One Ecosystem 9: e122079
doi: 10.3897/oneeco.9.e122079

OPEN ACCESS

Methods

A GIS methodology for mapping regional and community vitality for Canada using the CanEcumene 3.0 Geodatabase with census data

Brian Eddy [‡, §]

‡ Natural Resources Canada, Corner Brook, Canada
§ Memorial University of Newfoundland, Corner Brook, Canada

Corresponding author: Brian Eddy (briang.eddy@nrcan-rncan.gc.ca)

Academic editor: Brian D. Fath

Received: 01 Mar 2024 | Accepted: 29 May 2024 | Published: 26 Jun 2024

Citation: Eddy B (2024) A GIS methodology for mapping regional and community vitality for Canada using the CanEcumene 3.0 Geodatabase with census data. One Ecosystem 9: e122079.
https://doi.org/10.3897/oneeco.9.e122079

Abstract

Many ecosystem-based management (EBM) and related applications require integrating geospatial information about socio-economic conditions of human populated areas within a study area. However, integrating socio-economic data in such a way that it can be related to ecological data is challenging due to issues associated with spatial representation between socio-economic and ecological data frameworks. In Canada, this problem is particularly acute given its large geographic size, diversity of environments and highly irregular population distribution. Although several indices have been developed for Canada related to well-being and vulnerability, their suitability for EBM-related applications is limited. This article presents a GIS-based methodology for mapping *regional and community vitality index* (RVI/CVI) for Canada using standard Census data integrated with the CanEcumene 3.0 Geospatial Database (GDB). The method uses percentile ranks of five sub-indicators of vitality covering population growth, age structure, education, employment and economic wealth. Results reveal a number of notable patterns and trends in socio-economic conditions across the country and across different types of communities and regions. Most notable are decreasing CVI values from economic core regions to rural and remote communities; decreasing scores from high population centres to lower populated areas and lower scores for Indigenous communities when compared with non-

Indigenous communities. A series of maps show variation in RVI/CVI values for specific locations with changes over time.

Keywords

Ecosystem-based Management (EBM), socio-economic, community vitality, regional vitality, well-being, vulnerability, GIS

Introduction

As part of a comprehensive and inclusive approach to Ecosystems-based Management (EBM), many ecosystem analysis and assessment applications require integrating geospatial information about socio-economic conditions of human populated areas within a study area (Kappel et al. 2006, Eddy et al. 2014). Common examples include natural resource and environmental management (NREM), ecosystems services assessment (ESA) and climate change impact and adaptation (CCA). Socio-economic data used in these applications need to be integrated with bio-physical data for a region in a spatially explicit manner to enable analysis of relationships between socio-economic and proximal environmental conditions. Integrating socio-economic data in such a way that it can be related to ecological data is is challenging due to geospatial frameworks used to collect socio-economic data do not often align with natural boundaries and patterns in a landscape (Cumming et al. 2006,Eddy et al. 2020).

In Canada, this problem is particularly challenging given its large geographic size, diversity of environments and highly irregular population distribution. Currently, several high priority policy issues require national level mapping of socio-economic conditions of communities across Canada including climate change adaptation (NRCan 2024), wildfire risk mapping (Erni et al. 2021, Erni et al. 2024) and cumulative effects assessment related to new natural resource development projects (Antwi et al. 2023, IAAC 2024). Ideally, information required for these applications about human populated areas would benefit from a spatial index of the overall state of socio-economic conditions of communities and regions. Such an index would help provide information for analyses of relationships that communities have with their proximal natural environments and further help contextualise important human-environment issues.

This article presents a GIS-based methodology for mapping *regional and community vitality index* (RVI/CVI) for Canada using standard census data integrated with the CanEcumene 3.0 Geospatial Database (GDB) (Eddy et al. 2023). First, concepts related to vitality and related applications for Canada are reviewed from which an alternative approach is described. Second, a GIS-based method for mapping both regional and community vitality is described in detail. Third, results are presented and discussed in terms of notable patterns observed, followed by conclusions that discuss the benefits and limitations.

Background

As with many regions around the world, researchers in Canada are tasked with developing a scientific information base to support decision-making for a host of concurrent EBM-related issues including, but not limited to, climate change adaptation (CCA) (NRCan 2024), regional and project-level impact assessments (RAs/IAs) (Noble and Harriman Gunn 2010, IAAC 2024) and natural hazards risk analysis and mapping (Erni et al. 2021, CFS 2024, PSC 2024). Although environmental and ecological dimensions may differ, a common requirement amongst these applications is the need to geospatially map and integrate socio-economic data with spatialised environmental and ecological data.

Different terms and concepts have been used to describe overall socio-economic conditions of communities including well-being, resilience, vulnerability, vitality and sustainability (Grigsby 2001, Beckley et al. 2002, Dale et al. 2010, Etuk and Acock 2016). How these terms are defined and applied is often contextual depending on the application and scope of analysis and methods used are diverse. The majority of studies tend to focus on individual communities and involve qualitative research methods and not on 'mapping' indicator values for all or many communities (Grigsby 2001, Beckley et al. 2002, Stedman et al. 2004, Pearce 2005, Dale et al. 2010, Lax and Krug 2013, Etuk and Acock 2016). Studies conducted in other countries that focus on quantitative methods use a variety of definitions and methods depending on the geographic and temporal scope and end-user requirements (Flanagan et al. 2011, Ahsan and Warner 2014, Pravitasari et al. 2018, Zeng et al. 2018, Wangdi 2022).

For our purposes, the aim is to develop a quantitative method for mapping socio-economic conditions for over three thousand communities in Canada over multiple time periods and at a scale that is amenable for integration into regional to national level EBM applications. Herein, the concept of 'vitality' is used to capture the overall state of socio-economic conditions of a community or region whereby communities with high vitality have a thriving economy and population and those with low vitality are struggling to survive. This is important for EBM-related applications as community relationships with proximal environmental conditions are often associated with the general vitality in the community. For example, communities and regions experiencing population decline, high unemployment with an aging demographic (low vitality) may be in greater need of resource development projects than those that are socially and economically thriving (high vitality) and such needs may be reflected in the values of the local stakeholders in a decision-making process.

Developing a method for mapping an index of regional and community vitality requires a more nomothetic approach than the idiographic approaches of the aforementioned studies (Sui and Turner 2022). The closest comparisons available for Canada include several studies focused on developing similar indicators for communities in Canada which use a variety of methods for different purposes (Table 1). The Canadian Index of Wellbeing (CIW) uses qualitative survey data, but does not provide data for all communities in Canada nor systematically cover multiple time periods (Scott 2016). A similarily named Canadian Well

Being Index (CWB) maps all communities in Canada at the census sub-division (CSD) level (ISC 2021). Developed primarily to compare conditions of Indigenous communities with non-Indigenous communities, the CWB is not suitable for integration with environmental and ecological datasets due to its use of standard CSD boundaries and its lack of a peer-reviewed methodology used in ranking variables (ISC 2016).

