

Geocoding COR*-Antwerpen Database*

Diogo Paiva

Early-stage researcher of the LONGPOP project at the International Institute of Social History of the Koninklijke Nederlandse Akademie van Wetenschappen (KNAW)

Contents

Summary	2
Goals.....	2
Data & Resources	2
Methodology.....	3
Results	6
Attachment 1 – Success per municipality	11
Attachment 2 – Websites used as sources information on old street names	13
Attachment 3 – Evolution of the Municipalities of Antwerp Arrondissement (1856-1930).....	14

* Thanks to Hideko Matsuo and Sam Jenkinson for their reviews and comments.

Summary

This report describes the process of geocoding the historical database of COR*-Antwerpen. This is part of the Expected Result 11.4 - *Building algorithms to implement a GIS based address system for the two databases that are the most suited to start with, result: adding variables with GIS coordinates.*

The COR*-Antwerpen database gathers the information on a sample of individuals whose first three letters of the name are COR. It provides individual data on biography, family relations and residential history from population records from the 19th and early 20th century. Addresses collected in this database amount to 14,536 records.

Information on residence is expressed as text strings of varied types: municipality, neighbourhoods, streets, and house numbers. Noteworthy is the fact that the entries of addresses have already underwent a process of standardization. Georeferenced data is limited to coordinates of municipalities. Therefore, this project intends to georeference in greater detail the residence data included in the database by adding a coordinate for each household representing the centroid of the street alongside a revised coordinate of the municipality. Inclusion of geocoded information in the forthcoming release of COR- is one of the most important asset to the data set allowing additional research pertaining spatial analysis.

Goals

Revise the standards of street names and municipalities already present in the database

Add coordinates to each address, in a twofold system (street centroid¹ and municipality²)

Data & Resources

The georeference process will use the table from the COR*- database named huisSAMEN2 extracted from August 2010 COR* release. This table contains several fields, including:

- ID code for **household** (identificatiehuis)
- Original and standardized **municipality name** (gemeente / gem)
- Original and standardized **Year** of the source where information was retrieved - *bevolkingsregister* (bevolkingsregister / BR)
- Original and standardized **Wijk name** (wijknummer / wk)
- Original and standardized **house number in wijk** (wijkhuisnummer / wkhsnr)
- Original and standardized **street name** (straat/ str)
- Original and standardized **street house number** (hsnr/ huisnummer)

Additionally, two other sets of data were gathered:

- Municipality borders of Antwerp arrondissement (1856-1930)³

¹ I.e., the median point of the line representation of the street.

² A point coordinate representing the location of the administrative-political centre of the municipality.

³ Vrielinck; S.; Wiedemann; T.; Deboosere; P., *HISGIS België 1800-2000*.

- Historical streets for the year 1898, from the GISTorical Antwerp (UAntwerpen/Hercules Foundation)⁴ project.

For the execution of this work, the following available resources were used:

- Google search engine, for searching webpages with information pertaining old street names (see Attachment 2)
- Google Maps and Esri's locator (ArcGIS Pro), for finding streets and other landmarks
- ArcGIS Pro was used to draw shapefiles (lines) of street layouts
- [Geopunt.be](https://geopunt.be)'s historical maps (Atlas der Buurtwegen (1841), Vandermaelen kaarten (1846 – 1854) and Popp kaarten (1842 – 1879)) to identify old streets that no longer exist and landmarks and their layout

Methodology

The process of georeferencing the addresses in the COR*-Antwerpen database involves performing several sub-processes of increasing complexity and time consumption. The basic philosophy is to link between textual information on addresses (municipality and street) and geographic data, in the most efficient and less time-consuming form. Therefore, primarily we use existing geo-databases and only for the records lacking information, a more thorough process is applied.

Given that the huisSAMEN2 table already provides standards for the original municipality (**gem**) and street (**str**) names derived from the sources, at first, a simple record linkage (exact matching) seemed to be enough to georeference most addresses. For the remaining addresses, further research would aid in defining the suitable coordinates.

