Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The use of a recurrent neural network model with separated time-series and lagged daily inputs for waste disposal rates modeling during COVID-19

Hoang Lan Vu a, Kelvin Tsun Wai Ng a,*, Amy Richter a, Golam Kabir b

a Environmental Systems Engineering, Faculty of Engineering and Applied Science, University of Regina, 3737 Wascana Parkway, SK S4S 0A2, Canada
b Industrial Systems Engineering, Faculty of Engineering and Applied Science, University of Regina, Saskatchewan S4S 0A2, Canada

ARTICLE INFO

Keywords:
COVID-19
Waste disposal behaviors
Solid waste management
Time series analysis
Recurrent neural network
Lagged inputs

ABSTRACT

A new modeling framework is proposed to estimate mixed waste disposal rates in a Canadian capital city during the pandemic. Different Recurrent Neural Network models were developed using climatic, socioeconomic, and COVID-19 related daily variables with different input lag times and study periods. It is hypothesized that the use of distinct time series and lagged inputs may improve modeling accuracy. Considering the entire 7.5-year period from Jan 2013 to Sept 2020, multi-variate weekday models were sensitive with lag times in the testing stage. It appears that the selection of input variables is more important than waste model complexity. Models applying COVID-19 related inputs generally had better performance, with average MAPE of 10.1%. The optimized lag times are however similar between the periods, with slightly longer average lag for the COVID-19 at 5.3 days. Simpler models with least input variables appear to better simulate waste disposal rates, and both ‘Temp-Hum’ (Temperature-Humidity) and ‘Temp-New Test’ (Temperature-COVID new test case) models capture the general disposal trend well, with MAPE of 10.3% and 9.4%, respectively. The benefits of the use of separated time series inputs are more apparent during the COVID-19 period, with noticeable decrease in modeling error.

List of Acronyms

ANN Artificial neuron network
GT Gamma test
IA Index of agreement
IQR Interquartile range
LSTM Long Short-Term Memory
MAE Mean absolute error
MAPE Mean absolute percentage error
MSE Mean square error
MSW Municipal solid waste
PCA Principal component analysis
RNN Recurrent Neural Network

1. Introduction

Rapid urbanization and population growth have presented significant challenges to municipal solid waste management across the globe (Alam & Qiao, 2020; Heidari, Yazdanparast & Jabbarzadeh, 2019; Lu, Li & An, 2020), particularly with respect to the overall system costs (Richter, Bruce, Ng, Chowdhury & Vu, 2017, 2018). The need for accurate waste modeling is imperative to the efficiency of disposal operations and the well-being of sanitation workers during the global COVID-19 pandemic (Vu, Ng, Richter, Karimi & Kabir, 2021). Waste generation characteristics and disposal behaviors are important for the planning and operation of any waste management system, and are of great practical interest and commonly reported (Azadi & Karimi-Jashni, 2016; Kumar & Samadder, 2017; Wu, Niu, Dai & Wu, 2020). Improving the accuracy of waste generation and disposal models facilitates efficient waste collection and improves efficiency of municipal waste management planning and operation (Hannan et al., 2020; Vu, Bolingbroke, Ng & Fallah, 2019a, 2020).

1.1. Machine learning and waste generation and disposal models

Due to the robustness of the method, machine learning approaches have been widely adopted for various environmental applications, including indoor air quality (Zhang, Li, Zhao & Rao, 2019), electricity consumption (Kim, Kim & Srebric, 2020), and urban heat island effects (Equere, Mirzaei, Riffat & Wang, 2021). Specifically, machine learning techniques have also been applied in waste management studies in the last decade. For example, waste generation rate in Mashhad, Iran, has

* Corresponding author.
E-mail address: kelvin.ng@uregina.ca (K.T.W. Ng).

https://doi.org/10.1016/j.scs.2021.103339
Received 29 June 2021; Received in revised form 4 September 2021; Accepted 4 September 2021
Available online 8 September 2021
been studied using various machine learning techniques (Abdoli, Nezhad, Sede & Behboudian, 2012; Noori, Abdoli, Ghasrodashti & Ghazizade, 2009). Table 1 presents recent literature on municipal solid waste (MSW) prediction using machine learning approaches, and among them, Artificial Neuron Network (ANN) and ANN-based models are one of the most popular analytical tools. Noori, Karbassi and Salman Sabahi (2010) applied principal component analysis (PCA) and Gamma Test (GT) techniques and they found that the PCA-ANN and GT-ANN models had better results compared to conventional ANN models. Shahabi, Saeed, Ahmed and Zabihi (2012) altered the number of neurons in the hidden layer to improve the accuracy of weekly waste generation. Vu, Ng and Bolingbroke (2019b) modeled the effects of lag times on weekly yard waste time-series models and found that modeling error reduced by 50% at optimal lag times. In a Danish study, Cubillos (2020) attempted to model waste generation at household levels using a Long Short-Term Memory (LSTM) Neural Network. Wu et al. (2020), on the contrary, explored regional scale ANN models on municipal solid waste generation in China. Niu, Wu, Dai, He and Wu (2021) adopted a LSTM for MSW forecasting and obtained satisfactory results. ANN-based models are versatile and applicable to many non-linear problems, provided that a good training data set is provided (Xu et al., 2021). Common inputs used in waste studies include socioeconomic variables such as earnings and income, education level, employment status, dwelling and household characteristics, workplace and demographic parameters (Kannangara, Dua, Ahmadi & Bensebaa, 2018; Kontokosta, Hong, Johnson & Starobin, 2018; Wu et al., 2020; Younes et al., 2015), or climatic variables such as temperature, humidity, wind speed, and precipitation (Cubillos, 2020; Kontokosta et al., 2018). Some studies utilized both socioeconomic and climatic variables on waste forecasting (Kontokosta et al., 2018; Vu et al., 2019b). The selection of input variables, however, appears case specific. For example, Wu et al. (2020) used only 7 socioeconomic parameters in their regional ANN models and obtained satisfactory modeling results in China. Kontokosta et al. (2018), on the other hand, concluded that climatic variables were vital features in their ANN waste prediction models at New York. Recently, Jassim, Coskuner and Zontul (2021) used various environmental related parameters such as annual tourist numbers, annual electricity consumption, and total annual CO2 emissions to model MSW generation rates in Bahrain. One difficulty associated with waste disposal rate modeling during