Table 1

Table 1.

Summary of recent efforts for socio-economic index database development for Canada. See text for discussion. (Note: PCA = Principal Component Analysis, CT = Census Tract, DA = Dissemination Area, CSD = Census Subdivision)

Source	**Database/Index**	**Purpose**	**Core Method**	**Spatial Units**	**Time Periods**	**Comments**
Chakraborty et al. (2020)	Social Vulnerability Index (SoVI)	Flood risk mapping	PCA on 49 variables	CT	2016	Data not available
Chan et al. 2015	Socio-Economic Status (SES) Index	Environmental Health	PCA on 22 variables	DA	2006	Data not mapped geographically
ISC (2021), ISC (2016)	Canadian Well Being (CWB) Index	Comparing Indigenous and non-Indigenous Communities	Variable Weights	CSD	2016, 2021	Method description incomplete
Journeay et al. (2022)	Canadian Social Vulnerability Model (CanSVM)	Natural hazards mapping	Hazards of place model	DA/CT	2016	Complex framework for specialists
Scott (2016)	Canadian Index of Wellbeing (CIW)	Social science research	Qualitative survey	Limited	various	This index is not mapped for communities
StatsCan (2019)	Canadian Index of Multiple Deprivation (CIMD)	Study of inequalities of socio-economic conditions	Factor analysis on 17 variables	DA	2016, 2021	Designed for social analysis

A Social Vulnerability Index (SoVI) developed by Chakraborty et al. (2020) for flood risk mapping uses dissemination areas (DAs) of the Canadian census framework. The SoVI index is an output of a principal component analysis (PCA) on 49 variables and, although result scores are available, analytical data are not accessible for evaluation. Similarly, Chan et al. (2015) use PCA on 22 census variables at the DA level to develop a Socio-Economic Status index (SES) for all communities in Canada, but these data were limited to 2006 and were not developed for mapping purposes. Journeay et al. (2022) also use a

PCA approach using data derived from the DA and Census tract (CT) level and mapped using a raster-based human settlement layer, but coverage is only for the year 2016. Although not directly related to community vitality, StatsCan (2019) developed a Canadian Index of Multiple Deprivation (CIMD) using factor analysis on 17 variables, but only for the years 2016 and 2021.

In addition to these noted limitations with these sources, it is worth noting that, while studies that map data at a dissemination area (DA) level data may be suitable for environmental applications (with some modifications), the use of PCA or other factor analysis methods limits accessibility for lay people in terms of their ability to understand how the indices were calculated. Additionally, because statistical distributions and properties of variables change from one census period to another, it may be difficult to normalise indices for time-series analysis.

For purposes of integration with EBM applications, several requirements are identified:

1. It would be helpful to have a standardised compound index that captures the overall state of socio-economic conditions of communities, with subdimensions that are easy to dissect and understand for lay people and practitioners.
2. Given how communities do not exist in isolation, it would also be useful to assess how communities compare within a region.
3. Another requirement is for such indicators to be provided as time-series to allow changes in socio-economic conditions to be analysed over time in relation to environmental changes or changes in economic activity related to natural resource development.

As the studies reviewed above fall short of meeting these requirements, this article presents an alternative approach to meet the stated requirements.

Approach

The approach taken here considers two aspects to ensure adequate integration with EBM and to meet the stated requirements listed above:

1. *Spatial Representation*: pertaining to spatial units used for communities and how they are represented in relation to spatial environmental and natural scientific data and
2. *Index Composition*: how well an index represents socio-economic conditions, the transparency of the method used and its understandability and relevance for lay people (public and stakholders, in addition to specialists).

1) <u>Spatial Representation:</u> One of the most critial requirements for EBM is the ability to relate index values to the physical environment of communities and regions, particularly in areas where there is natural resource development and/or environmental conditions of concern. Integrating socio-economic data in such a way that it can be related to the environment is not a trivial process and the spatial unit for representing communities is of

critical importance (Cumming et al. 2006, Eddy and Dort 2011, Eddy et al. 2020). Studies reviewed above use standard census frameworks as the spatial unit of analysis. This is problematic because the euclidean geospatial frameworks used to collect census data do not often align with non-euclidean natural boundaries and patterns related to the physical geography and environment of regions where communities are situated. Failure to address this issue could result in violating fundamental cartographic principles resulting in misrepresentation of communities and regions and misinterpretation of results (MacEachren et al. 2005).

To address these aspects, this study makes use of the CanEcumene 3.0 Geodatabase (GDB) (Eddy et al. 2023). This database (GDB) was developed using natural boundaries for all communities in Canada (Eddy et al. 2020) (Fig. 1). The spatial unit of analysis in the CanEcumene GDB can be either point or area representations depending of the scale of analysis and requirements. Area representation of communities are represented by natural boundaries derived from satellite DMSP 'night-lights' imagery for the year 2010. Identifiers from Census sub-divisions (CSDs) were attached to each point and polygon feature representing communities thereby providing an alternative spatial framework from which to map and analyse Census variables and derivatives. Data can be mapped using either point or polygon features, but unlike the use of Census polygon features, CanEcumene data do not obscure other environmental data when displayed simultaneously. Additionally, because CanEcumene data contain accurate locations of communities and in contrast to choropleth mapping with standard Census boundaries, data can be used for spatial interpolation for a more accurate representation of regional patterns.

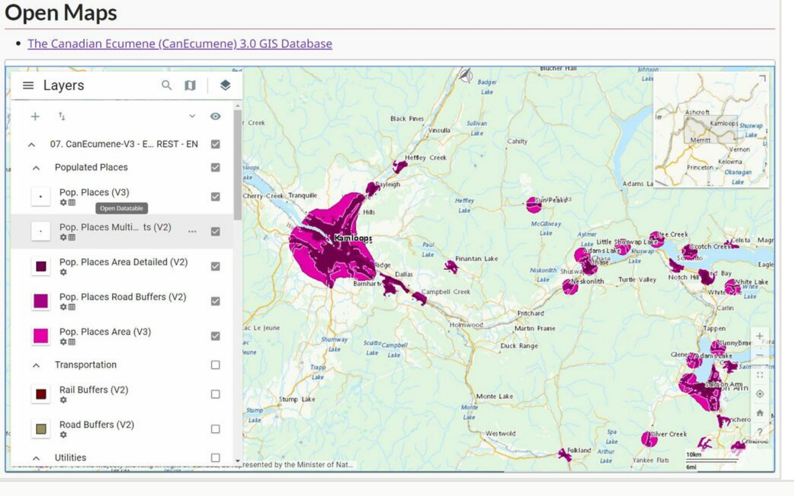

Figure 1.