However, it was noted that the normalization process executed to provide standards in **gem** and **str** fields was imperfect. For example: "Driesch", "Driessche", "Drieschstraet" and "Dries Straat" in Antwerp were all being standardized as "Driesstraat" while the correct spelling is "Dries"; more significantly, "Hagelkruisstraat", "Hagelkruis" (both as "Hagelruis"), "Gr. Hagelkruisstraat" (as "Klein Hagelkruis") were being divided into two different streets although all are forms of "Groot Hagelkruis" and the latter was transformed from Groot ("Gr.") to Klein. Therefore, a revision of the standards provided in huisSAMEN2 was necessary as an intermediary step between exact matching and further research, for those addresses that failed to link directly.

Moreover, the geographic data available (GISTorical Antwerp 1898 street shapefile) only covered the city of Antwerp and some of its immediate surroundings. Considering this gap of information, an effort was done to provide spatial information that covered the entire Antwerp arrondissement. This resulted in the creation of two other shapefiles (lines): one to link current streets with huisSAMEN2 table; and another for historical streets (i.e. no longer existing). Figures 1 and 2 show an example of how the landscape of the former town of Oorderen completely changed with the development of Antwerp's port area.

⁴ We would like to thank Iason Jongepier and the GISTorical Antwerp project for providing access to historical data used in the georeference process of COR*-Antwerpen database.

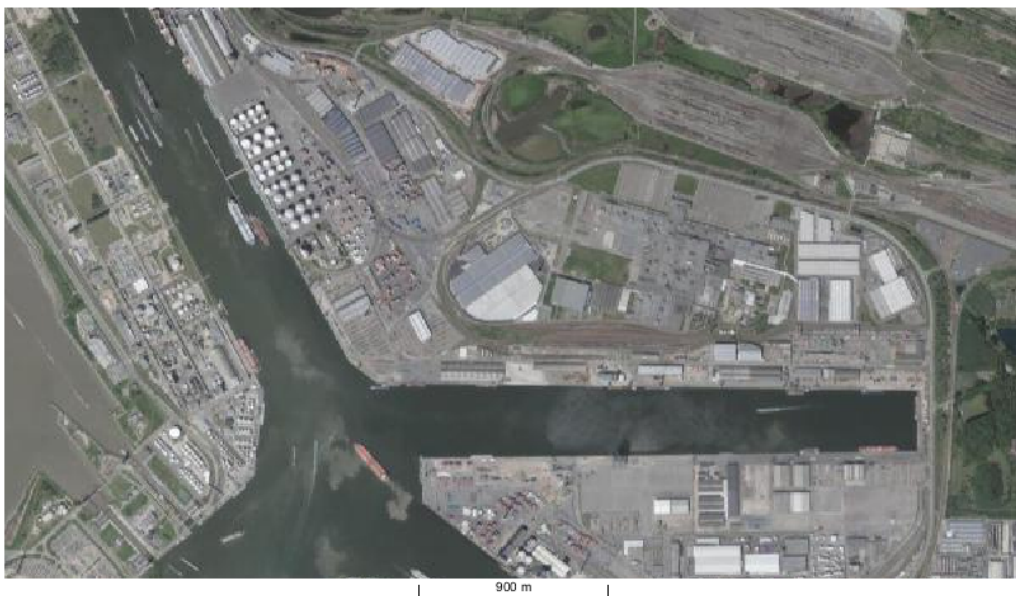


Bron: www.geopunt.be

10/05/2019

1

Figure 1 - Oorderen area 19th century (Van Popp map – source: www.geopunt.be)



Bron: www.geopunt.be

10/05/2019

1

Figure 2 - Oorderen area today (source: www.geopunt.be)

Finally, the process can only be complete with the use of conversion tables. These act as dictionaries between the original name in huisSAMEN2 (**straat**) and an alternative standard (in case the street still exists today with the same name - **fix_str**) or the corresponding modern name (if the street name in the sources is outdated - **hist_str**). The latter is based on online research (diverse media, including blogs, forums, and historical maps) and consultation of a docx file provided by GISTorical Antwerp project containing an extensive list of old street names (mostly for Antwerp and its surroundings)⁵ based on Robert vande Weghe's *Geschiedenis van de Antwerpse straatnamen* (1977). For the revision of street name standards, Google (Google Maps) and Esri's (ArcGIS Pro) search engines were used. A last table was created which manually assigns coordinates to specific streets (in the case that it was not possible to draw the street in the shapefile, but there is a textual description of a broad localization – e.g., between the known streets A and B) or landmarks (using historical maps).

