

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

ARTICLE INFO

Article history:

Received 15 March 2017
 Revised 22 November 2017
 Accepted 30 November 2017
 Available online xxx

Keywords:

Municipal solid waste
 Forecasting
 Machine learning
 Modeling
 Neural networks
 Decision trees

ABSTRACT

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.

© 2017 Published by Elsevier Ltd.

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).

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

Download English Version:

<https://daneshyari.com/en/article/8869886>

Download Persian Version:

<https://daneshyari.com/article/8869886>

[Daneshyari.com](https://daneshyari.com)