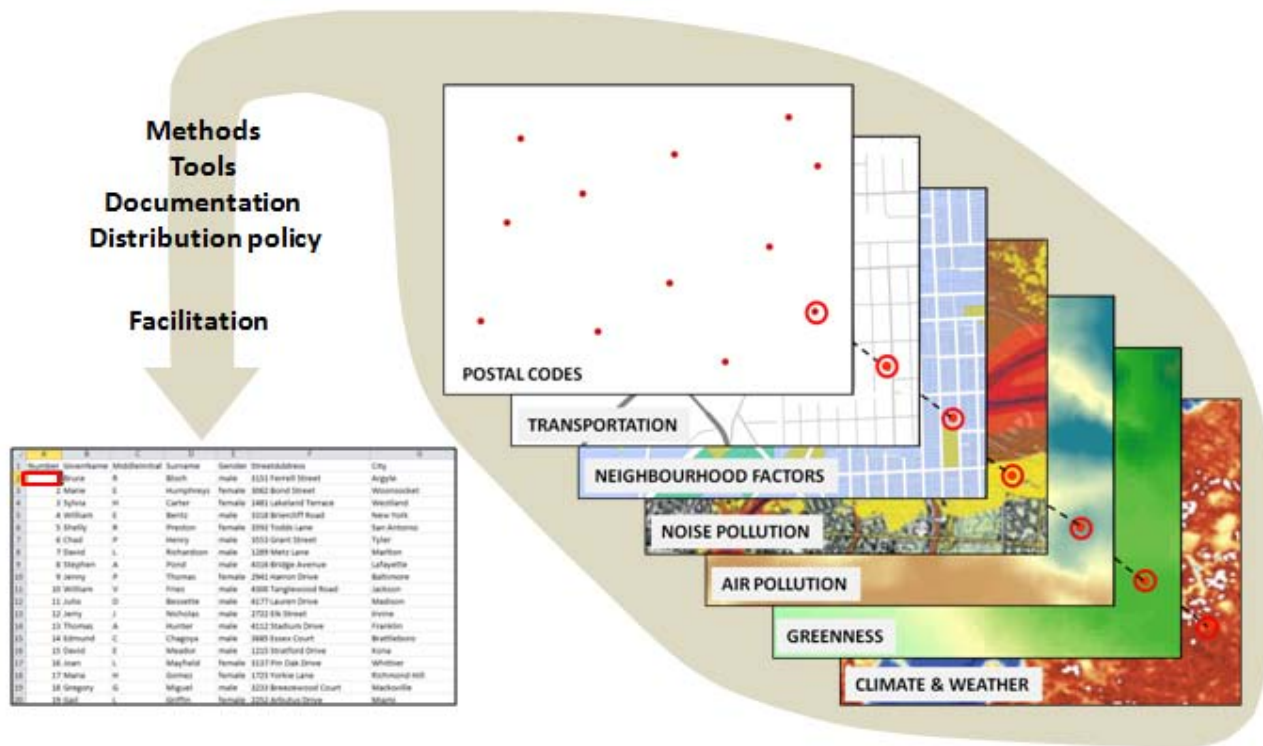




Navigating the Data Merge – September 28 | 2017

MEETING START: 9:00AM PACIFIC | 12:00 noon EASTERN

**FOR AUDIO – PLEASE CALL:
1-877-558-8690**





Navigating the Data Merge – September 28 | 2017



chat...



info@canue.ca



Navigating the Data Merge – September 28 | 2017




UberConference mahdi

Dialpad, Inc. [US] | https://www.uberconference.com/Canue

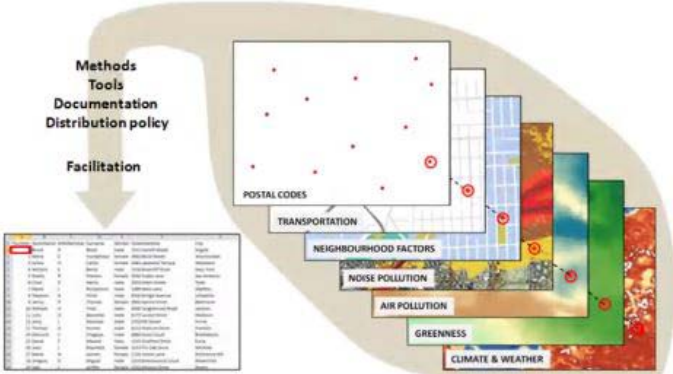
CONFERENCE INFO AUDIO OPTIONS Back to Conference

Navigating the Data Merge – September 26 | 2017

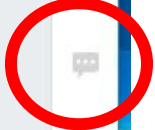


MEETING START: 9:00AM PACIFIC | 12:00 noon EASTERN

FOR AUDIO – PLEASE CALL:
1-877-558-8690



Canadian Urban Environmental Health Research Consortium 1





Navigating the Data Merge – September 28 | 2017



UberConference x

← → ↻ Dialpad, Inc. [US] | https://www.uberconference.com/Canue

CONFERENCE INFO AUDIO OPTIONS Back to Conference

September 25, 2017
00:00:00

Chat Viewers (2)

The organizer has enabled chat logging.

Hi, my name is ...

Methods
Tools
Documentation
Distribution policy
Facilitation

FOR AUDIO – PLEASE CALL:
1-877-558-8690

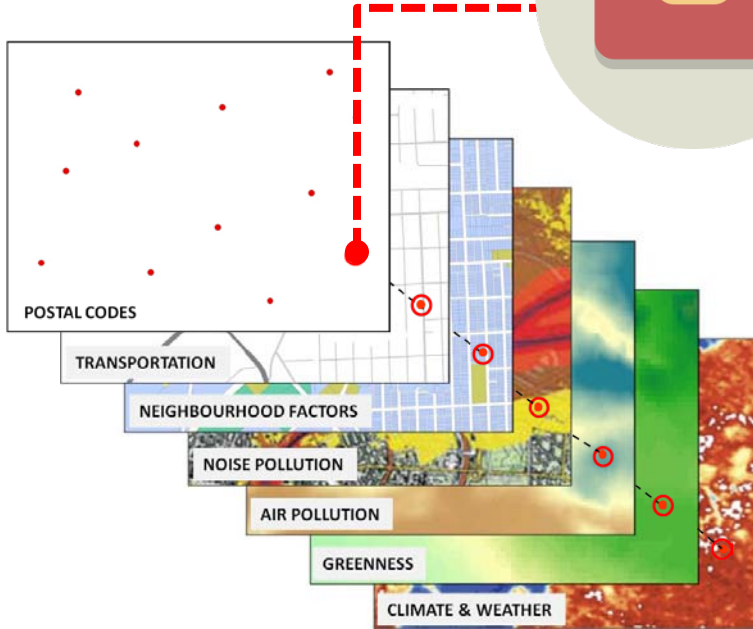
POSTAL CODES
TRANSPORTATION
NEIGHBOURHOOD FACTORS
NOISE POLLUTION
AIR POLLUTION
GREENNESS
CLIMATE & WEATHER

Canadian Urban Environmental Health Research Consortium



Navigating the Data Merge – September 28 | 2017

Tools



Data Sets

1	A	B	C	D	E	F	G
2	Number	GivenName	MiddleInitial	Surname	Gender	StreetAddress	City
3	1	Bruce	R	Bloch	male	3151 Ferrell Street	Argyle
4	2	Marie	E	Humphreys	female	3062 Bond Street	Woonsocket
5	3	Sylvia	H	Carter	female	1481 Lakeland Terrace	Westland
6	4	William	E	Bentz	male	3318 Briercliff Road	New York
7	5	Shelly	R	Preston	female	3592 Todds Lane	San Antonio
8	6	Chad	P	Henry	male	3553 Grant Street	Tyler
9	7	David	L	Richardson	male	1289 Metz Lane	Marlton
10	8	Stephen	A	Pond	male	4316 Bridge Avenue	Lafayette
11	9	Jenny	P	Thomas	female	2941 Harron Drive	Baltimore
12	10	William	V	Fries	male	4300 Tanglewood Road	Jackson
13	11	Julio	D	Bessette	male	4177 Lauren Drive	Madison
14	12	Jerry	J	Nicholas	male	2722 Elk Street	Irvine
15	13	Thomas	A	Hunter	male	4112 Stadium Drive	Franklin
16	14	Edmund	C	Chagoya	male	3685 Essex Court	Brattleboro
17	15	David	E	Meador	male	1215 Stratford Drive	Kona
18	16	Joan	L	Mayfield	female	3137 Pin Oak Drive	Whittier
19	17	Marla	H	Gomez	female	1723 Yorlie Lane	Richmond Hill
20	18	Gregory	G	Miguel	male	3223 BreezeWood Court	Macksville
21	19	Gail	L	Griffin	female	2252 Arbutus Drive	Miami

Data Delivery and Merging



Discussion



Navigating the Data Merge – September 28 | 2017





English | Français   



CANUE ABOUT DATA AND TOOLS RESEARCH PARTNERS RESOURCES EVENTS CONNECT

STRATEGIC PLAN
LEADERSHIP & SPECIALISTS



Welcome to **The Canadian Urban Environmental Health Research Consortium**

Every location in Canada can be described by a complex set of environmental factors – the amount of nearby traffic, local air quality, access to greenspaces, opportunities for walking and cycling, the amount of noise and light pollution, to name a few.

The Canadian Urban Environmental Health Research Consortium will gather and develop measures of environmental factors for every neighbourhood across Canada.

Search

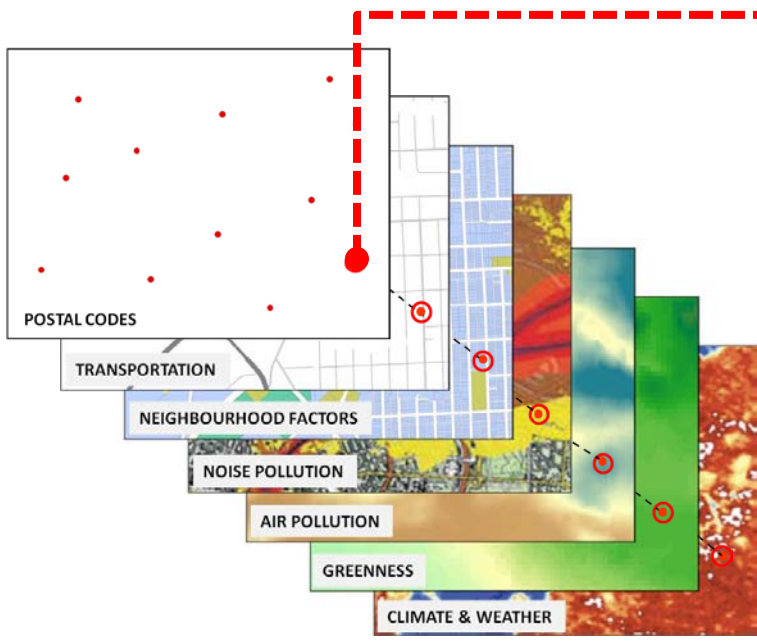
NEWS

UPCOMING EVENTS

Navigating the Data Merge | September 20 | 26 | 28 | 2017



Navigating the Data Merge – September 28 | 2017



Data Sets

1	A	B	C	D	E	F	G
2	Number	GivenName	MiddleInitial	Surname	Gender	StreetAddress	City
3	1	Bruce	R	Bloch	male	3151 Ferrell Street	Argyle
4	2	Marie	E	Humphreys	female	3062 Bond Street	Woonsocket
5	3	Sylvia	H	Carter	female	1481 Lakeland Terrace	Westland
6	4	William	E	Bentz	male	3318 Biercliff Road	New York
7	5	Shelly	R	Preston	female	3592 Todds Lane	San Antonio
8	6	Chad	P	Henry	male	3553 Grant Street	Tyler
9	7	David	L	Richardson	male	1289 Metz Lane	Marlton
10	8	Stephen	A	Pond	male	4316 Bridge Avenue	Lafayette
11	9	Jenny	P	Thomas	female	2941 Harron Drive	Baltimore
12	10	William	V	Fries	male	4300 Tanglewood Road	Jackson
13	11	Julio	D	Bessette	male	4177 Lauren Drive	Madison
14	12	Jerry	J	Nicholas	male	2722 Elk Street	Irvine
15	13	Thomas	A	Hunter	male	4112 Stadium Drive	Franklin
16	14	Edmund	C	Chagoya	male	3685 Essex Court	Brattleboro
17	15	David	E	Meador	male	1215 Stratford Drive	Kona
18	16	Joan	L	Mayfield	female	3137 Pin Oak Drive	Whittier
19	17	Marla	H	Gomez	female	1723 Yorkie Lane	Richmond Hill
20	18	Gregory	G	Miguel	male	3223 Breezeswood Court	Macksville
21	19	Gail	L	Griffin	female	2252 Arbutus Drive	Miami

Data Delivery and Merging



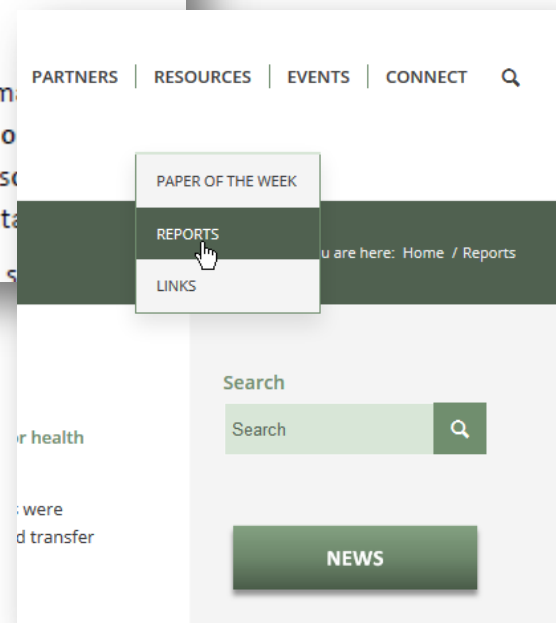
CANUE Health Data Holder Survey

Understanding technical, procedural and functional needs of major health data holders

*Prepared by: Nancy Meagher (Lumine Research Inc.) with input from Dany Doiron (Maelstrom Research/CANUE)
August 2017*

Background

The Canadian Urban Environment Health Research Consortium (CANUE) has as a mission to develop measures of environmental factors such as pollution, green space, traffic and noise for every neighborhood in Canada. They are collecting data from a wide range of sources and will be developing standardized measures as well as accompanying data documentation. The team has as an objective of maximizing the use of the data, and are therefore s...





DATA MERGING STAKEHOLDER INTERVIEWS

Objective:

To identify data transfer model that works best for Canadian health data holders



7 adult cohorts



6 birth/mother-child cohorts



9 admin data platforms

Key results:

Preference for bulk data transfers on routine basis.

