

## Frequently Asked Questions

### 1. Which postal code data source is CANUE using?

There are two main sources of postal code data files – the Postal Code CanMap Suite from DMTI Spatial (DMTI Spatial, 2015) and Statistic Canada’s Postal Code Conversion File (PCCF) dataset (Statistics Canada, 2016). Neither data set is necessarily better than the other, each offer data in a slightly different format, including different data fields, as well as geographic coordinates for the 6-digit postal codes. At CANUE we are using DMTI’s postal code files due to the simpler licensing.

### 2. When multiple postal code locations exist for a single postal code, how does CANUE determine which location to use?

We see that in some cases, a single postal code can have multiple locations and cover a whole town, especially in rural areas – like this example of Loretto Ontario (Figure 1). Both PCCF and DMTI include multiple locations for the same postal code in cases like this – shown here as the red dots on the left, that fall within the postal code area shown on the right – 121 in total.



Figure 1. Postal codes with multiple locations (red dots) and the single link indicator (blue dot)

This makes it difficult to assign exposure accurately - we could measure exposure at every location shown, but they would all have the same postal code, and we would not be able to tell which location was the most accurate for the participant, just by postal code. Many software programs will take the



---

first value encountered in the list. Some studies have handled this by taking an average of the exposure from all locations with the same postal code. Either way, some exposure measurement error will occur.

Both PCCF and DMTI also include a 'single link indicator' (SLI) that identifies the postal code that is most representative of the population location – shown as the blue dot on the left figure. At CANUE, we use only the postal codes identified as the single link so there is a one-to-one match. Again, some exposure measurement error will occur.

### **3. Which years are the postal codes available for?**

Currently, for all years from 1983 to 2015. Most of the exposure data we are acquiring or developing are available for many years, going back to the 1980s in some cases. To make the most of the data, CANUE has extracted exposure values to every 6-digit postal code in Canada, year by year. For example, in 2015 this represents over 850,000 unique locations across Canada.

### **4. Are rural postal codes included in CANUE datasets?**

Yes. When possible, all 6-digit postal codes in Canada, for all urban areas and rural areas are used to extract data from environmental exposure surfaces.

### **5. How will CANUE deal with changes in postal codes?**

When new postal codes are created and old postal codes retired, we will extract and update annual postal code files e.g. 2016, 2017 and the related precision metrics. These will be made available in future updates.

We compared the geographic coordinates for each single link postal code over time. There were no changes in location for the single link postal codes between 1983 and 2015. This means that the geographic coordinates for each postal code included do not change over time, and there are no cases where a postal code was located in one place, and then subsequently move to another location.

### **6. Can I filter data by region?**

Yes. We include data fields that can be used to filter postal code locations based on province, or common place name, for example the city of NORTH YORK or municipality such as TORONTO. This will allow researchers to quickly subset data for local or regional analysis.

### **7. What are the spatial boundaries or footprint represented by each postal code location?**

There is no single representative areal measure of the spatial footprint represented by each postal code. Postal codes are point locations and do not have defined boundaries – especially in an historical context.

**8. How can I use the distance to nearest postal code measure to better understand spatial error in the postal code data?**

To address the issue of spatial representativeness of the postal code locations, we have calculated a distance to nearest postal code metric (NEAR\_DIST), which provides an indirect measure that approximates the area represented by each postal code and that is also reflective of population density. We also include the 6-digit postal code of that nearest neighbour (NEAR\_PC).