Detailed view of the CanEcumene 3.0 GDB for an area in southern British Columbia showing both larger and smaller communties and their spatial representation. Each community is represented by multiple natural boundary types for use with different scales of analysis and representation (this view captured from the OpenMaps implementation at https://open.canada.ca/data/en/dataset/3f599fcb-8d77-4dbb-8b1e-d3f27f932a4b).

2) Index Composition: Several aspects are considered in composing indicators and indices for EBM applications. These include: a) how the terms *indicator* and *index* are defined, b) their *purpose* and how they are used and c) how they are *constructed*.

a) The term *indicator* has been defined in many different ways in many disciplines, including EBM (Heink and Kowarik 2010). While there may not be a universally accepted definition of *indicator*, most define it in relation to the use of data or measurements that are used to assess or measure the *performance* of some phenomenon (ibid.). It is therefore important to distinguish indicators from *primary measurements* which are defined simply as "*the process of associating numbers with physical quantities and phenomena*" (Encyclopedia Britannica 2023). Primary measurements, or raw data, do not inherently carry decision value, whereas indicators are data that *also* carry decision value and index is commonly a compound indicator comprised of two or more indicators.

b) The second aspect consideres the *purpose* or intended usage of the indicator or index along with its user base. Many scientific or technical indicators are developed for analysts and specialists (such as studies listed above), while others are developed for use by the general public and lay people. Indicators and indices aimed for public use require additional consideration with regard to their ease of understanding and decision value for a wider audience (Eddy et al. 2014). This has implications for the choice of analytical method used in that statistical methods need to be transparent and easy to communicate and understand, while remaining scientifically robust.

c) Both a) and b) are carried forward in considering how indicators are *constructed*. This choice of data variables, spatial and temporal coverage, analytical method and presentation and communication of results all need to be taken into consideration as a whole. From this vantage point, indicators may be composed through either *top-down* or *bottom-up* processes. In a top-down process, indicators and indices are first identified through either an end-user consultation or from another form of decision requirements analysis, followed by the search and collection of required data. In a bottom-up process, indicators are first identified from analysis of existing primary data and defined on the basis of exploratory analysis, such as was done with the studies that incorporate PCA above. The results are then assessed for their decision-value.

Here, a *hybrid* approach is used that involves statistical analysis of *existing primary data* while keeping in mind end-user needs, which, in this case, include both analysts and the public in EBM and related applications. From an end-user perspective, the requirement is to have a set of indicators and overall index for all communities in Canada and for multiple time periods, that are representative of key dimensions of socio-economic conditions and are easy to understand and evaluate. This study benefits from the availability of the CanEcumene 3.0 GDB. In addition to the advantages the CanEcumene offers in terms of spatial representation, this database also contains a selection of eight core socio-economic variables (Table 2) used to characterise communities as 'human habitats' (Eddy and Dort 2011). These variables cover five census periods from 2001 to 2021 which allows mapping time-series changes in human habitat conditions. For efficiency and to avoid duplication of effort, the approach taken here is to make use of this existing data first and build upon the work of Eddy and Dort (2011) before collecting new data.

Table 2.

Selected socio-economic variables used to characterise communities as 'human habitats' (after Eddy and Dort (2011)). See Table 3 for derived indicators.

Symbol	Field	Description
P_a	TotPop{PYR}	Total Population (Previous Census Year)
P_b	TotPop{CYR}	Total Population (Current Census Year)
Y	TotYut{YR}	Total Youth (< 15 yrs)
S	TotSen{YR}	Total Seniors (> 65 yrs)
P	TotPost{YR}	Total Post-Secondary Education
F	TotLForce{YR}	Total Labour Force
U	TotLFUNM{YR}	Total Unemployed
D	TotValDwell{YR}	Total Value of Dwellings

The objective of this study, therefore, is to develop and assess a method for mapping an index of Vitality for both communities (as a Community Vitality Index, CVI) and regions (as a Regional Vitality Index, RVI) for all of Canada using the socio-economic variables and time periods available in the CanEcumene GDB.

Method

The method developed involves a four step process using community point reference and associated attribute data, a polygonal boundary file that represents the extent of the Canadian ecumene and the socio-economic tables, all extracted from the recently upgraded CanEcumene 3.0 Geodatabase (Fig. 2). The socio-economic tables contain eight variables described in Table 2 and indicated as HHI (Human Habitat Indicator) variables in Fig. 2. GIS and data files are available in Eddy et al. (2023) (CanEcumene 3.0 GDB) and Eddy (2024) (Regional and Community CVI GDB).

Step 1: Deriving Indicator Variables

The first step involves deriving a set of community vitality indicators (CVI) from the HHI variables listed in Table 2. Table 3 provides a list of five indicators along with the fomulae and a rationale for selection. Each indicator represents a sub-dimension of an overall index covering Population Change, Age Structure, Education, Employment and Economy. All indicators are calculated for each community and for each census year from 2001 to 2021 at 5-year intervals to support time-series and trend analyses. A summary of statistical properties for each variable and year is provided in Table 4.

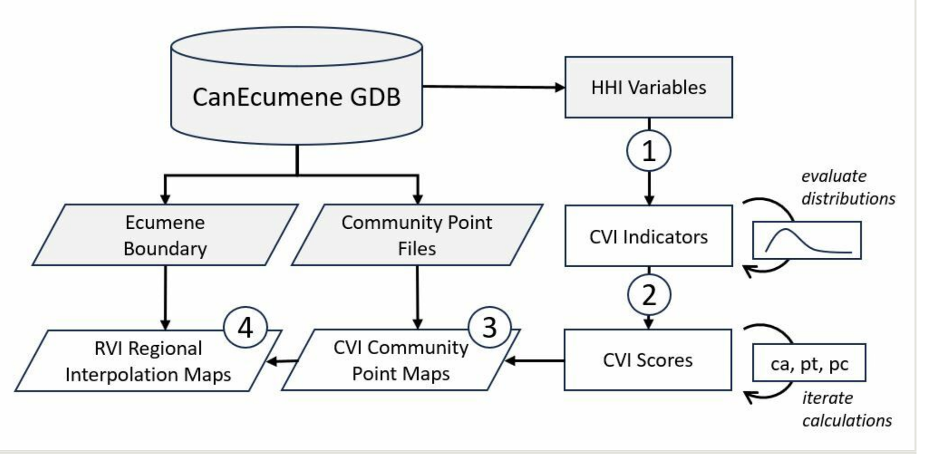

Figure 2.

Process model showing the four (4) steps involved in mapping community vitality index (CVI) and regional vitality index (RVI) data using the CanEcumene 3.0 GDB for Canada. Shaded symbols indicate existing data files and non-shaded symbols indicate new data files. See text for complete description of processing steps.

Table 3.