Interest in a hybrid model, with “high interest” data being transferred routinely and other less “in-demand” data being left at CANUE central to be pulled as needed



DATA MERGING STAKEHOLDER INTERVIEWS

Key results (cont'd):

Incorporate delivery of CANUE data into existing access processes



Privacy and legal considerations are not likely to be barriers to uptake



Clear metadata is essential!





Two scheduled yearly environmental exposure data deliveries:



FALL – National datasets

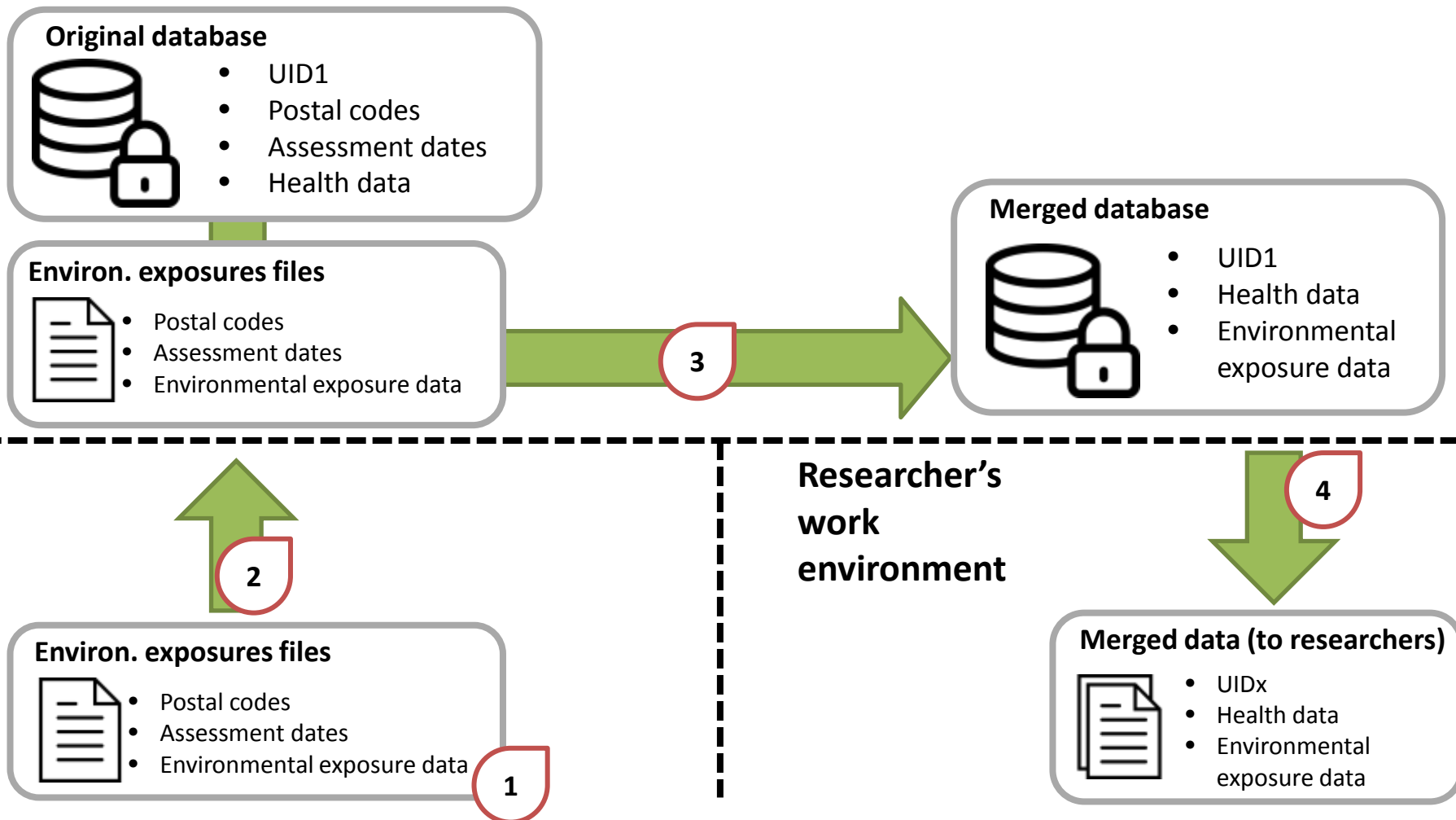


SPRING – regional datasets





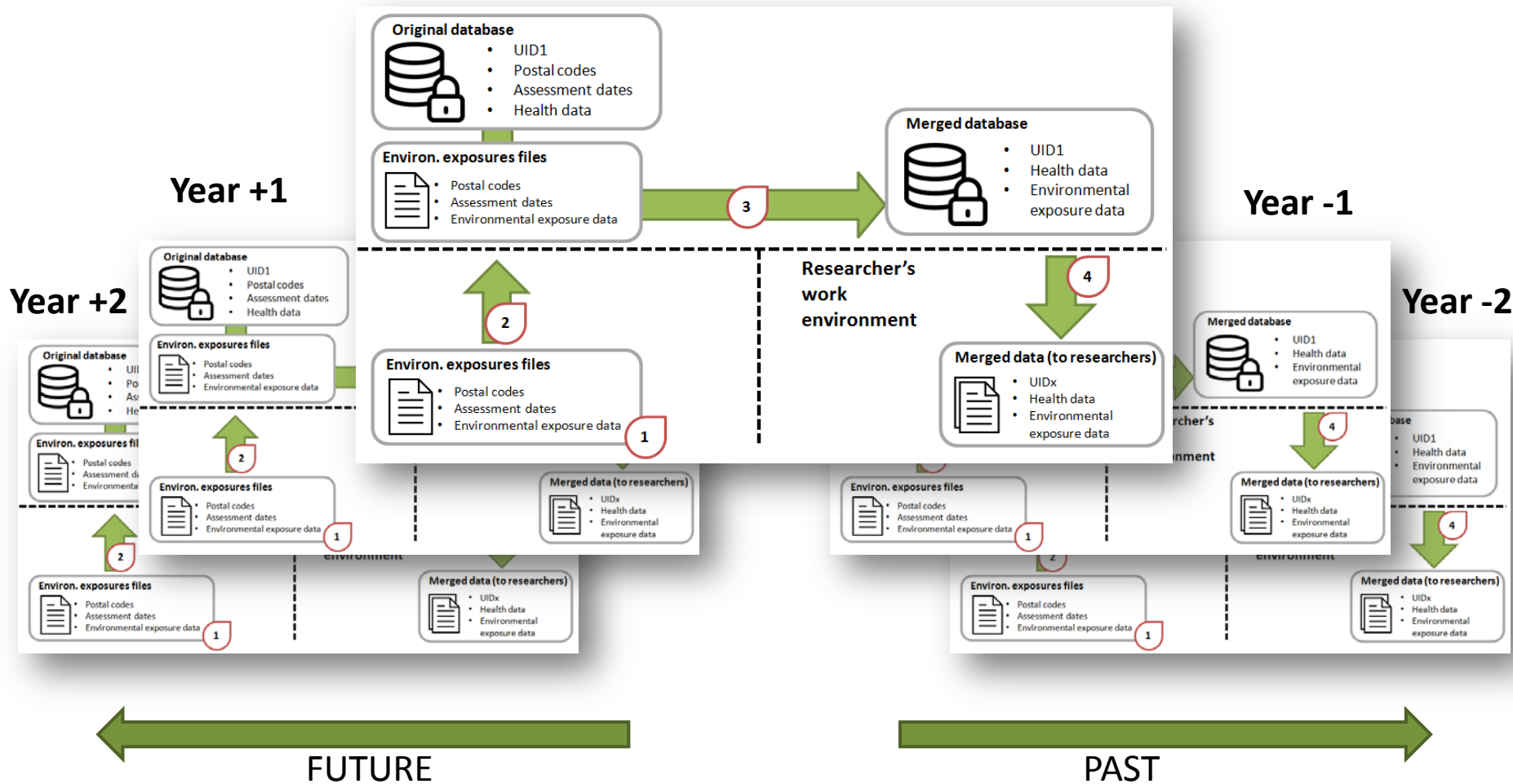
PROPOSED DATA MERGING PROCESS





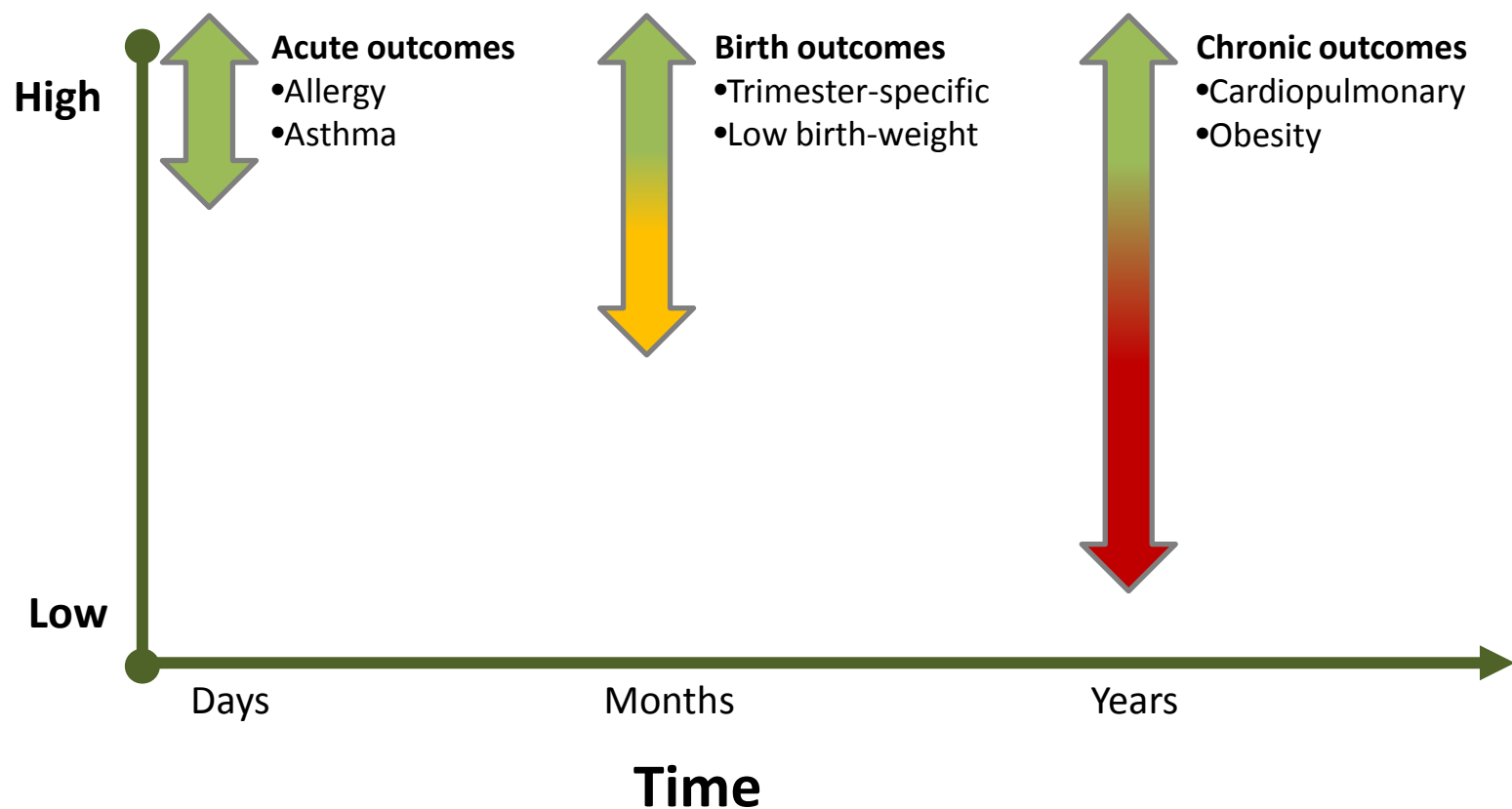
Navigating the Data Merge – September 28 | 2017

