ADVANCED NUMERICAL MODELING TECHNIQUES
FOR MODERN WASTE MANAGEMENT SYSTEMS

A Thesis
Submitted to the Faculty of Graduate Studies and Research
In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

in
Environmental Systems Engineering
University of Regina

By
Hoang Lan Vu
Regina, Saskatchewan
January, 2019

Copyright 2019: H.L. Vu
Hoang Lan Vu, candidate for the degree of Doctor of Philosophy in Environmental Systems Engineering, has presented a thesis titled, *Advanced Numerical Modeling Techniques for Modern Waste Management Systems*, in an oral examination held on December 12, 2018. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

External Examiner: *Dr. Zhong Li, McMaster University*

Supervisor: Dr. Tsun Wai Kelvin Ng, Environmental Systems Engineering

Committee Member: Dr. Joseph Piwowar, Geography and Environmental Studies

Committee Member: Dr. Yee-Chung Jin, Environmental Systems Engineering

Committee Member: Dr. Amornvadee Veawab, Environmental Systems Engineering

Chair of Defense: Dr. Karen Meagher, Department of Mathematics & Statistics

*via Zoom conferencing

**Not present at defense
Abstract

This thesis presents original results to the advancement of numerical modelling of a modern waste management system with respect to generation, collection, and disposal of non-hazardous solid waste. The first part of the thesis specifically look at lag times relating to variables that attempt to predict municipal yard waste generation using machine learning approaches. Weekly averaged climatic and socio-economic variables are screened through correlation analysis and the significant variables are then used to develop yard waste models. These models then utilize artificial neural networks where the variables are time lagged a different number of weeks. Optimal lag times for each model varied from 1-11 weeks. The best model used both the ambient air temperature and population variables, in a model with 3 layers, 11 neurons in the hidden layer, and an optimal lag time of 1 week. A mean absolute percentage error of 18.72% was obtained at testing stage. One model saw a 55.4% decrease in the mean squared error at training, showing the value of lag time on the accuracy of weekly yard waste prediction models.

The second part of the thesis focuses on geospatial modelling of a dual phase waste collection. A model integrating the handcart pre-collection phase and truck collection phase was proposed. Temporary collection points were first identified using both the maximize coverage and minimize facility location-allocation tools from a list of candidate temporary collection points and constraints. A total of 30 scenarios were considered in order to investigate the interrelationships between the model parameters, with respect to the total operation costs and maintenance system costs. The scenario with 11 temporary collection points and a maximum handcart collection distance of 500 m gave the lowest overall cost in the study area. The results suggest a single temporary
collection point in the study is able to serve about 2,590 people in an area of 0.11 km². It is found that the number and distribution of temporary collection points greatly affected the cost effectiveness in both pre-collection and collection phases. In the third part of thesis, landfill gas data was collected at semi-arid landfills, and curve fitting was carried out to find optimal k and L₀ or DOC values using LandGEM, Afvalzorg Simple, and IPCC first order decay models. Model parameters at each landfill were estimated and compared using default values. Methane generation rates were substantially overestimated using default values (with percentage errors from 55 to 135%). The mean percentage errors for the optimized k and L₀ or DOC values ranged from 11.60% to 19.93% at the Regina landfill, and 1.65% to 10.83% at the Saskatoon landfill. Finally, the effect of different iterative methods on the curve fitting process was examined. The residual sum of squares for each model and iterative approaches were similar, with the exception of iterative method 1 for the IPCC model. The default values in these models fail to represent landfills located in cold semi-arid climates. The fourth part of the thesis focuses on the development of a systematic approach for modelling of WMS. ANN time series was first applied to forecast the amounts of recyclables and garbage in the year 2023 at the target study area. MAPE of 10.92% to 16.51% were obtained for the forecast. Both the amount of recyclables and garbage appeared to decrease with time. Truck travel distance of the optimized routes were found sensitive to the composition and density of the materials. The use of dual-compartment trucks reduced total travel distances by 10.30% to 16.00%. However, single-stream trucks were likely to be more efficient in terms of total collection time.
Acknowledgement

I would like to express my sincere gratitude to my supervisor Dr. Kelvin Ng for his excellent guidance, patient, encouragement and support during my PhD study. I would also like to thank the comprehensive exam and thesis defense committees for providing comments and suggestions to improve my doctoral thesis. I am also grateful to the research group (past and present) for their comments and suggestions to improve my research work. I would like to thank and acknowledge the VIED 911 fellowship, the Faculty of Graduate Studies and Research, the Faculty of Engineering and Applied Science, the Natural Science and Engineering Research Council of Canada for their financial supports. I would also like to thank the City of Regina, the URENCO Hai Phong for providing the data for my studies. Finally, I would like to thank my parents, family and friends for their support and encouragement.
# Table of Contents

