Waste Management 74 (2018) 3–15

Contents lists available at ScienceDirect

Waste Management

journal homepage: www.elsevier.com/locate/wasman

Modeling and prediction of regional municipal solid waste generation and diversion in Canada using machine learning approaches

Miyuru Kannangara *, Rahul Dua, Leila Ahmadi, Farid Bensebaa

Energy, Mining and Environment, National Research Council Canada, 1200 Montreal Road, Ottawa, ON K1A 0R6, Canada

A R T I C L E   I N F O

Article history:
Received 15 March 2017
Revised 22 November 2017
Accepted 30 November 2017
Available online 6 December 2017

Keywords:
Municipal solid waste
Forecasting
Machine learning
Modeling
Neural networks
Decision trees

A B S T R A C T

The main objective of this study was to develop models for accurate prediction of municipal solid waste (MSW) generation and diversion based on demographic and socio-economic variables, with planned application of generating Canada-wide MSW inventories. Models were generated by mapping residential MSW quantities with socio-economic and demographic parameters of 220 municipalities in the province of Ontario, Canada. Two machine learning algorithms, namely decision trees and neural networks, were applied to build the models. Socio-economic variables were derived from Canadian Census data at regional and municipal levels. A data pre-processing and integration framework was developed in Matlab® computing software to generate datasets with sufficient data quantity and quality for modeling. Results showed that machine learning algorithms can be successfully used to generate waste models with good prediction performance. Neural network models had the best performance, describing 72% of variation in the data. The approach proposed in this study demonstrates the feasibility of creating tools that helps in regional waste planning by means of sourcing, pre-processing, integrating and modeling of publically available data from various sources.

1. Introduction

The amount of municipal solid waste (MSW) generated is increasing rapidly due to urbanization and population growth, presenting unique opportunities and challenges (Korai et al., 2017). Waste can be recycled into various industries at global scale, creating jobs and boosting economies (Xu et al., 2017). MSW can be used as fuel in waste-to-energy plants, converting waste to electricity while mitigating the environmental impacts (Cucchiella et al., 2017).

Sustainable waste management efforts are however hampered by many challenges. Scarcity and reliability of available data is a major challenge in planning (Mrayyan and Hamdi, 2006), in implementing sorting technology and deploying information systems that support waste management (Hannan et al., 2015; Vitorino de Souza Melaré et al., 2017). Lack of data can be attributed to waste management infrastructure and practices. Waste is not measured at user or disaggregate levels and is managed by different channels involving several stakeholders, making the data collection and compilation difficult (Beigl et al., 2008). The data scarcity is a critical issue in Canada, where vast geographical areas are involved and waste is managed by a large number of parties including provinces, territories, local authorities and private contractors. For example, many Canadian northern and remote communities lack proper waste management practices and have little or no data regarding waste (Environment and Climate Change Canada, 2017). With accurate data, MSW could be a viable alternative source of energy to replace currently used diesel power generation.

Recent efforts have been made to increase the availability of Canadian waste data to facilitate waste management planning and operation. A Canadian database, called Biomass Inventory Mapping and Analysis Tool (BIMAT), is under development to provide information on the location-based waste biomass availability across Canada (Agriculture-Canada, 2017). The data for BIMAT has been collected by the tedious process of conducting industrial surveys. In addition to waste biomass data, a centralized database containing MSW data at disaggregate municipal levels is very useful. However, collection of such data is challenging due to sheer number of jurisdictions involved and varying waste management practices across Canada.

2. Waste data modeling

Advanced modeling is one of the main approaches available for estimating the amount of waste (Li et al., 2017). There is already a

* Corresponding author.
E-mail address: miyuru.kannangara@nrc-cnrc.gc.ca (M. Kannangara).

https://doi.org/10.1016/j.wasman.2017.11.057
0956-053X/

large body of publications. We will refer only to few recent publications with highly relevant methodologies. Input-output models have been used to estimate the waste generation of a country using a top-down material flow approach based on national supply-use economic data (Joosten et al., 1999; Reynolds et al., 2016). Despite having reliable data for modeling, they can only be used at county or state level. Time series data has been used to model recurring seasonal waste generation and dynamic waste generation patterns. Navarro-Esbrí et al. (2002) modeled daily and monthly residential waste data using seasonal autoregressive moving average (SARIMA) modeling technique for long term forecasting. In addition to autoregressive techniques, artificial intelligence methods such as artificial neural networks (ANNs) have been used to model time series data (Abbasi and El Hanandeh, 2016). Regardless of the modeling technique, the time series models require significant past data. Dyson and Chang (2005) used system dynamic modeling to forecast MSW for a fast growing region with limited past data. System dynamic approach relies on modeling cause and effect relationships between waste, socio-economic, managerial and planning parameters (Kollikkathara et al., 2010). Both time series and system dynamics models focus on a single composite region without considering cross-sectional data across many regions or households.

Multiple linear regression (MLR) approach has been often used to model cross-sectional data by establishing statistical relationships between socio-economic explanatory variables and waste generation. Data for regression studies have been reported to be obtained from survey sampling (Bandara et al., 2007; Benítez et al., 2008; Monavari et al., 2012) or from authorities, including municipal councils and statistical agencies (Ali Abdoli et al., 2011; Beigl et al., 2004; Daskalopoulos et al., 1998). Waste data collected at disaggregate levels such as households, counties or municipalities show high variance due to heterogeneity between modeling units. MLR models were able to describe around 50% variation at household level (Benítez et al., 2008) and around 49% variation at county or municipal level (Bach et al., 2004; Hockett et al., 1995), as given by coefficient of determination (R^2). Lower R^2 value may be attributed to failure to find relevant explanatory variables, limitations of MLR model structure and data inconsistencies. Beigl et al. (2004) improved the MLR model structure by grouping the data into prosperity categories and fitting an MLR model for each group separately. Panel data models, which consider cross sectional heterogeneity and dynamic aspects of time series data, have been used by Arbulú et al. (2015) for waste generation prediction, with improved R^2 estimated at 64%.

Machine learning (ML) or artificial intelligence approaches have been increasingly used for waste modeling. ANN approach has been frequently used (Antanasijević et al., 2013; Sodanil, 2014; Zade and Noori, 2008). Other methods based on support vector machines (Abbasi and El Hanandeh, 2016) and decision tree based methods (Johnson et al., 2017) have also been used. They have been successfully applied for waste modeling using cross-sectional data (Jahandideh et al., 2009), time series data (Abbasi and El Hanandeh, 2016; Sodanil, 2014) and panel data (Azadi and Karimi-Jashni, 2016; Johnson et al., 2017). ML algorithms adjust both model structure and parameters to fit data and hence are usually better at modeling complex non-linear behavior than regression methods. ANN models were reported to be better at predicting than MLR models using pure cross-sectional data (Jahandideh et al., 2009) and panel data (Azadi and Karimi-Jashni, 2016) because of their ability to model non-linear behavior. Neural network models are however black-box models and hence they do not provide insight into explanatory variables that cause waste generation. Abbasi and El Hanandeh (2016) compared different ML algorithms for time-series waste modeling, however there is limited information on effect of ML algorithm application to cross-sectional or panel data.

In this work, we developed waste models using cross-sectional residential waste data from 220 municipalities in Ontario, Canada for 9 consecutive years. Socio-economic and demographic parameters of the municipalities were used as predictor variables. The intended application of the models was the prediction of waste data across Canadian municipalities to facilitate waste management planning, particularly in places where waste data is non-existing. We have selected ML approach for modeling because it has better prediction abilities than regression models. Also, the interpretation of the model coefficients to understand waste generation patterns was not a requirement. Two types of machine learning algorithms, namely decision trees and neural networks have been compared in terms of prediction ability. A framework for data collection, processing and integration was proposed to generate datasets with sufficient data quantity and quality for modeling.

3. Methodology

3.1. Waste data availability

The residential waste generation and diversion data from province of Ontario has been used in this study. Waste management in Ontario is mandated by the provincial government. Local municipalities are responsible for developing their own waste management programs with curbside collection, depot collection and pay-as-you-throw collection methods. Waste material streams collected in Ontario include garbage, Bluebox recyclables (Paper, glass metal, plastic and textile), leaf and yard waste, organic household waste, waste electronic equipment, bulky waste, hazardous waste and numerous other waste streams. However, municipalities are not mandated to implement programs to collect all the waste streams except garbage and Bluebox streams. As a result, the recycling rates vary largely across the province. The amount of organics recycled in Ontario increased by 35% from 2010 to 2014 as increasing number of municipalities implemented programs to collect organic waste. Still, only 48% of municipalities recycled organic waste in 2014 (Waste Diversion Ontario, 2014).

Waste Diversion Ontario (WDO), a provincial government agency, provides coordination and allocates resources for waste management programs implemented by local municipalities. Municipalities are required to report waste stream quantity information, number of households serviced and waste management expenditure data each year during the municipal data call survey carried out by WDO. WDO uses this information to generate waste diversion rates and best practices. This data is publicly available. Although the municipal data call generates a compiled database of waste data, they are many sources of error due to self-reporting by the municipalities. Municipalities may keep erroneous records, over-or under-estimate the waste quantities or misinterpret survey questions. WDO performs reconciliation and validation of gathered data to detect errors and outliers. Municipalities with unusual data are audited. In this study, we have assumed that the data has sufficient quality and variation to describe cross-sectional differences among municipalities so that it can be used to develop cross-sectional data models. The waste quantities analyzed for modeling in this study are shown in Table 1.