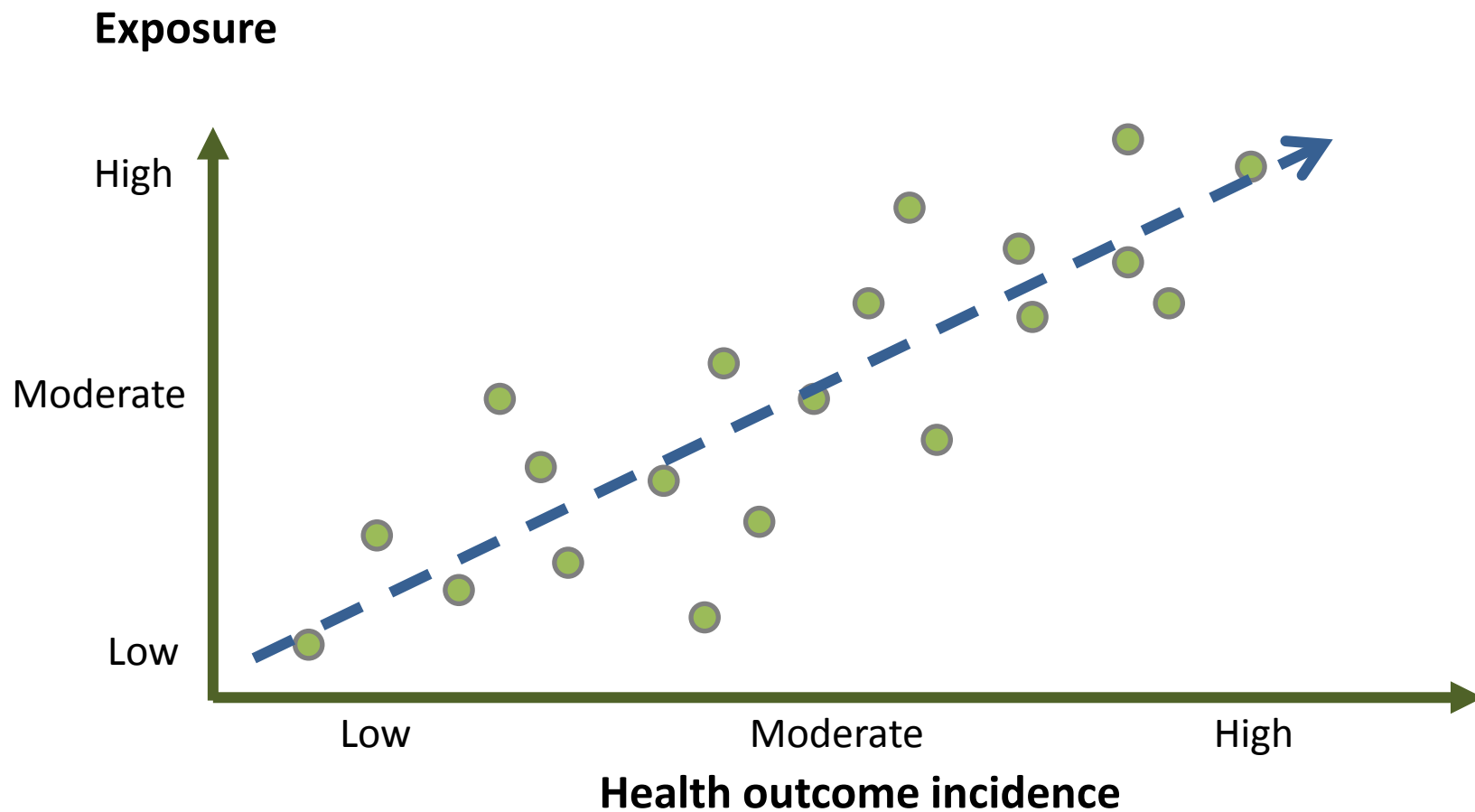
CURRENT YEAR

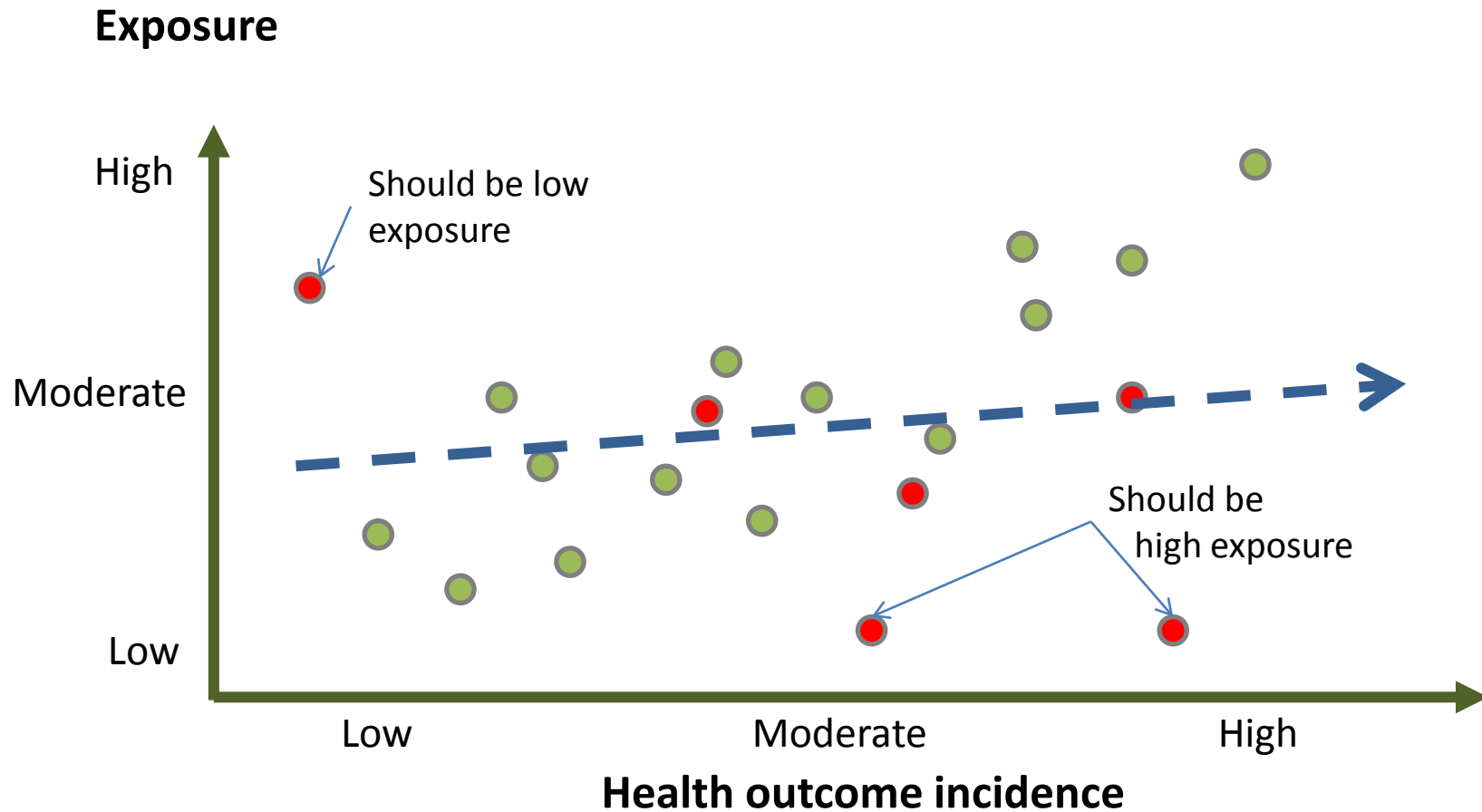




Accuracy of exposure using current postal code









Navigating the Data Merge – September 28 | 2017



1991 CanCHEC

The current version of the 1991 Canadian Census Health and Environment Cohort (CanCHEC) was derived on linkage-based data by compiling socio-economic, demographic, cancer, mortality, and place of residence data reported on T1 (1991) for 2.5 million Canadians over an eventual 20-year follow-up period.

The 1991 Canadian Census Health and Environment Cohort contains variables measuring population characteristics, place of residence, cancer incidence, and mortality. These variables were obtained from six different data sources:

- 1991 Canadian Census
- 1984-2011 Health Statistics
- 1981-1983 T1 Census
- 1969-1991 Census
- 1992-2010 Census
- 1991-2011 Census

CanCHEC (Canadian Census Health and Environment Cohort)

Description

2001 CanCHEC

The 2001 Canadian Census Health and Environment Cohort (CanCHEC) is a large population-based dataset that allows for the examination of mortality and cancer incidence by census characteristics (i.e. ethnocultural and socioeconomic factors).

The 2001 CanCHEC was created by linking data from the following different data sources:

- 2001 Census of Population
- T1 Universe Files (place of residence only)
- Canadian Mortality Database

2.5 million records

3.5 million records



300,000 Canadian residents

35-69 years of age

\$150M+





Navigating the Data Merge – September 28 | 2017



Health
Canada

Santé
Canada



Statistics
Canada

Statistique
Canada



**Canadian Health
Measures Surveys**

(2007-09, 2009-11,
2012-13, 2014-15)

~23,000 individuals

Reconstruct residential history using federal records

Five datasets over five years:

2016-17

2017-18

2018-19

2019-20

2020-21



Statistics
Canada

Statistique
Canada

SDLE – Social Data Linkage Environment

<https://www.statcan.gc.ca/eng/sdle/index>

The screenshot shows the Statistics Canada website for the Social Data Linkage Environment (SDLE). At the top left is the Statistics Canada logo. To the right is a search bar with the text "Search website" and a magnifying glass icon. Below the logo is a breadcrumb trail: "Home → Social Data Linkage Environment (SDLE)".

The main content area has a heading "Social Data Linkage Environment (SDLE)" followed by a horizontal line. Below this is a section titled "Expanding data potential" with a "[No Title]" label. The text reads: "The Social Data Linkage Environment (SDLE) at Statistics Canada promotes the innovative use of existing administrative and survey data to address important research questions and inform socio-economic policy through [record linkage](#)." Below this is another section titled "Protecting personal information" with the text: "Statistics Canada takes your confidentiality very seriously. Under the *Statistics Act*, all information provided to Statistics Canada is protected under the *Access to Information Act* and the *Privacy Act*." A sidebar on the left contains a table of contents for the SDLE page.

Social Data Linkage Environment (SDLE)
Overview
DRD linkage status
Getting started
Frequently asked questions
Definitions
More information



Statistics
Canada

Statistique
Canada

HEALTH DATABASE

Individual identifiers
and unique id

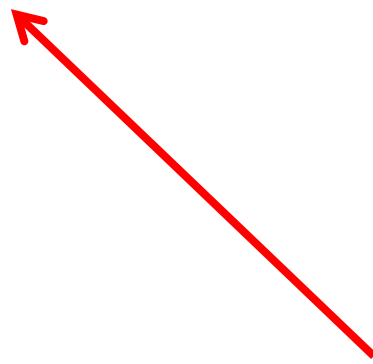
CANUE

exposure data by
postal code



SDLE

Residential history
created and exposure
assigned by postal code



exposure and unique id



Navigating the Data Merge – September 28 | 2017




Unique ID	V1	V2	V3	Y1 pcode	Y1 Exp	Y2 pcode	Y2 Exp	Y3 pcode	Y3 Exp	...	Y20 pcode	Y20 Exp
1	.	.	.	V9B 3A3	12	V9B 3A3	11	V0S 1N0	8		V1N 3A5	
2	.	.	.	B0T 1WO		B0T 1WO		B0T 1WO			B0T 1WO	
3	.	.	.	P5A 2B6		P5A 2B6		P5A 2B6			P3X 9A6	

2000		2001		2002	
	Exp		Exp		Exp
V9B 3A3	12	V9B 3A3	11	V9B 3A3	5
V0S 1N0	5	V0S 1N0	6	V0S 1N0	8
P5A 2B6	8	P5A 2B6	9	P5A 2B6	2
P3X 9A6	3	P3X 9A6	3	P3X 9A6	10



Navigating the Data Merge – September 28 | 2017



Unique ID	V1	V2	V3	Pcode 1	Date 1	Pcode 2	Date 2	Pcode 3	Date 3	...	Pcode x	Date x
1	.	.	.	V9B 3A3	06-Jul-00	V0S 1N0	Dec 5 2001	V0S 1N0				
2	.	.	.	B0T 1W0	15-Jan-00							
3	.	.	.	P5A 2B6	June 6 2001						P3X 9A6	03-Feb-15

6 months

2000	Exp
V9B 3A3	12
V0S 1N0	5
P5A 2B6	8
P3X 9A6	3

2001	Exp
V9B 3A3	11
V0S 1N0	6
P5A 2B6	9
P3X 9A6	3

11 months
1 months

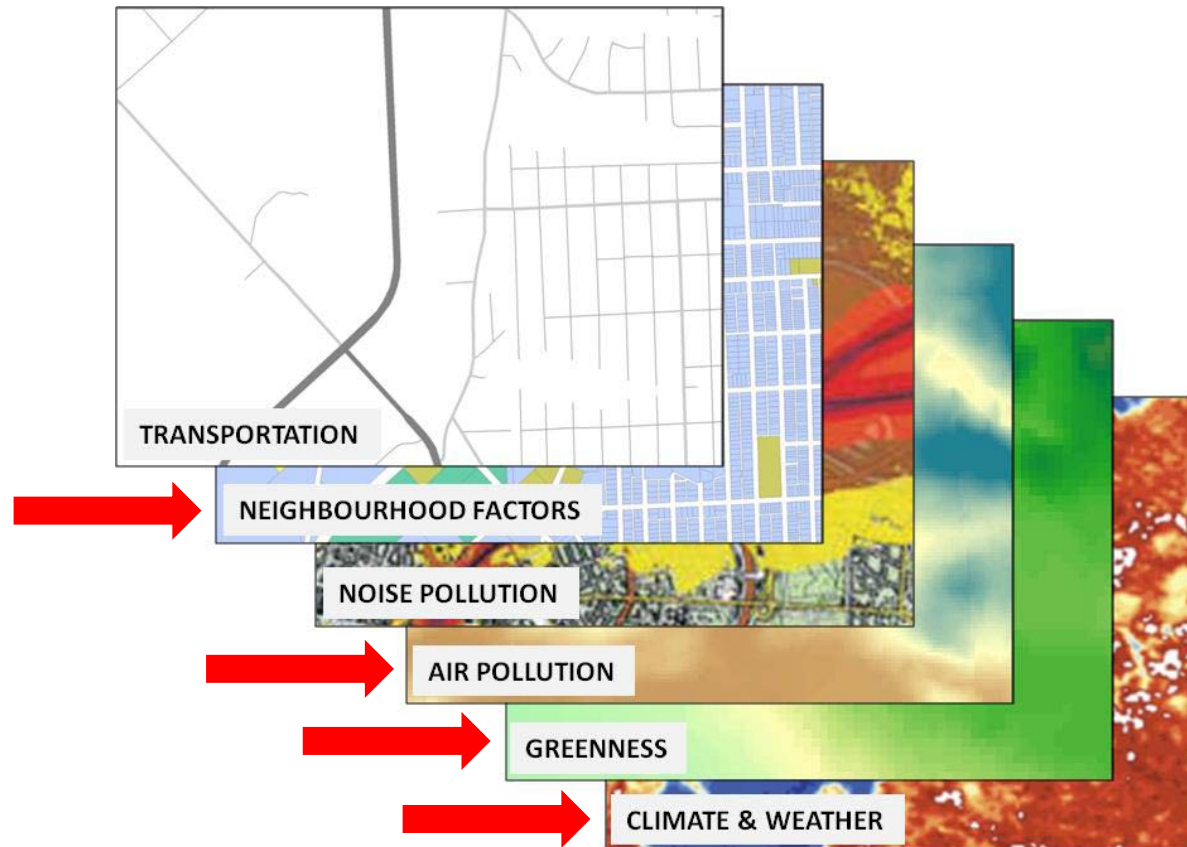
2002	Exp
V9B 3A3	5
V0S 1N0	8
P5A 2B6	2
P3X 9A6	10

12 months

Questions or comments?



Navigating the Data Merge – September 28 | 2017





Navigating the Data Merge – September 28 | 2017

Data Format

Text Import Wizard

This screen lets you specify the format of the data you are importing.

Delimiters

- Tab
- Semicolon
- Comma
- Space
- Other:

Data preview

	PDC	MULT_FLAG	NEAR_DIST	NEAR_LONG	NEAR_LAT	NEAR_DIRECT	NEAR_POSTAL	PM25
76	20	1	27223	-82.6760194	46.37052428	39.24918614	P5A1A1	8
88	0	1	9446	-79.09664637	45.36303375	-31.76528635	P1H2J6	6
26	21	1	42	-73.43937235	46.02429409	136.2262156	J6E5E8	11
47	20	1	4946	-81.70993159	44.11308432	33.51808138	N2Z2X7	13
59	21	1	22	-124.8027613	49.26435815	-155.3288701	V9Y5V3	3
99	0	1	5580	-79.22783696	48.50364129	117.0732511	J0Z1W0	5

Buttons: Cancel, < Back, Next >, Finish

Comma-separated values (CSV)

Data in tabular format, field separated listing of data

Files easily imported a database, Excel spreadsheet, GIS software etc.