Abstract .................................................................................................................................................. i
Acknowledgement ............................................................................................................................... iii
Table of Contents ................................................................................................................................... iv
List of Tables ........................................................................................................................................ viii
List of Figures ...................................................................................................................................... x
List of Abbreviations and Symbols ..................................................................................................... xii

1. Introduction ....................................................................................................................................... 1
   1.1 Yard Waste Generation Model ....................................................................................................... 2
      1.1.1 Machine learning approaches ................................................................................................. 3
      1.1.2 Effects of lag on ANN model ................................................................................................. 4
      1.1.3 Yard waste generation ............................................................................................................... 5
   1.2 Municipal Solid Waste Collection Model ..................................................................................... 6
      1.2.1 The first phase (pre-collection phase) ................................................................................. 8
      1.2.2 The second phase (collection phase) ...................................................................................... 9
      1.2.3 The first and the second phases (pre-collection and collection phases) ......................... 10
   1.3 Municipal Landfill Gas Model .................................................................................................... 11
      1.3.1 First-order-decay gas generation ......................................................................................... 12
      1.3.2 Accuracy and reliability of the gas model ........................................................................... 12
      1.3.3 The parameters $k$ and $L_0$ ............................................................................................. 13
   1.4 Waste Generation and Collection Route Design ...................................................................... 16
      1.4.1 Residential waste generation rate and composition ......................................................... 17
      1.4.2 Single and multiple stream residential waste collection ............................................... 19

2. Time-lagged effects on ANN Municipal Yard Waste Prediction ..................................................... 21
   2.1 Objectives of Chapter Two .......................................................................................................... 21
   2.2 Study area and background ........................................................................................................ 21
   2.3 Methodology ............................................................................................................................... 24
      2.3.1 Data collection and analysis ............................................................................................... 26
         2.3.1.1 Waste data sets .......................................................................................................... 26
2.3.1.2 Climatic parameters ................................................................. 27
2.3.1.3 Socio-economic parameters ....................................................... 27
2.3.1.4 Data variability and correlation analysis ...................................... 28
2.3.2 Lag time and ANN model development and optimization ............... 28
2.3.3 Models performance assessment .................................................. 30
2.4 Results and discussion ................................................................... 31
2.4.1 Data analysis and screening .......................................................... 31
2.4.2 Optimal lag times and data overfitting ......................................... 41
2.4.3 Performance of models at the testing stage ................................... 41
2.4.4 The optimal weekly yard waste ANN prediction models ............... 46
3. Parameter Interrelationships in a Dual Phase Collection Model ......... 52
3.1 Objectives of Chapter Three ............................................................ 52
3.2 Study area and existing waste collection system .............................. 53
3.3 Methodology .................................................................................. 56
3.3.1 Geodatabase development and field data ...................................... 58
3.3.2 Location-allocation of TCPs .......................................................... 59
3.3.3 Handcart route optimization .......................................................... 61
3.3.4 Truck route problem and constraints .......................................... 62
3.3.5 Comparison among different scenarios using cost estimates ........... 66
3.4 Results and discussion ................................................................... 68
3.4.1 Waste temporary collection points for different scenarios ............. 68
3.4.2 VRP Solution for handcarts ............................................................ 72
3.4.3 VRP Solution for the trucks ........................................................... 75
3.3.4 Effects of the number and spatial distribution of TCPs on handcart and truck costs ........................................................................... 78
3.4.5 Total capital, operating, and maintenance costs ............................ 80
4. Model Optimization for Landfills located in Cold Semi-arid Climates .. 85
4.1 Objectives of Chapter Four .............................................................. 85
4.2 Methodology .................................................................................. 85
4.2.1 The landfills and the waste records ............................................. 85
4.2.2 FOD models and default model parameters ................................ 87
4.2.3 Iterative methods and curve fitting by minimizing RSS values .............. 89
4.3 Results and discussion .............................................................................. 90
  4.3.1 Methane estimates using default k and L₀ (DOC) values ....................... 90
  4.3.2 Iterative approaches .............................................................................. 94
  4.3.3 Optimal FOD gas model parameters .................................................... 100
  4.3.4 Comparison of the selected optimal k and L₀/DOC values and literature .... 103