Five community vitality indicators (CVIs) derived from primary HHI variables provided in Table 1, along with formulae and rationale for selection. Note elements of formulae pertain to those listed in Table 2.

Symbol	**Indicator**	**Description**	**Formula**	**Rationale**
Pop	Pop{YR}	Population (% 5 yr. Change)	$= (P_b/P_a)-1$	Population growth indicates positive vitality and decline indicates negative vitality.
Age	Age{YR}	Age Structure (Seniors/Youth)	$= S/Y$	Communities with Age > 1.0 indicate insufficient youth to replace seniors (declining vitality) and values < 1.0 indicate communities with more youth than seniors (increasing vitality). (Note: the inverse of this indicator is taken in the calculation of percentile ranks)
Edu	Edu{YR}	Education (% with Post-Secondary Education)	$= P/Pb$	A larger proportion of residents with post-secondary education indicates positive vitality over lower proportions (negative vitality).
Emp	Emp{YR}	Employment (% Employed)	$= 1-(U/F)$	Taken as the inverse of unemployment rate, higher employment rates indicate positive vitality.
Eco	Econ{YR}	Economic Wealth (Total value of Dwellings/capita)	$= D/P_b$	The total value of all dwellings per capita indicates relative overall wealth in a community as a result of increasing economic activity. Higher values indicate higher vitality and lower values indicate lower vitality.

Table 4.

Statistical summary of RVI/CVI indicator variables derived from the CanEcumene 3.0 GDB. Note: Econ values expressed in standardised 2001 $CDN.

Indicator	Valid	Median	Mean	Minimum	Maximum	Std. Deviation	Skewness	Std. Error of Skewness
PopCh01	2875	-0.02	-0.01	-0.76	0.98	0.12	1.72	0.05
PopCh06	2794	-0.01	0.00	-0.96	0.98	0.15	0.63	0.05
PopCh11	2819	0.00	0.05	-0.99	17.36	0.62	19.04	0.05
PopCh16	2903	0.00	0.06	-1.00	30.41	0.88	23.77	0.05
PopCh21	2999	0.01	0.05	-1.00	65.51	1.27	46.37	0.04
Age01	2851	0.72	0.80	0.01	6.50	0.56	1.96	0.05
Age06	2819	0.89	0.99	0.02	9.00	0.74	2.95	0.05
Age11	2847	1.07	1.23	0.00	15.00	1.01	4.03	0.05
Age16	2963	1.27	1.54	0.00	79.00	1.92	23.04	0.04
Age21	2941	1.47	1.87	0.00	34.00	1.86	5.06	0.05
Edu01	2892	0.31	0.31	0.02	0.81	0.10	0.22	0.05
Edu06	2836	0.32	0.31	0.02	0.84	0.10	-0.20	0.05
Edu11	2904	0.29	0.25	0.00	0.98	0.17	-0.17	0.05
Edu16	3000	0.37	0.35	0.00	0.88	0.12	-0.56	0.04
Edu21	2987	0.37	0.36	0.00	1.00	0.12	-0.65	0.04
Emp01	2898	0.89	0.85	0.14	1.00	0.14	-1.46	0.05
Emp06	2867	0.92	0.87	0.17	1.00	0.12	-1.59	0.05
Emp11	2276	0.91	0.88	0.17	1.00	0.12	-1.48	0.05
Emp16	2975	0.90	0.86	0.00	1.00	0.11	-1.60	0.04
Emp21	2950	0.90	0.88	0.20	1.00	0.09	-1.75	0.05
Econ01	2536	28,772	33,077	1,776	229,166	18,803	2.29	0.05
Econ06	2504	38,257	47,388	1,029	341,772	32,016	2.56	0.05
Econ11	1916	51,884	60,401	374	353,028	37,851	1.93	0.06
Econ16	2558	76,626	99,230	3,835	2,312,977	102,631	8.01	0.05
Econ21	2570	78,169	107,765	1,456	11,793,103	96,553	4.20	0.05

The five indicators represent major dimensions of a communities vitality, all of which are positively oriented whereby lower values are regarded as indicators of lower vitality and higher values are indicative of higher vitality (with the exception of Age in which the inverse values are used to indicate higher vitality). Pearson-r coefficients show low to moderate correlation amongst indicators suggesting they are sufficiently independent (Table 5). This is partly reflected by the variability and direction of skewness of variables, which are observed to shift amongst census periods (Table 4). Several trends are worth mentioning. All variables exhibit an increase in mean and median values along with increasing spread of the distribution indicated by increases in minimum, maximum and standard deviation values (Table 4). Overall, although Canada's population is growing, the increase in average age structure ratio beyond values of 1.0 indicates an aging population. Increases in education, employment and residential real estate values/capita indicate continuous growth and development. These statistical characteristics of the data are of critical importance in selecting an appropriate method for standardised scoring of variables when transforming primary data values into indicator and index scores.

Table 5.

Pearson r coefficients for five regional and community vitality indicator variables showing low to moderate correlation values.

	PopCh01		**Age01**		**Edu01**		**Emp01**		**Econ01**
1. PopCh01	—								
2. Age01	-0.19	***	—						
3. Edu01	0.01		0.21	***	—				
4. Emp01	0.1	***	0.18	***	0.41	***	—		
5. Econ01	0.4	***	0.01		0.62	***	0.35	***	—
	PopCh06		**Age06**		**Edu06**		**Emp06**		**Econ06**
1. PopCh06	—								
2. Age06	-0.08	***	—						
3. Edu06	0.07	***	0.3	***	—				
4. Emp06	0.06	***	0.08	***	0.4	***	—		
5. Econ06	0.32	***	0.01		0.49	***	0.31	***	—
	PopCh11		**Age11**		**Edu11**		**Emp11**		**Econ11**
1. PopCh11	—								
2. Age11	-0.08	***	—						
3. Edu11	0.02		0.04		—				
4. Emp11	-0.02		0.11	***	0.3	***	—		
5. Econ11	0.07	**	-0.01		0.57	***	0.23	***	—
	PopCh16		**Age16**		**Edu16**		**Emp16**		**Econ16**

	PopCh01		Age01		Edu01		Emp01		Econ01
1. PopCh16	—								
2. Age16	-0.03		—						
3. Edu16	0.01		0.22	***	—				
4. Emp16	0.03		-0.01		0.43	***	—		
5. Econ16	0.03		0.32	***	0.37	***	0.15	***	—
	PopCh21		**Age21**		**Edu21**		**Emp21**		**Econ21**
1. PopCh21	—								
2. Age21	-0.01		—						
3. Edu21	0.02		0.24	***	—				
4. Emp21	0.01		-0.13	***	0.29	***	—		
5. Econ21	0.06	**	0.19	***	0.36	***	0.13	***	—
* p < 0.05, ** p < 0.01, *** p < 0.001									

Step 2: Calculating Percentile Ranks (Pr)

Three measures of standard scoring were considered including standardised z-scores (z), s-scores (s) and percentile ranks (Pr) (Table 6). The benefits of the z-score is it provides a useful standard measure of a value in terms of its positive or negative distance from the mean value. However, for skewed distributions, it is necessary to transform the data to a normal distribution. While it is possible to used z-scores to compare different variables, it is not useful for variables that have skewed distributions that vary over multiple time periods, such as the case with the CVI variables described above. In contrast to z-scores, which provide standard values in the (-4,+4) range, s-scores provide values in the (0,1) interval. However, using standard s-scores results in the same problem when working with distributions that have changing skewness of observations over multiple time periods. It is also not possible to directly compare s-scores calculated from non-normal (skewed) distributions for different variables and over different time periods.