Among the waste quantities reported in Table 1, total residential MSW and total Bluebox waste can be considered as most reliable as it has been subjected to a number of data adjustments and reconciliations by WDO. These adjustments include replacing with past average values when no data is reported or unrealistically higher values are reported and calculating garbage tonnes based on average values for missing households when data for only some

Table 1
Waste quantities analyzed in the study.

Waste quantity	Description	Available years
Total residential MSW	All the residential waste generated, reported as recycled and landfilled	2006–2014
Blue box waste	Waste collected in residential Blue box program. Available data inventory includes paper (printed paper, OCC – old corrugated containers & OBB – Old box board, mixed paper, Polycoat), plastics (PET, HDPE, plastic film, tubes & lids, polystyrene, mixed paper), metal (steel and aluminum) and glass (flint and colored)	2002–2014
Leaf and yard waste	Waste collected in yard waste programs. Available data inventory includes yard waste, leaves, Christmas trees, and bulky yard waste	2002–2014
Household kitchen organics	Organic food waste from households mainly collected by Green bin program	2002–2014

households are available. The total residential MSW is calculated as the sum of the waste disposed and waste recycled. Paper, plastic, metal, glass, leaf/yard, and kitchen organic waste are reported as provided by municipalities without any adjustments.

3.2. Selection of socio-economic parameters

Population is unarguably the dominant parameter that determines the total waste generation and it has been used in long-term time series forecasting models as a predictor (Ali Abdoli et al., 2011; Daskalopoulos et al., 1998). Population has been incorporated into dependent waste quantity variables in cross-sectional studies, resulting in per capita waste, under the assumption that per capita waste generation does not depend on the size of the population (Hockett et al., 1995). We chose the number of households in a municipality as the normalization variable to calculate waste per household because the actual number of households served each year by the waste collection programs was available. It was assumed that the size of a municipality does not affect per household waste quantities and other socio-economic parameters.

Socio-economic factors that measure economic affluence and quality of life have often been found to increase the waste generation. They include income level (Monavari et al., 2012), GDP (Beigl et al., 2004), consumer expenditure (Daskalopoulos et al., 1998) and purchasing power indices (Bach et al., 2004), which are positively correlated with waste generation. Ghinea et al. (2016) reported that demographic age structure strongly influenced the waste generation. Age structure has a profound effect on most aspects of the society such as spending, savings, labor force participation and taxation as reported by Lindh (2003). It is also an ideal variable for forecasting because it picks up medium term trends and varies slowly, increasing projection accuracy. The size of household has also been reported as a significant factor, having a negative correlation, in studies based on actual waste surveys (Bandara et al., 2007; Benítez et al., 2008). Type of occupation (Bach et al., 2004) and employment (Bandara et al., 2007) have also been reported to influence waste generation. Arbulú et al. (2015) reported that tourism arrivals and expenditure to affect waste generation. However, such information is generally not available at municipal level considered in this study.

Waste recycling is largely affected by policy, recycling program and social norm variables as they influence incentives and intentions for recycling. Sidique et al. (2010) reported that variable pricing for waste disposal is a significant factor in increasing recycle because households have to pay for the amount of waste they are disposing. Implementation of Pay-As-You-Throw waste policies has been reported to positively influence the recycling rate (Puig-Ventosa, 2008; Starr and Nicolson, 2015). Recycling program related variables such as accessibility and convenience are also important in determining recycling rates. Installation of convenient source segregation equipment in households (Bernstad, 2014) and increasing curbside collection (Abbott et al., 2011) have been reported to significantly increase the recycling rates due to increased access. Proximity to central collection depots and provision of clear instructions also increased recycling (Rhodes et al., 2014). Social norms such as attitudes towards impact of recycling, behaviors of other households and trust in the system have also been reported to affect the recycling (Hage et al., 2009; Miliute-Plepiene et al., 2016). Some previous studies identified that socio-economic parameters such as age, education level (Sidique et al., 2010) and income (Grazhdani, 2016) influence the recycling rate. This may arise from correlation between socio-economic and behavioral or social norm variables. For example, older people may have more time to engage in recycling activities and educated public is more aware of environmental issues (Sidique et al., 2010).

The main criteria for selection of socio-economic variables for this work were significance as reported in previous studies and availability at municipal level. The best data source that matched these requirements was the Canadian census program, which is performed nationwide at 5 year intervals. Most of the socio-economic parameters relevant to waste generation such as income, age structure, household size, occupation and employment can be obtained using census data at municipal level. It also contains parameters relevant to recycling such as education levels, age and income levels. Despite the significance to recycling rate, the variables related to policy, program and social norms were not included in modeling because such variables are difficult to obtain and usually are not available on a country-wide basis. Use of census data guarantees the availability of socio-economic parameters even for the most remote communities in Canada. This is important as remote and Northern communities in Canada are the places where waste data availability is truly limited.

The most recent census data was available from census year 2011, followed by years 2006 and 2001. Census reporting level at subdivisions corresponded to the municipality level of MSW data reporting. This was verified by comparing the population values reported in MSW data and census subdivision data. The level of details in Census data allows calculating the same socio-economic parameter in many ways. This level of details has undergone changes over the years. The socio-economic parameters were defined so that they can be calculated consistently using data structures in census 2001, 2006 and 2011. When more than one definitions were possible (e.g. percentage population over 55 years, over 45 years and so on), the definition with highest correlation with waste data was selected. Table 2 shows the list of socio-economic parameters selected for analysis. Derivations and descriptions of the parameters are given in Appendix B. In addition to the parameters selected according to significance from previous studies, two other parameters describing dwelling and work place characteristics were added as they were reported in separate areas in the census.

As socio-economic variables were only available at 5 year intervals, linear interpolation and extrapolation were used to estimate annual values from 2002 to 2014. The income values were adjusted for inflation using annual consumer price index of Ontario. Linear interpolation has been reported to provide sufficient reliability

Table 2
The list of socio-economic parameters and their abbreviations.

Socio-economic parameter	Census areas	Abbreviation
Fraction of population over 45 years	Demographics	POP
Median personal income	Earnings and income	INC
Fraction of population with no high school graduation certificate, diploma or degree,	Education	EDU
Employment rate	Employment	EMP
Fraction of people employed by agriculture, resource based, manufacturing or construction industries	Industries and occupations	IND
Fraction of owned dwellings (versus rented)	Dwelling characteristics	DWE
Fraction of one person households	Household characteristics	HHCHAR
Fraction of population worked at a usual work place	Work place	WORKP

for short length data imputation problems (Junninen et al., 2004). Further, socio-economic data from census are expected to change gradually and without random variation. Other methods such as polynomial fitting are not expected to increase the accuracy as there is only 3 relevant data points (years 2001, 2006 and 2011). Linear interpolation is the only available option for total MSW generation dataset, which was reported from 2006 to 2014, having only 2 corresponding census years (years 2006 and 2011).

A correlation analysis was performed between MSW quantities and socio-economic variables as a preliminary screening of socio-economic parameters for the modeling. Linear correlation coefficients between each variable were calculated as given in Eq. (1) (Albright et al., 2010).

$$\rho(A,B) = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{A_i-\mu_A}{\sigma_A}\right)\left(\frac{B_i-\mu_B}{\sigma_B}\right) \quad (1)$$

$\rho(A,B)$ is the correlation coefficient between variables A and B, N is the number of observations, μ_A and σ_A are mean and standard deviation of variable A, respectively, and, μ_B and σ_B are mean and standard deviation of variable B. The correlation coefficient has been used in previous studies as a ranking method for selecting features for ML algorithms (Guyon and Elisseeff, 2003). Variables were included in ML model according to their rank until considerable improvements in prediction accuracy is achieved. Predictor variables that are highly correlated with each other were excluded to avoid multi-collinearity problems.

3.3. Data preprocessing

A number of data preprocessing steps were required to prepare and transform the collected data into variables suitable for modeling and analysis. They include loading the data into suitable data structures, derivation of socio-economic parameters, transformation of data, filtering to remove outliers, filtering and correction of the city and municipality names and final integration of the data into combined datasets. Fig. 1 outlines the data collection and preprocessing steps as well as modeling and analysis steps, showing the overall methodology of the study. Matlab® technical programming language was utilized in all the steps of the project. Matlab® scripts were developed to automate the data loading, preprocessing and integration.

The datasets considered for modeling viz. MSW generation, paper diversion, leaf and yard waste diversion and kitchen organic waste diversion, were normalized using number of households, hence providing waste quantities per household. Some studies have considered population as normalization parameter to calculate per capita waste quantities (Bach et al., 2004; Beigl et al., 2004). In the current study, the number of households was selected as normalization parameter due to two reasons. First, the household numbers served by each waste collection program were available along with waste quantity data. Second, the number of households served by some collection programs was less than the total number of households in the city.

Interquartile range (IQR) filtering was used to detect and remove outliers from datasets. Outliers can arise from reporting discrepancies from the municipalities, incorrect estimates or unusual changes in waste generation patterns. The upper and lower limits of the valid data range per municipality were determined as $Upper\ Limit = Q_3 + IQR \times 1.5$ and $Lower\ Limit = Q_1 - IQR \times 1.5$. Q_1 and Q_3 are the first and third quartiles of the dataset, respectively. IQR is the interquartile range. The data points, which lay outside of the upper and lower limits, were considered as outliers and were filtered out.