5. Effects of Waste Characteristics on Optimal Collection Routes ............... 107
  5.1 Objectives of Chapter Five ..................................................................... 107
  5.2 Selection of target area and waste collection routes .............................. 107
  5.3 Methodology .......................................................................................... 110
    5.3.1 Data collection and analysis .............................................................. 112
      5.3.1.1 Waste data sets ........................................................................ 112
      5.3.1.2 Road network data sets, target areas and number of stops .......... 114
    5.3.2 ANN waste generation forecast .......................................................... 114
    5.3.3 Models performance assessment ........................................................ 115
    5.3.4 Waste characteristics forecast ............................................................ 116
    5.3.5 GIS - Network Analysis - VRP models .............................................. 119
      5.3.5.1 Model development .................................................................. 119
      5.3.5.2 Model result comparison ............................................................ 122
  5.4 Results and discussion .......................................................................... 123
    5.4.1 Future waste forecast and performance of ANN models at the testing stage 123
    5.4.3 GIS - Network Analysis - VRP solutions ........................................... 126
      5.4.3.1 Effects of waste mass and waste composition on optimized GIS truck travel distances ......................................................... 126
      5.4.3.2 Comparison of total truck travel distances and collection time ...... 131
      5.4.3.3 Air emissions from waste collection trucks at the study area ......... 134

6. Conclusion from the studies .................................................................... 137
  6.1 Effects of time lag on ANN Yard Waste Modeling ................................. 137
  6.2 Dual phase Municipal Solid Waste Collection ....................................... 138
  6.3 Landfill Gas Model Optimization ............................................................. 139
  6.4 Waste Characteristics and Collection Route Optimization ...................... 141
References............................................................................................................................................. 143
Appendices............................................................................................................................................. 162
A1 – Chapter Two .................................................................................................................................. 162
A2 – Chapter Three ................................................................................................................................. 168
A3 – Chapter Four ................................................................................................................................. 172
A4 – Chapter Five ................................................................................................................................. 174
List of Tables

Table 1-1: First order decay (FOD) landfill gas model parameters from landfills with different climatic conditions (Vu et al., 2017) ........................................................................................................ 15
Table 2-1: Weekly variable statistics (October 2004 - April 2018) (Vu et al., 2019) .......................... 33
Table 2-2: Correlations of the input variables and weekly municipal yard waste from October 2004 to April 2018 (without the effect of lag time) (Vu et al., 2019) .................. 35
Table 2-3: The proposed weekly municipal yard waste prediction models (Vu et al., 2019) ........................................................................................................................................ 40
Table 2-4: Comparison of each model’s performance at their optimal lag time (Vu et al., 2019) ........................................................................................................................................ 43
Table 2-5: Performance of the optimized yard waste models (Vu et al., 2019) .......................... 48
Table 3-1 Model parameters and constrains (Vu et al., 2018b) ........................................................ 65
Table 3-2: Results of VRP for trucks and handcarts (Vu et al., 2018b) ........................................ 74
Table 3-3: Pickup waste volume balance of the five truck routes in selected scenarios
(D_{max}= 400 m, TCP= 11, 13, and 15) (Vu et al., 2018b) ............................................................. 77
Table 3-4: Truck travel distances compared between the optimized case (D_{max}= 500, TCP= 11) and the status quo (Vu et al., 2018b) ................................................................. 82
Table 4-1: Measured methane generation data and estimates from FOD models using default parameters (Vu et al., 2017) ........................................................................................................ 93
Table 4-2: Optimal FOD landfill gas model parameters using different iterative approaches at Regina (Vu et al., 2017) ........................................................................................................ 95
Table 4-3: Mean percentage error between estimated rates using optimal parameters and actual gas rates (Vu et al., 2017) ........................................................................................................ 102
| Table 5-1: Waste generation in the study area | 113 |
| Table 5-2: Waste composition in the study area | 118 |
| Table 5-3: Major inputs to the VRP solution | 121 |
### List of Figures

Figure 2-1: Location of the study area (After National Geographic et al., 2011, WU, 2018, and City of Austin, 2018c, Vu et al., 2019) ................................................................. 23

Figure 2-2: Methodology flow chart for Chapter two (Vu et al., 2019) ...................... 25

Figure 2-3: Correlation coefficient between the input variables and the yard waste variable at different lag times (Vu et al., 2019) ................................................................. 38