For these reasons, the percentile rank (Pr) measure is deemed most appropriate for our purposes. The rationale in using the percentile rank method is based on the assumption that there is no absolute measure of overall vitality and that it is only possible to compare how communities are performing relative to one another. Pr values range in the (0,1) interval and can be expressed as percentages indicating how a community ranks in terms of its relative position to other communities. A value of 0.25 for variable *M*, for example, means a communty's data value is higher than 25% of other communities and 75% less than the remaining communities. Taking an average of Pr scores from multiple variables will indicate an overall Index score for a particular community. The Pr measure is not sensitive to outliers and does not require normalisation. The only caveat with the Pr measure is that ranked values are on an unequal interval. This is deemed to be not a major issue with the data variables described above as few data values are affected due to the skewness of the distributions.

Table 6.

Common measures used in standard scoring. (Note: x_i = variable value, u = mean, SD = standard deviation, n = number of observations, CF_x = *cumulative frequency of* x_i *, F = Frequency of* x_i) .

Measure	Formula	Comments
z-Score (*z*)	$z = (x_i - u)/SD$	Sensitive to outliers, requires normalisation, equal intervals.
Standard Score (*s*)	$s = (x_i - min)/(max - min)$	Sensitive to outliers, may require normalisation, equal intervals.
Percentile Rank (*Pr*)	$Pr = (CF_x-(0.5xF_x))/n$	Not sensitive to outliers, does not require normalisation, unequal intervals.

Percentile rank (Pr) scores are first calculated for each indicator and each year and for three categories: *ca* - all of Canada, *pt* - Provinces/Territories and *pc* - Population Class. Scores are calculated for the three categories to allow for more equitable comparisons, based on province or territory and population size classes (from hamlets to major cities) in addition to comparisons of all communities in Canada. For example, a community may score a low value when compared against all other communities in Canada, but may have a higher score for its province or territory or its population class. This process is repeated for all variables, categories and census years. Finally, overall CVI Index scores are calculated by taking the average of indicator percentile rank values of the three categories (ca, pt, pc) for each census year.

The result of these first two steps is three new sets of tables (indicators, percentile ranks and CVI scores) indexed by community for each census year allowing analysis of scores, based on individual categories or as an aggregate CVI index score. All tables can be related back to the master populated places table in the CanEcumene 3.0 Geodatabase for mapping and further analysis using additional descriptive attributes.

Steps 3 and 4: Mapping the Results

The third and fourth steps involve mapping the results in two forms: Step 3) as individual community vitality index (CVI) point maps and Step 4) as regional spatial (raster) interpolations of CVI point data to produce regional vitality index (RVI) maps. Mapping CVI scores as points simply requires joining the CVI scores tables to the populated places point file in the CanEcumene GDB, then mapping the data according to the preferred legend and style. Mapping regional RVI scores requires performing a spatial interpolation of the point data used in creating the CVI point maps. For this process, the CVI point scores are used as the input into the interpolation and an ecumene extents boundary file is used to limit the extent of the interpolation. Limiting the extent of the interpolation prevents values from being projected on to geographic regions without permanent settlements or infrastructure. The spatial interpolation parameters used are the same as those used for interpolations in

Eddy et al. (2020) (Supplement A). The following section discusses the results using a selection of maps along with statistical graph outputs.

Results

Results are presented in two forms corresponding to Steps 3 and 4 above:

1. CVI *point* maps (Fig. 3),
2. RVI *interpolation* maps (Fig. 4) and
3. Statistical Box and Whisker plots (Fig. 5).

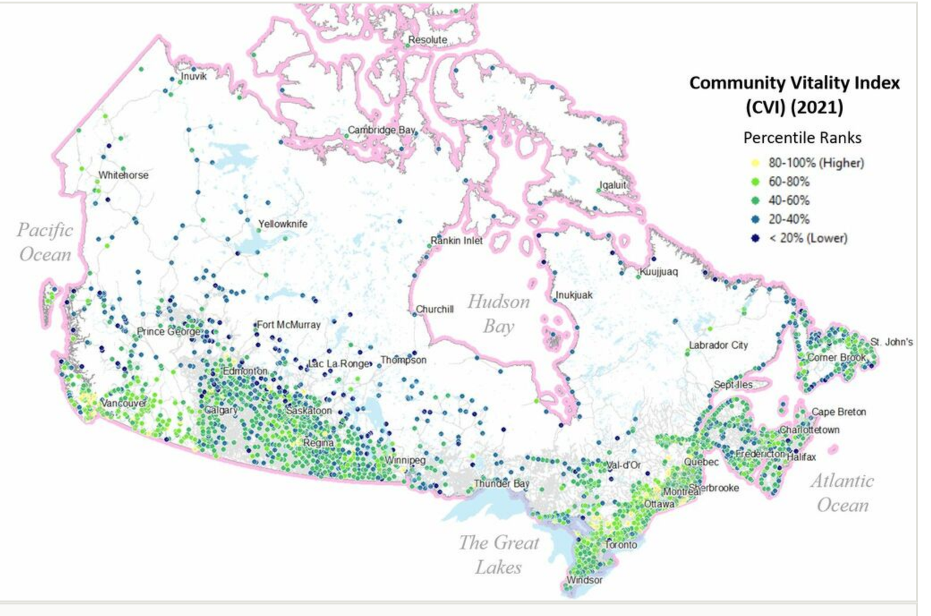

Figure 3.

Community Vitality Index (CVI) 'point map' for all of Canada for Census year 2021. Note: CVI results displayed in quintiles.

(Note: maps are available for viewing on-line at https://open.canada.ca/data/en/dataset/461123f1-1370-4709-aeda-639783ee8455).