| Study area            | Model used                  | Input/Target                                                                 | Findings                                                                 |
|-----------------------|-----------------------------|------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| Noori et al. (2010)   | Mashhad, Iran              | ANN, PCA-ANN, and GT-ANN models                                             | The PCA-ANN and GT-ANN models have more effective results than the ANN model. |
| Shahabi et al. (2012) | Squez, Iran                | ANN                                                                          | Creating the models and changing the number of neurons in the hidden layers, the optimal number of neurons was found. |
| Antanasijevic, Popovic, Redic and Ristic (2013) | 26 countries in Europe | Back-projection, general regression GRNN                                      | GRNN model proved to be significantly better than the more traditional BP model. |
| Shamshir, Mokhtar, Abdulai, Komoo and Yahay (2014) | Malaysia                  | ANN, genetic algorithm, response surface method                              | Combined ANN and BSM to predict or forecast solid waste generation and optimize the cost of waste collection and transportation |
| Younes et al. (2015)  | Malaysia                   | Modified Adaptive Neural Inference System (MANFIS)                          | The best input variables were people in age groups 0–14, 15–64, and 65 years, and the best model structure was 3 triangular fuzzy membership functions and 27 fuzzy rules |
| Azadi and Karimi-Jahni (2016) | 20 cities in Iran      | ANN, MLR                                                                     | ANN is better than MLR in predicting the mean seasonal municipal solid waste generation rate |
| Abbasi and Hanamdeh (2016) | Australia                | SVM, ANFIS, ANN, kNN                                                        | SVM reliably predicted monthly MSW generation, ANFIS predicted the most accurate forecasts of the peaks, kNN was successful in the monthly averages of waste quantities prediction |
| Kumar and Samadder (2017) | India                     | MLR                                                                          | R2 was 0.782 for biodegradable waste generation rate and 0.676 for non-biodegradable waste generation rate |
| Kannangara et al. (2018) | Ontario, Canada         | Decision trees and neural networks                                          | Neural network models had the best performance |
| Kontokosta et al. (2018) | New York, US             | Gradient boosting regression trees and neural network                        | Weather variables were important features for all prediction models |
| Vu at al. (2019b)     | Austin, Taxas, US         | ANN time series                                                              | Lag times affected model performance |
| Cubillos (2020)       | Herning, Denmark          | Long Short-Term Memory (LSTM) Neural Network                                 | The model can combine predictions of different households |
| Wu et al. (2020)      | 26 cities in China        | ANN                                                                          | MSW prediction taking into account regional difference |
COVID-19 is the fluidity of the situation (Richter, Ng, Vu & Kabir, 2021b). Regulatory guidelines, consumer behaviors, social practices, and public acceptance are evolving with the spread of virus (Richter, Ng, Vu & Kabir, 2021a). Conventional ANN modeling uses weekly, monthly, and annual input data, making them less applicable for waste forecasting during COVID-19. We instead propose to use daily inputs in a Recurrent Neural Network (RNN) model, an ANN-based tool, for waste disposal rate modeling.

1.2. Study objectives, novelty, and contributions

The objectives of the present study are to (i) develop mixed waste disposal RNN time series models using lagged inputs in a Canadian capital city, (ii) construct waste disposal rate models using COVID-19 related variables to address waste disposal behaviors during the pandemic, (iii) examine the use of lagged inputs and distinct datasets on RNN time-series modeling. The effect of lag time on waste generation was first explicitly studied by Vu et al. (2019b), this work adds to the literature regarding the use of variable lag times on RNN time-series modeling. Conventional ANN-based time series models inherently assume an immediate causal effect between input variables and the output parameter. This assumption may not be realistic when considering that COVID-19 symptoms and subsequent patient treatment take weeks (Wong, Yuan, Haderlein, Jones & Washington, 2021; Zhu et al., 2021).

Unlike other ANN-based waste modeling studies unitizing weekly or monthly inputs, daily socio-economic and climatic input variables were used in the present work. Nabavi-Pelesaraei, Bayat, Hosseinzadeh-Bandbafha, Afrasyabi and Berrada (2017) successfully conducted a life cycle and energy flow assessments of municipal solid waste in Iran using daily waste data. Some socioeconomic variables, such as disposal rates in a city landfill, change markedly depends on the day of the week (weekday or weekend/holiday). The periodicity of these
socioeconomic inputs would introduce unnecessary bias to the modeling results. As such, the 7.5-year waste disposal dataset (January 2013 to September 2020) is divided into three distinct sets (weekday, weekend, and week-long) for model construction and result comparisons. Vu et al. (2021) modeled waste disposal rates during COVID by considering multiple waste streams separately and obtained promising results. It is hypothesized that the use of separated weekday and weekend time series may improve modeling accuracy. The proposed modeling approach (with the use of daily values and lagged inputs, together with distinct time-series) requires more work; however, they are more appropriate with respect to the study objectives. The use of lagged input variables and distinct time-series are original and they fill the knowledge gap in ANN-based modeling. Instead of estimated waste generation rates, recorded waste disposal rates at landfill were used in this study to provide more reliable modeling inputs.

2. Methodology

An original analytical approach for RNN waste disposal rate modeling during a pandemic is proposed. It is important to note that the proposed framework was developed for waste disposal rate estimation only, and changes in waste generation behaviors were not explicitly considered.

