuracy) and how to cope with varying measurement errors. Section 4 concludes the work by discussing some implications of the present study for academic issues and practical aspects of land surveying. Section 5 delineates some possible (future) lines of academic research and technical implementation.

2. Methodology

2.1. Study Area

Italy is administratively partitioned into 3 geographical areas (North, Center, South), 20 Nuts-2 administrative regions, and more than 100 Nuts-3 provinces, covering a total surface area of nearly 301,330 km^2 [28]. While the regions remained basically stable for a relatively long time in Italy, provinces ranged from slightly more than 90 to 110 over the study period. In this work, we considered the administrative regional boundaries according to the last population census, dated 2021, and the corresponding provincial boundaries and local administrative setting, more than 8000 governing units, i.e., municipalities [29]. We extracted the needed information by referring to stable administrative boundaries from territorial databases publicly released by the Italian National Statistical Institute (Istat) in order to facilitate calculations and make statistical analysis fully comparable [30]. From a contextual perspective, Italy displays important territorial disparities in economic growth [31], social development [32], and natural resource availability [33] and thus serves as a relevant case study to address the intrinsic limitations of land cadasters and the complex process of the implementation of spatially explicit public registers with fully documented errors.

2.2. Data Sources

To assess the quality of a building register, given the intrinsic difficulty of verifying each individual building in an ortho-photo, separate scrutiny of a reference archive is needed for comparison purposes [34]. In Italy, there is no single archive that can be used as a full reference in terms of both completeness and positional accuracy [35]. Therefore, it is necessary to simultaneously construct a statistical registry from different archives based on a mixture of administrative and cartographical sources [36]. These sources, routinely

used in spatial planning, have undergone rigorous testing in their production process [37]. After the publication of the 2011 population census results, a new (centralized) building register was implemented for official statistics' purposes (i.e., assisting census operations in the 2020s). This phase was particularly intricate and required the integration of various archives at Istat [21]. Specifically, an integrated system of registers was established. For our purposes, the integration between the land cadaster and the address registry was assumed to be of paramount importance. From 2018 to 2020, all of the buildings recorded in the land cadaster were automatically geo-referenced, having performed a series of automatic and manual quality controls, e.g., [38].

The process implemented various software tools (including, but not limited to, Geographical Information System suites such as ArcMap and QGIS) and extensive visual checks. However, integrating different sources poses challenges related to conflation, duplication, and overlap among sparse data [39], which are combined into a single archive with homogeneous characteristics. The technical details of this complex operation for Italy were provided in a previous work by Salvucci and Salvati [21]. The main building data source alternatives (or supplements) to the land cadaster in Italy include (i) independent (and mostly decentralized) administrative sources individually produced by regions, provinces, or municipalities for their specific governance purposes; (ii) the National Geoportal (geospatial database) released by the Italian Ministry of the Environment for centralized purposes of territorial planning; (iii) data from the Civil Protection Agency for surveillance and national security purposes [21].

Administrative regions have a particular interest in mapping buildings because of their own administrative duties in responding to various landscape and environmental constraints introduced through national or regional laws in Italian spatial planning over time [25]. As a result, the Regional Technical Maps (hereafter, Carte Tecniche Regionali, CTR) identified individual buildings and provide, even now, extensive spatial coverage (Table 1), albeit with some territorial heterogeneities because of slightly different field operations carried out in the Italian regions [26]. The National Geoportal (GPN) is a website of the Ministry of the Environment where buildings from cadastral sources have been identified for provincial town capitals, particularly within the boundaries of urban centers [27]. It represents only a small portion of the cadastral information. Moreover, for Italian buildings, the National Civil Protection Authority recently made available a geodatabase containing 'urban structural aggregates' covering the entire national territory [21]. These geometries were derived from buildings in various regional topographic databases and have been converted into polygons representing contiguous buildings, each assigned a progressive number for regional identification, e.g., [8].

To precisely evaluate data quality, we used a unified Building Open Data Archive (BODA), which has already been employed in similar works to assess the completeness of OpenStreetMap databases [24]. The process of cartographical integration into a unique database takes account of two research dimensions: (a) data resolution, defined as the ability to identify individual buildings, and (b) data completeness. Based on these considerations, a completely new and updated BODA was implemented using all of the above-mentioned sources (Regional Technical Maps, National Geoportal, Civil Protection, OpenStreetMaps), from which the geometric elements of the buildings were derived in four sequential steps. First, all buildings from the land cadaster were included in the new BODA archive. Second, buildings from regional maps that do not intersect with any buildings from the land cadaster were also included. Third, buildings from the GPN that do not intersect with any buildings from the established archive in the previous operational step were included. Fourth, structural aggregates from Civil Protection data that do not intersect with any buildings from the established archive in the previous operational step were finally consolidated in the new database.

Table 1. A list of regional administrative sources for the construction of a Building Open Data Archive (Boda); all links below accessed in June 2024.