Besides the 1898 streets of Antwerp database (1) provided by the GISTorical Antwerp project, two other sets (lines) were created: actual streets that still exist today (2); and historical streets that existed in the past, based on old maps (3). For each street of these sets a medium point was obtained. Afterwards, these datasets were exported to three .csv files (that include street name, municipality name, longitude and latitude of the medium street point).

Using the geographic data and the two conversion tables, the process performs rounds that link the street names present in huisSAMEN2 with those in the aforementioned geographic three datasets. Table 1 shows the order of the process sequence as it geocodes the different addresses. The huisSAMEN2 variables that contain names of streets (**str** and **straat** from the COR database plus the added alternatives from conversion tables, **fix_str** and **hist_str**) are matched with names contained in the spatial datasets. After going through in total 12 steps, links were sought manually. Finally, for the records successfully linked street coordinates are added (**s_lon** and **s_lat**).

Geographic data Variable used	GISTorical 1898 street layer	Antwerp Actual street layer	Historical street layer	Manual coordinates
COR*-Antwerpen standardized street names (str)	1	2	3	-
COR*-Antwerpen original street name (straat)	4	5	6	13
Alternative street standards, provided by conversion table (fix_str)	7	8	9	-
Historical converted names, provided by conversion table (hist_str)	10	11	12	-

Table 1 - Order of geocoding process for streets

⁵ Covers parts of Berchem, Hoboken, Borgerhout and Oosterweel.

Results

Assigning coordinates for the municipalities was a very quick and simple process, only requiring minor revisions to the municipalities' names. The centres of the municipalities were defined by extracting the coordinates of the municipal capitals in the [geonames.org](https://www.geonames.org) database.

The overall success in georeferencing the addresses in COR*-Antwerpen database, at street level, was around 85%. Figure 3 shows the evolution of the georeferencing process in its different stages using the COR*-Antwerpen database's street standards, COR*-Antwerpen database's original street names, alternative street standards, conversions of historical street names into current names and manual coordinate assignment, sequentially.