2.1. Study area and the COVID-19 pandemic

The capital city of Saskatchewan, Regina, was selected as the study area. The city is a typical mid-sized prairie city with a population of 236,500 (Statistics Canada, 2016). The average temperature in Regina was 3.4 °C during the 7.5-year study period (January 2013 to September 2020) (WU, 2020). The Regina landfill is the sole municipal landfill in the area and accepts wastes from nearby towns. Canadian landfill design embraces a wide variety of design principles as each jurisdiction have slightly different regulations and design guidelines (Richter, Ng & Falah, 2019). The Regina landfill has an active gas management system (Bruce, Ng & Richter, 2017, 2018) and a comprehensive groundwater monitoring program (Pan, Ng & Richter, 2019b, 2019a). The landfill is open 7 days/week in summer and 6 days/week in winter. The Regina landfill receives mixed solid waste, construction and demolition waste, asphalt, grit, and treated biomedical wastes (City of Regina, 2020). Mixed solid waste represents over 62% of the waste stream disposed of at the landfill by weight, and is considered in the present study. Mixed solid waste consists mostly wastes collected from residential, industrial, commercial and institutional sources. The average mixed waste disposal in Regina was 440.2 tonnes/day during the study period.

The first confirmed COVID-19 case was reported in Regina on March 12th, 2020 (Government of Saskatchewan, 2020a), and a provincial state of emergency was declared six days later on March 18th. During the initial lockdown phase, most bars and restaurants in the city were closed except for takeout service. According to Goddard (2020), online grocery shopping in Canada has increased significantly, and consumer stockpiling behavior was observed. Unemployment rates in Saskatchewan increased from 6.0% in January 2020 to 12.5% in May 2020 (Government of Saskatchewan, 2020c). The number of COVID-19 cases in Saskatchewan has stabilized, and the government has partially lifted restrictions according to the provincial “re-open Saskatchewan” plan. On May 29th, 2021, the total number of COVID-19 cases in the city was 11,636 (Government of Saskatchewan, 2020b).

The details of the methodology are shown in Fig. 1, as separately discussed in the following sub-sections from Sections 2.2 to 2.5.

2.2. Data collection and processing

A 7.5-year dataset of daily mixed waste disposal at Reinga landfill from January 1, 2013 to September 12, 2020 was collected, verified, and consolidated. Furthermore, three groups of input variables are considered, namely one socioeconomic variable (unemployment rates), three climatic variables (temperature, humidity, wind speed), and four COVID-19 related variables (number of new test positive cases, total COVID-19 cases, active cases, COVID-19 patients in hospital). Climatic variables were collected from Weather Underground (Weather Underground (WU) 2020), and COVID-19 variables were collected from official records published by the Government of Saskatchewan (2020b). Socioeconomic and climatic variables were widely applied to predict waste generation (Abbasi, Rastgoo & Nakisa, 2018; Johnson et al., 2017; Kannangara et al., 2018; Vu et al., 2019b), whereas the COVID-19 related variables were specifically selected in this study to capture the scale of the pandemic in Saskatchewan. Studies suggested that the pandemic may have affected waste disposal behavior in the city (Richter et al., 2021a; 2021b). Evidence from Spain suggests that people have been using more single use products and personal protection equipment during the COVID-19 pandemic, according to Kalina and Tilley (2020). Similar findings were observed in South Korea (Rhee, 2020).

Mixed waste disposal of extremely low value (less than one tonne/day) were removed. These are likely due to landfill scale calibrations or emergency disposals during landfill closure. Upper and lower boundaries derived by Interquartile Range (IQR) were used to eliminate the outliers Eqs. (1)-(3). Data points that were higher than the upper bound or below the lower bound were removed (Fallah, Ng & Vu & Torabi, 2020; Kannangara et al., 2018; Niu et al., 2021). If the computed lower bound is negative, then one tonne/day is taken as the lower bound.

\[
\text{IQR} = Q_3 - Q_1
\]  

(1)

\[
\text{Upperbound} = Q_3 + \text{IQR} \times 1.5
\]  

(2)

\[
\text{Lowerbound} = Q_1 - \text{IQR} \times 1.5
\]  

(3)

Where: \(Q_1=\)First Quartile of the data set; \(Q_3=\)Third Quartile of the data set.

2.3. Study periods, separated time series, and variable screening

Two periods for time series were specifically defined for modeling purposes: the 7.5-year entire study period (January 1st, 2013 to September 12th, 2020) and the 6-month COVID-19 period (March 18th, 2020 to September 12th, 2020). The periods were selected based on the availability of landfill disposal data. A provincial state of emergency was declared on March 18th, 2020 and was selected to define the starting date of the COVID-19 period in Regina. Richter et al. (2021) examined historical waste disposal rates and identified temporal variations in disposal behaviors at Regina. To overcome the periodicity of the waste disposal rates and to minimize unnecessary input fluctuations, three distinct time-series were considered: “week-long”, “weekday”, and “weekend”. The “week-long” set contains continuous daily disposal data with no distinction between workdays and weekends. The “weekday” set consists of truncated time-series from Mondays to Fridays in a given period, and the “weekend” set consists of truncated time-series from only Saturdays and Sundays in a given period.

Correlation analysis of the mixed waste disposal rates, as well as the identified 8 socioeconomic, climatic, COVID-19 related variables were conducted for both the entire study period and the COVID-19 period to examine the interrelationships between the variables. The input variables having the highest Pearson correlation coefficients with \( p < 0.05 \) during their respective periods were selected as the core input variables to build the models. On the other hand, highly intercorrelated input pairs with a large correlation coefficient were identified. Only one of the correlated inputs will be used for model building to avoid multicollinearity of the inputs. In the current study, the correlation coefficient cutoff of |0.95| was adopted, similar to Abdoli et al., 2011 and Vu et al., 2019b.
2.4. Development of mixed waste disposal prediction models

RNN, an ANN-LSTM for time series, was applied to develop different mixed waste disposal rate models. The models were created using Tensorflow 2.0, with modifications of an open-source code by Valkov (2019). Preliminary trials were used to select the ranges of number of hidden layers, lag time of input variables, as well as the number of neurons in the hidden layer, as further discussed in Sections 2.4.2 and 2.4.3. The dropout regularization technique was adopted to prevent overfitting of the models. The ratio of the training data set and testing data set was selected at 80:20. Similar ratio was also applied in Azadi and Karimi-Jashni (2016), and Kannangara et al. (2018). A total of 104 scenarios were developed and simulated, including 48 scenarios on effects of lagged inputs on model performance, 30 scenarios on waste disposal behaviors during the COVID-19 period; and 26 scenarios on the use of distinct data sets of RNN time-series modeling. There were three types of models developed: weekday, weekend, and week-long models (Fig. 1). The “weekday” model used the daily data on weekdays and the “weekend model” used the daily data on weekends. The “week-long” model used the continuous daily input data including both weekday and weekend data.