Code	Region	Buildings
1	Piedmont	https://www.geoportale.piemonte.it/cms/progetti/progetto-mosaicatura-catastale (accessed on 1 August 2024)
2	Aosta Valley	https://mappe.partout.it/pub/geonavitg/geodownload.asp?carta=CTR (accessed on 1 August 2024)
3	Lombardy	https://www.geoportale.regione.lombardia.it (accessed on 1 August 2024)
4	Trentino Alto Adige	http://www.openkat.it/ (accessed on 1 August 2024)
5	Veneto	https://idt2.regione.veneto.it/idt/downloader/download (accessed on 1 August 2024)
6	Friuli Venezia Giulia	http://irdat.regione.fvg.it/consultatore-dati-ambientali-territoriali/search (accessed on 1 August 2024)
7	Liguria	https://www.regione.liguria.it/open-data/item/7099-carta-tecnica-regionale-1-5000-dal-2007-ii-edizione-3d-db-topografico.html (accessed on 1 August 2024)
8	Emilia Romagna	https://geoportale.regione.emilia-romagna.it/download/download-data?type=dbtopo (accessed on 1 August 2024)
9	Tuscany	http://www502.regione.toscana.it/geoscopio/ (accessed on 1 August 2024)
10	Umbria	http://www.umbriageo.regione.umbria.it/pagina/fabbricati-sistema-ecografico-catastale-regione-um (accessed on 1 August 2024)
11	Marche	https://www.regione.marche.it/Regione-Utile/Paesaggio-Territorio-Urbanistica-Genio-Civile/Cartografia-regionale/Repertorio/Carta-tecnica-numerica-110000/opendata (accessed on 1 August 2024)
12	Latium	http://dati.lazio.it/catalog/it/dataset/carta-tecnica-regionale-2002-2003-5k-roma (accessed on 1 August 2024)
13	Abruzzo	http://opendata.regione.abruzzo.it/content/dbtr-regione-abruzzo-scala-15000-edizione-2007-formato-shp (accessed on 1 August 2024)
14	Molise	Not available
15	Campania	http://sit.cittametropolitana.na.it/downloads.php (accessed on 1 August 2024)
16	Apulia	http://www.sit.puglia.it/portal/portale_cartografie_tecniche_tematiche/Download (accessed on 1 August 2024)
17	Basilicata	http://dati.regione.basilicata.it/catalog/dataset/database-topografico-tema-edificato (accessed on 1 August 2024)
18	Calabria	/ftpopendata/CartaTecnicaRegionale5KPDF (accessed on 1 August 2024)
19	Sicily	https://www.sitr.regione.sicilia.it/download/download-carta-tecnica-2000/ (accessed on 1 August 2024)
20	Sardinia	http://www.sardegnageoportale.it/areetematiche/databasegeotopografico/ (accessed on 1 August 2024)

2.3. Logical Framework and Data Analysis

Accuracy is the first foundational element of data quality and involves comparing populations of spatial elements to assess the degree of coverage [40]. In the practical aspects of constructing the registry, the concept of accuracy is broad, and the main question is, 'what geographical elements should be included?' The definition of a building can vary across different sources that may be used to populate the geographic component of the registry [41]. The definition provided in census manuals may not align with that provided for buildings in land cadasters, which mainly serve fiscal purposes [42]. Official

statistics require that the register's individual record univocally identify the statistical unit of the building [43]. Conversely, from the cadaster's perspective, a single record may be divided into multiple parts based on, e.g., ownership, as in the case of semi-detached houses [44].

In accordance with the methodology proposed by Hecht [45], various indicators of data accuracy were preliminarily calculated in the present study, specifically by implementing three sequential steps. The first involves comparing the count of elements present in different databases for reference areas, e.g., the administrative regions (Nuts-2 level in our case). This count does not guarantee that the two archives refer to the same entities. The second step prescribes counting the centroids of the reference database that overlap with buildings in the land cadaster. This comparison is basically more detailed than the first step while not ensuring that elements found in the same locations are exactly the same entity. The third step entails a comparison of the overlapping building surface between the two archives. The comparison is based on the total built-up area, and the two indicators are defined as (i) the number of buildings surveyed in the land cadaster compared to the number of buildings recorded in the BODA in the same area and at the same time point and (ii) the total extent (square meters) of buildings surveyed in the land cadaster compared to total square meters recorded in the BODA in the same area and at the same time point.

In accordance with the operational criteria of the Civic Numbers Survey [21], with a total survey being part of census operations, it is recommended to eliminate all graphical elements with an area of less than 20 square meters, as these are likely not statistical units but technical volumes that fall outside our scope of observation. If subsequent checks on data integration reveal the routine use of such volumes, they could be reintegrated into the dataset individually [46]. If census sections are chosen as the areal unit for this comparison, trends over time can finally be verified by comparing the content of the current register with the content of the previous census(es) [47].

3. Results

The present study provides the preliminary results of an empirical exercise dealing with the complex issue delineated in the methodological section above, considering operations relevant to official statistics. While not all of the proposed approaches and metrics illustrated above were performed or can be computed with the information sources available for Italy, a comparison between the main outcomes from the new Italian building register validated for official statistics' purposes (BODA) and the land cadaster is presented, initially focusing on the two indicators mentioned in the methodological section. In the following, a preliminary quantification of the role of additional archives acting as the informative 'donor' conferring new buildings' records (not covered by land cadaster) is provided. The final aim of this step was to delineate the most relevant information sources providing original information that were scrutinized and successfully verified to be compliant with official statistics' standards and thus incorporated into the new building archive.

Table 2 shows the main indicators calculated for the entire national territory. Considering the geographical boundaries of the administrative regions (Nuts-2 level of Eurostat nomenclature), the situation appears quite complex, since the regional data sources originally used to populate the BODA (see Table 1) are the basis of largely different outcomes when compared with the aggregated outcomes (statistics) from the land cadaster. It is likely that the most relevant information stemming from this descriptive analysis is that the total extent of built-up areas in Italy estimated from BODA coverage is 24% larger than the equivalent figure estimate from the land cadaster. Presumably, regional authority maps informing the BODA include (and account for) slightly different settlement objects/elements compared with those under the coverage of land cadaster. Interestingly, the largest discrepancies between land cadaster and Boda sources at the regional level were observed in Sicily (82%), Latium (52%), and Campania (49%), three areas of the country with big

cities (Rome, Naples, Palermo, Catania), a huge pressure on new settlements because of a population increase and intense land speculation, and some evidence of informal buildings, specifically realized in the 1950s and the 1960s, that are likely not completely covered in the traditional land cadaster.