Data Documentation and User Agreements

DATA FILE INFORMATION	
CANUE Data Catalogue ID:	
Date File Generated:	
File Name:	
File Type:	
Software and Version Used to Create This:	
Beginning Date:	
End Date:	
Sampling frequency of data in file:	
Sampling interval of data in file:	
Data acquisition sampling interval:	
Detection limits:	
Number of Data Records:	
Availability Status:	
Maintenance Description:	
Log of Changes:	
Conditions of Use:	
Data Sharing Restrictions:	
Data Comment:	
Data Comment:	
Data Comment:	
Data Comment:	
Data Comment:	
Data Set Title:	
Theme Keywords:	
Place Keywords:	
File Content Description:	Enter a short, accurate explanation of the dataset in the box below. Be informative as this is used for the description on the web page index. Write in box or cut and paste information. Insert rows as required to extend box.
Recommended Citation:	
Recommended Acknowledgment:	
DATASET CONTACTS	
Principal Investigator	Co-Investigator(s)
Email:	

METADATA

Data Sharing and End Use Agreement

USER AGREEMENT

1. Purpose of Agreement:

One of the goals of CANUE is to develop a properly curated set of geospatial data for distribution and use to consortium members for environmental exposure analysis. The purpose of this agreement is to allow Authorized Users to use data available from the CANUE Data Catalogue and to permit the dissemination of any resulting data products or publications, while restricting Authorized Data Users from transferring raw data from the CANUE Data Catalogue to those who are not Authorized Data Users.

2. Summary of DO's and DON'Ts

DO	DO NOT
Responsibly use the data for your own analysis, research, and policy-making as per data use restrictions for the named upon dataset	Lend, rent, sell, or give away downloaded data products to non-authorized users

Original Investigation

Association of Neighborhood Walkability With Change in Overweight, Obesity, and Diabetes

Maria I. Creatore, PhD; Richard H. Glazier, MD; Rahim Moineddin, PhD; Ghazal S. Facit, MPH; Ashley Johns, MSc; Peter Gozdyra, MA; Flora I. Matheson, PhD; Vered Kaufman-Shriqui, PhD; Laura C. Rosella, PhD; Doug G. Manuel, MD, PhD; Gillian L. Booth, MD

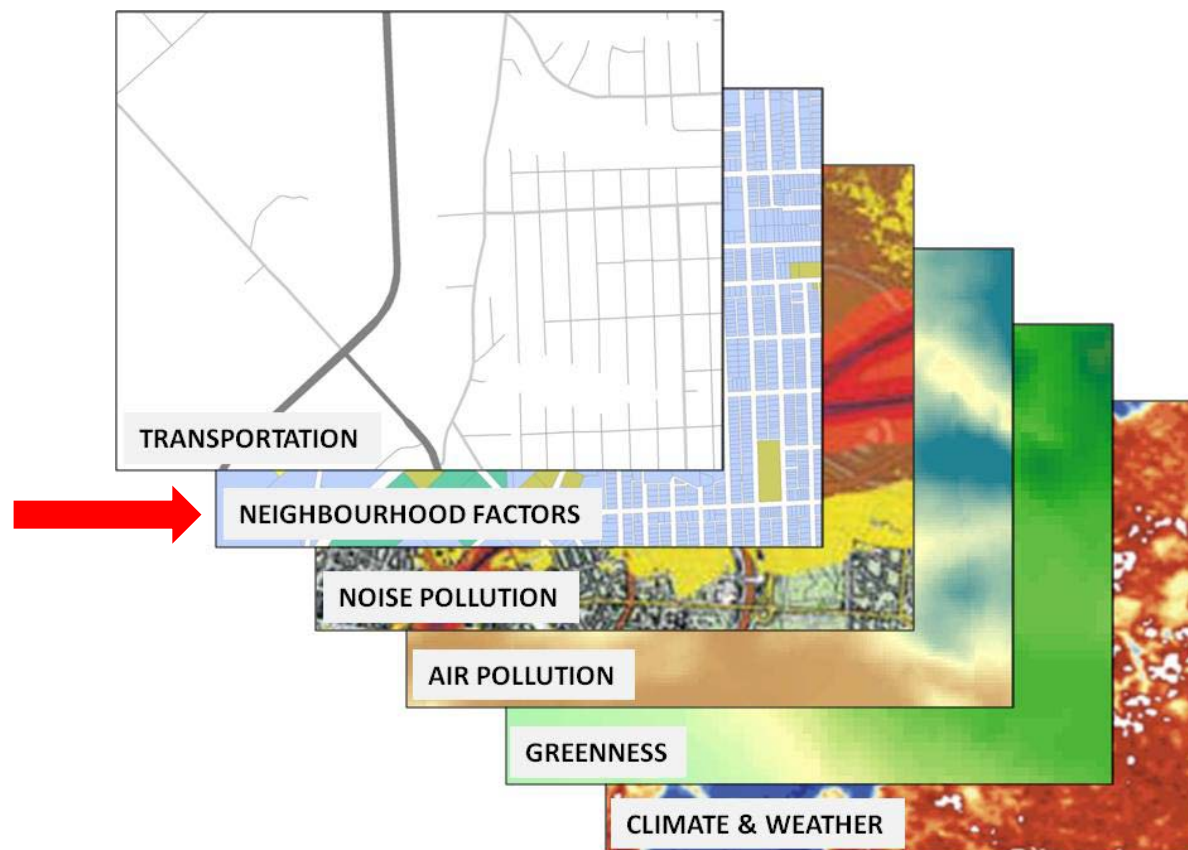
IMPORTANCE Rates of obesity and diabetes have increased substantially in recent decades; however, the potential role of the built environment in mitigating these trends is unclear.

OBJECTIVE To examine whether walkable urban neighborhoods are associated with a slower increase in overweight, obesity, and diabetes than less walkable ones.

Editorial page 2175
Author Video Interview and JAMA Report Video at jama.com
Supplemental content at jama.com



Navigating the Data Merge – September 28 | 2017

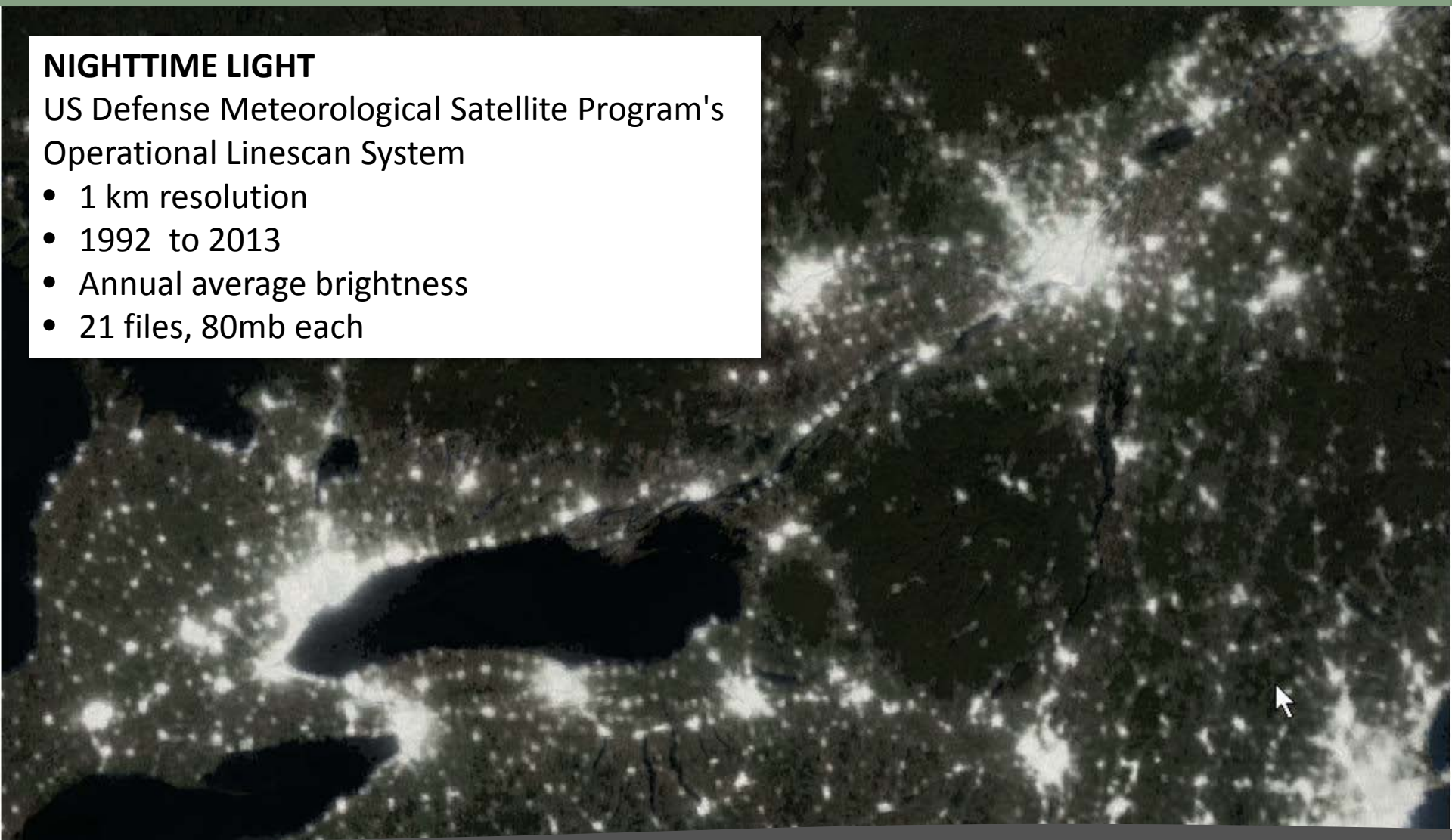




NIGHTTIME LIGHT

US Defense Meteorological Satellite Program's
Operational Linescan System

- 1 km resolution
- 1992 to 2013
- Annual average brightness
- 21 files, 80mb each





Navigating the Data Merge – September 28 | 2017

PAMPALON INDEX

Developed by R. Pampalon, available from INSPQ website

- Provided for census dissemination areas
- Census years - 1991, 1996, 2001, 2006, 2011
- 14 metrics including Deprivation index - material factor score; Deprivation index - social factor score, quintile position and percentile for each
- 5 files, < 100 mb each

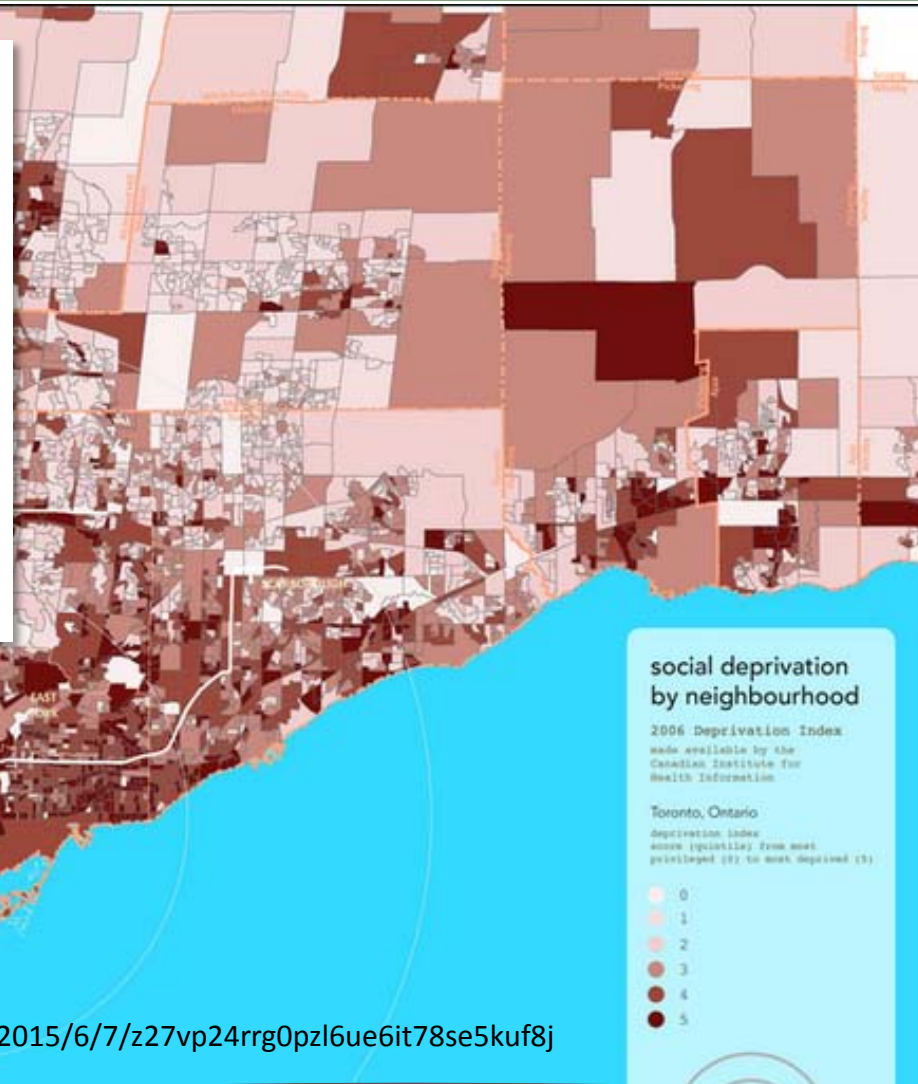
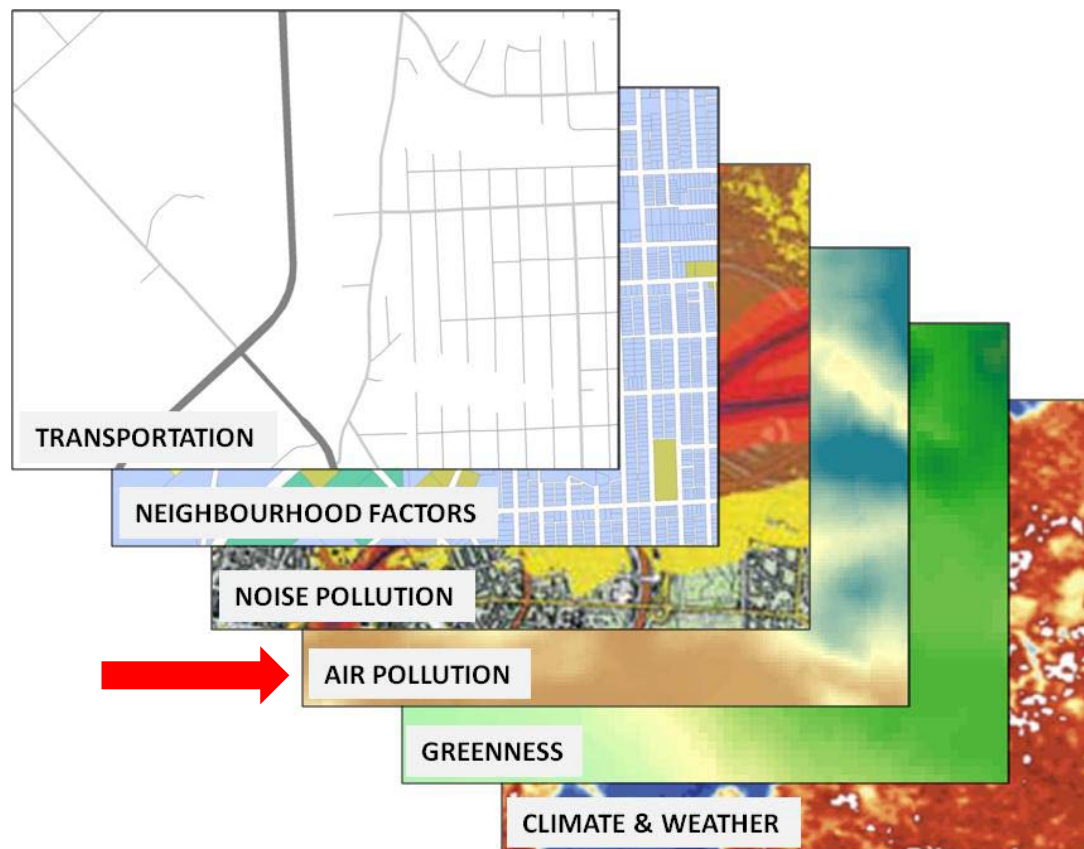