2.4.1. Effects of lags on input variables

A total of four weekday multi-variate models were developed for the entire 7.5-year study period (Fig. 1). The lag time of the models ranged from 1 - 12 days. One hidden layer with 128 neurons was used for the

---

**Fig. 2.** Time-series mixed solid waste disposal data in Regina over the entire study period (Jan 2013-Sep 2020) (a) Week-long; (b) Weekdays only; (c) Weekends only.
The accuracy of the multi-variate models was compared and their model structure studied with respect to lag time (Output 1, Fig. 1).

2.4.2. Effects of COVID-19 on waste disposal behaviors using weekday data only

In this part, the weekend set is ignored and only the more representative weekday set is considered. Two periods (the entire study period and the 6-month COVID-19 period) are considered for waste disposal prediction in the COVID-19 period, as separately discussed below.

### 2.4.2.1. Models using the entire study period data set

A total of five models were developed for waste disposal prediction using the entire dataset: a single feature model and four multi-variate models (Output 2a, Fig. 1). The single feature model (SingleWaste1) had four layers: an input layer, two hidden layers, and an output layer. The number of neurons of each hidden layer was selected at 600. On the
other hand, a single hidden layer with 128 neurons is adopted for the four multi-variate models. Both single-hidden layer and double hidden layers are commonly applied in ANN-based waste studies (Xu et al., 2021). The narrower range of lag times (1–6 days) are studied for both single and multi-variate models. The disposal rates obtained from the single feature model and the four multi-variate models were then compared with those of the models using the COVID-19 period data set (Output 2a vs. 2b, Fig. 1).

2.4.2.2. Models using the 6-month COVID-19 period data set

A total of four models were developed using the COVID-19 period data set: a single feature model and three multi-variate models (output 2b, Fig. 1). The single feature model (SingleWaste2) had four layers: an input layer, two hidden layers, and an output layer. The number of neurons of each hidden layer was selected at 600. The defined COVID-19 period was about 6 months, and a shorter range of lags (1–6 days) were studied. The structure of the three multi-variate models (number of hidden layers and neurons) using the COVID-19 period data was identical to that of the multi-variate models using the entire dataset (Fig. 1).

2.4.3. Effects of distinct time series on model accuracy using the optimized model

In this part, the model accuracy between the use of a continuous time-series and the two distinct time series covering the entire study period was examined. The analysis is conducted using the optimized models in the entire study period (output 1, Fig. 1). Three different RNN models were developed to simulate daily mixed waste disposal in the entire study period: week-long, weekday, and weekend models. The week-long model used the continuous daily input data including weekday and weekend data with a total of 2390 daily disposal data. The weekday model consists of 2002 data points, and the weekend model consists of 388 data points. The results from the optimized weekday model (output 3a, Fig. 1) and weekend model (output 3b, Fig. 1) are then stitched together to form a complete set, and then compared with the optimized week-long waste model (output 3c, Fig. 1). All models have one hidden layer with 128 neurons. The lag time of the weekday and week-long models ranged from 1 - 12 days, whereas the lag time of the weekend models varied from 1 to 2 weeks (Fig. 1).

2.5. Model performance

Mean absolute error (MAE), mean square error (MSE), mean absolute percentage error (MAPE), correlation coefficient (R), and index of agreement (IA) were adopted to assess model performances in this study. MAE, MSE, MAPE are common model error indicators used in waste studies (Abdoli et al., 2011; Abbasi & Hanandel, 2016; Kannangara et al., 2018; Vu et al., 2019b; Cubillos, 2020), and they are selected here to facilitate rapid comparison with literature. R and IA measure the goodness of fit between the predicted and actual values. Values close to unity suggest perfect agreement. Both R and IA are common indicators of ANN-based model performance (Adamović, Antanasijević, Ristić, Perić-Grujić & Pocajt, 2018; Radović, Antanasijević, Perić-Grujić, Ristić & Pocajt, 2018; Coskuner, Majeed, Jassim, Zontul & Karateke, 2020; Fallah et al., 2020). These indicators are computed using the following equations:

\[ \text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \]  
\[ \text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \hat{Y}_i| \]  
\[ \text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100 \]  
\[ R = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y}) \times (\hat{Y}_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n} (Y_i - \bar{Y})^2 \times \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}} \]  
\[ IA = 1 - \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} (Y_i - \bar{Y}_a)^2 + \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y}_a)^2} \]  

Where: \( n \): Number of data points  
\( Y_a \): Actual mass of mixed waste disposal  
\( \hat{Y}_a \): Simulated mass of mixed waste disposal  
\( \bar{Y} \): The mean actual mass of mixed waste disposal  
\( \bar{Y}_a \): The mean simulated mass of mixed waste disposal

3. Result and discussion

3.1. Waste disposal data characteristics and the use of distinct time series

Fig. 2 shows Regina’s historical waste disposal trends. Seasonal variations in waste disposal are clearly observed in Regina, with more mixed waste being disposal of in summers compared to winters (Fig. 2a). Regina residents produce more waste from gardening and outdoor activities during summers. Although less obvious, seasonal variations are also observed in the weekend data (Fig. 2c). Furthermore, when comparing Figs. 2b and 2c, different disposal behaviors between weekday and weekends in Regina are revealed. This observation supports the use of distinct time series as input for RNN waste disposal rate modeling. The use of distinct time series helps to reduce unnecessary biases and uncertainties in the modeling, as discussed further in Section 3.5. Data variability and trends are important in time series modeling, and details of the data sets are separately discussed below.