Although the land cadaster was a centralized authority in Italy operating under unique rules and common guidelines all over the country, other reasons for such a discrepancy could possibly be traced to the different operational powers of some regional branches of the authority, especially when a large building stock is to be censused in a relatively short time (as occurred in the 1950s and the 1960s with the intense growth of the metropolitan poles mentioned above). Smaller discrepancies were observed in some Central Italian regions, such as Abruzzi, Molise, Umbria, and Tuscany, dominated by rural settlements and some medium-density urban settlements. The discrepancy in the built-up area between the two information sources in these regions was systematically below 10% (higher in Boda, lower in land cadaster). Moving out from the surface area, the difference in the building stock (i.e., the estimated number of buildings) according to the two sources is clearly smaller (6.7%). In the regions mentioned above, the discrepancy was lower in Sicily (30%) and Latium (12%) and had the reverse sign in Campania (−6%). In contrast, the highest discrepancy level was observed in Friuli Venezia Giulia, a region in Northern Italy. In general terms, a differential discrepancy across regions was not observed, and a supposed gradient from Northern to Southern regions was not recorded, failing to support an unverified assumption that data quality from the archives of Northern Italian regions can be higher than their counterparts reflective of Southern regions.

A key indicator for assessing the overall quality of cadastral buildings' positioning is accuracy, an indicator computed as the ratio between the intersection (Boda–cadaster) built-up area and the built-up area effectively surveyed in the land cadaster. At the national level, 63.5% of the cadaster's surface area is common to both databases. However, at the regional level, the situation is much more complex. For instance, in Emilia Romagna, built-up areas from regional maps exceed those in cadaster maps in terms of objects/elements (+12%), with the Boda surface area being 22% larger than that in the cadaster, a sign that building blocks are drawn differently within the two sources. Conversely, only 59% of built-up areas overlap in Sicily, and the number of buildings in the Boda is significantly higher than in the land cadaster. Friuli Venezia Giulia is the region with the largest disparity in the number of buildings (Boda surveyed 45% more buildings than land cadaster); however, considering that the overall difference in the built-up area is 24% and the accuracy level is 71%, the huge discrepancy in the building stock is likely due to the different methods of drawing buildings and merging them with adjacent ones. A moderate difference between the Northern and Southern regions was recorded when considering total accuracy. The highest accuracy was similarly observed in a region of Northern Italy, Emilia Romagna (78%); Central Italy, Umbria (77%); and Southern Italy, Molise (76%). However, considering aggregate values of accuracy by macro-regions, Northern Italy is systematically above 70%, Central Italy is slightly below 70%, and Southern Italy is slightly above 65%.

Table 2. A comparison (based on indicators) between the information content of the land cadaster and Boda by administrative region in Italy.

Land Code	Administrative Region	Boda (m^2) Surface Area	Cadaster (m^2) Surface Area	No. Buildings (Boda)	No. Buildings (Cadaster)	Intersection Area (m^2)	Difference (m^2) between Boda and Cadaster	Difference (no. Build.) Boda vs. Cadaster	Accuracy
		(a)	(b)	(c)	(d)	(e)	$((a - b)/b) \times 100$	$((c - d)/d) \times 100$	$(e/b) \times 100$
20	Sardinia	117,997,448	98,066,496	513,365	684,868	62,274,348	20.32	−25.04	63.50
19	Sicily	528,491,616	290,210,848	3,236,704	2,473,018	172,431,760	82.11	30.88	59.42
18	Calabria	168,693,520	129,288,872	1,052,487	1,097,207	86,677,328	30.48	−4.08	67.04
17	Basilicata	65,117,308	45,351,400	373,077	368,758	28,675,412	43.58	1.17	63.23
16	Apulia	286,672,992	240,625,168	1,523,100	1,522,340	165,759,712	19.14	0.05	68.89
15	Campania	389,408,320	260,497,424	1,467,142	1,565,318	157,948,608	49.49	−6.27	60.63
14	Molise	27,931,490	26,675,126	159,655	234,373	20,337,080	4.71	−31.88	76.24
13	Abruzzi	110,239,912	101,602,072	586,517	682,748	76,177,552	8.50	−14.09	74.98
12	Latium	385,837,920	253,106,224	1,779,197	1,593,053	164,478,176	52.44	11.68	64.98
11	Marche	132,732,496	114,961,736	560,871	691,995	78,655,576	15.46	−18.95	68.42
10	Umbria	76,429,072	73,302,024	496,034	500,483	56,605,780	4.27	−0.89	77.22
9	Tuscany	259,194,496	247,479,792	2,058,064	1,665,684	169,979,328	4.73	23.56	68.68
8	Emilia Romagna	398,939,968	356,597,184	2,158,413	1,768,253	276,977,856	11.87	22.06	77.67
7	Liguria	79,973,288	70,779,688	541,241	599,796	52,516,100	12.99	−9.76	74.20
6	Friuli Ven. Giulia	138,695,824	111,529,800	951,570	658,198	79,282,136	24.36	44.57	71.09
5	Veneto	469,670,144	420,581,472	2,557,998	2,083,220	299,694,272	11.67	22.79	71.26
3	Lombardy	740,176,128	656,650,752	3,509,140	3,098,234	481,084,544	12.72	13.26	73.26
2	Aosta Valley	13,508,645	11,243,489	81,784	115,017	7,561,153	20.15	−28.89	67.25
1	Piedmont	446,466,592	364,934,048	1,571,266	2,205,359	265,508,544	22.34	−28.75	72.76
Italy		4,836,177,382	3,873,485,703	25,177,627	23,607,928	2,702,625,265	24.85	6.65	69.77