Fig. 3 is a CVI point map that shows individual CVI scores for each community according to percentile ranks in quintiles (maps for all five census periods are provided in Suppl. material 1). In the on-line mapping interface of these data, users may dynamically zoom and pan the map and query an individual community to retrieve CVI scores for each indexing category (*ca, pt, pc*) plus the average CVI score (Eddy 2024). Although visualising the results in this form has some limitations due to the overlap of multiple points in the higher populated areas of southern Canada, several patterns can be seen. CVI scores are generally higher in the vicinities of major cities, such as the region between Montreal and Toronto and southern British Columbia. This is in contrast to northern and remote communities that generally show lower CVI scores and mixed CVI scores in the rural areas of the prairie and Atlantic regions.

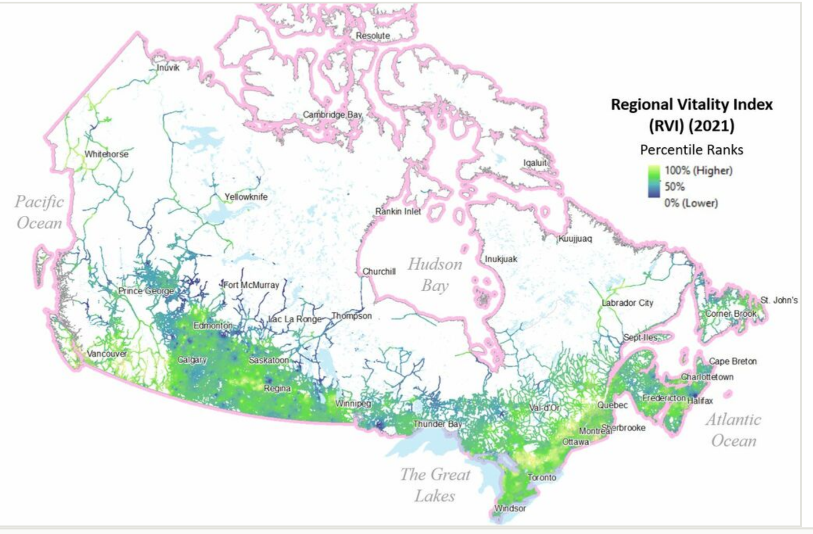

Figure 4.

Regional Vitality Index (RVI) 'interpolation map' of all of Canada for Census year 2021. Note: RVI results displayed as continuous percentile rank values.

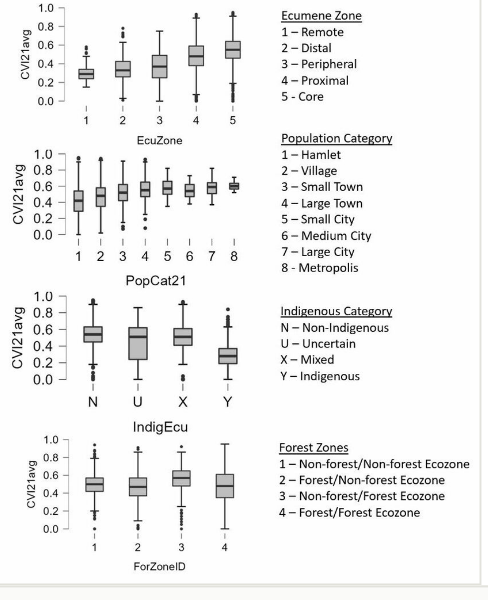

Figure 5.

CVI results for multiple dimensions of community classification displayed as Box and Whisker plots: 1) Ecumene Zone, 2) Population class, 3) Indigenous classification and 4) Forest Zone.

The spatial interpolation map shown in Fig. 4 provides an alternative means of visualising the results as a Regional Vitality Index (RVI) map (maps for all five census periods are provided in Suppl. material 2). This map reveals an important aspect about community and regional development across the country that generally follows Tobler's Law (Tobler 1970), whereby communities that are closer to each other tend to have similar values from those further apart. Many areas show this tendency for regions or clusters of communities to have similar scores, illustrating how no community is an island unto itself, but is part of a region of similar socio-economic conditions. This is evident in the higher scores in the vicinity of larger urban areas, while distal and remote regions clearly show significantly lower RVI scores than the southern regions. Regions surrounding larger northern communities of Whitehorse, Yellowknife and Labrador City are notable exceptions. This pattern is worth noting as, in addition to being subject to inequalities of social and economic development, many of these northern communities are Indigenous communities located in areas of higher risk to increased wildfire and other effects of climate change (Erni et al. 2021, Erni et al. 2024).

As with the points map in Fig. 3, some regions show mixed RVI scores illustrating some variablity in socio-economic conditions. Suppl. material 2 contains RVI maps for all five census periods (2001-2021) which reveal how RVI values change over time. One pattern worth noting is the decrease in RVI values in the prairie regions and increase in RVI values in the Atlantic regions from 2001 to 2021. It is beyond the scope of this work to comment further on reasons for the geographic patterns observed in the results. However, there are many patterns worthy of further investigation. Most notable are the numerous marginal communities and regions with lower CVI scores in both Figs. 3 and 4 in how they reveal clear evidence of the idea of 'left-behind places' (Fiorentino et al. 2024).

Other trends worth noting are shown in box and whisker plots in Fig. 5 based on aggregations by a number of community dimensions extracted from the CanEcumene Populated Places master table. As observed in both of the point CVI and the interpolated RVI maps in Fig. 3 and Fig. 4, a clear decreasing trend in CVI scores is evident amongst Ecumene Zones from core economic areas (southern) to distal and remote regions (northern). A similar trend is osbserved with Population Category values where there is a clear decreasing trend from metropolitan and larger cities to small towns, villages and hamlets, although the range in distribution of values increases with this trend. Two other dimensions for Indigenous communities and Forest Zones highlight two categorical differences worth noting. Indigenous communities have noticeably lower CVI scores compared with non-Indigenous communities and communities categorised as mixed Indigenous and non-Indigenous have similar CVI scores as communities categorised as non-Indigenous. For Forest Zones, communities located in forested areas show slightly lower CVI scores than non-forested communities. The full series of box and whisker plots provided in Suppl. material 3 for all Census years (2001-2021) reveal little change in these trends over a 25 year period.

Conclusions

The method developed here for mapping a Regional and Community Vitality Index (RVI/CVI) for Canada has both advantages and limitations. One of the main advantages is the method is more straightforward and computationally simpler than comparable methods. It uses only eight primary Census variables that are measured consistently over time to characterise communities as 'human habitats' (Eddy and Dort 2011). These variables are used to derive five sub-indicators of vitality pertaining to: 1) population growth, 2) age structure, 3) education, 4) employment and 5) economic wealth.

The use of percentile ranks is a computationally efficient metric that is easily understood by non-experts. RVI/CVI values can be dissected to examine which sub-indicators are performing poorly, thereby identifying policy objectives for communities or regions. The use of the CanEcumene 3.0 GDB with its use of natural boundaries as opposed to administrative or census boundaries ensures a naturalised geospatial representation of communities that is more amenable to integration with environmental datasets. As the method follows a naturalised approach, regional RVI patterns can be derived from spatial interpolation of CVI values offering a continuous (non-discrete) representation of the variability of socio-economic conditions across Canada. However, as noted in Eddy et al. (2020), the CanEcumene 3.0 GDB may be limited for local scale applications in some regions and does not provide details *within* major urban centres.