3.4. Training and validation of models

During the modeling phase, a function (model) that maps socio-economic variables to each MSW quantity variables was identified using the prepared datasets. Machine learning algorithms can produce over-fitted models, leading to poor performance with unseen data samples. Therefore, it is required to test the model with unseen data. This was achieved by dividing the dataset into training set and test set. The training set was used to build the model while testing set was used to check for over-fitting. Previous studies commonly used 80:20 (Azadi and Karimi-Jashni, 2016) or 85:15 (Pao and Chih, 2006) training to testing partition ratios. In this study, we evaluated five partition ratios, namely 60:40, 70:30, 80:20, 85:15 and 90:10, to randomly divide the dataset into training and testing sets. A total of 100 such random partitions were generated for each ratio and a model was trained for each dataset. The ratio that produced minimum average model error was selected. The predictive performance of the model was measured by the mean squared error (MSE), mean absolute percentage error ($MAPE$) and coefficient of determination (R^2) as given by Eqs. (2)–(4), respectively (Friedman et al., 2009).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2 \quad (2)$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{Y_i-\hat{Y}_i}{Y_i}\right| \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \quad (4)$$

n is the number of observations, $\hat{Y}_i$ is the predicted value by the model, Y_i is the observed value, $\bar{Y}$ is the mean value of waste quantity. MSE was also derived as a percentage value by obtaining its square root, the root mean squared error ($RMSE$), and then normalizing it using the mean of the observations ($\bar{Y}$), producing the normalized root mean squared error ($NRMSE$).

The MSE, $MAPE$ and R^2 were calculated for both training and testing datasets, providing training and testing performances. Usually, the training error is lower than the test error because model parameters and structure are adjusted to the training dataset.

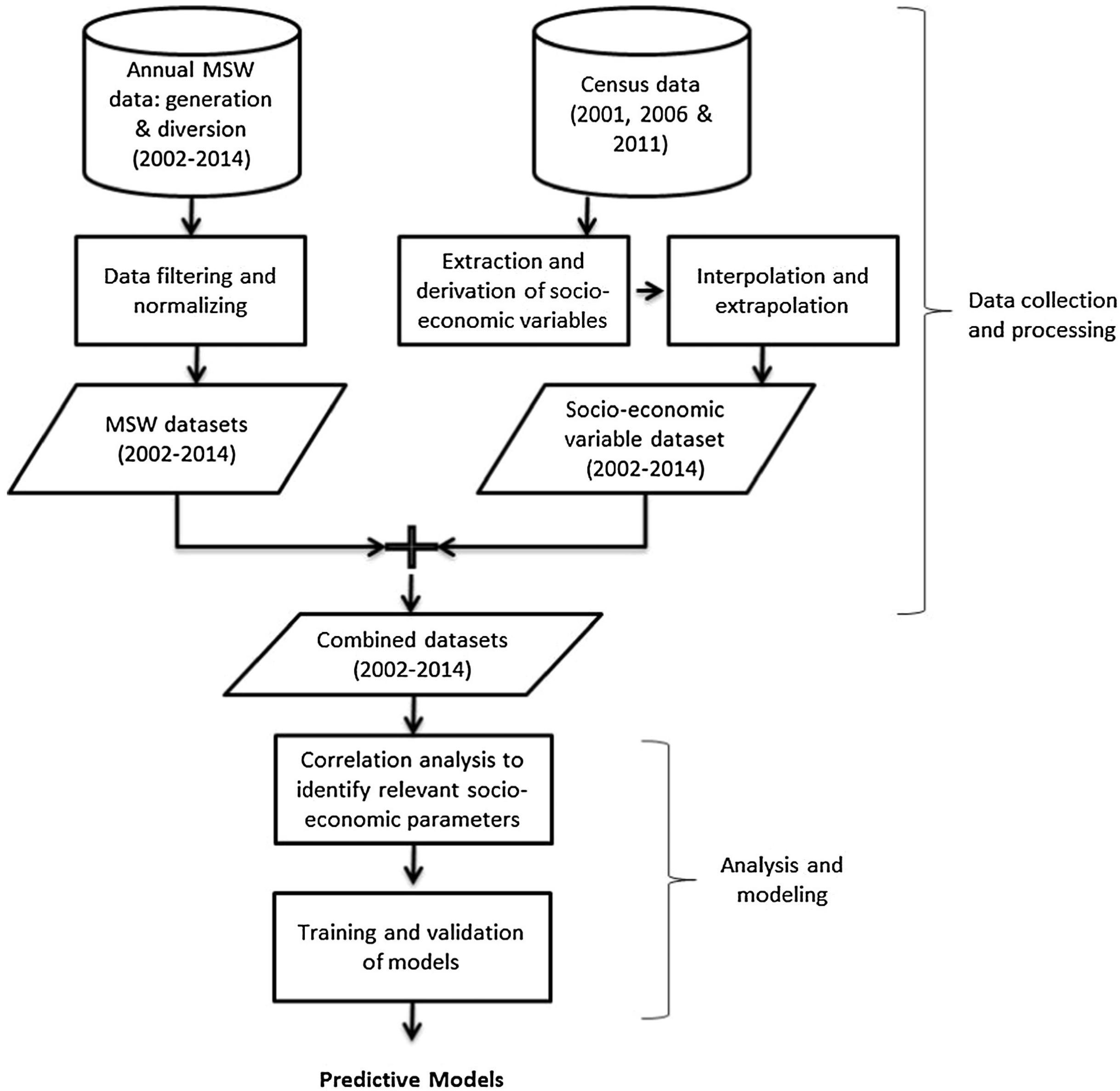

Fig. 1. Overall schematic of the methodology of the study.

The optimal model complexity, which provides highest prediction performance for unseen data, was achieved by adjusting the machine learning algorithm parameters to minimize both training and testing errors. A model was trained and tested for each randomly divided testing and training sets. The mean and standard deviation of performance indices were calculated and reported.

3.5. Machine learning approaches

The artificial neural networks are inspired by the function of human brain to learn complex relationships when presented with data. They are powerful methods of leaning complex non-linear relationships (Friedman et al., 2009). In this study, we used single hidden layer feed forward neural network architecture, which has been often used in previous studies (Ojha et al., 2017). This network architecture is composed of three layers; an input later, a hidden layer and an output layer. The neurons in input and output layers handle the input and output variables of the model while neurons in hidden layer are adjusted to map inputs to outputs during the learning process (Abbasi and El Hanandeh, 2016). The output of each neuron (Y_i) is calculated by taking the weighted sum of inputs, $z = \sum w_{ij}x_j$, and applying the sigmoid transfer function to the weighted sum, $Y_i = \frac{1}{1+\exp(z)}$. x_j is the j^{th} input to the neuron. The weight coefficients (w_{ij}) are assigned random initial values and then adjusted during the training phase to minimize the sum of squared differences between computed and actual output. The training of neural network was performed using Levenberg–Marquardt back-propagation algorithm, which has been reported to deliver stable and fast results (Yu and Wilamowski, 2011).

The choice of initial values for weight coefficients may lead to local minima during training, resulting in sub-optimal models. This has been corrected by iterating the model building process for 100 times with different random initial values for weight coefficients in each run. According to testing performance, the best 25 iterations were selected and their performances were averaged. A similar approach has been used to eliminate sub-optimal neural network models in a previous study (Azadi and Karimi-Jashni, 2016). The number of neurons in the hidden layer was varied to find the optimal model complexity that minimizes the prediction error. Main advantage of neural network approach compared to other machine learning approaches is its ability to model complex non-linear behavior (Dreiseitl and Ohno-Machado, 2002). However, neural networks are black box models, which do not allow interpretation. They can also be unstable and highly sensitive to changes in training data (Cunningham et al., 2000).

Decision tree methods employ recursive binary partitioning of the predictor variable data space into homogeneous multiple regions and fitting a linear constant to each region. These regions are called nodes, having a root node at top of the tree, internal nodes inside the tree and leaf nodes at the bottom (Tayefi et al., 2017). The model value of each region of a decision tree with M partitions is given by $f(x) = \sum_{1}^{M} c_m I(x \in R_m)$. c_m is the constant value of partition m, I is an indicator function returning 1 or 0 depending on the predictor variable (x) criteria and R_m is the predictor variable space (Friedman et al., 2009). The partitions are created by binary splits of predictor variables to minimize the sum of squared differences between computed and actual value. As the concurrent determination of the best splits for all the nodes is computationally difficult, the tree building occurs using the greedy algorithm, which performs locally optimal split at each node. Tree building starts by splitting the total dataset into two parts by a single variable value to minimize the squared error. The resulting datasets are again split into smaller datasets. This process continues until a stop criterion is reached. The depth of the tree can be controlled by defining number of data points in a leaf node so that splitting does not occur beyond certain leaf node size.