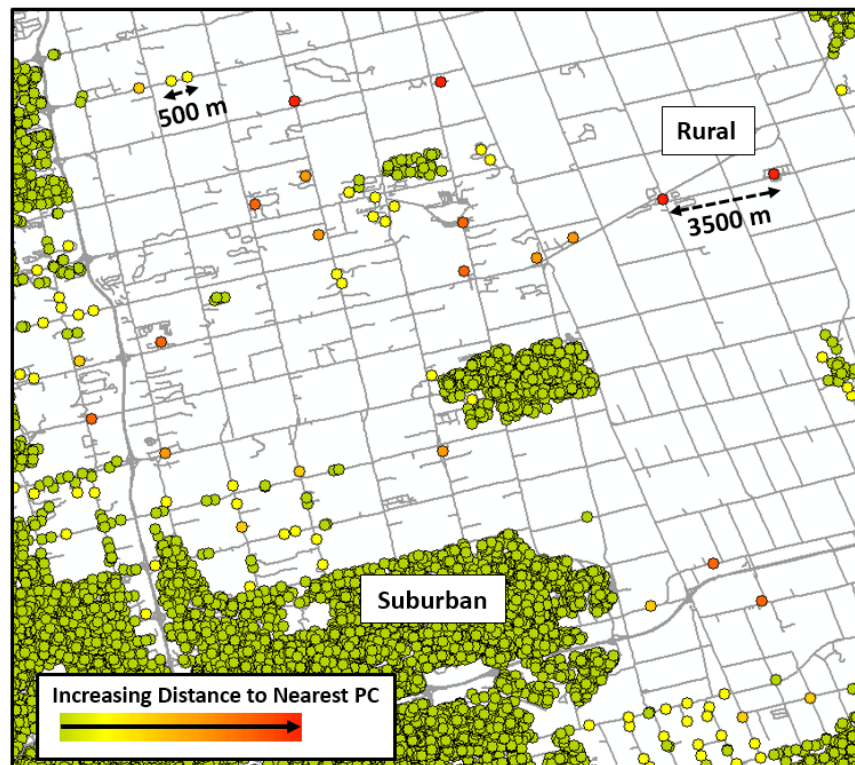


Figure 2. Distance to nearest postal code in suburban versus rural areas

In using the NEAR\_DIST data field you will notice some distances = 0. Why? There are situations where postal codes share the same specific x, y geographic co-ordinates, sometimes referred to as stacking. For example, in large apartment or office buildings different floors in the same building can be assigned separate postal codes but they share the same x, y coordinates. A nearest neighbour analysis of postal code locations would return a value = 0 in situations like this.

In terms of exposure error, when the distance from one postal code to the nearest postal code is small (e.g. a few hundred metres), there will be much lower exposure error as in urban areas. When the distance is larger, such as in rural areas in Figure 2, there may be more exposure error. While we can't calculate exactly how much error there is, the distance to nearest postal code metric allows the researcher to decide which participants they might want to exclude, depending on how much potential error they are willing to accept, and this may depend on the spatial resolution of the exposure data they are working with.

Another way of looking at this is by adding the count of possible locations for a postal code as well (PC\_COUNT). In a dense urban area, there may be only the one location, but in the previous example, there were 121 locations – so a high number here indicates more possibility for exposure error.

### 9. What are there differences between the location of 6-digit postal codes in the CanMap Postal Code Suite from DMTI Spatial and Statistic Canada's PCCF data?

DMTI and PCCF postal code data files do not always use the same geographic location for a postal code. For example, in a comparison within Oakville, Ontario, the distance between DMTI and PCCF postal codes is typically less than 250 m within a dense urban neighbourhood, and often much less, but differences might be larger especially a suburban neighbourhood or rural area (Figure 3).

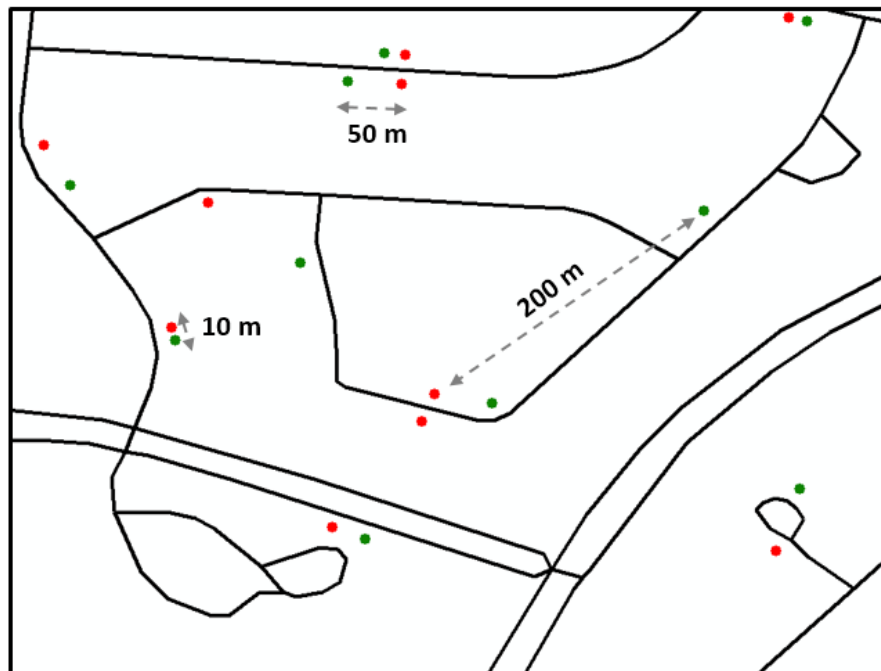


Figure 3. Comparing Statistic Canada's PCCF and DMTI Postal Code Single Link (SLI) Locational Differences, Oakville, ON

### 10. Can I mix exposure metrics calculated using PCCF postal codes locations with exposure measures calculated using DMTI postal code locations?

