Data Quality Statement

1991 Population Ecumene of Canada

1.0 Purpose of the Ecumene

The 1991 version of the population ecumene of Canada was generated to meet the needs of some of the clients of the Geography Division, Statistics Canada. The ecumene was developed in consultation with these clients and then finalized for release as a standard Geography Division product. The ecumene can be used in a variety of thematic mapping applications which focus on the spatial variation of characteristics of the Canadian population.

1.1. Definition of the Ecumene

The population ecumene represents the land base of significant permanent human settlement in Canada. In general, ecumenes are used by cartographers to limit the display of choropleth or dot symbols based on data on whole geostatistical units to those portions that actually exhibit the relevant land use. This provides a more realistic representation of the data and avoids the misinterpretation that can result from data symbolization over large regions that are not actually devoted to the land use in question.

2.0. Purpose of the Data Quality Statement

The data quality statement provides information to allow users to evaluate the quality of the product for their particular use. Geography products are assessed in relation to relevant aspects of spatial data quality and every product is accompanied by a data quality statement.

2.1. Lineage

Lineage includes descriptions of the source material from which the data were derived and the methods of derivation, including the dates of the source material and all transformations involved in producing the final digital product.

Source Files:

- 1991 EA National Digital Boundary File

The 1991 EA Digital Boundary File for Canada was created over a period of 3 years. The Federal Electoral District (FED) was used as the processing unit since all EAs must respect FED boundaries and the FED is the principle unit for organizing the census collection operations. First preliminary digital EA boundaries were created (by FED), which were subsequently updated due to municipality boundary changes, updates to the street network, and changes in dwelling counts. The 1991 EA boundaries were finalized with Census, June, 1991. Some EAs may have been "split" during the census collection process if the number of dwellings was greater than expected. These EA splits are included in the EA boundary file. The EA boundary files for each of the 295 FEDs were completed and verified, and finally joined together to create the national file.

- 1991 GAF Population Counts

The population counts were derived from the 1991 Census. The data were collected by census representatives for each enumeration area; the enumeration area counts were then tabulated based on the enumeration area's relationship to higher-level geographic areas. Data for enumeration areas flagged as incompletely enumerated Indian reserves were not included for any of the higher-level geographic levels.

- 1991 Digital Cartographic File, Census Division (CD)

The CD digital cartographic boundaries were obtained from the 1991 CSD Digital Boundary File, which in turn were derived from the 1991 Enumeration Area (EA) Digital Boundary File. The CD Digital Cartographic Files were created by aggregating the CSD cartographic boundaries.

- 1991 Digital Cartographic File, Shoreline and Lakes

The shorelines and lakes were obtained from the 1986 CARTLIB file, which in turn was derived from the 1981 CD CARTLIB file. The 1981 file was digitized from mylar EMR base maps at 1:2,000,000 (for the 10 provinces) and 1:5,000,000 (for the two territories). The 1986 file was enhanced by digitizing 1:250,000 NTS maps for shorelines where there was a high concentration of small CSDs; This ensured that all CSD computed centroids were contained within the CSD polygons. Other enhancements included extracting the Toronto, Montreal and Quebec shoreline (digitized at 1:50,000) from the 1986 CT CARTLIB files. Finally, for the national 1991 CARTLIBs, the northern shoreline of Laval was added by digitizing the shoreline from the 1991 Street Network File for Laval. The 1986 and 1991 enhancements also involved projection transformations from Transverse Mercator to Lambert Conformal Conic.

2.2. Method of Derivation

Using ARC/INFO®, Version 6.1, 1991 GAF Population Counts by Enumeration Area (EA) were related to the 1991 Enumeration Area Digital Cartographic File. This procedure joins the 1991 population count to the digital boundary file. ARC/INFO® maintains the area of a polygon in square meters. A conversion factor is then used to calculate square kilometres for each EA polygon ($1m^2$ = .000001 km^2). Population density is then calculated (population / km^2).

A population density of 0.4 persons per km^2 (approximately 1 per square mile) has often been used as a population threshold by Geographers. Using ARC/INFO®, Version 6.1, all EAs with a population density of greater than or equal to 0.4 persons per square kilometre were selected.

The boundaries of the selected EAs were then generalized from a hand drafted position to improve the cartographic appearance of the ecumene. The generalized boundaries were then digitized. In the sparsely populated areas of the country (the northern parts of the provinces, and north of 60 degrees latitude), generalization was accomplished by selecting EAs that met the density criterion, aggregating these areas for cartographic purposes where necessary, and locating these pockets to the nearest recognizable population centre (e.g. a small town). This procedure ensured that all ecumene pockets would be visible for small scale thematic mapping, and that there would be at least one ecumene pocket in each Census Division (CD).

Once generalization of ecumene boundaries was completed, the ecumene "in" polygons were "clipped" against the National CD coverage. Any resulting slivers that occurred due to the lineage of the EA derived ecumene and the "streamed" version of National CD boundaries were removed. As a result, sliver boundaries that were approximately 175,000 m^2 were eliminated manually through ARC/EDIT commands.

2.3. Logical Consistency

Logical consistency describes the fidelity of relationships encoded in the data structure of the digital spatial data (ie., how well elements of the data structure follow the rules imposed on them).

Automated processing through ARC/INFO® software verified that all boundary polygons were topologically correct, that they closed with no overlaps, overshoots or slivers. Lines are intersected only where intended and no lines are entered twice.

2.4. Positional Accuracy

Positional accuracy is the difference between the "true" position of a feature in the real world and the "estimated" position sorted in the digital file.

Positional accuracy depends on the quality of the source material used and the process used to create the files (ie. manual drafting, digitizing etc.). No numerical measurement of positional accuracy has been made for this product. (See 1991 Digital Boundary File User Guide for further information on the positional accuracy of the source files used).

With respect to the ecumene boundaries, the areal extent and position of the ecumene would have been influenced to a far greater extent by the density criterion used to select EAs than by the positional error associated with either the hand drafted generalization or subsequent digitizing. For example, if a density of 0.5 persons per km^2 were used to select EAs, this would have removed a fairly large portion of south central Saskatchewan from the present ecumene.

2.5. Attribute Accuracy

Attribute accuracy refers to the accuracy of the non-positional information attached to each polygon.

The 1991 CD DCF contains a 4 digit PROV/CD code used to identify each polygon in the file. The field is labelled CD and is defined as an integer.

To verify the accuracy of the attribute data attached to each of the cartographic polygons, the 4 digit CD code were plotted and verified against the 1991 SPR/CD reference map.

2.6. Completeness

Completeness expresses the degree to which the geographic entities (features) are captures according to the data capture specifications.

The 1991 CD Digital Boundary File contains 296 polygons representing 290 census divisions (CDs) which corresponds to the national coverage.

A breakdown of the number of CDs by province/territory, the number of polygons per CD, and other relevant information regarding the content of the file is given in appendix A.

The 1991 CD Population Ecumene Boundary File contains 581 polygons representing 290 census divisions (CDs) which corresponds to the national coverage.

2.7. Limitations on Use

This ecumene has been designed for the purpose of displaying all of Canada on map of relatively small scale. The ecumene may not be suitable for mapping smaller areas at larger scales, especially in the sparsely populated areas of the country where some positional bias exists due to the aggregation of small ecumene pockets.