We used Classification And Regression Tree (CART) algorithm to build the decision tree model. As with all greedy algorithms, the CART algorithm may find locally optimal results. However, it has been reported to perform very closely to globally optimal trees (Murthy and Salzberg, 1995). To reduce the impacts of local optima, 100 decision trees were trained by randomly partitioning dataset into training and test sets, providing different data for training set in each run. Average errors and performance indicators were reported. Decision tree models are non-parametric models. This allows them to be used with different types of data (continuous, categorical) and requires minimum data transformations (Figueira et al., 2017). Decision tree models tend build large complex trees with increasing dataset size. This can cause overfitting when training data with noise is provided (Pal and Mather, 2003). Further, decision trees has limited ability to handle non-linear data compared to other algorithms such as neural networks (Tso and Yau, 2007). The main differences in two algorithms are summarized in Table 3.

4. Results and discussion

4.1. Characteristics of datasets

Table 4 summarizes the datasets after pre-processing. MSW, Paper, Leaf & Yard, Kitchen organic data were sourced from WDO website, data call section (Waste Diversion Ontario, 2014). Data for socio-economic parameters were obtained from 2001 census (Statistics Canada, 2001), 2006 census (Statistics Canada, 2006) and 2011 national household survey/census (Statistics Canada, 2011). Histograms of waste quantity data, namely MSW, Paper, Leaf & yard, Kitchen organic are given in Appendix A.

The MSW dataset contained the most complete waste data for the reported years. It contained 1553 data points after data conditioning and integration operations. The histogram of MSW per household data is shown in Appendix A, Fig. A1. The overall trends of MSW generation in Ontario were evaluated by plotting total yearly MSW and average MSW per household as shown in Fig. 2. Average MSW per household was calculated by dividing total MSW generation by total number of households for respective years.

The total MSW generation, as shown in Fig. 2, has gradually increased since 2010. In contrast, the average MSW per household shows a gradual decline over the years. Large variations in waste generation patterns were not observed. The Bluebox and paper diversion dataset contained 1867 data points after data conditioning and integration. The histogram of paper diversion data is shown in Appendix A, Fig. A2. Fig. 3 shows the yearly total Bluebox and paper waste quantities diverted in Ontario and Bluebox and paper waste diversion per household. The total Bluebox waste and paper waste are closely related because around 75% of Bluebox waste diverted consists of paper. Both quantities show stagnation in terms of total and per household quantities after year 2008.

The leaf & yard and kitchen organic datasets contained comparatively lower number of data, 765 and 263 data points, respectively. Most of the municipalities in Ontario are yet to implement programs to collect organic waste and hence do not report such data. The histograms of leaf & yard and kitchen waste (shown in Appendix A, Figs. A3 and A4) indicate large number of smaller

Table 3
The main differences in decision tree and neural network approaches.

	Decision trees	Artificial neural networks (ANN)
Advantages	– Able to handle categorical variables – Data transformation is not required	– High learning capacity, able to model complex non-linear data
Disadvantages	– Limited learning capacity – Tends to over-fit with noisy data	– Requires data conditioning – Model is difficult to interpret

Table 4
The datasets generated after conditioning and integration.

Dataset	Response variable	Predictor variables	Number of data points
MSW-Census	Residential MSW generation per household (2006–2014)	Socio-economic variables from census (2006–2014)	1553
Bluebox/paper-Census	Residential Bluebox and paper waste diversion per household (2002–2014)	Socio-economic variables from census (2002–2014)	1867
Yard-Census	Residential leaf and yard waste diversion per household (2002–2014)	Socio-economic variables from census (2002–2014)	765
Kitchen-Census	Residential organic kitchen waste diversion per household (2002–2014)	Socio-economic variables from census (2002–2014)	263

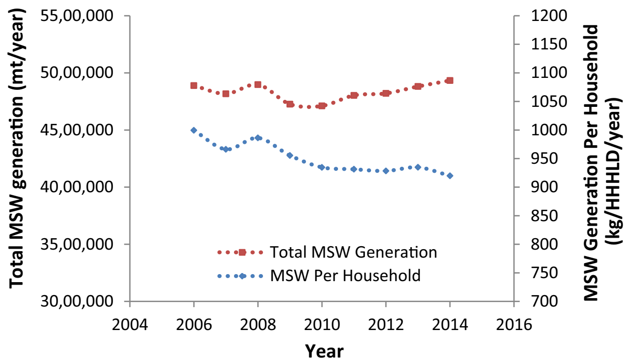

Fig. 2. Total MSW generation and average MSW per household for province of Ontario.

values. This may be due to reporting errors and absence of data reconciliation performed by WDO. Fig. 4 shows that both leaf & yard and kitchen organic waste diversion have increased steadily over the years. This is expected as increasing number organic waste recycling programs have been implemented over the years. Due to limited data availability and the absence data reconciliation, the leaf & yard and kitchen organic waste datasets were deemed not suitable for modeling and were not analyzed further.

4.2. Data correlations

The main purpose of the correlation analysis was to find and rank socio-economic variables that have high correlation with response variable and weak or no correlation with each other. Such variables carry the highest amount of relevant information to the model. This was determined by the correlation matrix, which provides the correlations between each variable pair. Correlation matrix for MSW generation per household (MSW), number of households (NHH) and socio-economic parameters is shown in Table 5. It is symmetric over the diagonal and hence only the bottom part is shown.

The MSW column of Table 5 shows how MSW generation per household is correlated with socio-economic variables. The

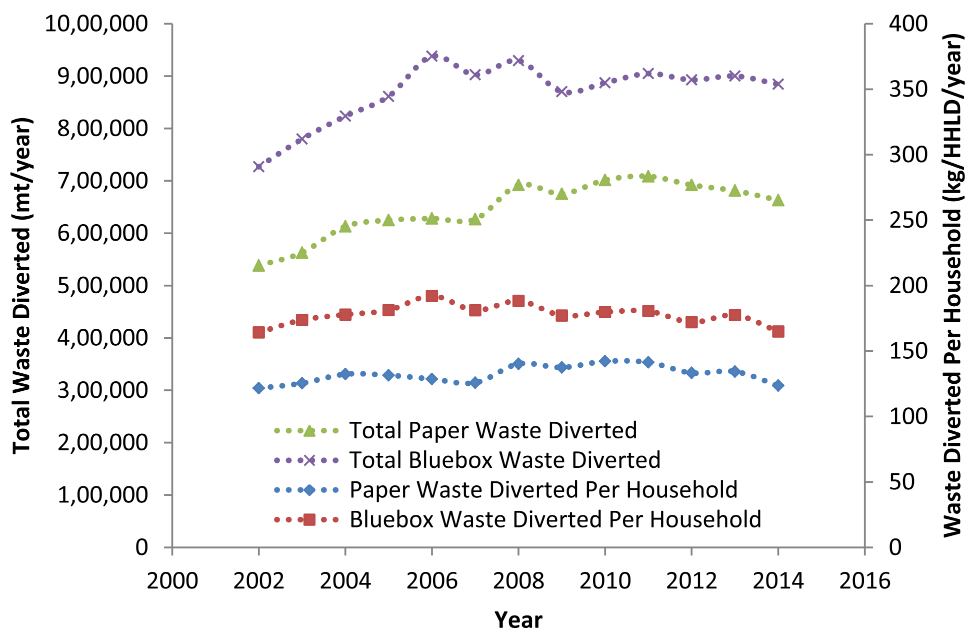

Fig. 3. Total Bluebox and paper waste diversion and Bluebox and paper waste diversion per household for province of Ontario.

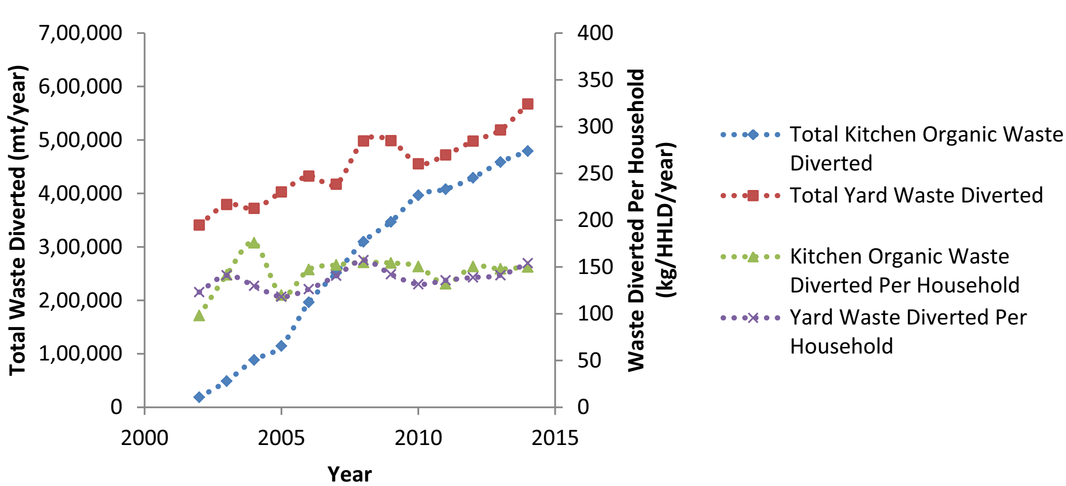

Fig. 4. Total leaf & yard and kitchen organic waste diversion and leaf & yard and kitchen organic waste diversion per household for province of Ontario.

Table 5
Correlation matrix for MSW generation per household (MSW) and socio-economic parameters.