Figure 2-4: Example of the two single variable model performances at different lag times: a) using R Value; b) using MSE (Vu et al., 2019) ................................................................. 45

Figure 2-5: A comparison of the two superior model’s (Temp and Temp-Pop) performance using MSE with a different number of neurons (Vu et al., 2019) ............. 47

Figure 2-6: Performance of the optimized Temp-Pop ANN model (a) Training stage from weeks 10/1/2006 to 9/30/2007; (b) Testing stage from weeks 4/2/2017 to 4/01/2018 (Vu et al., 2019) ........................................................................................................................................ 50

Figure 3-1: Map of the study area (After OpenStreetMap) ...................................... 55

Figure 3-2: Diagram of Location-Allocation and VRP solutions using GIS (Vu et al., 2018b) ........................................................................................................................................ 57

Figure 3-3: Results of Location-Allocation: a) candidate temporary collection points; b) minimize facility function with $D_{\text{max}}$ of 300 m; c) maximize coverage function with the number of temporary collection points equal to 7 and $D_{\text{max}}$ of 300 m, d) 400 m, and e) 500 m; f) maximize coverage with number of temporary collection points of 25 and $D_{\text{max}}$ of 300 m (Vu et al., 2018b) ........................................................................................................................................ 69

Figure 3-4: Percentage of coverage with respect to the number of sources and the maximum distances to be covered by handcarts (Vu et al., 2018b) .................... 71
Figure 3-5: Comparison of (a) handcart and truck collection costs; (b) system collection cost among different scenarios (Vu et al., 2018b) ................................................................. 79

Figure 3-6: Comparison of the existing truck routes (a, c, e, g, i) and the proposed truck routes (b, d, f, h, j) for the optimized scenario (D_{max} = 500, TCP=11) (Vu et al., 2018b). 83

Figure 4-1: Comparisons of actual and estimated gas data using optimal model parameters at Regina landfill (Vu et al., 2017) .................................................................................................................. 97

Figure 4-2: Effects of iterative methods on landfill gas generation characterises using (a) LandGEM, (b) Afvalzorg, and (c) IPCC model (Vu et al., 2017) ................................................. 99

Figure 4-3: Reported (a) k and (b) L_0 values in Saskatchewan and other studies (Vu et al., 2017) ........................................................................................................................................ 105

Figure 5-1: Study area (After City of Austin, 2018d, 2018e) ................................................. 109

Figure 5-2: Methodology flow chart of Chapter five ............................................................. 111

Figure 5-3: Results of ANN: a) MAPE of recyclables; b) MAPE of garbage; c) Future recyclables generation rate; d) Future garbage generation rate ............................................. 125

Figure 5-4: Travel distances: a) of recyclables trucks; b) of garbage trucks; c) of dual-compartment trucks in four sub-areas ........................................................................ 128

Figure 5-5: Example of recycling material truck routes in 2023: a, c, e) Routes using waste composition in 2014; b, d, f, g) Routes using waste composition in 2023 (Maps used were from © OpenStreetMap contributors, 2018) ................................................................. 130

Figure 5-6: Comparison among different scenarios in term of a) Total truck travel distance; b) Total collection time ................................................................................................. 133

Figure 5-7: CO_2, and NO_x emissions from the truck in different scenarios ....................... 135
**List of Abbreviations and Symbols**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>Artificial neural networks</td>
</tr>
<tr>
<td>CAA</td>
<td>Conventional and Arid Area</td>
</tr>
<tr>
<td>CP</td>
<td>Collection Point</td>
</tr>
<tr>
<td>$D_{\text{max}}$</td>
<td>Maximum Distance from a source to a collection point</td>
</tr>
<tr>
<td>DOC</td>
<td>Degradable Organic Carbon</td>
</tr>
<tr>
<td>$\text{DOC}_f$</td>
<td>Fraction of Degradable Organic Carbon</td>
</tr>
<tr>
<td>F</td>
<td>Fraction of Methane</td>
</tr>
<tr>
<td>FOD</td>
<td>First Order Decay</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross Domestic Product</td>
</tr>
<tr>
<td>GHG</td>
<td>Green House Gas</td>
</tr>
<tr>
<td>GIS</td>
<td>Geography Information System</td>
</tr>
<tr>
<td>IPCC</td>
<td>Intergovernmental Panel on Climate Change</td>
</tr>
<tr>
<td>HH</td>
<td>Household</td>
</tr>
<tr>
<td>k</td>
<td>First Decay Constant</td>
</tr>
<tr>
<td>$L_0$</td>
<td>Methane Generation Potential</td>
</tr>
<tr>
<td>LFG</td>
<td>Landfill Gas</td>
</tr>
<tr>
<td>MRF</td>
<td>Material Recovery Facility</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MAPE</td>
<td>Mean Absolute Percentage Error</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>MCF</td>
<td>Methane Correction Factor</td>
</tr>
<tr>
<td>MSW</td>
<td>Municipal Solid Waste</td>
</tr>
<tr>
<td>NAR</td>
<td>Nonlinear Autoregressive</td>
</tr>
<tr>
<td>NARX</td>
<td>Nonlinear Autoregressive with External Input</td>
</tr>
<tr>
<td>O&amp;M</td>
<td>Operating and Maintenance</td>
</tr>
<tr>
<td>RSS</td>
<td>Residual Sum of Squares</td>
</tr>
<tr>
<td>SCADA</td>
<td>Supervisory Control and Data Acquisition</td>
</tr>
<tr>
<td>TPC</td>
<td>Temporary Collection Point</td>
</tr>
<tr>
<td>VRP</td>
<td>Vehicle Routing Problem</td>
</tr>
<tr>
<td>WMS</td>
<td>Waste Management System</td>
</tr>
</tbody>
</table>
1. Introduction