The method described here uses five sub-indicators of vitality representing five key dimensions of overall socio-economic conditions. All five indicators are used in the final calculations of the CVI values. Although the five sub-indicators used are shown to be independent, if warranted, conducting a sensitivity analysis may be worth exploring in future research whereby variables may be added or removed to test the effect on results. However, although the current method and selection of variables may be expanded upon, caution should be taken to ensure the method remains understandable to end users.

Other limitations are more precautionary than actual limitations as such. It is important to appreciate that RVI/CVI scores are *not absolute* measures of vitality, rather are *relative* indicators. This has implications for how both the data and maps are interpreted. As there is no absolute measure of vitality, it is more appropriate to think of the comparative RVI/CVI scores as *higher or lower* vitality and not *high* or *low* vitality. Communities and regions with lower vitality scores may exhibit strong vitality in ways that are not captured by the indicators used in this method. Conversely, communities and regions high vitality scores may have weaker vitality in other ways or locations of lower vitality that are not captured by the scale and resolution of these data.

Most importantly, as with many socio-economic indicators and indices, the RVI/CVI scores should not be interpreted as a measure of *well-being* or *resilience* of communities and regions as these aspects can only be assessed using idiographic qualitative methods at a local scale (Beckley et al. 2002, Stedman et al. 2004). It is possible the RVI/CVI data and maps may be used as a surrogates for *sustainability* or *vulnerability* in certain contexts. Regions and communities with higher scores are more likely to have the resources to implement sustainability policies and practices. Conversely, RVI/CVI scores may be inversely proportional to the vulnerability of regions or communities in that lower scores are more likely an indication of lower adaptive capacity. In all cases, care must be taken in the interpretation and application of these RVI/CVI maps and data.

Acknowledgements

Acknowledgements go to the reviewers who provided valuable feedback that contributed to improvements of the manuscript and to senior management and the Cumulative Effects Program of the Canadian Forest Service (CFS) for funding and support. Special acknowledgement to Martin Lefebvre with the GeoDiscovery section of the Canadian Centre for Mapping and Earth Observation (CCMEO) at Natural Resources Canada for assistance in publishing the RVI/CVI data on-line.

Funding program

Cumulative Effects Program of the Canadian Forest Service (CFS), Natural Resources Canada.

Hosting institution

Natural Resources Canada.

Ethics and security

The work herein was conducted under compliance with the Values and Ethics Code of the Government of Canada. This study did not require ethics clearance.

Author contributions

The author is the principal investigator and conducted all research contained in this article including data collection, analysis, method development, results and writing of the manuscript.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Ahsan MN, Warner J (2014) The socioeconomic vulnerability index: A pragmatic approach for assessing climate change led risks–A case study in the south-western coastal Bangladesh. International Journal of Disaster Risk Reduction 8: 32-49. https://doi.org/10.1016/j.ijdrr.2013.12.009
- Antwi EK, Boakye-Danquah J, Owusu-Banahene W, Dabros A, Eddy IM, Silver DA, Abolina E, Eddy B, Winder R (2023) Risk assessment framework for cumulative effects (RAFCE). Frontiers in Environmental Science 10 https://doi.org/10.3389/fenvs.2022.1055159
- Beckley T, Parkins J, Stedman R (2002) Indicators of forest-dependent community sustainability: The evolution of research. Forestry Chronicle 78 (5): 626-636. https://doi.org/10.5558/tfc78626-5
- CFS (2024) Canadian Forest Service, Protecting Communities. https://natural-resources.canada.ca/our-natural-resources/forests/wildland-fires-insects-disturbances/forest-fires/protecting-communities/13153. Accessed on: 2024-2-04.
- Chakraborty L, Rus H, Henstra D, Thistlethwaite J, Scott D, et al. (2020) A place-based socioeconomic status index: Measuring social vulnerability to flood hazards in the context of environmental justice. International Journal of Disaster Risk Reduction 43 https://doi.org/10.1016/j.ijdrr.2019.101394
- Chan E, Serrano J, Chen L, Stieb D, Jerrett M, Osornio-Vargas A, et al. (2015) Development of a Canadian socioeconomic status index for the study of health outcomes related to environmental pollution. BMC Public Health 15 (1). https://doi.org/10.1186/s12889-015-1992-y
- Cumming GS, Cumming DH, Redman CL (2006) Scale Mismatches in Social-Ecological Systems: Causes, Consequences, and Solutions. Ecology and Society 11(1) (14). URL: https://www.ecologyandsociety.org/vol11/iss1/art14/
- Dale A, Ling C, Newman L, et al. (2010) Community Vitality: The Role of Community-Level Resilience Adaptation and Innovation in Sustainable Development. Sustainability 2 (1): 215-231. https://doi.org/10.3390/su2010215
- Eddy B, Dort A (2011) Integrating Socio-Economic Data for Integrated Land Management (ILM): Examples from the Humber River Basin, Western Newfoundland. GEOMATICA 65 (3): 283-291. https://doi.org/10.5623/cig2011-044
- Eddy B, Hearn B, Luther J, van Zyll de Jong M, Bowers W, Parsons R, Piercey D, Strickland G, Wheeler B, et al. (2014) An information ecology approach to science–policy integration in adaptive management of social-ecological systems. Ecology and Society 19 (3). https://doi.org/10.5751/es-06752-190340
- Eddy B, Muggridge M, LeBlanc R, Osmond J, Kean C, Boyd E, et al. (2020) An Ecological Approach for Mapping Socio-Economic Data in Support of Ecosystems Analysis: Examples in Mapping Canada's Forest Ecumene. One Ecosystem 5 https://doi.org/10.3897/oneeco.5.e55881
- Eddy B (2024) Regional and Community Vitality Index of Canada. Natural Resources Canada. Release date: 2024-2-14. URL: https://open.canada.ca/data/en/dataset/461123f1-1370-4709-aeda-639783ee8455
- Eddy BG, Muggridge M, LeBlanc R, Osmond J, Boyd E (2023) The Canadian Ecumene (CanEcumene) 3.0 Geodatabase (GDB). Natural Resources Canada, GeoDiscovery.