	MSW	NHH	POP	INC	EDU	EMP	IND	DWE	HHCH	WORKP
MSW	1.00									
NHH	0.18	1.00								
POP	−0.56	−0.20	1.00							
INC	0.26	−0.15	−0.25	1.00						
EDU	−0.14	0.11	−0.05	−0.48	1.00					
EMP	0.37	0.13	−0.66	0.53	−0.15	1.00				
IND	0.05	−0.06	−0.05	0.00	−0.12	0.06	1.00			
DWE	−0.24	−0.03	0.21	0.15	−0.04	0.15	0.15	1.00		
HHCH	−0.01	0.17	0.24	−0.32	0.06	−0.48	−0.19	−0.75	1.00	
WORKP	0.22	0.16	−0.16	0.28	−0.14	0.03	−0.26	−0.38	0.33	1.00

ranking of the socio-economic variables from largest to smallest was POP, EMP, INC, DWE and WORKP, EDU, IND and WORKP. Scatter plots of POP, INC, EMP and DWE vs. -MSW further illustrate these correlations as shown in Fig. 5. The parameter POP, the fraction of population over 45 years, had the largest correlation coefficient of −0.56. This negative relationship between POP and MSW can be seen in Fig. 5, leading to the observation that older populations create less waste. The positive trends of INC-MSW and EMP-MSW are consistent with previous studies (Bandara et al., 2007). INC-EMP, POP-EMP and DWE-HHCH parameter pairs were found to be correlated with each other. Using those pairs in models was avoided to overcome multi-collinearity problems. Additionally, the number of households in a municipality (NHH) was not found to be correlated with other parameters, validating our assumption that the size of a city does not affect per household waste quantities or socio-economic parameters.

Similar to MSW generation data, a correlation analysis was performed using Paper waste diversion per household (Paper) and the results are shown in Table 6. No socio-economic parameters were found to have correlation coefficients considerably larger than that of others. The ranking of the socio-economic variables from largest to smallest was DWE, POP, WORKP, INC, EDU, EMP, HHCH and IND.

4.3. Modeling and prediction of MSW generation and paper diversion

The socio-economic parameters were included in the model one by one according to their ranking. Inclusion of highly inter-correlated socio-economic variables was avoided. Table 7 shows the variation of *MSE* of training and testing according to the socio-economic parameters included. Training error decreased when more socio-economic parameters were included because of the increase of the model complexity. However, the testing error, which tests the model performance for new data, ceased to decrease after the parameter DWE was included. So, we concluded that it is sufficient to include POP, INC and DWE in MSW generation predictive models and addition of more predictors does not improve the model. This analysis was performed using neural network algorithm, with 15 neurons in the hidden layer. Similar results were obtained using decision tree algorithms.

Once the socio-economic parameters to be included in the model were chosen, the data partition ratio was varied to determine the optimal division of data. Table 8 shows the variation of training and testing errors for different data partition ratios. The error decreased with increasing testing set size until 80:20 ratio. For small datasets, a decrease in error is expected with increasing

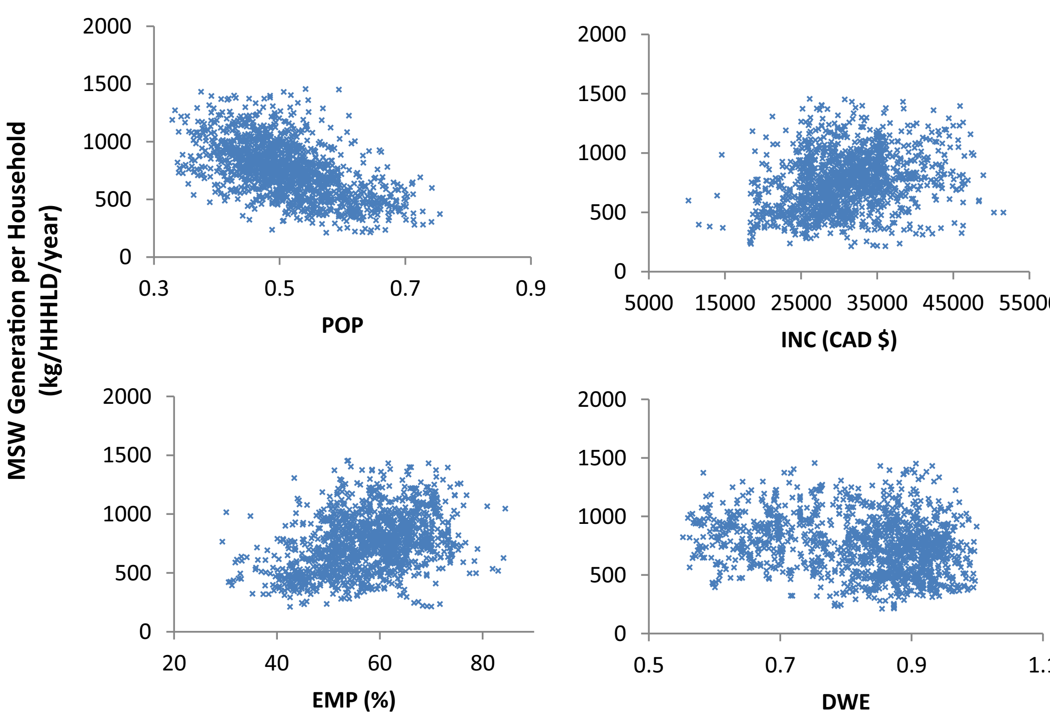

Fig. 5. Scatter plots showing correlation between and MSW generated per household and POP, INC, EMP and DWE.

Table 6
Correlation matrix for Paper diversion per household (Paper) and socio-economic parameters.

	Paper	NHH	POP	INC	EDU	EMP	IND	DWE	HHCH	WORKP
Paper	1.00									
NHH	0.14	1.00								
POP	−0.30	−0.19	1.00							
INC	0.26	−0.15	−0.27	1.00						
EDU	−0.25	0.13	0.02	−0.51	1.00					
EMP	0.20	0.12	−0.64	0.53	−0.19	1.00				
IND	−0.06	−0.06	0.06	0.04	−0.09	0.03	1.00			
DWE	−0.32	−0.03	0.26	0.12	−0.02	0.16	0.19	1.00		
HHCH	0.11	0.18	0.25	−0.33	0.09	−0.52	−0.16	−0.73	1.00	
WORKP	0.28	0.16	−0.26	0.31	−0.15	0.02	−0.25	−0.46	0.33	1.00

Table 7
Variation of the model error in terms of MSE according to the socio-economic variables included.

Model	Training *MSE* (kg/HHLD/year)^2	Testing *MSE* (kg/HHLD/year)^2
POP	22,549	30,723
POP, INC	20,978	28,996
POP, INC, DWE	18,219	25,907
POP, INC, DWE, WORKP	17,485	25,695
POP, INC, DWE, WORKP, EDU	16,357	25,628

Table 8
Model error with data partition ratio.

Training and testing data partition ratio	Training *MSE* (kg/HHLD/year)^2	Testing *MSE* (kg/HHLD/year)^2
60:40	20,306	28,419
70:30	19,059	26,786
80:20	17,761	24,901
85:15	18,219	25,907
90:10	18,257	26,651

training set size because model keeps learning more about underlying structures as more data is fed to the model. However, this argument is not valid for large datasets because the model has enough data to learn and feeding more data into the model does not improve its accuracy. In such cases, increasing the testing set size may reduce the testing error as more representative testing data can be provided for model validation. We assumed that the MSW generation model had enough data to learn at 80:20 ratio as there was no improvement in model errors by further increasing the training set size. The ratio of 80:20 was used in building MSW generation models. Similar results were obtained for Paper diversion models.

The model's ability to learn underlying structures is also dependent on the complexity of the model as allowed by the machine learning algorithm. The complexity of decision tree and neural network models can be controlled by minimum leaf size and number of neurons in hidden layer, respectively. Fig. 6 shows the average training and testing errors for the decision tree models the complexity was varied by changing the minimum leaf size. Minimum leaf size restricts the number of observations in each leaf node. Larger leaf size reduces the depth of the tree, thus resulting in a simpler model while a smaller leaf size increases the depth of the tree, producing a more complex model (Xue et al., 2015).

The optimal model complexity was around minimum leaf size of 30, producing a minimum MSE of 29,864. It is also important to look at the standard deviation of the error as it is an indication of the stability of models, particularly with small datasets. As shown by the error bars, the standard deviation of testing error did not increase considerably with increasing model complexity, indicating the model has generalized well to the data. The

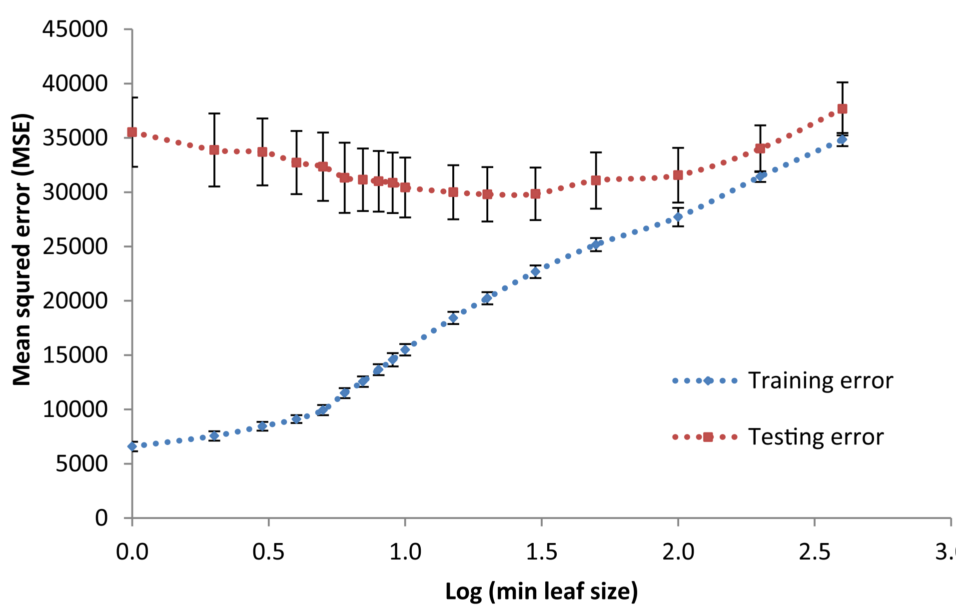

Fig. 6. Average training error and testing error vs. decision tree model complexity for MSW generation.

complexity of neural network models was varied by changing the number of neurons in the hidden layer. The average training and testing errors and their standard deviations of neural network models are shown in Fig. 7.