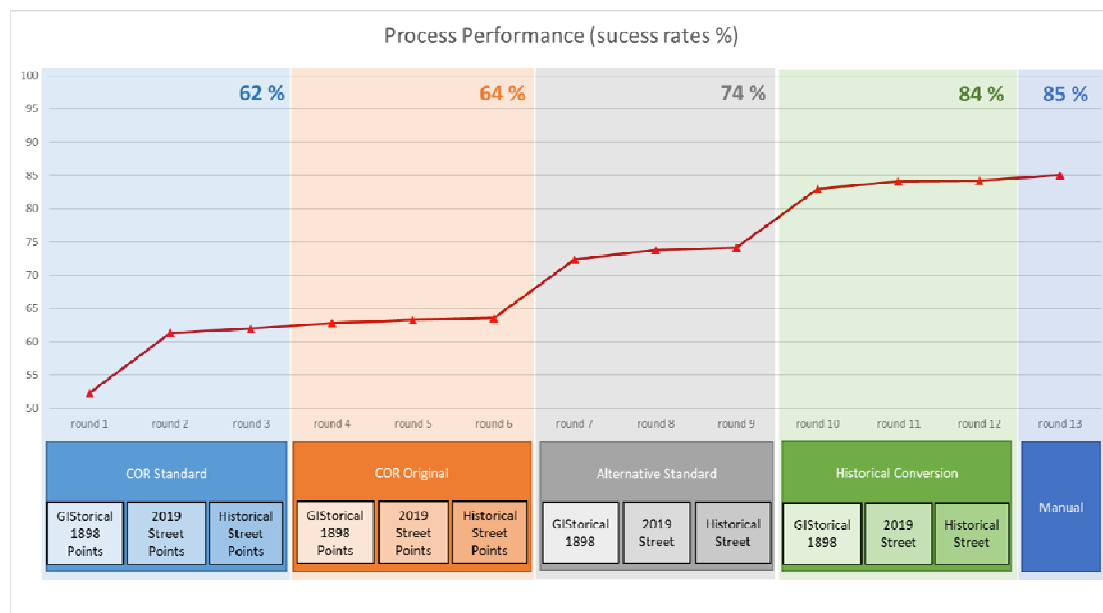


Figure 3 - Process performance

The best results are found in the municipalities of Antwerpen (96%), Hoevenen (86%) and Essen (85%), while the lowest are Sint-Job-in-'t-Goor and Halle, both with 0% found addresses (Figure 4). Given the amount of addresses from Antwerp (68 % of total addresses in COR*-Antwerpen database) and the success rate of the georeferencing process it is noteworthy the decisive contribution of Antwerp to the general high success rate. The accessed resources (GISTorical Antwerp layer file and document on old street names) and a better quality of the sources relating to Antwerp explain this. Nevertheless, for the remaining municipalities, the average success rate is expressively lower, reaching only 64%.

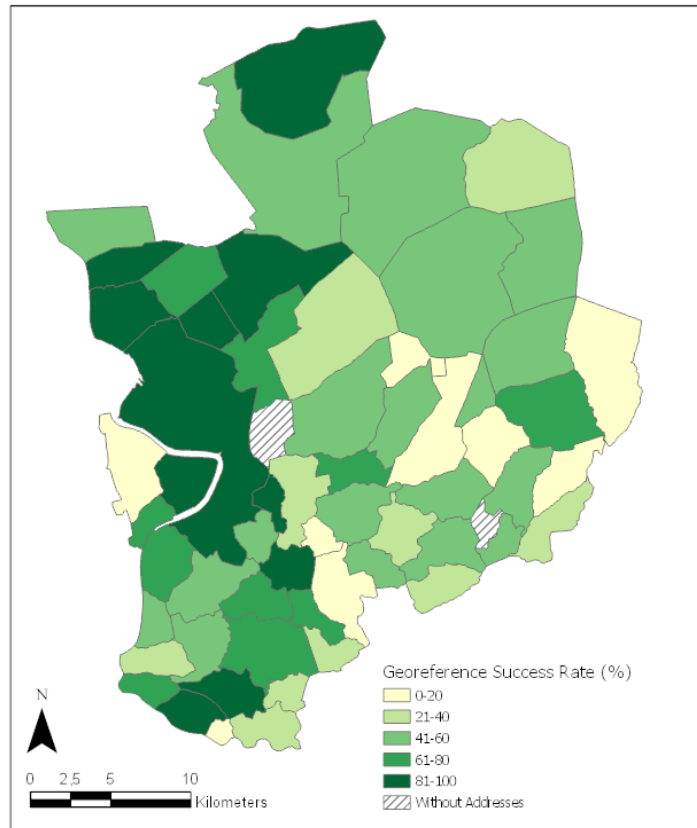
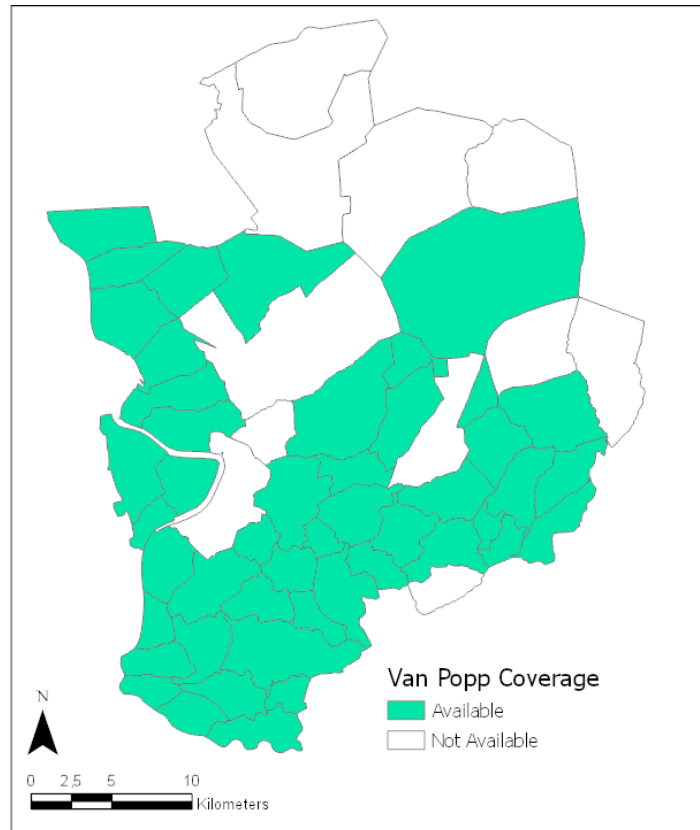