Solid waste generation has rapidly increased in the past decades, probably due to population bloom. In 2016, there was about 2.01 billion tonnes waste generated over the globe or about 0.74 kg/capita·day (World Bank, 2018). Solid waste has negatively impacted on the environment and human receptors as it can cause air, water, as well as soil pollutions (Alam and Ahmade, 2013). An efficient and effective waste management system helps to mitigate the negative impacts on the environment, prevent diseases, reduce raw materials, and save energy.

The key components of a modern waste management system (WMS) include waste generation, on-site storage, collection, transport, and permanent disposal. Operation of a WMS is expensive, and ways to improve the effectiveness and efficiency of the system is of great interest to the engineers and regulators. Advanced numerical techniques and tools help us to optimize WMS and to propose alternative solution using a fraction of time and cost compared to field study. Original numerical techniques and modelling approaches are developed and reported in this thesis. The present work contributes to the development of a sustainable WMS by improving the modelling techniques with respect to three specific functional elements: generation, collection and disposal.

Waste generation is important as it affects all of the functional elements of a WMS. Inaccurate prediction of generation rates may lead to overestimate or underestimate of capacity of bins, containers, collection trucks, transfer stations, recovery facilities, and landfills. Waste collection is also important in a modern WMS as the expenditure on waste collection and transfer typically accounted for 50% - 85% of the total budget (Dogan and Suleyman, 2003; Ghose et al., 2006; Sanjeevi and Shahabudeen, 2016;
Tavares et al., 2009). Land disposal or landfilling is the ultimate step in all modern WMSs. Residuals from incineration plants, composting centers, material recovery facilities are sent to landfills for permanent disposal. This thesis specifically addresses these three key components of WMS in order to better understand our modelling techniques and to improve the efficiency and effectiveness of our WMS.

The thesis includes six chapters. Chapter one introduces the background information, and provides a comprehensive review of the WMS literature (waste generation, collection, and disposal). Chapter two presents the development of a machine learning generation model for yard wastes. Chapter three discusses the advances of a dual-phase waste collection system using geospatial analysis. Chapter four compares and optimizes first-order landfill gas generation models for waste disposal sites. Chapter five combines machine learning generation model and geospatial analytical techniques to establish an original WMS framework. Chapter six summarizes the key findings of the four studies.

1.1 Yard Waste Generation Model

Municipal solid waste generation forecasting plays an important role in a sustainable and effective WMS (Intharathirat et al., 2015; Ghinea et al., 2016; Abbasi and Hanandeh, 2016; Azadi and Karimi-Jashni, 2016; Kumar and Sadmadder, 2017). If generation forecasts are imprecise, the improper design or inefficient operation of waste management facilities and systems may occur (Buenrostro et al., 2001; Intharathirat et al., 2015; Abbasi and Hanandeh, 2016). These errors can be minimized by using accurate and theoretically sound waste generation prediction models.
In recent decades, waste forecasting models using different approaches and techniques have been proposed and developed, including trend analysis and various regression techniques (Bridgwater, 1986; Chang and Lin, 1997; Daskalopoulos et al., 1998; Rimaiitye et al., 2012; Oribe-Garcia et al., 2015; Wang et al. 2016; Richter et al., 2017; Chowdhury et al., 2017), multiple linear regression (Abdoli et al., 2011; Azadi and Karimi-Jashni, 2016; Kumar and Samadder, 2017), and alternative methods such as moving average curves (Matsuto and Tanaka, 1993). Recently, artificial intelligent and machine learning methods have been widely applied in the field of solid waste management (Noori et al., 2010; Shahabi et al., 2012; Antanasijevic et al., 2013; Younes et al., 2015; Azadi and Karimi-Jashni, 2016; Abbasi and Hanandeh, 2016; Younes et al., 2016; Kannangara et al., 2018).