The optimal model complexity was achieved at around 40 neurons in the hidden layer with minimum testing of 20,106, lower than the minimum MSE of decision tree model. Stability of neural network models with increasing complexity was slightly lower than decision tree models as indicated by the error bars in Fig. 7.

The socio-economic parameters were included in the paper diversion model one by one according to their ranking. The parameters that improved the model were POP, INC and EDU. The optimal decision tree model complexity was at minimum leaf size of 8 with average *MSE* value of 1127 while the optimal neural network model complexity was at 40 neurons in hidden layer with average *MSE* value of 890. Table 9 summarizes and compares the predictive performance of the models for at their optimal complexities in terms of *RMSE*, *MAPE* and R^2 for both training and testing data.

MSW generation models showed the lowest out of sample prediction errors of 16–23% as given by the *MAPE* and *RMSE* of testing dataset. Although both neural networks and decision tree models showed similar error rates, the neural networks models were better at describing the variation in data as given by R^2 values. The neural network MSW generation models was better at prediction than similar previous studies based on socio-economic variables using cross sectional data (Benítez et al., 2008; Hockett et al., 1995) and panel data (Beigl et al., 2004). The R^2 values obtained for neural network waste generation models in this study, 0.83 and 0.72 for training and testing data, was comparable with previous studies using neural network approach (Azadi and Karimi-Jashni, 2016), confirming better prediction abilities of neural networks.

The decision tree models however did not perform well in describing the variation in the data compared to neural networks. This may be attributed to ability of neural network models to learn complex non-linear behavior and to optimize model parameters. Decision tree models can only produce a single model for a given training data due to their non-parametric nature while neural networks can produce multiple models for a given dataset by varying initial weight coefficients, hence optimizing to find best model. Advanced decision tree models such as boosted trees and random forests have been reported to overcome this drawback by building multiple trees and combining their output (Friedman et al., 2009). Future work may focus on using advanced decision tree methods for waste modeling.

The paper diversion models showed higher errors (32–36%) and poor R^2 values around 0.31–0.35 compared to MSW models. We believe that the main reason for the higher error was the lack of suitable predictor variables included in the paper diversion models. The socio-economic and demographic variables from census data, which was considered due to country-wide availability, were not sufficient in describing the variation in paper recycle rate. Therefore, future work will address the issue of finding recycling policy and program related parameters that consistently available at municipal level to build waste diversion models with better prediction.

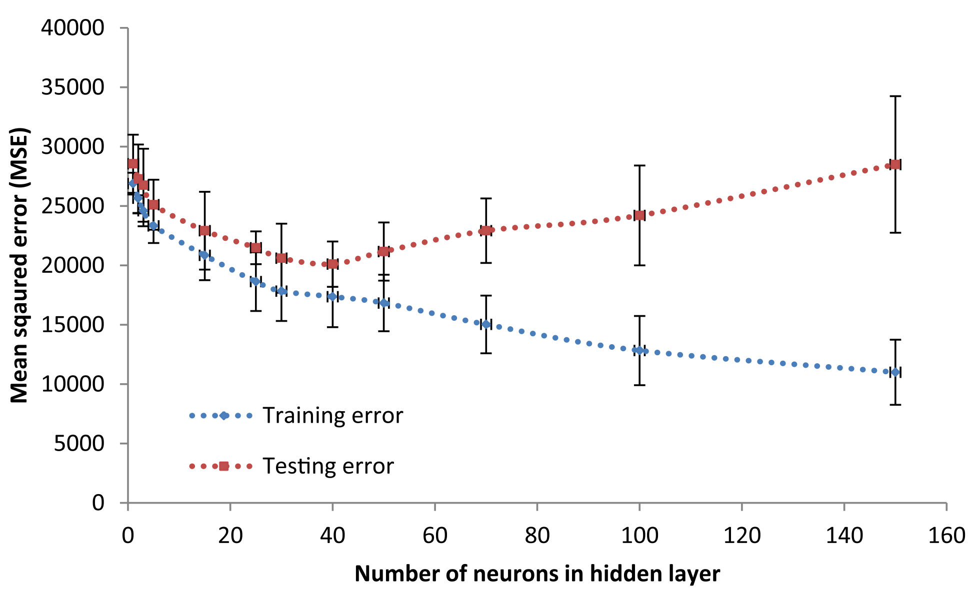

Fig. 7. Average training error and testing error vs. neural network complexity for MSW generation.

Table 9
Predictive performance of optimal models in terms of *MSE*, *RMSE* and *MAPE*.

Model	Training			Testing		
	$\frac{RMSE}{y}$ (%)	*MAPE* (%)	R^2	$\frac{RMSE}{y}$ (%)	*MAPE* (%)	R^2
MSW-Census (Decision Trees)	20	18	0.63	23	19	0.54
MSW-Census (Neural Networks)	18	15	0.83	20	16	0.72
Paper-Census (Decision Trees)	31	29	0.42	36	34	0.31
Paper-Census (Neural Networks)	27	26	0.56	32	32	0.35

5. Conclusions

In this study, we developed models to predict MSW generation and paper diversion of a given region in Canada. The developed models were able to predict the MSW generation and diversion of a given region in Canada using socio-economic and demographic parameters at municipal level. The data for modeling was an integrated dataset based on yearly residential solid waste generation and paper diversion data of 220 municipalities in Ontario and socio-economic and demographic data from Canadian census program. Artificial neural network (ANN) and decision tree machine learning algorithms were used to build the models. The neural network approach created superior models than decision tree approach. It produced MSW generation models with 72% accuracy for out of sample data. Paper diversion models however lagged in performance only describing 32–36% of out of sample data accurately. This was because variables explaining recycle behavior were not captured in socio-economic variable dataset derived from Canadian census program. The results demonstrate that given sufficient socio-economic explanatory variables, the machine learning methods can produce models with high accuracy for waste prediction applications.

The current waste generation model can be applied across Canada because the input socio-economic parameters can be consistently generated using census data, which is available to all the municipalities in Canada. Ability to predict waste generation provides municipalities a unique opportunity to plan and optimize their waste management operations. This work also provides critical information to Canadian remote and Northern communities where there is limited past data.

Acknowledgements

This work was funded by the Program of Energy Research and Development (PERD), an inter-departmental federal program operated by Natural Resources Canada (NRCan). We would like to thank Ms. Maria Constantinou of Waste Diversion Ontario (WDO) for providing additional data and information. We also thank the anonymous reviewers for their inputs to improve the article.

Appendix A. Histograms of waste quantity datasets

See Figs. A1–A4.

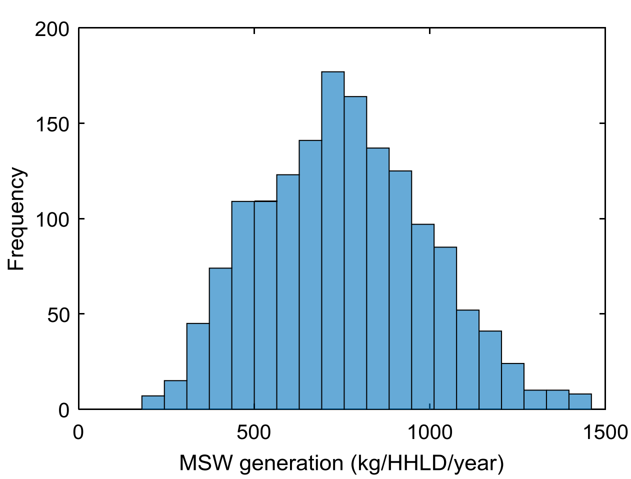

Fig. A1. Histogram of MSW generation per household and year.

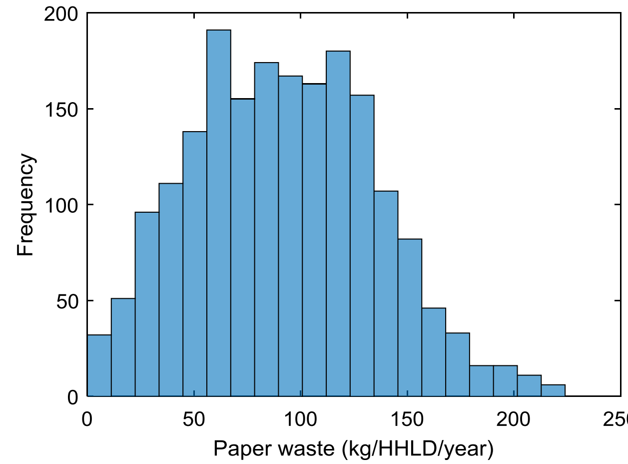

Fig. A2. Histogram of paper diversion per household and year.