Figure 4 - Rates of successfully georeferenced addresses per municipality (borders of 1930)

Furthermore, the variation on the available information about addresses outside of Antwerp conducted to an uneven success. For instance, historical maps are provided by geopunt.be. However, the most detailed (Van Popp 1842-1879), and often the only source for locating streets, does not cover all of the Antwerp arrondissement (Figure 5). Another determinant factor in not obtaining a higher success in georeferencing the addresses was the industrial and urban development of Belgium during the 19th and 20th centuries, which had a significant impact on the country's landscape. The combination of large morphological changes in urban and rural settings, related to both population and industrial growth, with not having access to historical maps from the end of the 19th century and early 20th century, resulted in a lack of information and consequently a lower amount of addresses that can be geocoded. For instance, the expansion of Antwerp's harbour completely razed whole villages in former Oorderen, Oosterweel and Wilmarsdonk municipalities.



*Figure 5 - Coverage of Van Popp maps
(borders of 1880, and also including Burcht, Zwijndrecht and west bank of Antwerp of 1930)*

A clear distinction in the process' performance relates with the urban or rural setting of the municipality (Figure 6). Addresses from rural municipalities are more frequently composed of just a location (name of the village) and a house number. In these cases, usually the location is the same of the municipality; therefore, the address will only have the municipality's coordinates. Additionally, small villages from the 19th century tend to present larger morphological changes than older towns and this leads to an increased difficulty in finding old and/or disappeared streets and their layouts. Take Sint-Job-in-'t-Goor in consideration, in the middle 19th century it was basically structured around Kerkstraat (nowadays Kerklei), with a few small alleys. While the main street still exists, with urban development, large buildings stand in place of those alleys and new streets were opened nearby. In contrast, in Deurne or Hoboken, examples of more developed towns in the 19th century the older street structure mostly remains in place today.

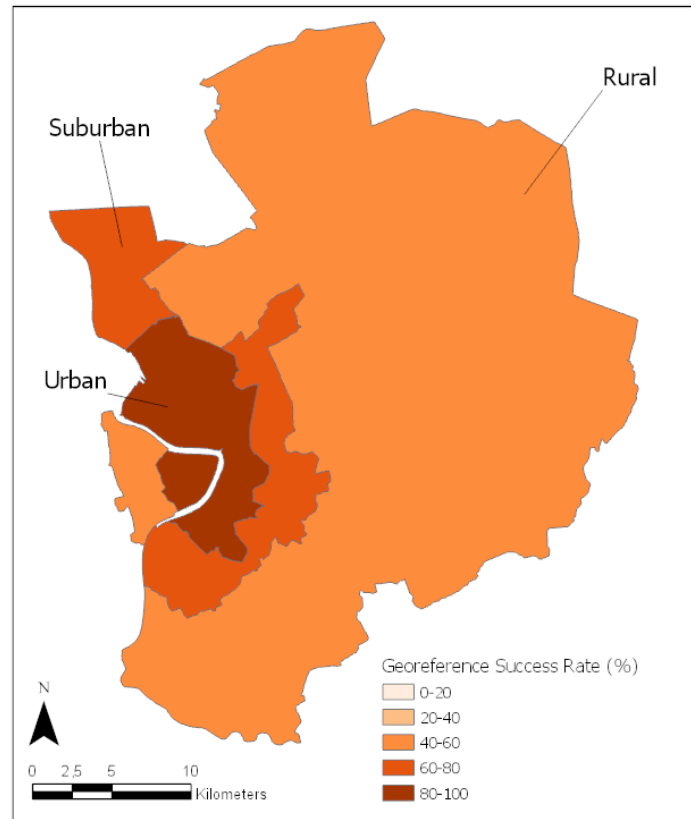


Figure 6 - Georeference Success Rate in Urban, Suburban and Rural settings⁶ (borders of 1930)

Figure 7 presents the geocoded COR*-Antwerpen database addresses for its closest coordinates (**n_lon** and **n_lat**). The predominance is clearly in the city of Antwerp, however the dispersion of addresses in other municipalities enables further research on migration movement between rural and urban settings, as well as allowing research into intra-city migration. Contextual data for neighbourhoods in Antwerp, for instance, permits comparing correlation between results of social mobility status already performed for the COR*-Antwerpen database with the new geographic data.