3.1.1. Weekday data set

The upper and lower bounds of the weekday data set were calculated at 815.17 tonnes/day and 63.21 tonnes/day, respectively, using Eqs. (2) and (3). Only a single outlier of 886.96 tonnes/day was eliminated from the weekday data set in the entire 7.5-year study period. There is no data point below the calculated lower bound. As shown in Table 2, the weekday data set after removing the outlier contained 2002 data points, with a mean value of 440.2 tonnes/day. The minimum value of 77.0 tonnes/day was observed on Good Friday (a Saskatchewan public holiday), March 30, 2018, when the landfill was partially closed. The maximum value of 811.7 tonnes/day was observed on Monday, July 8, 2019, probably due to the extra waste originating from road trips, camping activities, and other public festivities following the Canada day long weekend. Significant variations in mixed waste disposal are observed in the weekday set, with a standard deviation of 133.7 tonnes/day.

3.1.2. Weekend data set

The upper and lower bounds of the weekend data set were computed and a total of 10 outliers were identified (six data points larger than 142.92 tonnes/day, and four data points below 1 tonne/day) in the weekend data set during the entire study period. After removal of the outliers, the weekend set contained 388 data points. The average mixed waste disposal on weekends was 42.72 tonnes/day (Table 2). Compared to the weekday set, the average mass of mixed waste disposal of the
weekend set was about 10 times lower. The city does not provide curbside waste collection on weekends, however, the landfill remains open on weekends (Saturdays and Sundays during summer, and Saturdays only during the winter) to serve individual vehicles or private contractors. The coefficient of variation of the disposal rate is, however, the largest in the study at 0.642. It is found that the weekend set contributes little to the waste disposal trend in Regina but presents added variability. No obvious difference is observed on the climatic data between the weekday and weekend sets, as average temperature, humidity, and wind speed were all similar (Table 2).

### 3.1.3. COVID-19 data set

The COVID-19 data set is a weekday set with 128 data points, covering the period from March 18th, 2020 to September 12th, 2020 (Table 2). The average mass of mixed disposal waste in the COVID-19 period was 488.27 (tonnes/day), slightly higher than the weekday set during the entire period. However, the overall effect of COVID-19 on waste disposal is not definite, as the COVID-19 period occurred in spring and summer, a period with historically higher waste disposal in Regina (Fig. 2). A noticeable decrease in the variability of the waste disposal data is observed, with a standard deviation of 81.09 tonnes/day, and a higher min-to-max ratio of 0.29 (Table 2). The coefficient of variation of the disposal rate in the COVID-19 period is the lowest in the present study (0.166). It appears that COVID-19 improved the consistency of the disposal behaviors by the Regina residents, probably due to newly established social norms (avoid large in-person gatherings and social distancing) and the widely adopted work-form-home practices. The average unemployment rates increased from 5.60 to 5.65% to 9.99% during the COVID-19 period (Table 2).

### 3.2. Correlation analysis and input parameters selection

#### 3.2.1. Entire study period from Jan 2013 to Sep 2020

Table 3 shows the correlation coefficients among different variables and waste disposal in the entire 7.5-year study period. It is observed that temperature had the highest correlation (0.643) with the mass of mixed waste disposal, followed by humidity (0.297), wind speed (0.191), the total cases of COVID-19 (0.150), and total test (0.126). Humidity and wind speed are both negatively related to waste disposal. It appears that less waste is received and disposed of at the landfill under adverse weather conditions (wind speed higher than 60 km/hr, thunderstorm, etc.). A weak but significant positive correlation is observed between total cases and temperature (0.322) and humidity (0.114).

### Table 3

| Waste disposal (tonnes/day) | Unemployment rate (%) | Temperature (°C) | Humidity (%) | Wind Speed (kmph) | New test (capita) | Total case (capita) | Active case (capita) | Patient in hospital (capita) |
|----------------------------|-----------------------|------------------|--------------|-------------------|------------------|---------------------|---------------------|-----------------------------|
| Waste variable             | 1.000                 |                  |              |                   |                  |                     |                     |                             |
| Income variable            |                       |                  |              |                   |                  |                     |                     |                             |
| Unemployment rate (%)      | 0.016                 | 1.000            |              |                   |                  |                     |                     |                             |
| Climatic variables         | 0.277                 | 0.152            | 1.000        |                   |                  |                     |                     |                             |
| Temperature (°C)           | 0.184                 | 0.237            | 0.517        | 1.000             |                  |                     |                     |                             |
| Humidity (%)              | 0.044                 | 0.004            | 0.018        | 1.000             |                  |                     |                     |                             |
| Wind Speed (kmph)         | 0.002                 | 0.379            | 0.065        | 0.516             | 1.000            |                     |                     |                             |
| COVID-19 variables        | 0.007                 | 0.409            | 0.066        | 0.136             | 0.001            | 0.521               | 0.459               | 0.548                       |
| New test (capita)         | 0.001                 | 0.525            | 0.175        | 0.182             | 0.003            | 1.000               |                     |                             |
| Total case (capita)       | 0.055                 | 0.640            | 0.202        | 0.223             | 0.012            | 0.916               | 1.000               |                             |
| Active case (capita)      | 0.002                 | 0.379            | 0.058        | 0.115             | 0.003            | 0.659               | 0.561               | 1.000                       |
| Patient in hospital (capita) | 0.007                 | 0.409            | 0.066        | 0.136             | 0.001            | 0.521               | 0.459               | 0.548                       |

Note: Statistically significant values with $p < 0.05$ are **bolded.** Statistically insignificant coefficients are in *italic.*
Table 5
Summary of daily waste disposal rate models during the entire study period (Jan 2013 to Sep 2020) and the COVID-19 period (Mar 2020 to Sep 2020).