A descriptive analysis of processed data also outlines a difference between large municipalities (i.e., major cities) in Italy and the rest of the country. For the whole country, the average accuracy level stands at around 70%; for major cities (Table 3), the accuracy is, on average, significantly higher, reaching 75%. This result is explained by the assumption that the areas with the highest land revenues (namely, major cities in Italy and, more generally, in Southern Europe) are those displaying the most accurate geographical positioning in the land cadaster. As a matter of fact, the most densely populated areas are those where the land cadaster is updated more frequently, resulting in continuously improved positional accuracy. Milan is the area with the highest value of accuracy (87%), followed by Bologna (84%), Turin (83%), and Genoa (82%), all located in Northern Italy. Florence and Venice, two cities with outstanding potential for tourism and second-home dwellings, displayed accuracy values of around 80%. Unexpectedly, Rome showed one of the lowest accuracy levels (68%), together with four medium cities in Southern Italy (Reggio Calabria, Cagliari, Catania, Messina). Naples, Bari, and Palermo showed intermediate accuracy values. Similar to what has been observed for the national building stock, the accuracy experienced a moderate decline when moving from Northern to Southern Italy.

The geographical process of integrating buildings' records from different sources/surveys involves addressing the issue of conflation through geographic integration. This is a process of vertical conflation, in which the land cadaster clearly serves as the main reference source, and buildings from other sources are added only if they do not intersect any buildings already present in the cadaster register. The order in which buildings are added is predetermined. In other words, the registry is built up according to a predefined hierarchy of sources that aims at enhancing the authority (and the documented quality) of the individual (producing) object/element while maximizing the completeness of geographic data. To assess the level achieved, the building footprint in total square meters was calculated for the total building stock. Table 4 highlights the percent composition of the final (integrated) building register's surface area. Starting from the land cadaster, whose building coverage is fully incorporated into the new official statistics' archive, all of the buildings present in other 'donor' sources that do not geometrically intersect with those already in the register were included as new buildings (i.e., not covered in the land cadaster).

By the end of the process, 82% of the surface area of the building stock surveyed in the final register had come from the cadaster, 12% from regional data, and 5% from Civil Protection structural aggregates, while only a minimal proportion had come from OpenStreetMap and the Ministry of the Environment sources. As a reference source, it was expected that the percentage of buildings sourced from the cadaster would be very high in almost all Italian regions. Only Basilicata, Campania, and Latium displayed percentages slightly below 80%. In these cases, the main contributor is regional cartography. Interestingly, the OpenStreetMap source contributed slightly and only in some Northern Italian regions and Tuscany in Central Italy. Considering big cities, the composition of the register was predominantly of cadastral origin. Reggio Calabria has the lowest percentage (82%). In contrast, Turin has 95% of the register's surface area derived from cadastral sources, indicating an almost perfect correspondence between the land cadaster and other information sources (Table 5). In both analyses, regional cartographies were confirmed to be the only relevant—and spatially homogeneous—source of buildings' information in addition to the land cadaster. The remaining sources proved to be marginally informative and possibly representative only in some specific contexts, demonstrating residual local heterogeneity that merits consideration in a specific step of the integration procedure aimed at producing a complete building archive for official statistics' purposes.

Table 3. A comparison (based on indicators) between the information content of the land cadaster and BODA by major city in Italy.

Land Code	City	Boda (m^2) Surface Area	Cadaster (m^2) Surface Area	No. Buildings (Boda)	No. Buildings (Cadaster)	Intersection Area (m^2)	Difference (m^2) between Boda and Cadaster	Difference (no. Build.) Boda vs. Cadaster	Accuracy
		(a)	(b)	(c)	(d)	(e)	$((a - b)/b) \times 100$	$((c - d)/d) \times 100$	$(e/b) \times 100$
1272	Turin	28,518,100	24,276,800	38,132	68,885	20.041.000	17.47	−44.64	82.55
10025	Genoa	14,209,800	14,070,800	64,755	68,216	11.500.600	0.99	−5.07	81.73
15146	Milan	33,504,700	31,250,800	96,123	82,330	27.301.400	7.21	16.75	87.36
27042	Venice	15,991,005	13,527,925	72,423	64,009	10.583.839	18.21	13.15	78.24
37006	Bologna	13,432,200	12,819,000	69,137	38,610	10.753.100	4.78	79.07	83.88
48017	Florence	13,385,500	12,819,400	95,448	64,591	10.352.000	4.42	47.77	80.75
58091	Rome	81,620,496	69,409,904	270,865	276,920	47.480.900	17.59	−2.19	68.41
63049	Naples	22,551,300	17,540,600	47,731	56,767	13.126.900	28.57	−15.92	74.84
72006	Bari	11,107,200	10,225,500	41,735	39,440	8.066.830	8.62	5.82	78.89
80063	Reggio Calabria	8,962,440	7,376,890	52,122	50,882	4.694.370	21.49	2.44	63.64
82053	Palermo	17,736,200	15,302,400	72,820	74,031	10.851.000	15.90	−1.64	70.91
83048	Messina	9,685,540	8,565,330	67,319	63,399	5.552.840	13.08	6.18	64.83
87015	Catania	13,168,700	11,664,900	93,996	45,942	8.031.680	12.89	104.60	68.85
92009	Cagliari	6,101,030	5,082,270	13,512	19,979	3.541.880	20.05	−32.37	69.69
Total Municipalities		289,974,211	253,932,519	1,096,118	1,014,001	191,878,339	14.19	8.10	75.56

Table 4. Percent coverage of integrated cartographic sources in the building register by administrative region in Italy.