⁶ Based on the division presented by Matthijs, K., & Moreels, S. (2010). The Antwerp COR*-database: a unique Flemish source for historical-demographic research. *The History of the Family*, 15(1), 111, Figure 1.

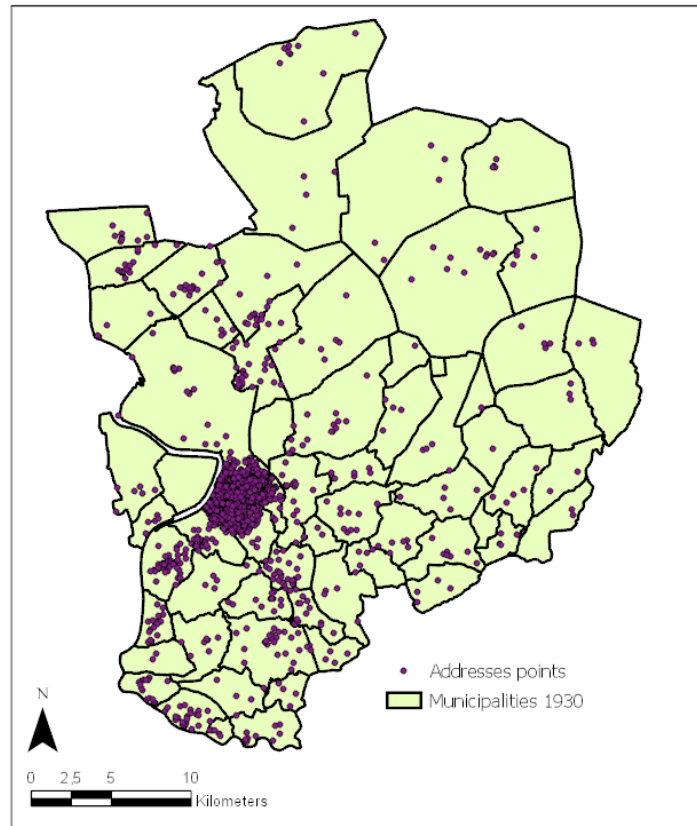


Figure 7 - Points representing COR*-Antwerpen addresses (borders of 1930)

Attachment 1 – Success per municipality

Municipality	Addresses	Addresses %	Georeferenced addresses at street level	Georeferenced %	Type of Setting
's-Gravenwezel	9	0.1	5	55.6	rural
Aartselaar	21	0.1	9	42.9	rural
Antwerpen	9,870	67.9	9,427	95.5	urban
Berchem	540	3.7	310	57.4	suburban
Berendrecht	107	0.7	87	81.3	suburban
Boechout	68	0.5	13	19.1	rural
Boom	248	1.7	205	82.7	rural
Borgerhout	1,320	9.1	1,088	82.4	suburban
Borsbeek	10	0.1	1	10.0	rural
Brasschaat	54	0.4	21	38.9	rural
Brecht	49	0.3	24	49.0	rural
Broechem	34	0.2	18	52.9	rural
Burcht	9	0.1	7	77.8	rural
Deurne	198	1.4	54	27.3	suburban
Edegem	32	0.2	22	68.8	rural
Ekeren	95	0.7	75	78.9	suburban
Emblem	12	0.1	4	33.3	rural
Essen	26	0.2	22	84.6	rural
Halle	1	0.0	0	0.0	rural
Hemiksem	62	0.4	28	45.2	rural
Hoboken	336	2.3	232	69.0	suburban
Hoevenen	42	0.3	36	85.7	rural
Hove	45	0.3	32	71.1	rural
Kalmthout	10	0.1	5	50.0	rural
Kapellen	104	0.7	85	81.7	rural
Kontich	82	0.6	57	69.5	rural
Lillo	61	0.4	51	83.6	suburban
Lint	23	0.2	9	39.1	rural
Loenhout	22	0.2	6	27.3	rural
Mortsel	92	0.6	77	83.7	rural
Niel	59	0.4	45	76.3	rural
Oelegem	34	0.2	14	41.2	rural
Oorderen	28	0.2	16	57.1	urban
Oosterweel	47	0.3	18	38.3	urban
Oostmalle	11	0.1	2	18.2	rural
Pulderbos	10	0.1	1	10.0	rural
Pulle	23	0.2	6	26.1	rural
Ranst	35	0.2	11	31.4	rural
Reet	6	0.0	5	83.3	rural
Rumst	45	0.3	16	35.6	rural
Schelle	24	0.2	5	20.8	rural