1.1.1 Machine learning approaches

Among these artificial intelligent and machine learning methods, the artificial neural network (ANN) method has been successfully applied in the forecasting of waste generation (Kolekar et al., 2016; Goel et al., 2017). A time series ANN model was used by Noori et al. (2010) to attempt the prediction of weekly solid waste generation rates in Iran. Principal component analysis and gamma test techniques were also applied to reduce the complexity of input variables and to improve the model’s performance (Noori et al., 2010). However, the effects of optimal lag time on the model’s performance was not considered. Another time series ANN model was also adapted by Abbasi and Hanandeh (2016) to forecast monthly waste generation in the Logan City council region in Australia. A total of 18-years of historical waste data was considered, and they
concluded that artificial intelligence approaches are promising and have good prediction performance (Abbasi and Hanandeh, 2016).

Antanasijevic et al. (2013) applied both back-propagation ANN and general regression neural network methods using various sustainability indicators (gross domestic product, domestic material consumption and resource productivity) to forecast annual waste generation in 26 European countries and found that the general regression neural network model performs better than the conventional back-propagation model. Azadi and Karimi-Jashni (2016) used feed forward ANN models with population, waste collection frequency, monthly temperature, altitude, and historical generation rates to predict seasonal waste generation for 20 cities in Iran, and noted that ANN models generally out-perform the multiple linear regression models in the projection of mean seasonal municipal waste generation rates. Feed forward ANN models were also utilized by Kannangara et al. (2018) to predict annual regional municipal solid waste generation and diversion rates in Canada. Residential waste quantities and a wide range of socio-economic factors such as average earnings and income, education levels, employment data, types of industries and occupations, dwelling and household characteristics, workplace and demographic parameters were incorporated in their models (Kannangara et al., 2018).

1.1.2 Effects of lag on ANN model

Time lag of the ANN was commonly applied in other fields such as meteorology, hydrology. For example, Silverman and Dracup (2000) noted that 1-year time lag was suitable for precipitation prediction for most climatic zones in California. Londhe and Narkhede (2018) found the benefit of using time lag of ANN to forecast stream flows in
India. However, in the field of solid waste management, the optimal time lag was not considered. All of these above studies in Section 1.1.1 attempted to improve the current ANN waste generation modelling techniques without considering the effects of an optimal lag time on the model’s performances. Their approaches were most appropriate for mixed waste prediction as the model inputs such as Gross Domestic Product (GDP) and family income are expected to have an immediate effect on the generation of consumer-related wastes such as packaging wastes and e-wastes. On the contrary, factors contributing to yard waste generation are time sensitive and the effects of a lag time must be explicitly considered in ANN prediction models. Among other things, the extent and rate of vegetation growth are heavily affected by climatic factors such as temperature as well as precipitation (Durand et al., 1999; Toledo et al., 2011), and there are considerable lags from the application of these climatic factors to the actual generation of yard waste. In the present study, the development of time series ANN models with lags were attempted and the results examined for the purpose of municipal yard waste prediction.

1.1.3 Yard waste generation

Yard waste accounted for 11-27% of the total municipal solid waste in urban centers within Australia, Canada, and United State of America (Abbasi and Hanandeh, 2016; City of Austin, 2018a; Tetra Tech EBA Inc, 2016). For example, the amount of US yard waste increased from 18.1 million tonnes in 1960 to 31.3 million tonnes in 2014 (USEPA, 2016), probably due to the population growth, and the urbanization of the rural areas. Landfill bans on yard waste were first introduced in the US in the 1990s due to concerns over yard waste’s use of valuable landfill space (US Composting Council, 2010) and the associated environmental impacts such as the generation of greenhouse
gases and acidic leachate (USEPA, 1991). As such, more US yard waste was diverted from landfill disposal and instead used as biomass for energy production (Shi et al., 2013) and as organic substrate for composting (Boldrin et al., 2011). In the US alone, the number of composting facilities increased from 650 in 1998 to more than 3,500 in 2010 (US Composting Council, 2010). An accurate municipal yard waste prediction model is therefore vital to the economic and environmental success of these waste-to-energy facilities.