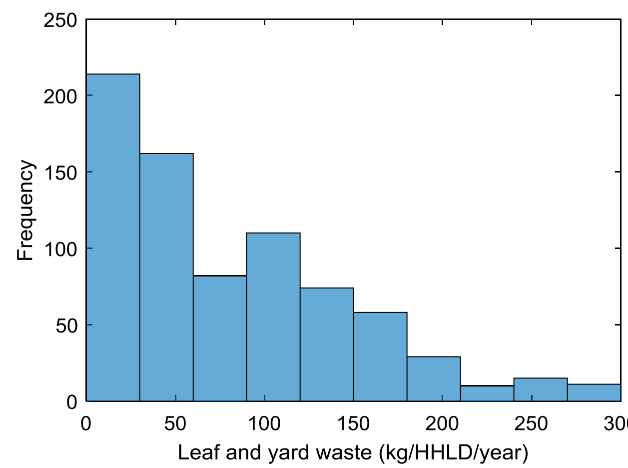

Fig. A3. Histogram of leaf & yard waste diversion per household and year.

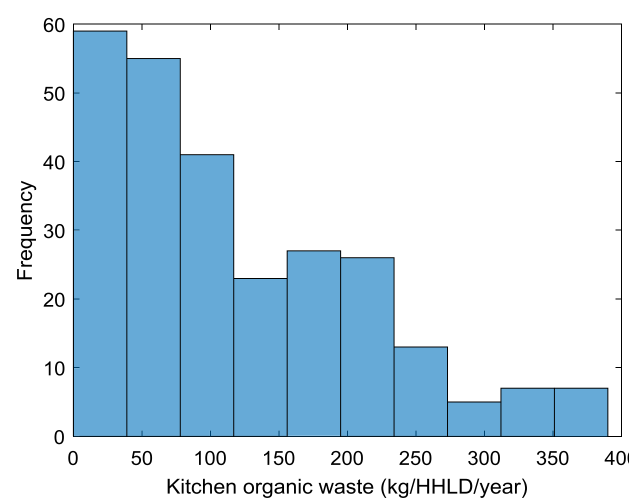

Fig. A4. Histogram of kitchen waste diversion per household and year.

Appendix B. Descriptions of demographic and socio-economic parameters

All the demographic and socio-economic parameters were calculated using Canadian census program data. The derivation of each parameter is briefly described below.

1. Fraction of population over 45 years (POP) was derived by adding population segments having ages 45–54, 55–64, 65–74, 75–84 and 85 or over, and dividing it by total population.
2. Median personal income (INC) was used as reported in Census data, which is the median value of all the incomes reported for selected jurisdiction.
3. Fraction of population with no high school graduation certificate, diploma or degree (EMP) was derived by dividing total population 15 years and over with no high school certificate, diploma and degree by total population over 15 years.
4. Employment rate (EMP) was used as reported in census data, which refers number of persons employed in the week prior to census day, expressed as a percentage of total population 15 years and over.
5. Fraction of people employed by agriculture, resource based, manufacturing or construction industries (IND) was derived by adding population segments employed by agriculture and other resource-based industries, manufacturing and construction industries and dividing in by total labour force.
6. Fraction of owned dwellings (versus rented) (DWE) was derived by dividing number occupied dwellings owned by a member of household by total number of dwellings.
7. Fraction of one person households (HHCHAR) was derived by dividing number of dwellings occupied by a single person by total number of dwellings.
8. Fraction of population worked at a usual work place (WORKP) was derived by dividing number of people who worked at a location other than their home by total people in the workforce.

References

Abbasi, M., El Hanandeh, A., 2016. Forecasting municipal solid waste generation using artificial intelligence modelling approaches. Waste Manage. 56, 13–22. https://doi.org/10.1016/j.wasman.2016.05.018.

Abbott, A., Nandeibam, S., O'Shea, L., 2011. Explaining the variation in household recycling rates across the UK. Ecol. Econ. 70, 2214–2223. https://doi.org/10.1016/j.ecolecon.2011.06.028.

Agriculture-Canada, 2017. Biomass Inventory Mapping and Analysis Tool [WWW Document]. <http://www.agr.gc.ca/atlas/bimat>.

Albright, S.C., Winston, W., Zappe, C., 2010. Data Analysis and Decision Making. Cengage Learning, Mason, OH.

Ali Abdoli, M., Falahnezhad, M., Behboudian, S., 2011. Multivariate economic approach for solid waste generation modeling: impact of climate factor. Environ. Eng. Sci. 28, 627–633.

Antanasijević, D., Pocajt, V., Popović, I., Redžić, N., Ristić, M., 2013. The forecasting of municipal waste generation using artificial neural networks and sustainability indicators. Sustain. Sci. 8, 37–46. https://doi.org/10.1007/s11625-012-0161-9.

Arbulú, I., Lozano, J., Rey-Maquieira, J., 2015. Tourism and solid waste generation in Europe: a panel data assessment of the Environmental Kuznets Curve. Waste Manage. 46, 628–636. https://doi.org/10.1016/j.wasman.2015.04.014.

Azadi, S., Karimi-Jashni, A., 2016. Verifying the performance of artificial neural network and multiple linear regression in predicting the mean seasonal municipal solid waste generation rate: a case study of Fars province. Iran. Waste Manage. 48, 14–23. https://doi.org/10.1016/j.wasman.2015.09.034.

Bach, H., Mild, A., Natter, M., Weber, A., 2004. Combining socio-demographic and logistic factors to explain the generation and collection of waste paper. Resour. Conserv. Recycl. 41, 65–73. https://doi.org/10.1016/j.resconrec.2003.08.004.

Bandara, N.J.G.J., Hettiaratchi, J.P.A., Wirasinghe, S.C., Pilapiiya, S., 2007. Relation of waste generation and composition to socio-economic factors: a case study. Environ. Monit. Assess. 135, 31–39. https://doi.org/10.1007/s10661-007-9705-3.

Beigl, P., Lebersorger, S., Salhofer, S., 2008. Modelling municipal solid waste generation: a review. Waste Manage. 28, 200–214. https://doi.org/10.1016/j.wasman.2006.12.011.

Beigl, P., Wassermann, G., Schneider, F., Salhofer, S., 2004. Forecasting municipal solid waste generation in major European cities. In: Pahl-Wostl, C., Schmidt, S., Jakeman, T. (Eds.), iEMSs 2004 International Congress: Complexity and Integrated Resources Management. Osnabrueck, Germany. Available at: <http://www.iemss.org/iemss2004/pdf/regional/beigfore.pdf>.

Benítez, S.O., Lozano-Olvera, G., Morelos, R.A., de Vega, C.A., 2008. Mathematical modeling to predict residential solid waste generation. Waste Manage. 28 (Supp.), S7–S13. https://doi.org/10.1016/j.wasman.2008.03.020.

Bernstad, A., 2014. Household food waste separation behavior and the importance of convenience. Waste Manage. 34, 1317–1323. https://doi.org/10.1016/j.wasman.2014.03.013.

Cucchiella, F., D'Adamo, I., Gastaldi, M., 2017. Sustainable waste management: waste to energy plant as an alternative to landfill. Energy Convers. Manage. 131, 18–31. https://doi.org/10.1016/j.enconman.2016.11.012.

Cunningham, Â., Carney, J., Jacob, S., 2000. Stability problems with arti ® cial neural networks and the ensemble solution 20, 217–225.

Daskalopoulos, E., Badr, O., Probert, S.D., 1998. Municipal solid waste: a prediction methodology for the generation rate and composition in the European Union countries and the United States of America. Resour. Conserv. Recycl. 24, 155–166. https://doi.org/10.1016/S0921-3449(98)00032-9.

Dreiseitl, S., Ohno-Machado, L., 2002. Logistic regression and artificial neural network classification models: a methodology review. J. Biomed. Inform. 35, 352–359. https://doi.org/10.1016/S1532-0464(03)00034-0.

Dyson, B., Chang, N. Bin, 2005. Forecasting municipal solid waste generation in a fast-growing urban region with system dynamics modeling. Waste Manage. 25, 669–679. https://doi.org/10.1016/j.wasman.2004.10.005.

Environment and Climate Change Canada, 2017. Solid Waste Management for Northern and Remote Communities: Planning and Technical Guidance Document.

Figueira, A. da C., Pitombo, C.S., de Oliveira, P.T.M. e S., Larocca, A.P.C., 2017. Identification of rules induced through decision tree algorithm for detection of traffic accidents with victims: a study case from Brazil. Case Stud. Transp. Policy 5, 200–207. https://doi.org/10.1016/j.cstp.2017.02.004.

Friedman, J., Hastie, T., Tibshirani, R., 2009. The Elements of Statistical Learning. Springer series in statistics Springer, Berlin, Stanford, CA.

Ghinea, C., Drăgoi, E.N., Comăniţă, E.-D., Gavrilescu, M., Câmpean, T., Curteanu, S., Gavrilescu, M., 2016. Forecasting municipal solid waste generation using prognostic tools and regression analysis. J. Environ. Manage. 182, 80–93. https://doi.org/10.1016/j.jenvman.2016.07.026.