IMAGE: <http://robertwalterj.ca/rwj-blog/2015/6/7/z27vp24rrg0pzl6ue6it78se5kuf8j>



Navigating the Data Merge – September 28 | 2017

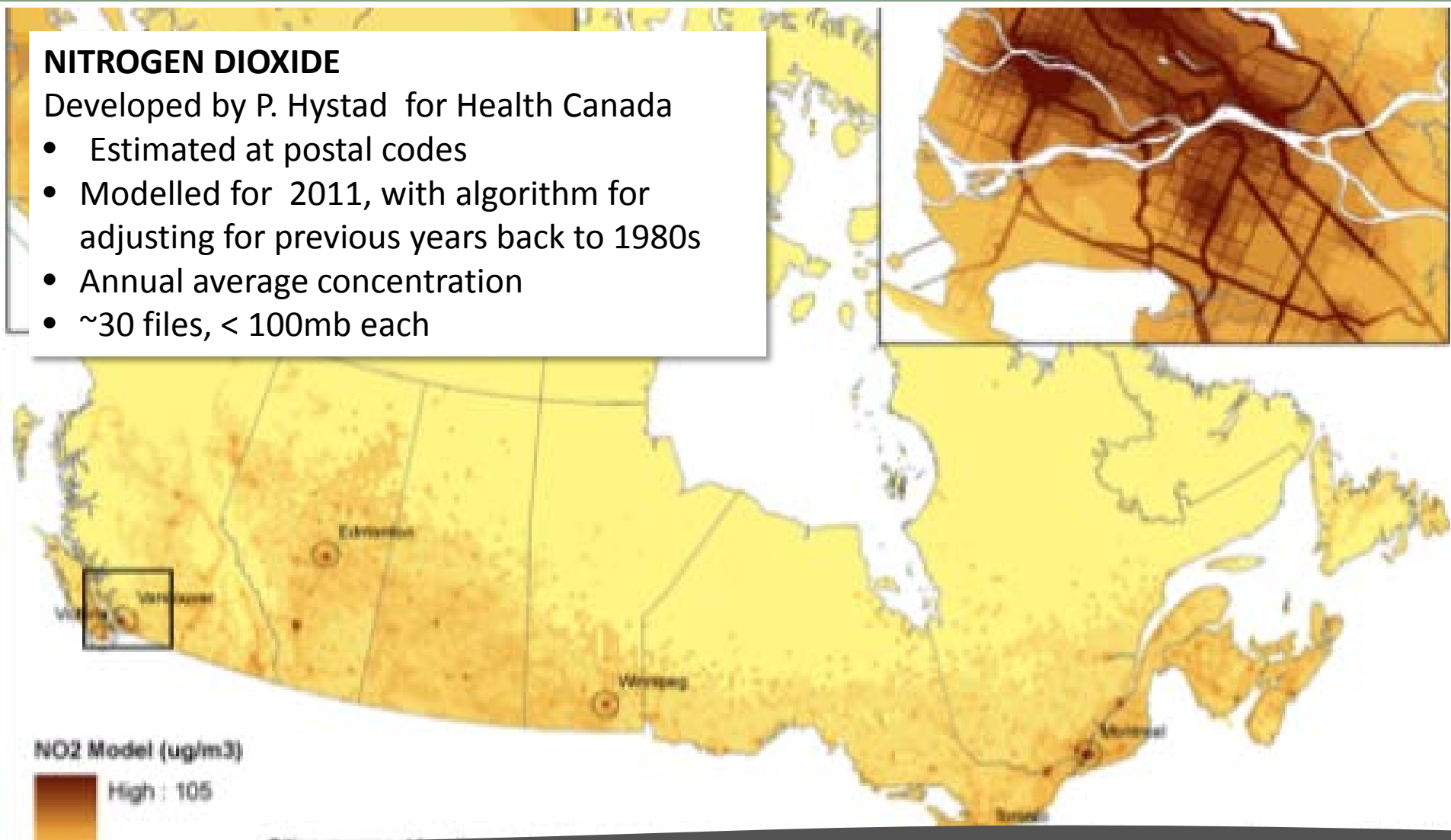




NITROGEN DIOXIDE

Developed by P. Hystad for Health Canada

- Estimated at postal codes
- Modelled for 2011, with algorithm for adjusting for previous years back to 1980s
- Annual average concentration
- ~30 files, < 100mb each

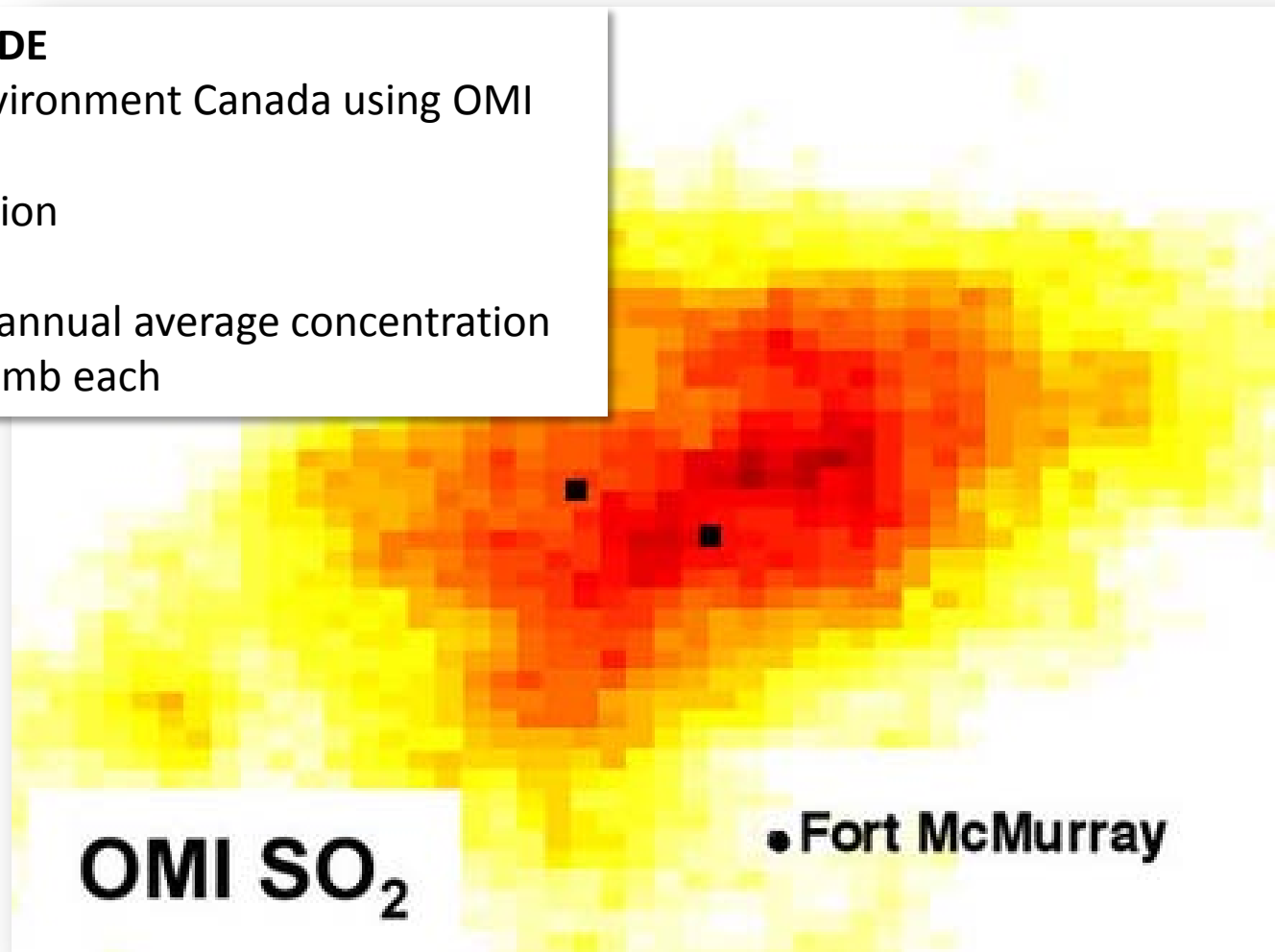




SULPHUR DIOXIDE

Modelled by Environment Canada using OMI satellite data

- 30km resolution
- 2005 - 2015
- One metric – annual average concentration
- 10 files, < 100mb each



<https://aura.gsfc.nasa.gov/science/feature-20120207.html>



FINE PARTICULATES

Developed by van Donkelaar and Martin et. al.
at Dalhousie University from NASA satellites

- 1km resolution
- 1998 - 2015
- Annual average concentration
- ~20 files, < 100mb each

Data SIO, NOAA, U.S. Navy, NGA, GEBCO
Image Landsat / Copernicus
Image IBCAO
Data LDEO-Columbia, NSF, NOAA