Schilde	15	0.1	2	13.3	rural
Schoten	35	0.2	20	57.1	rural
Sint-Job-in-'t-Goor	5	0.0	0	0.0	rural
Sint-Lenaarts	14	0.1	6	42.9	rural
Stabroek	117	0.8	78	66.7	rural
Terhagen	47	0.3	7	14.9	rural
Viersel	7	0.0	3	42.9	rural
Vremde	11	0.1	5	45.5	rural
Waarloos	27	0.2	6	22.2	rural
Westmalle	12	0.1	6	50.0	rural
Wijnegem	26	0.2	16	61.5	rural
Wilmarsdonk	23	0.2	7	30.4	urban
Wilrijk	91	0.6	11	12.1	suburban
Wommelgem	65	0.4	33	50.8	rural
Wuustwezel	17	0.1	7	41.2	rural
Zandhoven	11	0.1	5	45.5	rural
Zandvliet	45	0.3	26	57.8	suburban
Zoersel	4	0.0	3	75.0	rural
Zwijndrecht	60	0.4	6	10.0	rural
Total	14,536	100.0	12,396	85.3	

Attachment 2 – Websites used as sources information on old street names

Municipality	Website
Berchem	<ul style="list-style-type: none"> • https://books.google.nl/books?id=aTWGDwAAQBAJ • https://nl.wikipedia.org/wiki/Mechelsepoort
Berendrecht	<ul style="list-style-type: none"> • https://www.nieuwsblad.be/cnt/blsde_02878604
Brasschaat	<ul style="list-style-type: none"> • http://www.marek.be/familie/fotos36.html • http://www.kvoo-antwerpen.be/
Borgerhout	<ul style="list-style-type: none"> • http://krugerpleinpeperbus.blogspot.com/2008/10/driehoekstraat.html • http://www.geschiedenisvanantwerpen.be/forum/viewtopic.php?f=4&t=326 • http://www.geschiedenisvanantwerpen.be/forum/viewtopic.php?f=4&t=376
Boechout	<ul style="list-style-type: none"> • https://www.hetstillepand.be/boechout.html
Edegem	<ul style="list-style-type: none"> • https://www.dbnl.org/tekst/_naa002198001_01/_naa002198001_01_0001.php#002
Ekeren	<ul style="list-style-type: none"> • https://sites.google.com/site/wereldoorlog2inekeren/h-l
Hoboken	<ul style="list-style-type: none"> • https://zoveelmeerhoboken.blog/2017/06/23/van-a-tot-zwaantje-een-eigenzinnig-abc-over-hoboken-6/ • http://zoveelmeerhoboken.blogspot.com/2016/03/van-tot-zwaantje-een-eigenzinnig-abc.html
Kontich	<ul style="list-style-type: none"> • https://inventaris.onroenderfgoed.be/erfgoedobjecten/101893
Mortsel	<ul style="list-style-type: none"> • https://www.erfgoedzuidrand.be/databank/mortsel-grensstraat-huidige-hendrik-kuijpersstraat
Niel	<ul style="list-style-type: none"> • http://www.fotoniel.be/archief/history_niel/gemniel/pagina11.html
Schoten	<ul style="list-style-type: none"> • https://www.schoten.be/nutskast-marktplein
Wilmarsdonk	<ul style="list-style-type: none"> • https://mapio.net/pic/p-18782700/
Zandvliet	<ul style="list-style-type: none"> • https://inventaris.onroenderfgoed.be/erfgoedobjecten/102504