Grazhdani, D., 2016. Assessing the variables affecting on the rate of solid waste generation and recycling: an empirical analysis in Prespa Park. Waste Manage. 48, 3–13. https://doi.org/10.1016/j.wasman.2015.09.028.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182. https://doi.org/10.1016/j.aca.2011.07.027.

Hage, O., Söderholm, P., Berglund, C., 2009. Norms and economic motivation in household recycling: empirical evidence from Sweden. Resour. Conserv. Recycl. 53, 155–165. https://doi.org/10.1016/j.resconrec.2008.11.003.

Hannan, M.A., Abdulla Al Mamun, M., Hussain, A., Basri, H., Begum, R.A., 2015. A review on technologies and their usage in solid waste monitoring and management systems: issues and challenges. Waste Manage. 43, 509–523. https://doi.org/10.1016/j.wasman.2015.05.033.

Hockett, D., Lober, D.J., Pilgrim, K., 1995. Determinants of per capita municipal solid waste generation in the Southeastern United States. J. Environ. Manage. 45, 205–217. https://doi.org/10.1006/jema.1995.0069.

Jahandideh, S., Jahandideh, S., Asadabadi, E.B., Askarian, M., Movahedi, M.M., Hosseini, S., Jahandideh, M., 2009. The use of artificial neural networks and multiple linear regression to predict rate of medical waste generation. Waste Manage. 29, 2874–2879. https://doi.org/10.1016/j.wasman.2009.06.027.

Johnson, N.E., Ianiuk, O., Cazap, D., Liu, L., Starobin, D., Dobler, G., Ghandehari, M., 2017. Patterns of waste generation: a gradient boosting model for short-term waste prediction in New York City. Waste Manage. 62, 3–11. https://doi.org/10.1016/j.wasman.2017.01.037.

Joosten, L.A.J., Hekkert, M.P., Worrell, E., Turkenburg, W.C., 1999. STREAMS: a new method for analysing material flows through society. Resour. Conserv. Recycl. 27, 249–266. https://doi.org/10.1016/S0921-3449(99)00009-9.

Junninen, H., Niska, H., Tuppurainen, K., Ruuskanen, J., Kolehmainen, M., 2004. Methods for imputation of missing values in air quality data sets. Atmos. Environ. 38, 2895–2907. https://doi.org/10.1016/j.atmosenv.2004.02.026.

Kollikkathara, N., Feng, H., Yu, D., 2010. A system dynamic modeling approach for evaluating municipal solid waste generation, landfill capacity and related cost management issues. Waste Manage. 30, 2194–2203. https://doi.org/10.1016/j.wasman.2010.05.012.

Korai, M.S., Mahar, R.B., Uqaili, M.A., 2017. The feasibility of municipal solid waste for energy generation and its existing management practices in Pakistan. Renew. Sustain. Energy Rev. 72, 338–353. https://doi.org/10.1016/j.rser.2017.01.051.

Li, Y., Zhou, L.W., Wang, R.Z., 2017. Urban biomass and methods of estimating municipal biomass resources. Renew. Sustain. Energy Rev. 80, 1017–1030. https://doi.org/10.1016/j.rser.2017.05.214.

Lindh, T., 2003. Demography as a forecasting tool. Futures 35, 37–48. https://doi.org/10.1016/S0016-3287(02)00049-6.

Miliute-Plepiene, J., Hage, O., Plepys, A., Reipas, A., 2016. What motivates households recycling behaviour in recycling schemes of different maturity? Lessons from Lithuania and Sweden. Resour. Conserv. Recycl. 113, 40–52. https://doi.org/10.1016/j.resconrec.2016.05.008.

Monavari, S.M., Omrani, G.A., Karbassi, A., Raof, F.F., 2012. The effects of socioeconomic parameters on household solid-waste generation and composition in developing countries (a case study: Ahvaz, Iran). Environ. Monit. Assess. 184, 1841–1846. https://doi.org/10.1007/s10661-011-2082-y.

Mrayyan, B., Hamdi, M.R., 2006. Management approaches to integrated solid waste in industrialized zones in Jordan: a case of Zarqa City. Waste Manage. 26, 195–205. 10.1016/j.wasman.2005.06.008.

Murthy, S.K., Salzberg, S., 1995. Decision tree induction: how effective is the greedy heuristic? In: First International Conference on Knowledge Discovery and Data Mining, pp. 222–227.

Navarro-Esbrí, J., Diamadopoulos, E., Ginestar, D., 2002. Time series analysis and forecasting techniques for municipal solid waste management. Resour. Conserv. Recycl. 35, 201–214. https://doi.org/10.1016/S0921-3449(02)00002-2.
Ojha, V.K., Abraham, A., Snášel, V., 2017. Metaheuristic design of feedforward neural networks: a review of two decades of research. Eng. Appl. Artif. Intell. 60, 97–116. https://doi.org/10.1016/j.engappai.2017.01.013.
Pal, M., Mather, P.M., 2003. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sens. Environ. 86, 554–565. https://doi.org/10.1016/S0034-4257(03)00132-9.
Pao, H.-T., Chih, Y.-Y., 2006. Comparison of TSCS regression and neural network models for panel data forecasting: debt policy. Neural Comput. Appl. 15, 117–123. https://doi.org/10.1007/s00521-005-0014-x.
Puig-Ventosa, I., 2008. Charging systems and PAYT experiences for waste management in Spain. Waste Manage. 28, 2767–2771. https://doi.org/10.1016/j.wasman.2008.03.029.
Reynolds, C., Geschke, A., Piantadosi, J., Boland, J., 2016. Estimating industrial solid waste and municipal solid waste data at high resolution using economic accounts: an input–output approach with Australian case study. J. Mater. Cycles Waste Manage. 18, 677–686. https://doi.org/10.1007/s10163-015-0363-1.
Rhodes, R.E., Beauchamp, M.R., Conner, M., deBruijn, G.-J., Latimer-Cheung, A., Kaushal, N., 2014. Are mere instructions enough? Evaluation of four types of messaging on community depot recycling. Resour. Conserv. Recycl. 90, 1–8. https://doi.org/10.1016/j.resconrec.2014.04.008.
Sidique, S.F., Joshi, S.V., Lupi, F., 2010. Factors influencing the rate of recycling: an analysis of Minnesota counties. Resour. Conserv. Recycl. 54, 242–249. https://doi.org/10.1016/j.resconrec.2009.08.006.
Sodanil, M., 2014. Artificial Neural Network-based Time Series Analysis Forecasting for the Amount of Solid Waste in Bangkok 16–20.
Starr, J., Nicolson, C., 2015. Patterns in trash: factors driving municipal recycling in Massachusetts. Resour. Conserv. Recycl. 99, 7–18. https://doi.org/10.1016/j.resconrec.2015.03.009.
Statistics Canada, 2011. 2011 Census Data Products [WWW Document]. <http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/index-eng.cfm>.
Statistics Canada, 2006. 2006 Census of Canada [WWW Document]. <http://www12.statcan.ca/census-recensement/2006/index-eng.cfm> (accessed 6.14.16).
Statistics Canada, 2001. 2001 Census of Canada [WWW Document]. <http://www12.statcan.ca/english/census01/home/index.cfm> (accessed 6.14.16).
Tayefi, M., Esmaeili, H., Saberi Karimian, M., Amirabadi Zadeh, A., Ebrahimi, M., Safarian, M., Nematy, M., Parizadeh, S.M.R., Ferns, G.A., Ghayour-Mobarhan, M., 2017. The application of a decision tree to establish the parameters associated with hypertension. Comput. Methods Programs Biomed. 139, 83–91. https://doi.org/10.1016/j.cmpb.2016.10.020.
Tso, G.K.F., Yau, K.K.W., 2007. Predicting electricity energy consumption: a comparison of regression analysis, decision tree and neural networks. Energy 32, 1761–1768. https://doi.org/10.1016/j.energy.2006.11.010.
Vitorino de Souza Melaré, A., Montenegro González, S., Faceli, K., Casadei, V., 2017. Technologies and decision support systems to aid solid-waste management: a systematic review. Waste Manage. 59, 567–584. https://doi.org/10.1016/j.wasman.2016.10.045.
Waste Diversion Ontario, 2014. Data Call Documents [WWW Document]. <http://wdo.ca/learn/documents> (accessed 6.15.16).
Xu, Z., Elomri, A., Pokharel, S., Zhang, Q., Ming, X.G., Liu, W., 2017. Global reverse supply chain design for solid waste recycling under uncertainties and carbon emission constraint. Waste Manage. 64, 358–370. https://doi.org/10.1016/j.wasman.2017.02.024.
Xue, D., Pang, F., Meng, F., Wang, Z., Wu, W., 2015. Decision-tree-model identification of nitrate pollution activities in groundwater: a combination of a dual isotope approach and chemical ions. J. Contam. Hydrol. 180, 25–33. https://doi.org/10.1016/j.jconhyd.2015.07.003.
Yu, H., Wilamowski, B.M., 2011. Levenberg–marquardt training. Ind. Electron. Handb. 5, 1.
Zade, J.G., Noori, R., 2008. Prediction of municipal solid waste generation by use of artificial neural network: a case study of Mashhad. Int. J. Environ. Res. 2, 13–22.