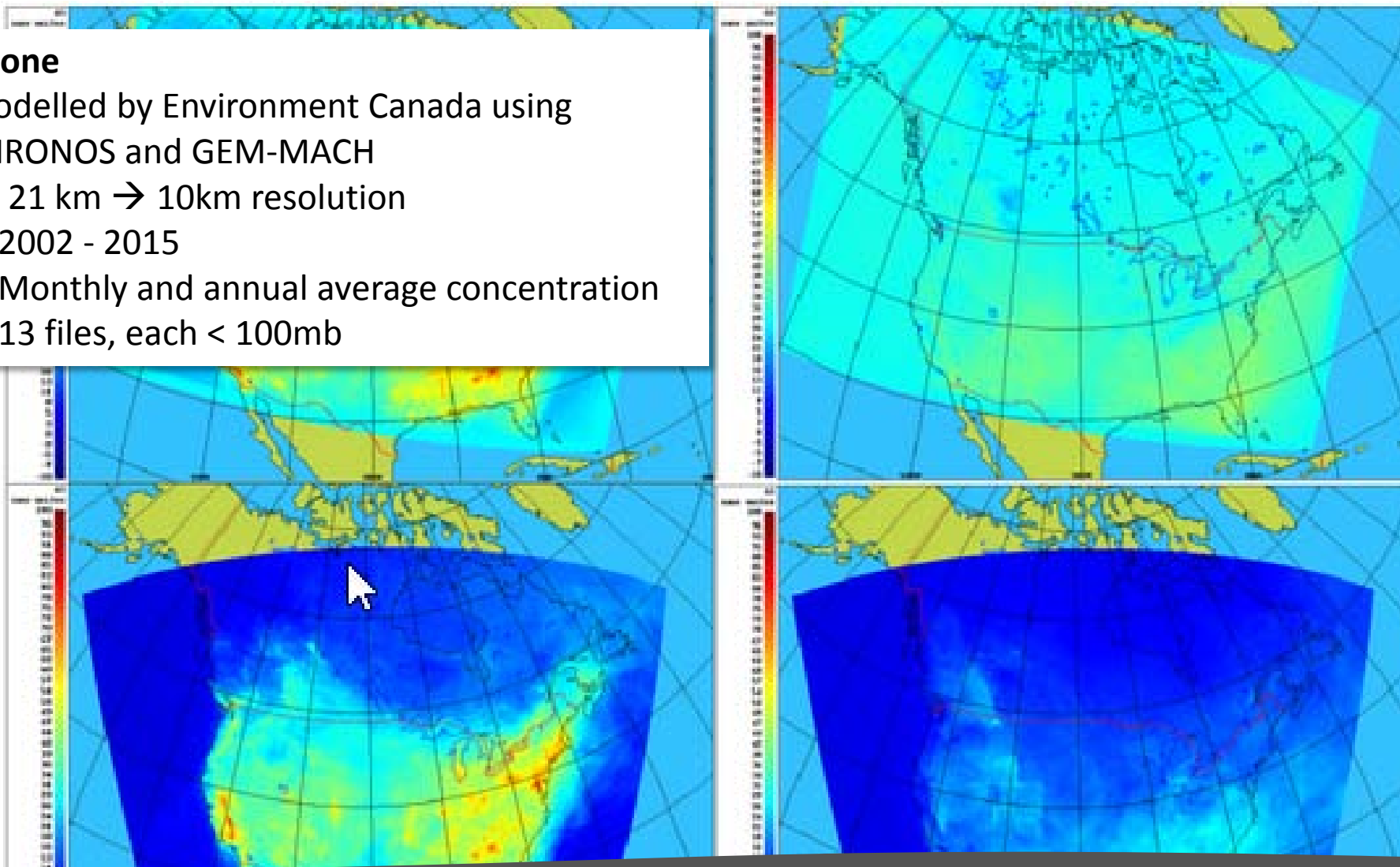
Google Earth



Ozone

Modelled by Environment Canada using
CHRONOS and GEM-MACH

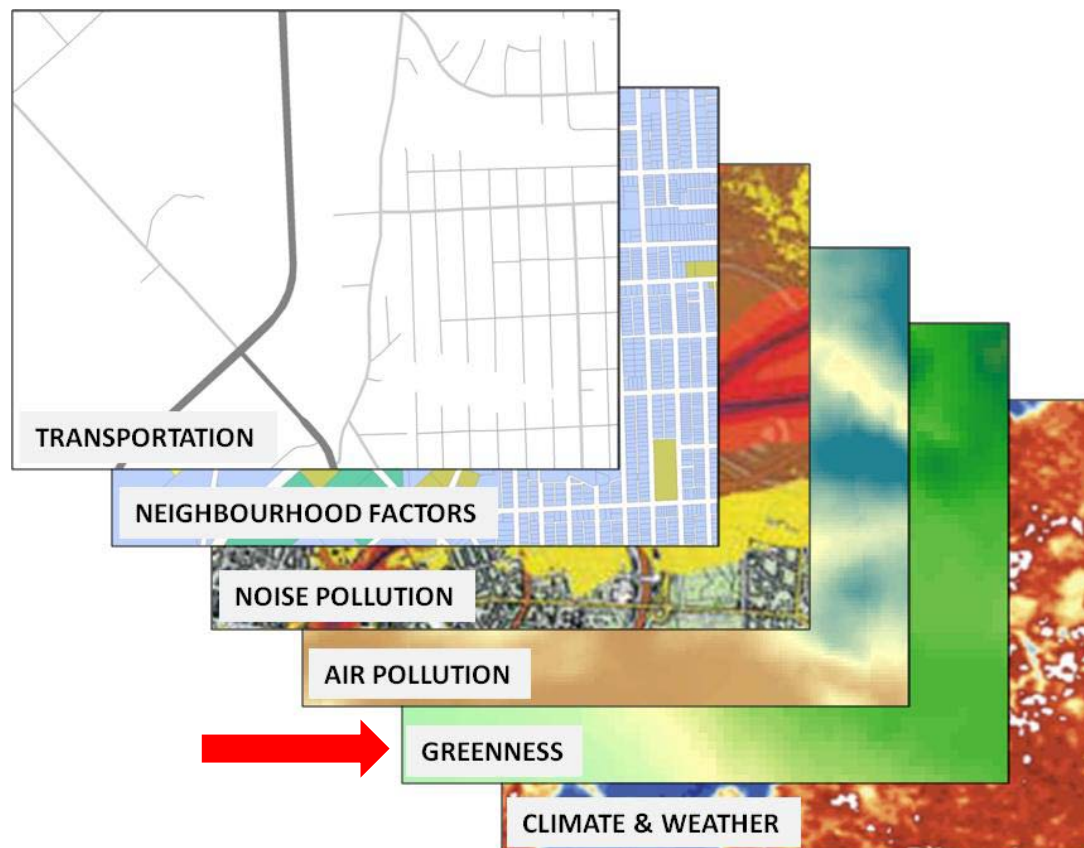
- 21 km → 10km resolution
- 2002 - 2015
- Monthly and annual average concentration
- 13 files, each < 100mb



CHRONOS



Navigating the Data Merge – September 28 | 2017

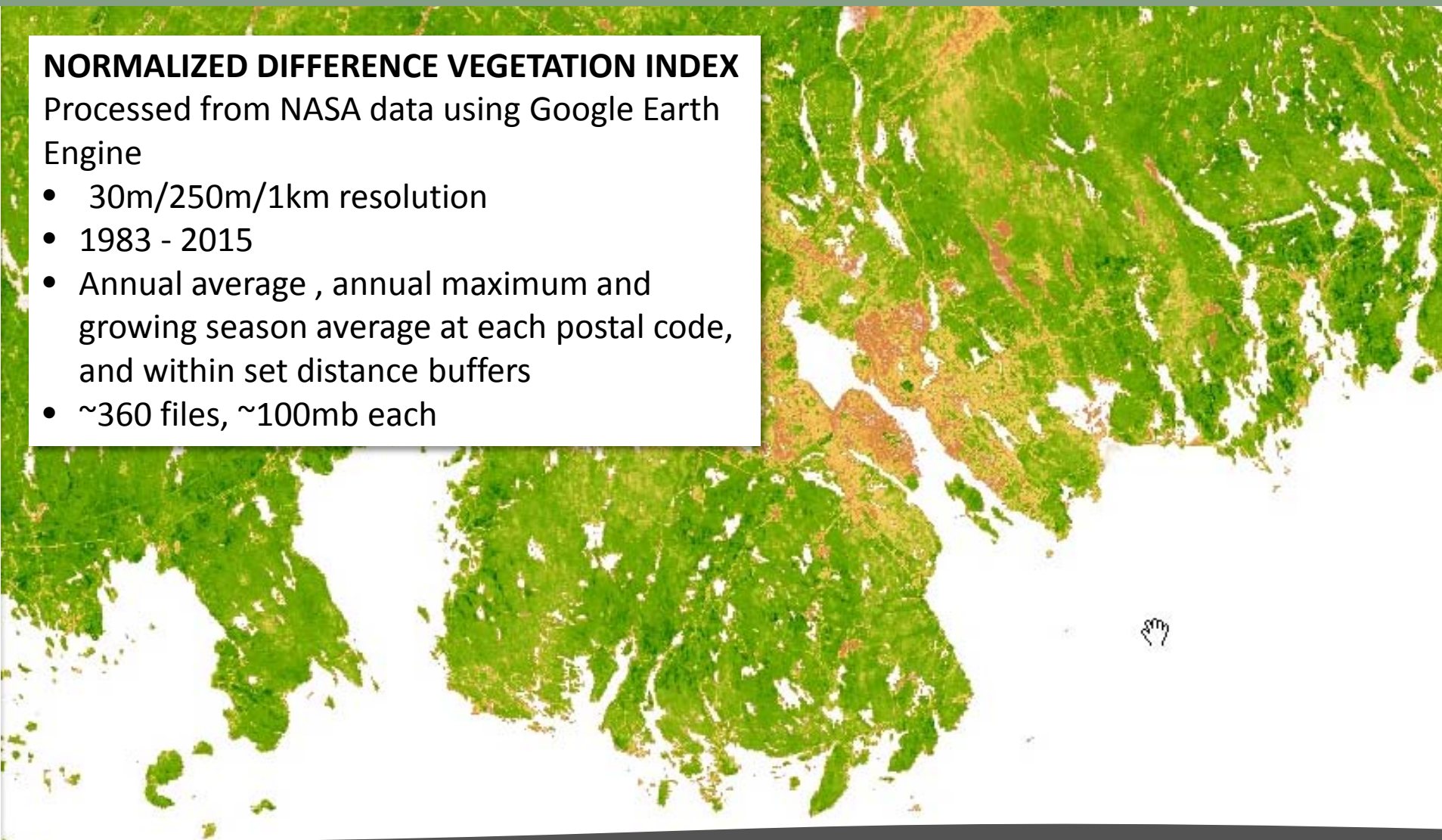


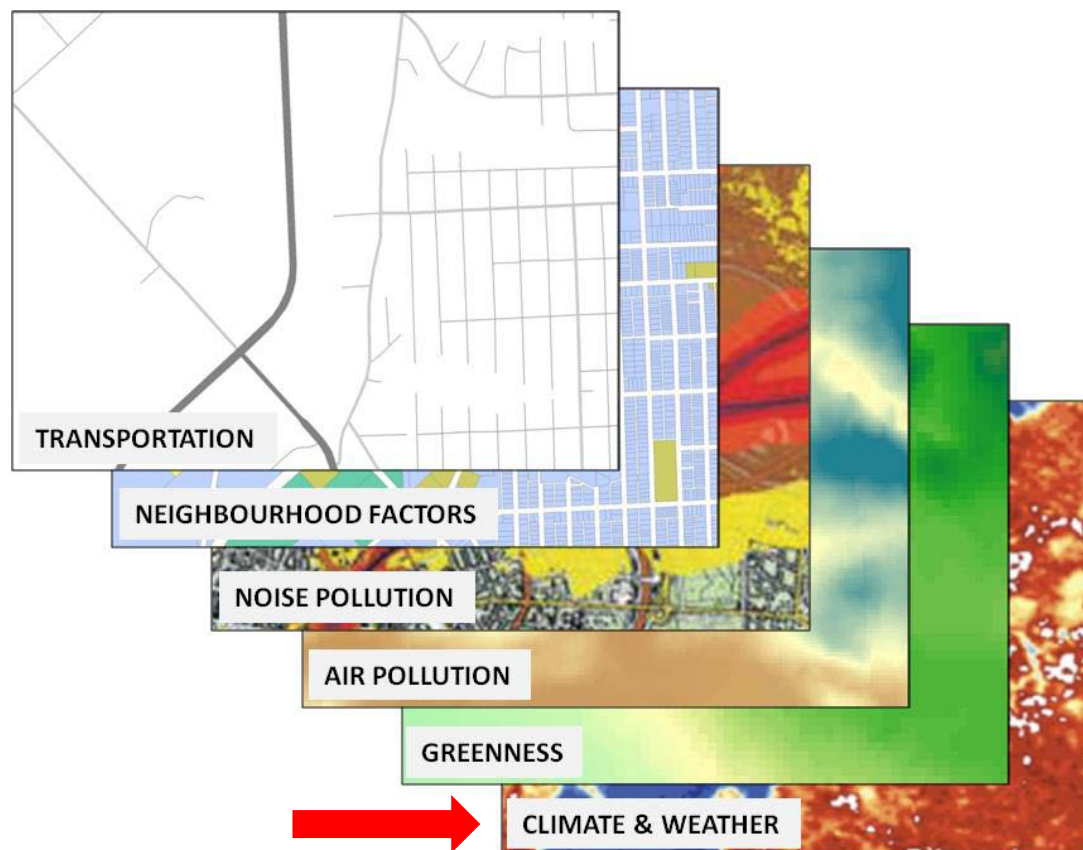


NORMALIZED DIFFERENCE VEGETATION INDEX

Processed from NASA data using Google Earth Engine

- 30m/250m/1km resolution
- 1983 - 2015
- Annual average , annual maximum and growing season average at each postal code, and within set distance buffers
- ~360 files, ~100mb each



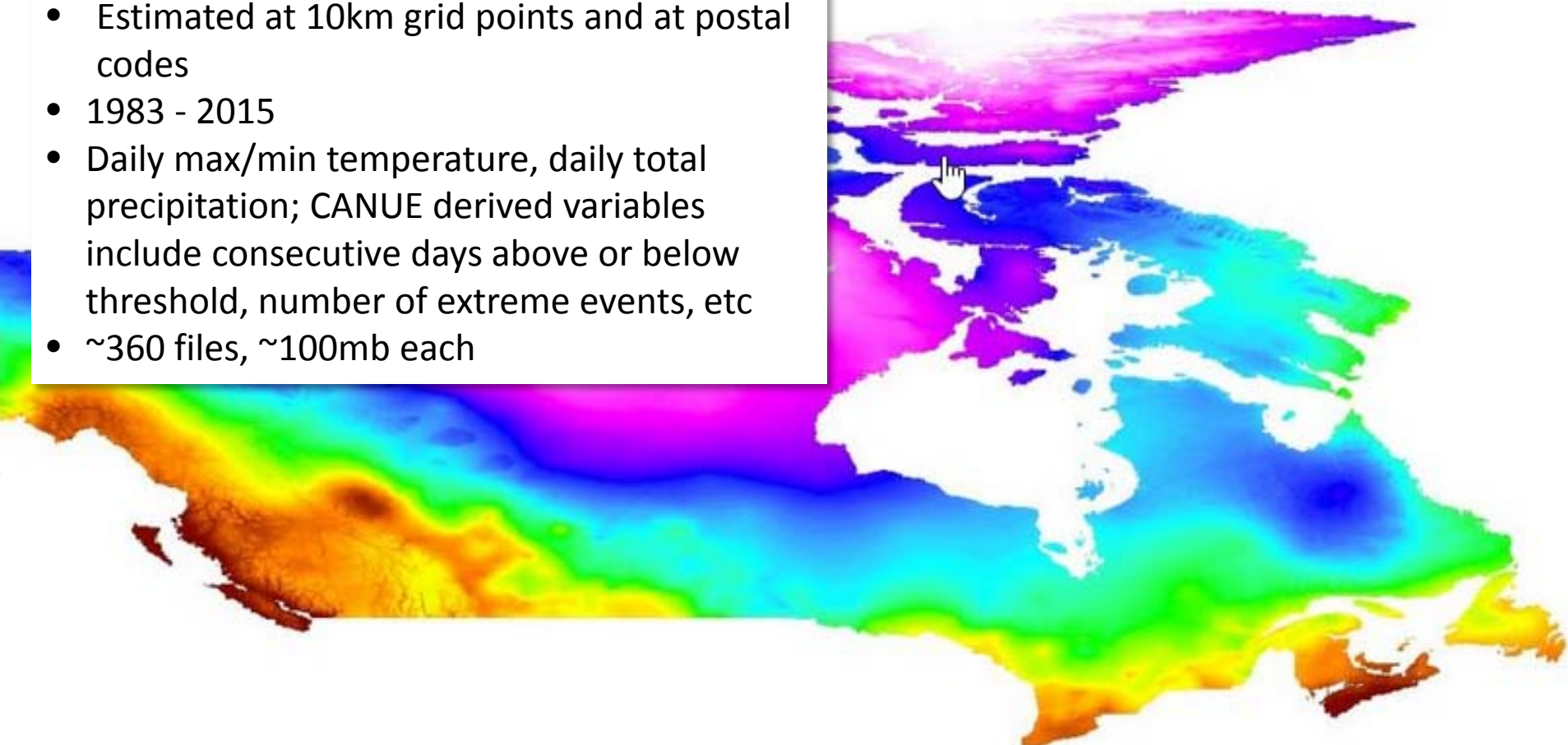




INTERPOLATED WEATHER STATION DATA

Processed from station observations by NRCAN

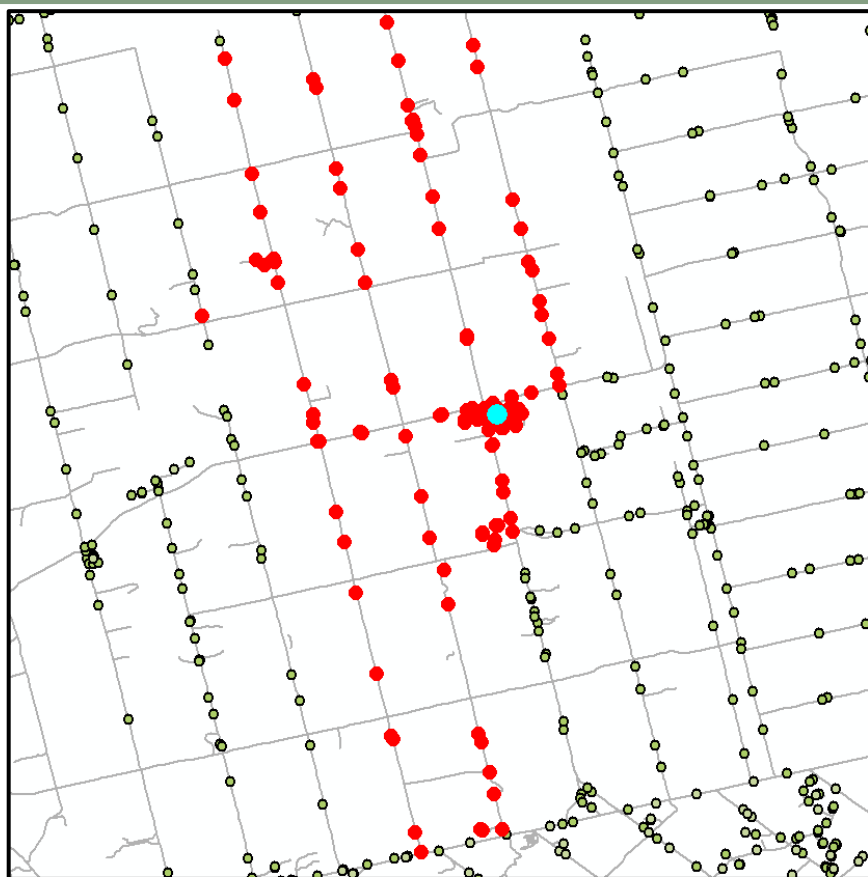
- Estimated at 10km grid points and at postal codes
- 1983 - 2015
- Daily max/min temperature, daily total precipitation; CANUE derived variables include consecutive days above or below threshold, number of extreme events, etc
- ~360 files, ~100mb each