| Types of models | Explanation and application | Model ID |
|-----------------|------------------------------|----------|
| Entire period from Jan 2013 to Sep 2020 | The daily mass of waste disposal in the period from Jan 2013 to Sep 2020 was selected to build the single feature model. This model will be applied for calculating waste disposal in the COVID-19 period. | SingleWaste1 |
| (1) Waste disposal prediction model with the input containing a single feature of mass of waste disposal | | |
| (2) Waste disposal prediction model with the inputs containing temperature and humidity at the study area | Two parameters including daily temperature and humidity in the period from Jan 2013 to Sep 2020 were selected for the model as they have the highest and second highest correlation coefficient with the waste disposal rates. This model will be used to predict (i) week-long waste disposal; (ii) weekday waste disposal in the entire study period and the COVID-19 period, and (iii) weekend waste disposal in the entire study period. | Temp-Hum |
| (3) Waste disposal prediction model with the inputs containing temperature, humidity, and unemployment rate at the study area | This model was created using the core daily temperature and humidity variables in the entire study period. The socioeconomic parameter of unemployment rates was added in the model inputs. This model will be used to predict weekday waste disposal in the entire study period and the COVID-19 period. | Temp-Hum-Unemp |
| (4) Waste disposal prediction model with the inputs containing temperature, humidity, and total COVID-19 cases at the study area | This model was created using the core daily temperature and humidity variables. The total number of COVID-19 cases was added in the model inputs as it has the highest correlation coefficient with the waste disposal compared to others in the COVID-19 variables group (Table 3). This model will be used to predict weekday waste disposal in the entire study period and the COVID-19 period. | Temp-Hum-Total Case |
| (5) Waste disposal prediction model with the inputs containing temperature, humidity, wind speed, unemployment rates, new tests and total COVID-19 cases at the study area | This model was created using the variables having six highest correlation coefficients with waste disposal. They are temperature, humidity, wind speed, unemployment rates, total cases, and new test cases. This model will be used to predict weekday waste disposal in the entire COVID-19 period. | Temp-Hum-Wind-Total Case-New Test-Unemp |
| COVID-19 period from Mar 2020 to Sep 2020 | The daily mass of waste disposal in the period from Mar 2020 to Sep 2020 was selected to build the single feature model. | SingleWaste2 |
| (6) Waste disposal prediction model with the input containing a single feature of mass of waste disposal | | |
| (7) Waste disposal prediction model with the inputs containing temperature and new test cases at the study area | Two parameters including daily temperature and number of new test cases in the COVID-19 period from Mar 2020 to Sep 2020 were selected for the model as they have the highest and second highest correlation coefficient with the waste disposal in this period. This model will be used to predict weekday waste disposal in the COVID-19 period. | Temp-NewTest |
| (8) Waste disposal prediction model with the inputs containing temperature, new test cases, and unemployment rates at the study area | This model was created using the core daily temperature and number of new test cases. The socioeconomic parameter of unemployment rates was added in the model inputs. This model will be used to predict weekday waste disposal in the COVID-19 period. | Temp-NewTest-Unemp |
| (9) Waste disposal prediction model with the inputs containing temperature, new tests, active cases, total cases, number of COVID-19 patients in hospitals and unemployment rates at the study area | This model was created using variables having six highest correlation coefficients with waste disposal as the model inputs. They are temperature, new tests, active cases, total cases, number of COVID-19 patients in hospitals and unemployment rates. This model will be used to predict weekday waste disposal in the COVID-19 period in the study area. | Temp-NewTest-ActiveCase-Total Case-Patient-Unemp |

and the mass of mixed waste disposal in Regina. New test is highly correlated with total cases (coefficient = +0.916) but is less than the threshold value of 0.95. Unemployment rates shared positive correlations with all four COVID-19 variables but failed to show a statistically significant correlation with waste disposal rate.

Temperature and humidity were selected as the core input variables to create various mixed waste disposal rate models in Regina. In the entire study period, four different models were developed including Temp-Hum, Temp-Hum-Unemp, Temp-Hum-Total Case, and Temp-Hum-Wind-Total Case-New Test-Unemp. The Temp-Hum models were applied to predict week-long, weekday, weekend waste disposal in the entire study period, and weekday waste disposal during the COVID-19 period in the study area, whereas the rest of models were only used to predict weekday mixed waste disposal.
3.2.2. COVID-19 period from Mar 2020 to Sep 2020

Table 4 presents the correlation matrix of the variables during the 6-month COVID-19 period. During this shorter period, temperature remained the highest correlated parameter with waste disposal, with a correlation coefficient of +0.206. Among the four COVID-19 variables, only 'new test' has shown a statistically significant, but negative correlation (−0.166) with waste disposal. It appears that the use of correlation results alone may not be sufficient to fully explain the effects of COVID-19 on waste disposal behavior. It is not clear why the number of new tests were found negatively correlated with unemployment rate during COVID-19 period. It may be due to the unprecedented plunge in unemployment rate in the province due to the first wave. The lockdown and the fear of the virus, may have, in part, caused sales of food services and drinking places subsector in Saskatchewan to reduce by 32.4% between March and February 2020 (Statistics Canada, 2020). Sharma et al. (2020) also reported difficulties in estimating food waste during nationwide lockdowns. In general, the relationships between the nine variables appear weaker during the COVID-19 period (Table 4) than that of the entire study period (Table 3), probably due to the shorter period (representing about 19.3% of total data).

Four waste disposal rate models were built in the COVID-19 period, including a single feature model (SingleWaste 2) using mixed waste disposal as the sole input, and three multi-variate models. Temperature and number of new test cases had the highest and second highest correlation with mass of mixed waste disposal. They were used as the core variables to develop the multi-variate models in the COVID-19 period (Table 5). There were three multi-variate models developed in this period including Temp-New Test, Temp-New Test-Unemp, and Temp-New Test-New Case-Active Case-Patient-Unemp. Table 5 summarizes the models in both periods.