Release date: 2023-12-01. URL: https://open.canada.ca/data/en/dataset/3f599fcb-8d77-4dbb-8b1e-d3f27f932a4b

- Encyclopedia Britannica (2023) *Measurement* (Definition of). https://www.britannica.com/technology/measurement. Accessed on: 2024-2-12.
- Erni S, Johnston LM, Boulanger Y, Manka F, Bernier P, Eddy BG, Christianson AC, Swystun T, Gauthier S (2021) Exposure of the Canadian Wildland-Human Interface (WHI) and population to wildland fire, under current and future climate conditions. Canadian Journal of Forest Research https://doi.org/10.1139/cjfr-2020-0422
- Erni S, Wang X, Swystun T, Taylor S, Parisien M, Robinne F, Eddy B, Oliver J, Armitage B, Flannigan M (2024) Mapping wildfire hazard, vulnerability, and risk to Canadian communities. International Journal of Disaster Risk Reduction 101 https://doi.org/10.1016/j.ijdrr.2023.104221
- Etuk L, Acock A (2016) Toward a rural community vitality measurement practice. Community Development 48: 1-13. https://doi.org/10.1080/15575330.2016.1251480
- Fiorentino S, Glasmeier AK, Lobao L, Martin R, Tyler P (2024) 'Left behind places': what are they and why do they matter? Cambridge Journal of Regions, Economy and Society 17 (1): 1-16. https://doi.org/10.1093/cjres/rsad044
- Flanagan B, Gregory E, Hallisey EJ, Heitgerd J, Lewis B (2011) A Social Vulnerability Index for Disaster Management. Journal of Homeland Security and Emergency Management 8 (1). https://doi.org/10.2202/1547-7355.1792
- Grigsby WJ (2001) Community vitality: Some conceptual considerations. Rural Development Paper 6: 1-16.
- Heink U, Kowarik I (2010) What are indicators? On the definition of indicators in ecology and environmental planning. Ecological Indicators 10 (3): 584-593. https://doi.org/10.1016/j.ecolind.2009.09.009
- IAAC (2024) Impact Assessment Agency of Canada Research Program. https://www.canada.ca/en/impact-assessment-agency/corporate/research-program.html. Accessed on: 2024-2-06.
- ISC (2016) Report on trends in First Nations communities, 1981 to 2016. https://www.sac-isc.gc.ca/eng/1345816651029/1557323327644. Accessed on: 2024-2-05.
- ISC (2021) The Community Well-Being index. Indigenous Services Canada. Online Map: https://www.sac-isc.gc.ca/SAC-ISC/CWB/index-map-en.html. URL: https://www.sac-isc.gc.ca/eng/1100100016579/1557319653695
- Journeay M, Yip JZ, Wagner CL, LeSueur P, Hobbs T (2022) Social vulnerability to natural hazards in Canada:an overview of methods and findings. Natural Resources Canada, Open File 8902. https://doi.org/10.4095/330295
- Kappel CV, Martone RG, Duffy JE (2006) Ecosystem-based Management (definition). Encyclopedia of Earth..
- Lax J, Krug J (2013) Livelihood assessment: A participatory tool for natural resource dependent communities. Johann Heinrich von Thünen Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries. Issue: 7. URL: https://ideas.repec.org/p/zbw/jhtiwp/7.html
- MacEachren A, Robinson A, Hopper S, Gardner S, Murray R, Gahegan M, Hetzler E (2005) Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. Cartography and Geographic Information Science 32 (3): 139-160. https://doi.org/10.1559/1523040054738936

- Noble B, Harriman Gunn J (2010) Regional Strategic Environmental Assessment for Integrated Land Management. Horizons: Special Issue on Sustainable Places.Policy Research Initiative, Ottawa, Canada. 10 (4): 106-112.
- NRCan (2024) Canada's National Climate Change Adaptation Platform. https://natural-resources.canada.ca/climate-change/canadas-climate-change-adaptation-platform/10027. Accessed on: 2024-2-06.
- Pearce C (2005) Natural resources and community vitality: A rural perspective. Journal of Ecosystems and Management https://doi.org/10.22230/jem.2005v6n2a318
- Pravitasari A, Rustiadi E, Mulya S, Fuadina L (2018) Developing Regional Sustainability Index as a New Approach for Evaluating Sustainability Performance in Indonesia. Environment and Ecology Research 6 https://doi.org/10.13189/eer.2018.060303
- PSC (2024) National Risk Profile: Strengthening Canada's All-Hazards Approach to Emergency Management, Public Safety of Canada (PSC). https://www.publicsafety.gc.ca/cnt/mrgnc-mngmnt/ntnl-rsk-prfl/index-en.aspx. Accessed on: 2024-2-02.
- Scott K (2016) Canadian Index of Wellbeing. Canadian Council on Social Development. URL: https://policycommons.net/artifacts/1968549/community-vitality/2720314/
- StatsCan (2019) The Canadian Index of Multiple Deprivation (CIMD), Statistics Canada Catalogue no. 45-20-0001. https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5274. Accessed on: 2024-2-01.
- Stedman RC, Parkins JR, Beckley TM (2004) Resource dependence and community well-being in rural Canada. Rural Sociology 69 (2): 213-234. https://doi.org/10.1526/003601104323087589
- Sui D, Turner M (2022) General theories and principles in geography and GIScience: Moving beyond the idiographic and nomothetic dichotomy. Annals of GIS 28 (1): 1-4. https://doi.org/10.1080/19475683.2022.2030939
- Tobler WR (1970) A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography 46 https://doi.org/10.2307/143141
- Wangdi T (2022) A Review: Bhutanese perspective of Community Vitality. Indonesian Journal of Education and Social Sciences 1 (2): 92-99. https://doi.org/10.56916/ijess.v1i2.228
- Zeng C, Song Y, He Q, Shen F, et al. (2018) Spatially explicit assessment on urban vitality: Case studies in Chicago and Wuhan. Sustainable Cities and Society 40: 296-306. https://doi.org/10.1016/j.scs.2018.04.021

Supplementary materials

Suppl. material 1: Community Vitality Index (CVI) (Point) Maps for Canada (2001-2021) doi

Authors: BG Eddy
Data type: Map images
Brief description: Series of maps showing Community Vitality Index (CVI) values for communities in Canada. These maps illustrate how individual CVI scores vary by region and proximity to major economic cores and also reveal changes over time.
Download file (1.25 MB)

Suppl. material 2: Regional Vitality Index (RVI) Interpolation Maps for Canada (2001-2021) doi

Authors: BG Eddy
Data type: Map images
Brief description: Series of maps showing spatially interpolated Regional Vitality Index (RVI) values for Canada. These maps illustrate how RVI scores vary by region and proximity to major economic cores and also reveal changes over time.
Download file (1.04 MB)

Suppl. material 3: Box and Whisker Plots of CVI Values for Selected Community Dimensions for 2001-2021 doi

Authors: BG Eddy
Data type: Graphs
Brief description: Series of Box and Whisker plots for a selection of community dimensions showing trends and variations amongst categories: 1) Ecumene Zone, 2) Population Category, 3) Indigenous Category and 4) Forest Zones. Interpretive notes are provided on each set of graphs.
Download file (489.04 kb)