Municipal yard waste contains mostly water and plant-based organic matter (Boldrin and Christensen, 2010). In Chapter two, the term yard waste is defined as the combination of grass, leaves, and small tree branches pruned throughout a growth cycle collected from a municipality.

1.2 Municipal Solid Waste Collection Model

Municipal solid waste collection is expensive and in some cities and towns 50%-85% of their total waste management expenditures are used for waste collection and transportation (Dogan and Suleyman, 2003; Ghose et al., 2006; Sanjeevi and Shahabudeen, 2016; Tavares et al., 2009). Enhancing collection efficiency by optimizing collection routes helps to reduce collection time, labour cost, and fuel consumption. Emissions from the collection vehicles will also be reduced and thus help lower the environmental footprint of the solid waste management system. Due to their practical importance, different optimization approaches and search techniques have been proposed when developing and evaluating waste collection systems. These approaches include the ant colony system (Karadimas et al., 2007), the construction method (Kim et al., 2006; Nuortio et al., 2006; Tung and Pinnoi, 2000), the chaotic particle swarm model (Son,
2014), the genetic method (Viotti et al., 2003), various heuristic methods (Cortinhal et al., 2016; Hemmelmayr et al., 2013), the meta-heuristic method (Benjamin and Beasley, 2013), and the Tabu search method (Bing et al., 2014; Krichen et al., 2014). Over the past decade, network analysis completed using Geographical Information Systems (GIS) with Dijkstra's algorithm and the Tabu search approach have been widely explored and applied to optimize the collection of municipal solid wastes and recyclables (Alvarez et al., 2008; Boskovic and Jovicic, 2015; Gallardo et al., 2015; Karadimas, and Loumos, 2008; Kanchanabhan et al., 2010; Nguyen et al., 2017; Son and Louati, 2016; Vijay et al., 2005; Vu et al., 2018a; Zamoranoa et al., 2009; Zsigraiova et al., 2013) due to the comparative advantages of better output visualization and easier data interpretation (Chalkias and Lasaridi, 2011; Malakahmad et al., 2014; Sanjeevi and Shahabudeen, 2016). The long list of studies shows the significance of waste collection optimization to a solid waste management system.

In many developing countries with transportation systems containing both narrow alleys and wider streets, accessing waste generation sites for collection and transportation generally consists of two distinct phases. In the first phase (pre-collection), materials from sources (generation sites) are collected by handcarts, cyclos, or smaller compacting vehicles and transported to temporary collection points (TCPs) for temporary storage. The size, number, and spatial distribution of these TCPs are constantly evolving. In the second phase (collection), the waste is then transported from the TCPs to a disposal site or processing facility using larger trucks. GIS based studies on waste collection can be generally divided into three categories: (i) studies of the first (pre-collection) phase, (ii)
studies of the second (collection) phase, and (iii) studies of both the first (pre-collection) and second (collection) phases.

1.2.1 The first phase (pre-collection phase)

Most of the GIS-based collection studies that focused on the pre-collection phase were about the design and operational parameters of the collection bin systems. These studies examined the optimal number, location, and size of bins for different collection systems (door to door, drop-off sites, curbside bins, etc.) and the collection frequencies (daily, weekly, etc.) using the GIS location-allocation and service area tools. Vijay et al. (2005) used the GIS location-allocation tool to determine the number, size, and location of bins taking into account the geographical features for a typical city in India. Using a spatial geodatabase, Karadimas and Loumos (2008) developed a GIS based model to calculate the optimal number of bins and their allocations for Athens, Greece. Additionally, Boskovic and Jovicic (2015) applied the service area tool of GIS to obtain the optimal TCP locations and the optimal number of bins for the City of Kragujevac, finding the required number of bins dropped by 30.0 - 33.5% compared to the non-optimized case. Further, this strategic bin placement led to a reduction in truck fuel consumption (Boskovic and Jovicic 2015). Fewer waste bins also bring down maintenance costs as the number of required truck stops fell (Karadimas and Loumos, 2008; Boskovic and Jovicic, 2015).