3.3. Effects of lag times on weekdays mixed waste modeling

Fig. 3 shows the stability of the multi-variate weekday models with different lagged inputs (1 to 12 days) using weekday data from the entire study period. In the training stage, it is clear that the performances of the Temp-Hum, Temp-Hum-Unemp, Temp-Hum-Total Case, and Temp-Hum-Wind-Total-Test-Unemp were relatively insensitive to lag times.
The average MAE ranged from 56.04 to 56.76 tonnes/day. No obvious difference in MAE is observed among models in the training stage. The average MAE of the four models increased slightly in the testing stage, ranging from 62.09 to 81.67 tonnes/day (Fig. 3b). Temp-Hum, the simplest model, has the slightest increase in average MAE in the testing stage, and is also less affected by uncertainties of inputs. On the contrary, the Temp-Hum-Total Case model was found very sensitive to lagged input parameters, with MAE ranging from 60.65 to 116.85 tonnes/day. This is in marked contrast with both Temp-Hum and Temp-Hum-Unemp models. A closer look at the tri-variate models Temp-Hum-Unemp and Temp-Hum-Total reveals that the selection of input variables is more important than model complexity. For example, the more complex hexa-variate model has the highest average MAE of 81.67 tonnes/day.

Fig. 4 further shows the effects of lagged inputs on the Temp-Hum model using different performance indicators. All models perform better in the training stage, and the model performance in the testing stage will be used to evaluate their relative performance. As shown in Fig. 4a, a lag time of 5 days and 10 days had similar MAPE (11.98% and 11.99%, respectively). A closer look at MSE, however, suggest that the model error is minimized at a lag time of 10 days at 6004 tonnes/day.

The Temp-Hum model generally simulates the waste disposal behaviors in Regina well. Irrespective of the lag times applied, $R$ was generally greater than 0.74 and $IA$ was greater than 0.83 in the testing stage (Fig. 4b). $R$ is also maximized (0.807) at a lag of 10 days. $IA$ values show similar result, with optimized lags of about 10–11 days ($IA = 0.866, 0.867$, respectively). By increasing the lag time from 1 day to the optimum of 10 day, the accuracy is improved and reduces the MSE by 21.8% (from 7020 to 5493 tonnes/day) in the training stage, and 25.3% (from 8033 to 6004 tonnes/day) in the testing stage (Fig. 4a).

### 3.4. Effects of COVID-19 on waste disposal behaviors

Table 6 compares the model performances at the optimized lag when using data from the entire study period and the COVID-19 period to simulate waste disposal rates during the COVID-19 period. In general,
the models applying COVID-19 period inputs had better performance than those of the models applying entire study period inputs. The average MSE of models applying entire study period inputs was 7770 tonnes/day. The average MSE of models applying COVID-19 period inputs was only about 3382 tonnes/day, or about 43.5% of its counterpart. The average MAPE of models applying COVID-19 period inputs was 10.1%, considerably lower than its counterpart at 14.5%. R and IA values of the models applying COVID-19 period inputs were also better (Table 6). The findings confirm that waste disposal behaviors during COVID-19 were changed, and the use of COVID-19 variables on waste modeling can better capture the spread and severity of the pandemic. The optimized lag times between the periods were however similar, with a slightly longer average lag for the COVID-19 period at 5.3 days.

In both periods, MSE was much higher for the hexa-variate models and the single feature models using historical disposal rate as sole input. On the other hand, the uni-variate and tri-variate models appear to simulate waste disposal rates better. The Temp-Hum model and the Temp-NewTest-Unemp model have the lowest MSE for the entire period and COVID-19 period, respectively. The lag times for both models were 5 days. Among all models in both periods, Temp-Hum-Wind-Total Case-New Test-Unemp has the highest MSE (15,573 tonnes/day), probably due to the high correlations between New Test and Total Case (Table 3).

Fig. 5 compares the actual disposal data during the COVID-19 Pandemic (Mar 2020-Sep 2020) with the simulated results from the optimized Temp-New model (using COVID-19 period inputs) and the Temp-Hum model (using data from the entire study period). Despite of the simplicity of the models, both were able to capture the average disposal rate in Regina during the COVID-19 period. Temp-New Test model (solid line, Fig. 5) generally predicts a higher disposal rate than Temp-Hum. Consistency between the models was highest when there

### Table 6

Comparison of optimized model performance using different inputs and study periods.

| Model                          | Optimal lag time (days) | MSE   | MAPE  | R    | IA    |
|-------------------------------|-------------------------|-------|-------|------|-------|
| **Applying weekday data inputs** |                         |       |       |      |       |
| **inputs in the entire period** |                         |       |       |      |       |
| Single feature model          | 5                       | 11,326| 17.28 | 0.49 | 0.73  |
| Multi-variate models          |                         |       |       |      |       |
| Temp-Hum                      | 5                       | 3581  | 10.30 | 0.69 | 0.85  |
| Temp-Hum-Unemp                | 2                       | 3741  | 10.63 | 0.65 | 0.82  |
| Temp-Hum-Total Case           | 5                       | 4629  | 12.34 | 0.59 | 0.72  |
| Temp-Hum-Wind-Total Case-New Test-Unemp | 6               | 15,573| 21.98 | 0.53 | 0.52  |
| Average                       | 4.6                     | 7770  | 14.50 | 0.59 | 0.73  |
| **Applying weeklong data inputs COVID-19 period** |                         |       |       |      |       |
| Single feature model          | 6                       | 4168  | 11.48 | 0.55 | 0.63  |
| Multi-variate models          |                         |       |       |      |       |
| Temp-NewTest                  | 5                       | 3078  | 9.37  | 0.69 | 0.80  |
| Temp-NewTest-Unemp            | 5                       | 2988  | 9.42  | 0.70 | 0.81  |
| Temp-NewTest-New Case-Patient-Active Case-Patient-Unemp | 5               | 2929  | 10.12 | 0.68 | 0.81  |
| Average                       | 5.3                     | 3382  | 10.10 | 0.66 | 0.76  |
was more scattering in the actual disposal data (from April 29th to Aug 5th). Although the models were able to capture the general disposal trend well, none of the optimized models could identify precisely the peaks and troughs on daily basis.