Land Code	Administrative Region	Land Cadaster	Ministry of Environment	Regions	OpenStreetMap	Civil Protection	Total
1	Piedmont	92.61	0.39	5.50	0.88	0.62	100
2	Aosta Valley	91.50	0.09	6.31	1.82	0.28	100
3	Lombardy	89.16	0.20	7.62	0.90	2.12	100
5	Veneto	90.55	0.10	5.88	3.48	0.00	100
6	Friuli Venezia Giulia	86.26	0.03	12.45	0.51	0.74	100
7	Liguria	86.55	0.60	3.91	3.25	5.69	100
8	Emilia Romagna	91.32	0.24	2.05	1.00	5.38	100
9	Tuscany	90.62	0.19	3.34	3.98	1.86	100
10	Umbria	94.67	0.12	5.07	0.00	0.14	100
11	Marche	90.03	0.23	7.11	0.00	2.63	100
12	Latium	69.46	1.03	8.27	0.00	21.23	100
13	Abruzzi	92.97	0.27	6.76	0.00	0.00	100
14	Molise	94.53	0.70	4.77	0.00	0.00	100
15	Campania	69.71	0.39	3.91	0.00	25.99	100
16	Apulia	83.16	0.51	16.33	0.00	0.00	100
17	Basilicata	69.72	0.16	30.11	0.00	0.00	100
18	Calabria	79.03	0.48	20.49	0.00	0.00	100
19	Sicily	56.06	0.37	43.57	0.00	0.00	100
20	Sardinia	80.31	0.14	19.54	0.00	0.00	100
Italy		81.81	0.34	12.14	0.94	4.77	100

Table 5. Percent coverage of integrated cartographic sources in the building register by major city in Italy.

Land Code	City	Land Cadaster	Ministry of Environment	Region	OpenStreetMap	Civil Protection	Total
1272	Turin	95.03	2.04	2.52	0.35	0.05	100
10025	Genoa	94.95	1.06	2.86	0.80	0.32	100
15146	Milan	93.32	1.71	4.14	0.72	0.10	100
27042	Venice	93.52	0.25	4.82	1.41	0.00	100
37006	Bologna	93.86	1.42	1.97	0.61	2.15	100
48017	Florence	93.39	0.76	1.99	2.70	1.16	100
58091	Rome	88.12	3.88	8.00	0.00	0.00	100
63049	Naples	83.54	2.49	9.64	0.00	4.33	100
72006	Bari	90.05	2.77	7.18	0.00	0.00	100
80063	Reggio Calabria	82.44	5.76	11.79	0.00	0.00	100
82053	Palermo	87.21	5.29	7.51	0.00	0.00	100
83048	Messina	87.60	5.16	7.24	0.00	0.00	100
87015	Catania	86.81	0.73	12.46	0.00	0.00	100
92009	Cagliari	86.32	1.14	12.54	0.00	0.00	100
Total		89.91	2.69	6.49	0.39	0.52	100

4. Discussion

The increasing use of geographic data in official statistics prompts a reflection on their intrinsic linkage within the qualitative levels of statistics derived from the positioning of geographic data [48]. From this perspective, our work addresses a topic of great interest: the harmonization of geographic information, particularly cadastral information related to buildings, for the generation of complete databases addressing official statistics'

requirements at the national level, e.g., [49]. From a modern perspective that considers well-defined reference systems and includes the corresponding quality measures, this is an operational issue of interest in a large number of countries [50]. In this context, two approaches can be considered to be of special interest: the first is the use of frameworks that propose a methodological contribution, i.e., how to match and make different cartographies and lists of alphanumeric information unique in a single geographic database that integrates the different components [51]. This is a very interesting subject, since, in the end, it is necessary to correct different types of errors that are not always homogeneous. These errors frequently derive from the fact that such cartographies usually have a complex history as a background, in which different methodologies, instruments, tools, and procedures are brought together [52]. The second approach is to check the real quality of the integrated information [53]. This is a scheme that may be methodologically articulated (namely, requiring the accurate definition of quality measures), but it is certainly not easy to apply because, as mentioned above, errors do not have to be continuous, and in some cases, we may even have specific errors that are difficult to detect and delineate [54].

The linkage between statistical units and territorial units involves the geocoding process, where the statistical unit is matched, for instance, to a census tract. It is essential to consider that a geographic datum is the statistical characteristic measured and recorded at a given location. This involves pairing a statistical unit that belongs to the collective under investigation with a given territorial unit. The latter, at various levels of resolution, determines the positioning of the surveyed statistical unit in the territory [1]. In the context of building registers, various statistical units, such as dwellings, buildings, and addresses, will be identified and located in a given territory with a specific position, determining the maximum level of resolution [2]. As the precision of the positioning decreases, these statistical units will be geocoded at the level of, e.g., parcels, census sections, cadastral sheets, and municipalities [6]. From a qualitative standpoint, this involves ensuring a series of considerations related to the statistical unit and its consistency with the related territorial unit [11].