Postal Code Files

- **Two sources: DMTI Spatial and StatsCan PCCF**
- **CANUE using DMTI**
- **Includes all 6-digit postal code locations, retired and active to 1983**
- **Both urban and rural postal codes**

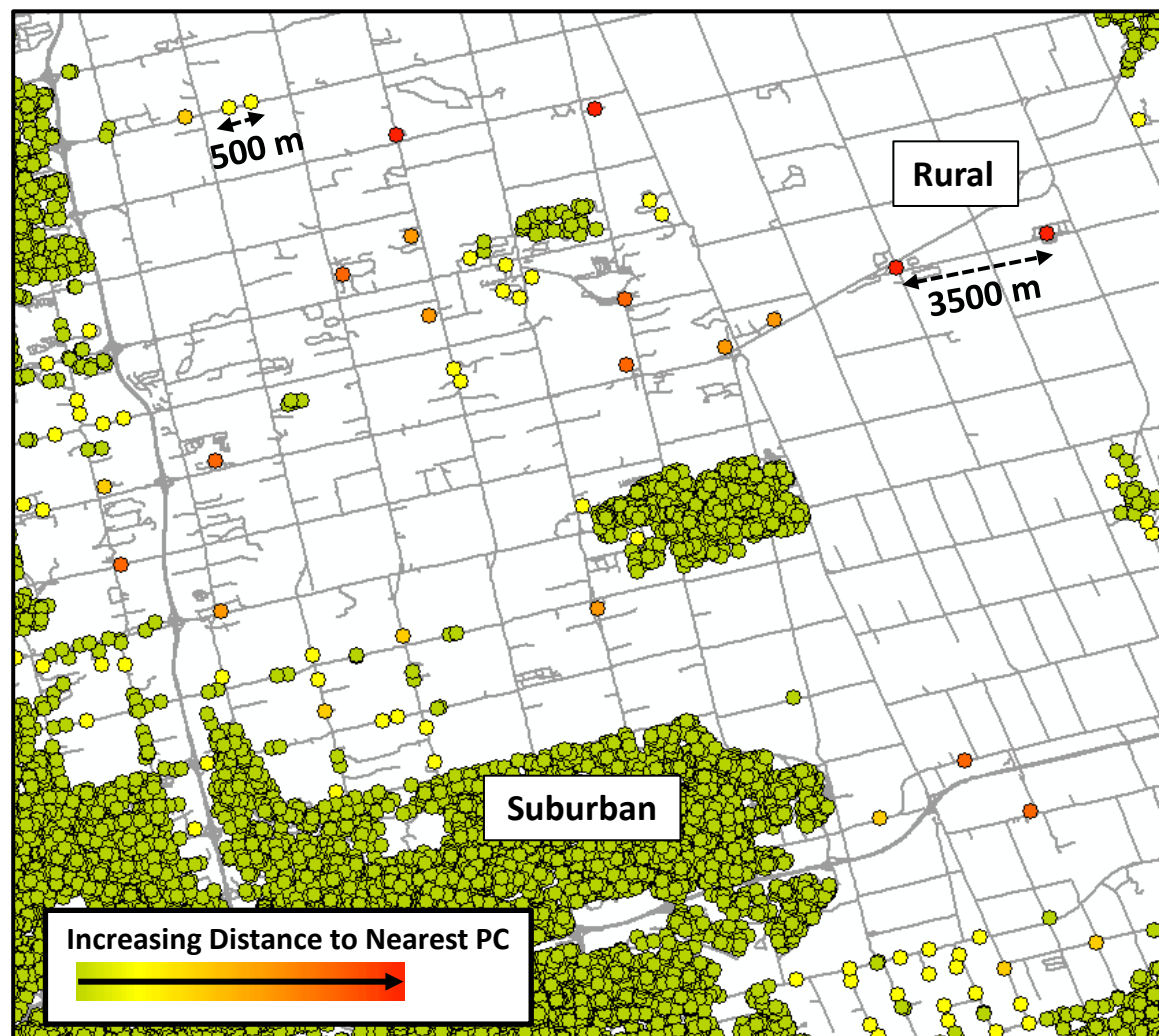


DMTI and PCCF sometimes contain more than one location for a postal code – especially in rural areas



Including some precision measures that will be useful for researchers:

- Distance to nearest postal code
- Indicator flag and total count of postal codes with multiple coordinate locations
- Delivery mode type



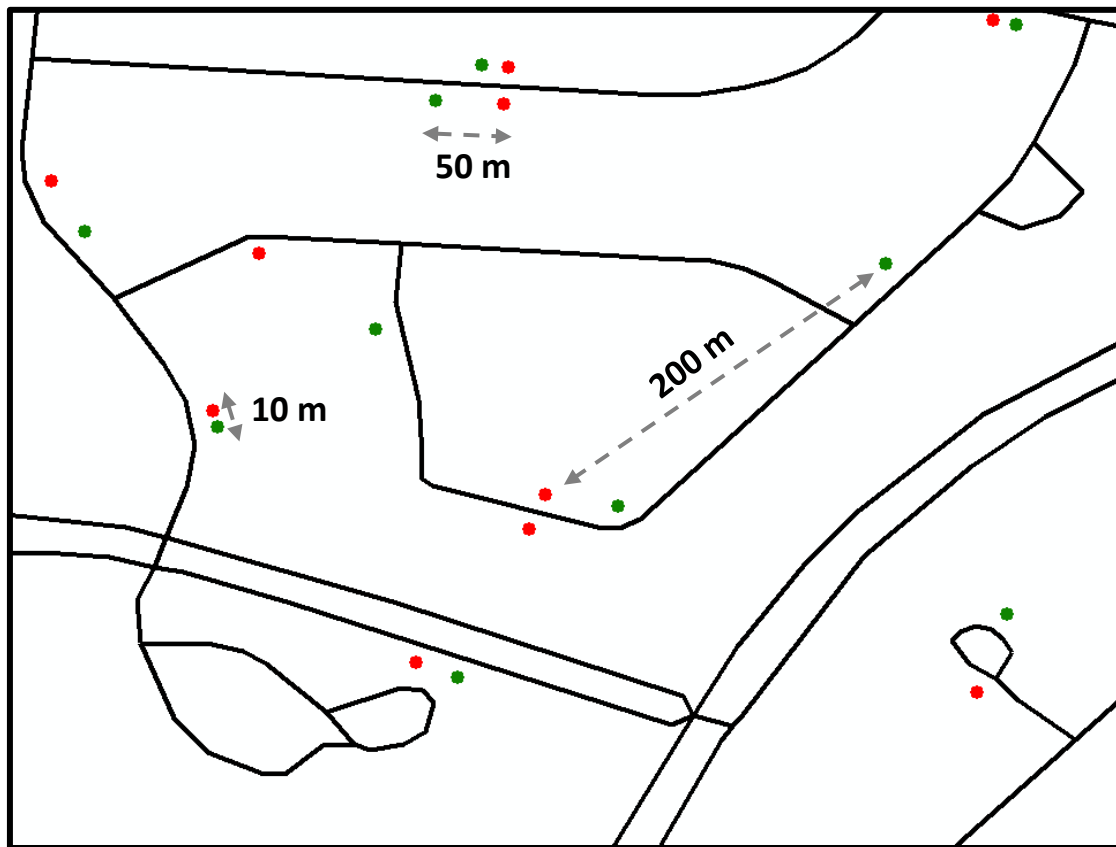


Navigating the Data Merge – September 28 | 2017



anced_PostalPoint_BirthRetire_SLI_v2002_Precision_ALL											
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	MASHTEUJATSH	MASHTEUJATSH	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

- Filter to remove postal codes serving general delivery mode e.g. serving a post-office box
- Filter to find postal codes serving apartment buildings or farms etc. for specific studies
- Filter or subset data using province, place name or municipality for local or regional studies



Comparing PCCF and DMTI Postal Code Single Link Indicator (SLI) Locational Differences, Oakville, ON

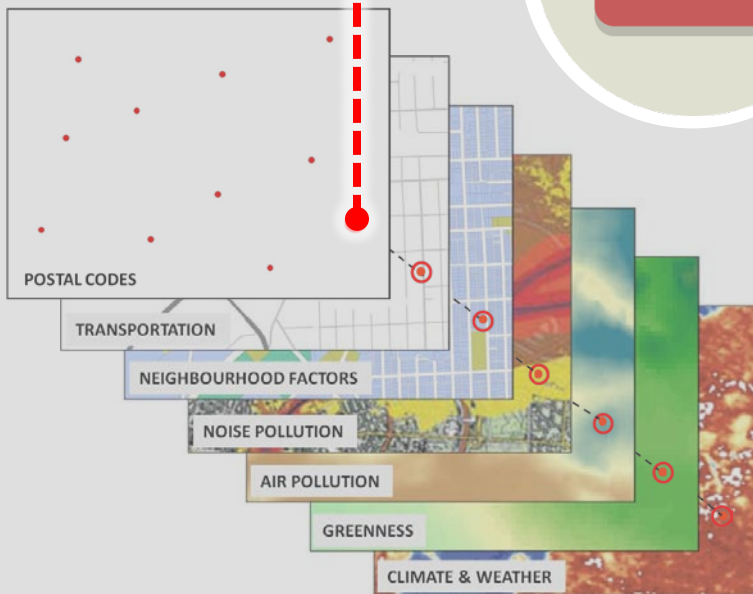
- DMTI is not exactly the same as PCCF
- Don't mix exposures based on PCCF with those based on DMTI.
- Developing a white paper on postal code file differences and guidance for use
- We can run most exposures for PCCF single locations if necessary.

Questions or comments?



Navigating the Data Merge – September 28 | 2017

Tools



	A	B	C	D	E	F	G
1	Number	GivenName	MiddleInitial	Surname	Gender	StreetAddress	City
2	1	Bruce	R	Bloch	male	3151 Ferrell Street	Argyle
3	2	Marie	E	Humphreys	female	3062 Bond Street	Woonsocket
4	3	Sylvia	H	Carter	female	1481 Lakeland Terrace	Westland
5	4	William	E	Bentz	male	3318 Briercliff Road	New York
6	5	Shelly	R	Preston	female	3592 Todds Lane	San Antonio
7	6	Chad	P	Henry	male	3553 Grant Street	Tyler
8	7	David	L	Richardson	male	1289 Metz Lane	Marlton
9	8	Stephen	A	Pond	male	4316 Bridge Avenue	Lafayette
10	9	Jenny	P	Thomas	female	2941 Harron Drive	Baltimore
11	10	William	V	Fries	male	4300 Tangleswood Road	Jackson
12	11	Julio	D	Bessette	male	4177 Lauren Drive	Madison
13	12	Jerry	J	Nicholas	male	2722 Elk Street	Irvine
14	13	Thomas	A	Hunter	male	4112 Stadium Drive	Franklin
15	14	Edmund	C	Chagoya	male	3685 Essex Court	Brattleboro
16	15	David	E	Meador	male	1215 Stratford Drive	Kona
17	16	Joan	L	Mayfield	female	3137 Pin Oak Drive	Whittier
18	17	Maria	H	Gomez	female	1723 Norlie Lane	Richmond Hill
19	18	Gregory	G	Miguel	male	3223 Breezeswood Court	Macksville
20	19	Gail	L	Griffin	female	2252 Arbutus Drive	Miami

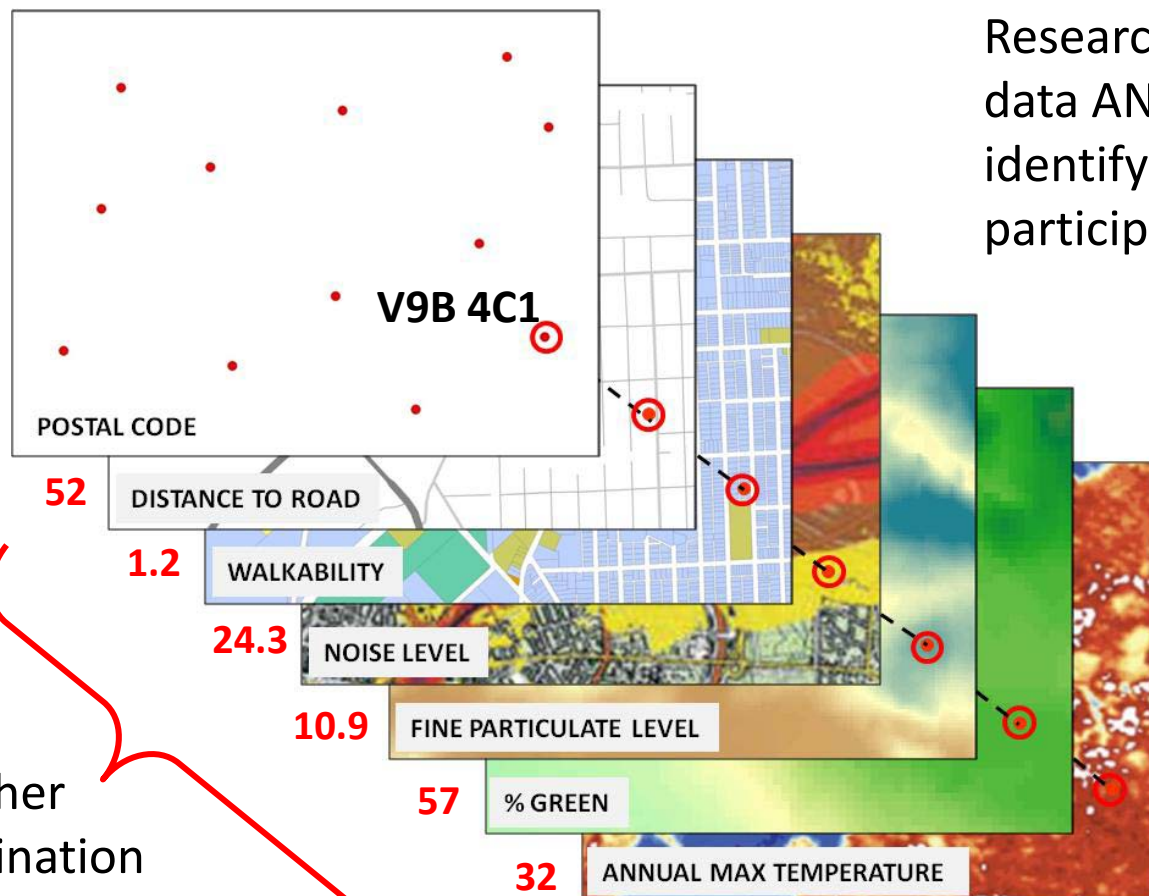
Data Delivery and Merging

Data Sets



Navigating the Data Merge – September 28 | 2017

Researcher with access to health data AND exposure maps could identify the postal code of the participant



No other combination of these values exists



Navigating the Data Merge – September 28 | 2017

CANUE tool development unique exposure combinations tool

This instruction is for installing Python and running the programming files on windows (assuming you already extracted the files in a folder). Linux machines have python installed (most of the time), and if not, the installation steps would be different.