3.5. The use of distinct time-series on RNN model accuracy

Three Temp-Hum models were developed and optimized using weekday, weekend, and week-long waste disposal data. The results of weekdays and weekends were stitched together to compared with the optimized week-long model. Fig. 6 shows the model accuracy of the optimized Temp-Hum models. The performance of both models is acceptable, with MAE generally less than 150 tonnes/day. Fig. 6a shows that the use of separated sets of time series data (blue circles, Fig. 6a) has resulted in approximately 1.45 times lower MAE than that of the week-long model at the training stage. The relatively high MAE shown in Fig. 6a occurred mostly on weekends in the winter, especially when the mass of mixed waste disposal dropped to lower than 15 tonnes/day (Fig. 2c).

Fig. 6b shows the performances of the models in the testing stage. It can be seen that the Temp-Hum model with separated time series had approximately 2.13 times lower MAE than that of the model using week-long data. The advantage of using separated time series for waste disposal is more apparent during the COVID-19 period, when the MAE is significantly lower (typically under 100 tonnes/day). It is probably due to the lower data variability, together with the reduction of unnecessary bias between workdays and weekends. The finding supports the use of separated data sets for time series modeling if strong data periodicity is observed.

Fig. 6. Performance of Temp-Hum model using continuous input set (week-long data) and separated sets (weekday and weekend): a) Training stage; b) Testing stage.
4.0 CONCLUSION
In this study, an original RNN time series modeling framework is developed to simulate waste disposal rate in Regina, the capital city of Saskatchewan. The objectives are to develop a RNN modeling framework using lagged climatic, socio-economical, and COVID-19 related inputs and to examine the use of distinct time-series on RNN modeling. The disposal behaviors of the residents during the pandemic are also simulated and compared with the historical data. Unlike other studies, this work explicitly examines the effects of the use of daily values, lagged inputs, and distinct time-series on RNN modeling. A total of 104 scenarios were considered. The proposed framework focuses on waste disposal rate modeling, and no attempt is made to include changes in waste generation behaviors.

Historical disposal trends reveal that disposal behaviors are different between weekdays and weekends. As such, distinct time series were used to reduce unnecessary biases and uncertainties in the waste disposal rate modeling. A noticeable decrease in variability of the waste disposal data is observed in the COVID-19 period, with a coefficient of variation of 0.166. It appears that COVID-19 improved the consistency of disposal behaviors.

Correlation results suggest that temperature and humidity are correlated with the mass of mixed waste disposal during the entire period. During COVID-19, however, temperature and new test cases correlated with the mass of mixed waste disposal during the entire period. It appears that COVID-19 improved the consistency of waste disposal data is observed in the COVID-19 period, with a coefficient of variation of 0.166. It appears that COVID-19 improved the consistency of disposal behaviors.

In general, models applying COVID-19 period inputs had better performance than those of the models applying entire study period inputs. The optimized lag times are however similar between the periods, with slightly longer average lag for the COVID-19 period at 5.3 days. Simpler models appear to simulate waste disposal rates more accurately.

Both Temp-Hum (entire period) and Temp-New Test (COVID-19 period) were able to capture the average disposal rate in Regina during COVID-19, however, both optimized models fail to fully capture the peaks and troughs of the waste disposal data. The benefits of the use of separated time series are more apparent during the COVID-19 period, when the MAE is significantly lower. This could be due to the lower waste disposal data variability, and the reduction of unnecessary bias.

Declaration of Competing Interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments
The research reported in this article was partly supported by a grant from the Natural Sciences and Engineering Research Council of Canada (ALLRP 551383-20). We would also like to thank City of Regina Environmental Services branch for supporting this project. The authors are grateful for their support. The views expressed herein are those of the writers and not necessarily those of our research and funding partners.

References
Abbaszadeh, M., & Hananadeh, A. E. (2016). Forecasting municipal solid waste generation using artificial intelligence modelling approaches. Waste Management, 56, 13–22. https://doi.org/10.1016/j.wasman.2016.05.018,

Abdoli, M. A., Nazad, M. F., Sede, R. S., & Behboudian, S. (2012). Longterm forecasting of solid waste generation by the artificial neural networks. Environmental Progress & Sustainable Energy, 31, 628–636. https://doi.org/10.1002/ep.10591,

Abdul, M. A., Falahnejhad, M., & Falahnejhad, S. (2011). Multivariate econometric approach for solid waste generation modeling: Impact of climate factors. Environmental Engineering Science, 28(9), 627–633. https://doi.org/10.1089/ees.2010.0234,

Adamovic, V. M., Antanasijevic, D. Z., Ristic, M. D., Perić-Grujić, A. A., & Pocajuć, V. V. (2018). An optimized artificial neural network model for the prediction of rate of hazardous chemical and healthcare waste generation at the national level. Journal of Analytical Chemistry and Waste Management, 20, 1736–1750. https://doi.org/10.1007/s10661-018-0741-6,

Alam, O., & Qiao, X. (2020). An in-depth review on municipal solid waste management, treatment and disposal in Bangladesh. Sustainable Cities and Society, 52, Article 101775. https://doi.org/10.1016/j.scs.2019.101775,

Antanasijević, D., Pocajuć, V., Popovíc, I., Redzic, N., & Ristic, M. (2013). The forecasting of municipal waste generation using artificial neural networks and sustainability indicators. Sustainable Cities, 8, 37–46. https://doi.org/10.1016/j.scs.2012.01-0161-09,

Abdoli, M. A., Falahnejhad, M., & Falahnejhad, S. (2011). Multivariate econometric approach for solid waste generation modeling: Impact of climate factors. Environmental Engineering Science, 28(9), 627–633. https://doi.org/10.1089/ees.2010.0234,