A specific assessment should be carried out for those buildings recorded in any specific survey that is not a land cadaster and absent from this last source [49]. These elements should not necessarily be interpreted as illegal/spontaneous/informal constructions (i.e., outside the scope of land cadaster). While these buildings may not be uncommon in the landscape of Southern Europe [34], a well-known region because of limited planning policies and limited regulations and controls between the 1950s and the 1970s [32], all of these elements should be properly identified to determine whether they can be primarily associated with the information available in the cadaster [14]. Lastly, it should be considered that a significant number of added elements from ancillary sources on a particular cadaster map sheet could (indirectly) document (systematic) poor positioning within the cadaster data source [27]. This discussion should benefit from the assumption that data quality is a substantial characteristic of any given geographical source, because the production process differs between a cadaster and crowdsourced geographic information, such as OpenStreetMap [54].

The assumption is that an OpenStreetMap user produces less accurate spatial data compared to those produced by cartographic (public) institutions, which have dedicated (and skilled) personnel (enumerators), appropriate field tools/instruments, and a clear official commitment to operate [53]. Additionally, geographical elements (e.g., buildings) derived from administrative processes and surveys (both statistical and non-statistical) are subject to a series of checks, while open data typically undergo no specific controls [44]. If this is acceptable, the quality assessment becomes more challenging when comparing a geographical element from a cadastral source with one from a regional administrative source [36]. Having the accuracy of the building register as the main operational target with official statistical purposes, the evaluation should prioritize the maximization of the specific use of cadastral sources, uniquely including references that serve as key linkage with other archives [38].

At present, it is difficult to propose a specific evaluation process for data quality characteristic of various and multiple geographical sources for buildings without considering the land cadaster as the basic source [25]. The traceability of the origin of any geographical element is a fundamental aspect of any registry and can be operationalized from multiple perspectives [46]. The first is the integration of (possibly divergent) information sources, which is a core dimension of the registry system [20]. Different integrated sources may (or may not) have a key link, namely, a primary key in the language of databases [48]. The absence of an individual building in the cadaster does not mean that real estate units cannot be associated with that parcel based on cadastral anchoring [50]. Therefore, while integrating (missing) geographical elements is crucial for completeness (and, thus, accuracy), it is essential to maintain a clear record of the 'lineage'. In particular, 'lineage' refers to the organization of information contained in metadata, with specific reference to (i) the producing entity of the cartography used, (ii) the original reference system and projection, (iii) the spatial characteristics of the original database, including coverage, scale, and resolution, and (iv) license of use.

In cases of integration, buildings will have derived cadastral codes, meaning they will only reference the cadastral parcel in which they are located [51]. With this perspective in mind, an evolutionary perspective aimed at keeping track of the development of the various components that have contributed to the creation of the registry can be adopted [26]. Another aspect to be taken into account is the need to establish (improve or consolidate) references to European regulations in terms of digital data management. In this regard, the INSPIRE directive marks some aspects that should be operationally considered not only in the field of geographic data but also in the specific dimension of official statistics and, more precisely, the ISO 19157/2013 standard (INSPIRE directive), generally delineating quality measures and their definition. The same applies to positional accuracy, for which we would like to recommend the use of the Accuracy Standards for Digital Geospatial Data (ASPRS) Guidelines, especially the most recent version. Operational means and tools should be developed from this perspective when integrating data sources that only partly meet such criteria, such as the land cadaster in Italy [21].

In conclusion, the main issue of this contribution is the need to envisage accuracy and completeness as two faces of the same medal but, at the same time, a very complex and multi-dimensional challenge, including problems of translation between different layers of information [39]. They may involve multiple criteria and technical aspects, including (different) projection systems and data that require extensive harmonization work. Completely automatic procedures allowing translation from the (more or less) complete knowledge of the source coordinate system to the target coordinate system are particularly relevant in this direction [37]. Finally, renewed dissemination eliciting a complete awareness of their importance in the world of official statistics is particularly appropriate, especially in Southern European statistical systems.

5. Conclusions

A reflection on the potential of using geographic analysis to improve official statistics and, thus, census operations can be derived from the comparative scrutiny of the empirical results presented in this article. These findings reveal evident complexity in the analysis plan when official statistics' criteria for completeness, accuracy, and precision have to be applied to (non-statistical) public sources of spatially explicit information. However, while individual indicators—appropriately presented and calculated here with the aim of a preliminary assessment of data source integration—may highlight significant differences, these tend to diminish with the level, intensity, and pervasiveness of data integration procedures. Therefore, a simplified, comparative procedure has been adopted in order to ensure the completeness of the survey for the construction of an integrated building registry. According to this approach, a reference source has been identified (Regional Technical Maps, in this case), and buildings from other (ancillary) sources have been added, controlling for the eventual overlap with those buildings previously included in

the statistical register. Those records displaying an (even minimal) overlap with entities already present in the register and previously 'donated' from the most representative source are to be removed. Overall, the register that was built after these operational steps has achieved acceptable quality standards. Based on these findings, our study finally encourages the implementation of new (automatic) geo-referencing techniques for the integration of buildings' official statistics and field surveys and the related (improved) control processes.