Attachment 3 – Evolution of the Municipalities of Antwerp Arrondissement (1856-1930)

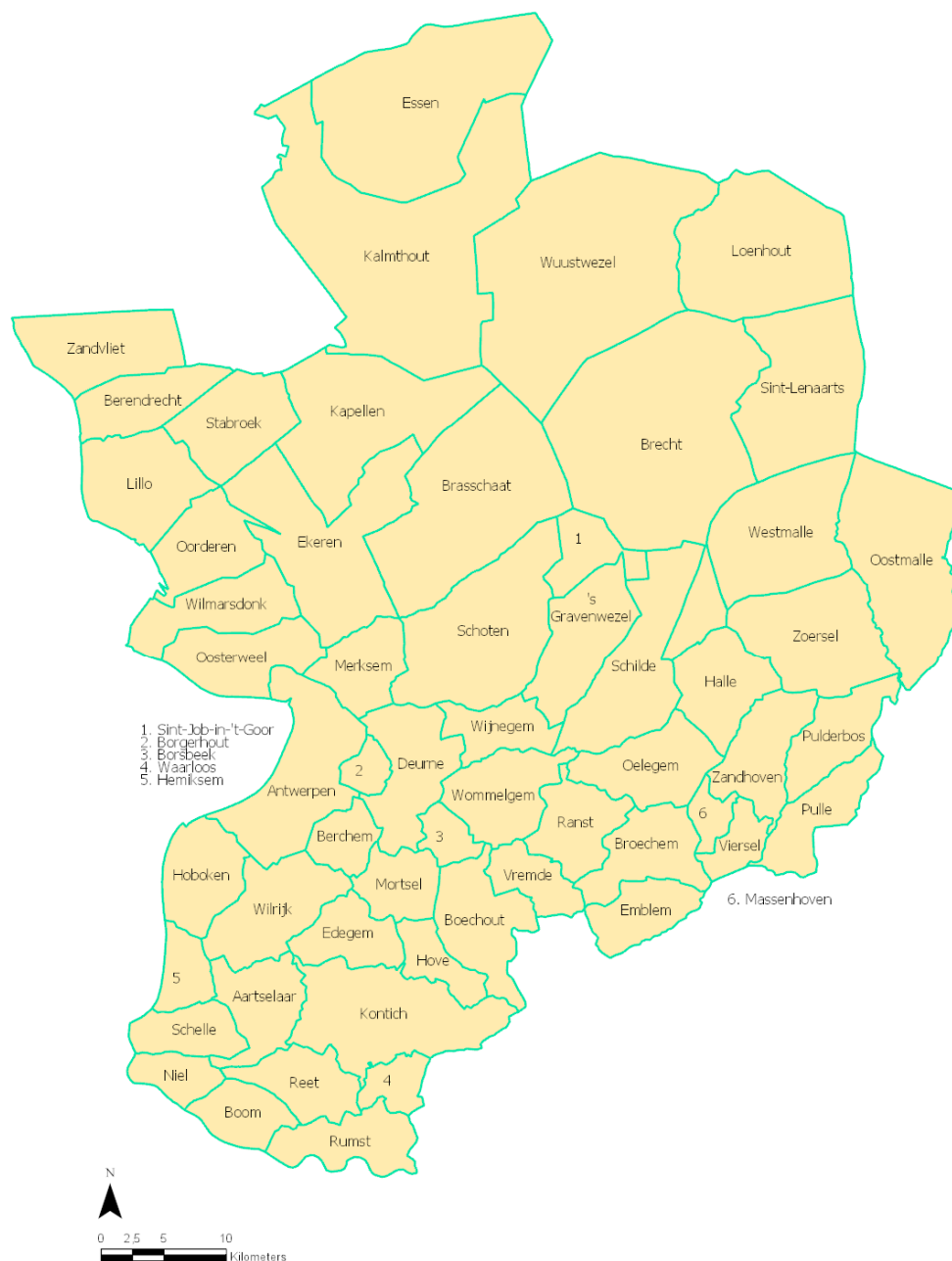


Figure 8 - Antwerp Arrondissement 1856



Figure 9 - Antwerp Arrondissement 1930