It depends on the spatial resolution of the exposure data and if the participant is in an urban area or a rural area, but in general – no, it is not advised. Exposure metrics from sources other than CANUE may have been calculated using the PCCF locations. These should not be mixed with CANUE exposures calculated using DMTI files because of the potential for different locations.



## Postal Code User Guide 2015-01-25

### 11. How can I use delivery mode type to better understand exposure at a postal code of a participant?

POSTAL	PROV	COMM_NAME	MUNICIPAL	BIRTH_DATE	RET_DATE	DOM	TOTAL_POC	POC_APART	POC_BUS	POC_HOUSE	POC_FARM
P0R1B0	ON	BLIND RIVER	BLIND RIVER	19830401	19000001	LB	2100	61	239	1786	14
P0A1H0	ON	DWIGHT	LAKE OF BAYS	19830401	19000001	RR	584	17	43	524	0
J6E5E7	QC	JOLIETTE	JOLIETTE	19830401	19000001	LC	21	9	10	2	0
N2Z2X3	ON	KINCARDINE	HURON-KINLOSS	19870101	19000001	RR	408	0	2	371	35
V9Y5V4	BC	PORT ALBERNI	PORT ALBERNI	19830401	19000001	LC	9	0	0	9	0
J0Z3G0	QC	RAPIDE-DANSEUR	RAPIDE-DANSEUR	19830401	19000001	LB	148	0	5	140	3
C0A1N0	PE	HUNTER RIVER	CHARLOTTE PARISH	19830401	19000001	RR	1160	0	44	1109	7
P0R1C0	ON	BRUCE MINES	PLUMMER ADDITIONAL	19830401	19000001	LB	757	0	21	660	76
G0W2H0	QC	MASHTEUATSH	MASHTEUATSH	19830401	19000001	RR	1001	32	56	911	2
G4Z2L6	QC	BAIE-COMEAU	BAIE-COMEAU	19830401	19000001	CF	7	0	7	0	0
C0A1S0	PE	MORELL	ST. PATRICK'S PARISH	19830401	19000001	LB	687	10	21	655	1
J2B2W8	QC	DRUMMONDVILLE	DRUMMONDVILLE	19830401	19000001	LC	10	4	0	6	0
P2N2B4	ON	KIRKLAND LAKE	KIRKLAND LAKE	19830401	19000001	LC	3	0	0	3	0
K0H2H0	ON	ODESSA	LOYALIST	19830401	19000001	LB	1487	72	30	1330	55

Table 1. Delivery mode type postal code measures

The delivery mode type e.g. POC\_APART, POC, FARM etc. in Table 1 are provided for each postal code in Canada and can be used by researchers to better understand how accurate the exposure at a postal code might be for a participant. For example, the delivery mode can be an “apartment building” in which case the exposure would be accurate for everyone in the building, or it might be ‘general delivery’ or ‘community mailbox’, in which case the exposure at that location may not be accurate for all the people who receive mail there. In rural areas, those postal codes servicing farms are also identified which may be of use for some studies.

Again, while we can’t calculate exactly how much error there is, delivery mode type also allows the researcher to decide which participants they might want to exclude, depending on how much potential error they are willing to accept, based on this precision measure. This may also depend on the spatial resolution of the exposure data they are working with.

Finally, there are postal codes in each year that do not have DOM\_DELMDE data associated with them e.g. approximately 26,500 in 2015. The same postal codes do not have point of call (POC) data associated with them either. This is a known issue. DMTI (2015) states that only active postal codes will have this data field populated.

#### References:

- [1] Postal Code Conversion File (PCCF), 2016. [computer file] Statistics Canada Catalogue no. 92-154-G
- [2] CanMap Postal Code Suite v2015.3. [computer file] Markham: DMTI Spatial Inc., 2015.