Author Contributions: Conceptualization, G.S. and D.S.; methodology, G.S.; software, M.M.; validation, M.M.; formal analysis, S.B.; investigation, K.R.; resources, A.M.; data curation, M.M.; writing—original draft preparation, A.S. and G.S.; writing—review and editing, M.M. and K.R.; visualization, M.M.; supervision, D.S.; project administration, S.B.; funding acquisition, S.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Official statistics released by multiple national and communitarian authorities, including the Italian National Statistical Institute (Istat), the National Institute for Environmental Research and Protection (Ispra), the European Environment Agency (Eea), and Eurostat, were used in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cosgrove, D. *Geography and Vision: Seeing, Imagining and Representing the World*; Bloomsbury: London, UK, 2012.
2. Muller, J.-C. The Concept of Error in Cartography. *Cartographica* **1987**, *24*, 1–15. [CrossRef]
3. Jun, C. Design of an Intelligent Geographic Information System for Multi-Criteria Site Analysis. *URISA J.* **2000**, *12*, 5–17.
4. Girres, J.F.; Touya, G. Quality assessment of the French OpenStreetMap dataset. *Trans. GIS* **2010**, *14*, 435–459. [CrossRef]
5. Mondzech, J.; Sester, M. Quality analysis of OpenStreetMap data based on application needs. *Cartogr. Int. J. Geogr. Inf. Geovisualization* **2011**, *46*, 115–125. [CrossRef]
6. Bluemke, M.; Resch, B.; Lechner, C.; Westerholt, R.; Kolb, J.-P. Integrating Geographic Information into Survey Research: Current Applications, Challenges and Future Avenues. *Surv. Res. Methods* **2017**, *11*, 307–327.
7. Fan, H.; Zipf, A.; Fu, Q.; Neis, P. Quality assessment for building footprints data on OpenStreetMap. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 700–719. [CrossRef]
8. Basiouka, S.; Potsiou, C.; Bakogiannis, E. OpenStreetMap for cadastral purposes: An application using VGI for official processes in urban areas. *Surv. Rev.* **2015**, *47*, 333–341. [CrossRef]
9. Turner, A. *Introduction to Neogeography*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2006. Available online: https://books.google.com.hk/books?id=oHgDv4feV-8C&printsec=frontcover&redir_esc=y#v=onepage&q&f=false (accessed on 1 September 2024).
10. Haklay, M.; Weber, P. OpenStreetMap: User-Generated Street Maps. *IEEE Pervasive Comput.* **2008**, *7*, 12–18. [CrossRef]
11. Spielman, S.E.; Logan, J.R. Using High-Resolution Population Data to Identify Neighborhoods and Establish Their Boundaries. *Ann. Assoc. Am. Geogr.* **2013**, *103*, 67–84. [CrossRef]
12. Kazemzadeh-Zow, A.; Zanganeh Shahraki, S.; Salvati, L.; Samani, N.N. A spatial zoning approach to calibrate and validate urban growth models. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 763–782. [CrossRef]
13. Pili, S.; Grigoriadis, E.; Carlucci, M.; Clemente, M.; Salvati, L. Towards sustainable growth? A multi-criteria assessment of (changing) urban forms. *Ecol. Indic.* **2017**, *76*, 71–80. [CrossRef]
14. Duvernoy, I.; Zambon, I.; Sateriano, A.; Salvati, L. Pictures from the other side of the fringe: Urban growth and peri-urban agriculture in a post-industrial city (Toulouse, France). *J. Rural Stud.* **2018**, *57*, 25–35. [CrossRef]
15. Zambon, I.; Colantoni, A.; Salvati, L. Horizontal vs vertical growth: Understanding latent patterns of urban expansion in large metropolitan regions. *Sci. Total Environ.* **2019**, *654*, 778–785. [CrossRef]
16. Sui, D.Z. The Wwikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. *Comput. Environ. Urban Syst.* **2008**, *32*, 29. [CrossRef]
17. Ciommi, M.; Chelli, F.M.; Salvati, L. Integrating parametric and non-parametric multivariate analysis of urban growth and commuting patterns in a European metropolitan area. *Qual. Quant.* **2019**, *53*, 957–979. [CrossRef]
18. Guo, J.Y.; Bhat, C.R. Modifiable Areal Units: Problem or Perception in Modeling of Residential Location Choice? *Transp. Res. Rec.* **2004**, *1898*, 138–147. [CrossRef]