Required Steps:

- 1- Install Python (skip this step if the Python is already installed)
- 2- Check the installed Python version
- 3- Run the correct programming file based on the Python version
- 4- Follow the program instructions

1- Install Python

If there is not any version of python installed on your computer (or you are not sure about that) keep reading, otherwise you can skip this step.

Open a command line (`cmd`). You can do that by any of the approaches mentioned in here:

<https://www.lifewire.com/how-to-open-command-prompt-2618089>

or

<https://www.howtogeek.com/235101/10-ways-to-open-the-command-prompt-in-windows-10/>

and then in the command line, navigate to the folder in which you extracted the files.

Recommended approach:

```

>>> Done!
Selected attributes: ['Max_100', 'Mean_500', 'Max_0']
Creating output file, containing number of repeats of each row based on the selected attributes.
This process might take a while, please wait...
Number of repeats is computed.
Writing output file...
All Done! Output file is stored in this folder as: NDUI_2004_output.csv
Have a good day :)

```

LONG	LAT	Mean_500	Max_0	Max_100	COUNT
-79.49	43.793	0.34	0.21	0.29	1
-82.9	46.181	0.61	0.55	0.71	21
-79.033	45.291	0.57	0.68	0.73	6
-73.44	46.025	0.22	0.12	0.29	8
-80.079	44.478	0.48	0.48	0.71	32
-81.744	44.076	0.65	0.6	0.72	17
-124.8	49.265	0.45	0.33	0.51	33
-79.295	48.526	0.5	0.6	0.69	13
-63.311	46.448	0.62	0.62	0.77	4
-75.92	45.31	0.39	0.25	0.33	3
-83.768	46.317	0.54	0.6	0.69	23
-72.233	48.574	0.31	0.01	0.75	1
-68.147	49.248	0.26	0.16	0.56	5

```

Mean_500
Max_0
Attribute list (type DONE when you are done!)
['Mean_500']
Attribute list (type DONE when you are done!)
['Mean_500', 'Max_0']
Attribute list (type DONE when you are done!)

```

784,000 records

27,000 unique

(3 percent)



Navigating the Data Merge – September 28 | 2017

CANUE tool development
temporal aggregation tool

Some internal – when base data very large
Plan to integrate into CANUE interface eventually

IndividualID	Date1	PostalCode1	Date2	PostalCode2	Date3	PostalCode3	Date4	PostalCode4	Date5	PostalCode5
1	7/23/1994	P9N1X9	12/23/2005	N3P1W8	11/1/2006	N1E7G5	7/29/2011	S6V0P8	11/1/2015	J9J1K4
2	6/6/2015	L5M6Z4	1/7/2015	L8E5V7						
3	5/27/1992	K1N9N1	1/5/1997	J1E2W6	12/29/2002	K8A5E2	7/14/2015	H1X2L7		
4	12/30/1992	K2A3J1	11/26/2006	M5M4E3	11/19/2011	V6P4B7	5/12/2015	N9B3Y7	6/19/2015	M2L2G1
5	4/28/2005	V4A1M3	10/19/2015	G5H3B6						
6	9/11/2005	M9M1E6	5/26/2006	M8Z5E3	7/13/2008	G6V0E7				
7	4/2/1999	T2A0R8	8/26/2006	J1K1W7	11/13/2011	E1N4V4	2/14/2012	V2M1E3		
8	3/5/1993	H1M3M5	3/19/1997	N6C1Z7	9/10/2006	N9J3V5	3/21/2012	V4N3C6		
9	8/17/1992	T2V1S2	7/7/1992	E1A1K3						
10	1/9/1984	S6H3Y1	4/28/2004	B2N2X9	1/18/2008	H9H4M5	6/27/2013	L2M0B5		
11										
12	5									
13	6									
14	9									
15	3									
16										
17	1									
18	3									
19	4									
20	11									

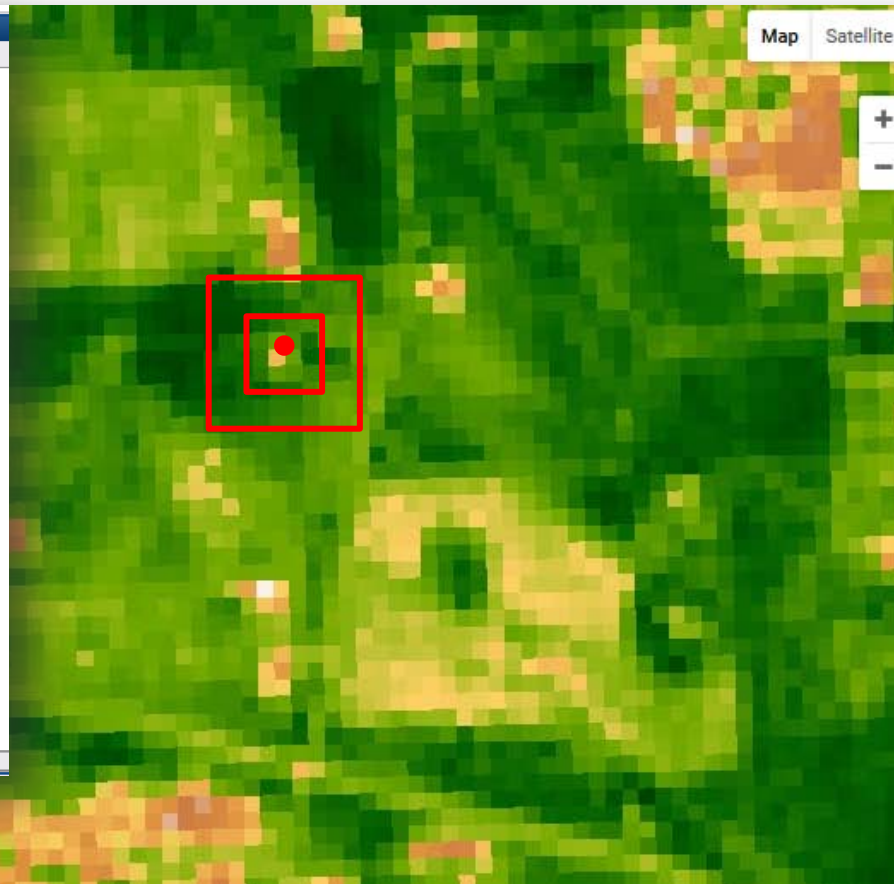
IndividualID	LivingYear	PostalCode	NDVI_max
1	2009	N1E7G5	0.5842
1	2010	N1E7G5	0.5827
1	2011	N1E7G5	0.5552
1	2011	S6V0P8	0.6651
1	2012	S6V0P8	0.6157
1	2013	S6V0P8	0.6159
1	2014	S6V0P8	0.6159
1	2015	S6V0P8	0.649
1	2015	J9J1K4	0.7158
2	2015	L5M6Z4	0.5274
2	2015	L8E5V7	0.6085

IndividualID	LT_Avg_NDVI_max
1	0.6123
2	0.5842
3	0.5827
4	0.5552
5	0.6651
6	0.6157
7	0.6159
8	0.6159
9	0.5842
10	0.5827
11	0.5552
12	0.6651
13	0.6157
14	0.6159
15	0.6159



CANUE tool development spatial aggregation tool

```
Editor - /mnt/canue/GitHub/GEE/GeoTIFF_Processing/compute_buffer_values.m
compute_buffer_values.m x +
35 - list_of_indices = find(postalCodes_to_tiles(:,tile_index)==1);
36
37 - start_long = temp_metadata.LongitudeLimits(1,1);
38 - start_lat = temp_metadata.LatitudeLimits(1,2);
39
40 - delta_long = temp_metadata.CellExtentInLongitude;
41 - delta_lat = temp_metadata.CellExtentInLatitude;
42
43 - for i = 1:size(bufers,2)
44 -     temp_buffer_in_meters = buffers(1,i);
45
46 -     % The resolution of Landsat images is 30 meters. That means
47 -     % each pixel covers 30 meters on the ground. I rounded down the
48 -     % result because we need to have integer values to work with
49 -     % pixels, and also, since the current pixel is also included in
50 -     % the window, rounding up makes more sense.
51 -     pixels_range = floor(temp_buffer_in_meters/30);
52
53 -     % For all the postal codes that are in connection with the
54 -     % current tile, we are interested to know information about the
55 -     % coverage area such as number of valid pixels (that are not NaN)
56 -     % and also sum or max of their values.
57 -     for j = 1:size(list_of_indices,1)
58 -         temp_index = list_of_indices(j,1);
59 -         temp_long = postalCodes(temp_index,1);
60 -         temp_lat = postalCodes(temp_index,2);
61
62 -         % These are the coordinates of the current postal code in
63 -         % the pixel grid of the tile.
64 -         long_index = ceil((temp_long - start_long)/delta_long);
65 -         lat_index = ceil(-(temp_lat - start_lat)/delta_lat);
```





Navigating the Data Merge – September 28 | 2017

CANUE tool development visualization and data extraction

Project Partners

BASIC PROJECT INFORMATION

Project Number (Name)	
Project Number	
Project Name	
Project Long Name	
Description	
Project Start Date	
Project Finish Date	
Organization	
Client Name	
Project Currency	

Download For Review
 Download For Update

Distribution Rule	
Project Type	
Project Template	
Operating Unit	
Project Currency	
Project Status	

CLASSIFICATIONS

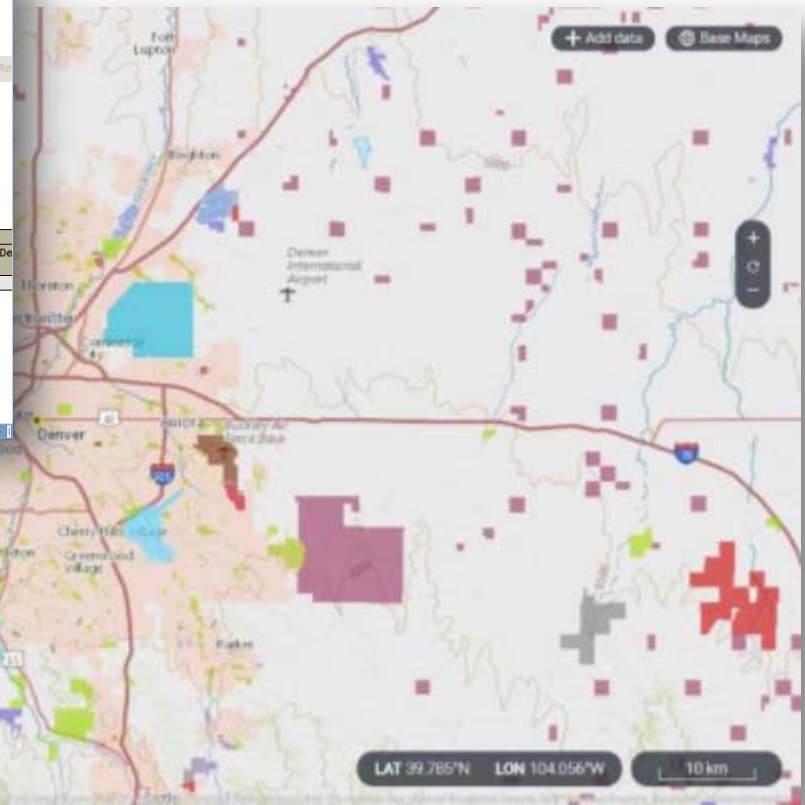
Category	Code

ATTACHMENTS

Creation Date	Category	De

KEY MEMBERS

Role	Employee Name	Start Date	End Date

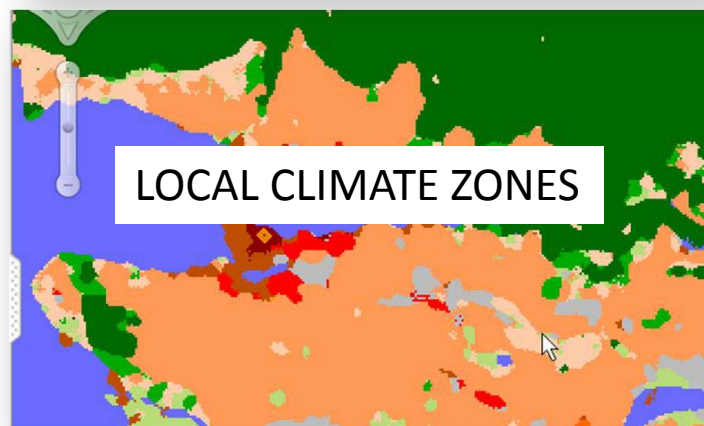


Questions or comments?

What's next:

- **Data ready to deliver in October**
- **Looking forward to working with health data holders to integrate CANUE data – tool development**
- **Form a stakeholder group that can help us meet the needs of health data holders:**
 - **Ideas for promoting use of CANUE data**
 - **Help resolve issues as we move forward**

In development....



and more....

Questions or comments?

THANK YOU !