19. Burrough, P.A.; McDonnell, R.A.; Lloyd, C.D. *Principles of Geographical Information Systems*; Oxford University Press: Oxford, MS, USA, 2015.
20. Salvucci, G.; Salvati, L. Official Statistics, Building Censuses, and Open Street Map Completeness in Italy. *ISPRS Int. J. Geo-Inf.* **2021**, *11*, 29. [CrossRef]
21. Salvati, L. Towards a Polycentric Region? The Socio-economic Trajectory of Rome, an 'Eternally Mediterranean' City. *Tijdschr. Voor Econ. En Soc. Geogr.* **2014**, *105*, 268–284. [CrossRef]
22. Thompson, M.M. *Maps for America: Cartographic Products of the U.S. Geological Survey and Others*; U.S. Government Print. Office: Washington, DC, USA, 1988.
23. Guzzo, F. Italian Land Cadastral Cartographic Representation of Public Waters for a Contribution to Territorial Governance: Learning from History. *UPLanD—J. Urban Plan. Landsc. Environ. Des.* **2018**, *3*, 45–68.
24. Brovelli, M.A.; Minghini, M.; Giori, G.; Beretta, M. Web Geoservices and Ancient Cadastral Maps: The Web C.A.R.T.E. Project. *Trans. GIS* **2012**, *16*, 125–142. [CrossRef]
25. Cina, A.; Manzino, A.M.; Manzino, G. Recovery of Cadastral Boundaries with GNSS Equipment. *Surv. Rev.* **2016**, *48*, 338–346. [CrossRef]
26. Bajocco, S.; Ceccarelli, T.; Smiraglia, D.; Salvati, L.; Ricotta, C. Modeling the ecological niche of long-term land use changes: The role of biophysical factors. *Ecol. Indic.* **2016**, *60*, 231–236. [CrossRef]
27. Salvati, L.; Ciommi, M.T.; Serra, P.; Chelli, F.M. Exploring the spatial structure of housing prices under economic expansion and stagnation: The role of socio-demographic factors in metropolitan Rome, Italy. *Land Use Policy* **2019**, *81*, 143–152. [CrossRef]
28. Carlucci, M.; Chelli, F.M.; Salvati, L. Toward a new cycle: Short-term population dynamics, gentrification, and re-urbanization of Milan (Italy). *Sustainability* **2018**, *10*, 3014. [CrossRef]
29. Ciommi, M.; Chelli, F.M.; Carlucci, M.; Salvati, L. Urban growth and demographic dynamics in southern Europe: Toward a new statistical approach to regional science. *Sustainability* **2018**, *10*, 2765. [CrossRef]
30. Zambon, I.; Benedetti, A.; Ferrara, C.; Salvati, L. Soil matters? A multivariate analysis of socioeconomic constraints to urban expansion in Mediterranean Europe. *Ecol. Econ.* **2018**, *146*, 173–183. [CrossRef]
31. Bajocco, S.; Dragoz, E.; Gitas, I.; Smiraglia, D.; Salvati, L.; Ricotta, C. Mapping forest fuels through vegetation phenology: The role of coarse-resolution satellite time-series. *PLoS ONE* **2015**, *10*, e0119811. [CrossRef]
32. Salvati, L.; Zitti, M.; Sateriano, A. Changes in city vertical profile as an indicator of sprawl: Evidence from a Mediterranean urban region. *Habitat Int.* **2013**, *38*, 119–125. [CrossRef]
33. Salvati, L.; Guandalini, A.; Carlucci, M.; Chelli, F.M. An empirical assessment of human development through remote sensing: Evidences from Italy. *Ecol. Indic.* **2017**, *78*, 167–172. [CrossRef]
34. Lü, G.; Batty, M.; Strobl, J.; Lin, H.; Zhu, A.-X.; Chen, M. Reflections and Speculations on the Progress in Geographic Information Systems (GIS): A Geographic Perspective. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 346–367. [CrossRef]
35. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; CRC: London, UK, 2019.
36. Guptill, S.C.; Morrison, J.L. *Elements of Spatial Data Quality*; Elsevier: Amsterdam, The Netherlands, 2013; ISBN 978-1-4832-8794-2.
37. Alexandroff, P. *Elementary Concepts of Topology*; Courier Corporation: New York, NY, USA, 2012.
38. Shi, W.; Fisher, P.; Goodchild, M.F. *Spatial Data Quality*; Elsevier: Amsterdam, The Netherlands, 2019; ISBN 978-1-134-51440-3.
39. Gieryn, T.F. What Buildings Do. *Theory Soc.* **2002**, *31*, 35–74. [CrossRef]
40. Conzen, M.R. *Thinking about Urban Form: Papers on Urban Morphology, 1932–1998*; Peter Lang: Frankfurt, Germany, 2004.
41. Hart, Q.J. GeoStreams: An Online Geospatial Image Database. *Diss. Abstr. Int.* **2006**, *68*.
42. Chrisman, N.R. The Error Component in Spatial Data. *Geogr. Inf. Syst.* **1991**, *1*, 165–174.
43. Hecht, R. Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 1066–1091. [CrossRef]
44. Ratcliffe, J.H. On the Accuracy of TIGER-Type Geocoded Address Data in Relation to Cadastral and Census Areal Units. *Int. J. Geogr. Inf. Sci.* **2001**, *15*, 473–485. [CrossRef]
45. Andersson, R.; Musterd, S. What Scale Matters? Exploring the Relationships between Individuals' Social Position, Neighbourhood Context and the Scale of Neighbourhood. *Geogr. Ann. Ser. B Hum. Geogr.* **2010**, *92*, 23–43. [CrossRef]
46. Wang, R.Y.; Reddy, M.P.; Kon, H.B. Toward Quality Data: An Attribute-Based Approach. *Decis. Support Syst.* **1995**, *13*, 349–372. [CrossRef]
47. Salvati, L.; Ferrara, A.; Chelli, F. Long-term growth and metropolitan spatial structures: An analysis of factors influencing urban patch size under different economic cycles. *Geogr. Tidsskr.—Dan. J. Geogr.* **2018**, *118*, 56–71. [CrossRef]
48. Craig, B.A.; Wahl, J.L. Cadastral Survey Accuracy Standards. *Surv. Land Inf. Sci.* **2003**, *63*, 87–106.
49. Hill, L.L. *Georeferencing: The Geographic Associations of Information*; Mit Press: Boston, MA, USA, 2009.
50. Femenia-Ribera, C.; Mora-Navarro, G.; Martinez-Llario, J.C. Advances in the Coordination between the Cadastre and Land Registry. *Land* **2021**, *10*, 81. [CrossRef]
51. Apostolopoulos, K.; Geli, M.; Petrelli, P.; Potsiou, C.; Ioannidis, C. A New Model for Cadastral Surveying Using Crowdsourcing. *Surv. Rev.* **2018**, *50*, 122–133. [CrossRef]
52. Ma, D.; Sandberg, M.; Jiang, B. Characterizing the heterogeneity of the OpenStreetMap data and community. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 535–550. [CrossRef]

53. Goodchild, M. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [CrossRef]
54. Ahmouda, A.; Hochmair, H.H.; Cvetojevic, S. Analyzing the effect of earthquakes on OpenStreetMap contribution patterns and tweeting activities. *Geo-Spat. Inf. Sci.* **2018**, *21*, 195–212. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.