In addition to the location and number of bins, the optimal distances from source to TCPs were reported in many GIS-based studies. The optimal distance between a waste generating source and a TCP varied with factors such as location, geographical features, building types, and waste composition. The reported optimal distance between a waste
generating source and a TCP was 75m for the city of Kragujevac (Boskovic and Jovicic 2015). A shorter distance of 20-30m was reported by Gallardo et al. (2015) for Castellón, Spain. For recyclables, a distance of 100-300m appeared to be appropriate for recycling drop-off sites in Castellón, Spain (Gallardo et al., 2015), whereas for commercially used paper, Alvarez et al. (2008) found the optimal distance in Madrid, Spain to be 60m. The variations of the optimal distances in the literature were expected as the distances were found to be highly site specific.

### 1.2.2 The second phase (collection phase)

GIS based tools have also been successfully applied across the globe to optimize the municipal waste process during the collection phase. Some of these studies focused on adjusting routes and minimizing travel distances (Ghose et al., 2006; Malakahmad et al., 2014; Sanjeevi and Shahabudeen, 2016) while others focused on the minimization of collection costs and pollutant emissions (Zsigraiova et al. 2013; Abdelli et al. 2016). In some Asian countries, municipal waste collection systems are often organized empirically and could greatly benefit from GIS tools and spatial modeling techniques. One study doing this in Vietnam includes work done by Son and Louati (2016) who created a generalized vehicle routing model with multiple transfer stations, gather sites, and inhomogeneous vehicles in time windows to optimize waste collection in Da Nang city, Vietnam. Another study in Vietnam was done by Nguyen et al. (2017) where they combined analytical-based and agent-based models to build a dynamic waste collection model for Ha Giang city, Vietnam. Nguyen et al.’s model (2017) showed a reduction of 11.3% in the total truck travel distance when compared with that of the existing collection system.
1.2.3 The first and the second phases (pre-collection and collection phases)

A number of studies have tried to utilize the advantages of combining the first and second phases (pre-collection and collection phases) when optimizing a dual-phase waste collection system. Zamorano et al. (2009) used GIS to optimize a waste collection scheme in Churriana de la Vega, Granada, Spain. Buffer zone radii from 50 m to 300 m were selected to analyze percentage coverage and it was reported that the optimal source-to-TCP distance was less than 75 m. The number of containers was reduced by 37.8% and the average truck travel distance was reduced by 40.6% when compared with those of the existing system (Zamorano et al. 2009).

Thanh et al. (2009) used MapInfo and GIS to create a model to optimize a waste collection system in Can Tho, Vietnam and found noticeable reductions in vehicle travel distance (19%) and collection time (12%). Arribas et al. (2010) built a GIS based model with local search technique to determine an optimal waste collection system for Santiago, Chile and found significant waste collection cost savings. Kanchanabhan et al. (2010) developed a GIS waste collection model for both pre-collection and collection phases at Pallavapuram municipality in India. The distances between TCPs varied from 100 m to 600 m, and the total number of TCPs was 118 for an area of 18 km² (Kanchanabhan et al., 2010). In their study, the Closest Facility and Vehicle Route Problem tools of the Network Analysis of GIS were used to identify the optimal TCPs and the routes for tricycles and trucks (Kanchanabhan et al., 2010).

Khan and Samadder (2016) applied GIS tools to develop a two-fold waste collection model for Dhanbad City, India, and reported promising cost savings with an optimal distance between a waste generating site and a TCP of 100 m. Erfani et al. (2017) used
Network Analysis - GIS to build an optimization model considering both bin locations and vehicle routes in an Iranian city and found an optimal range of distances between a source and a bin for low- and medium-rise apartments from 150 m to 180 m. According to Erfani et al. (2017), the daily collection tours were reduced by 12.5% and the required crew numbers were reduced by 41.7%.

1.3 Municipal Landfill Gas Model

Canada has one of the highest per capita waste generation rates in the world (Bruce et al., 2016; Wang et al. 2016, Richter et al., 2017). Statistics Canada reported the 2010 national per capita waste generation rate at about 965 kg/year (2010). Canada has adopted numerous programs in past decades to promote waste minimization, and to assist in the development of more sustainable solid waste management methods (Pan et al. 2018; Richter et al. 2018). Despite many efforts, waste recycling and composting are less common, and landfilling remains the predominant municipal non-hazardous solid waste treatment method in Canada. According to Statistics Canada (2010), permanent waste disposal such as landfilling accounted for over 75% of all solid waste managed. The reliance on landfilling as the primary treatment method is more pronounced in Saskatchewan, a Canadian Prairie province with a cold semi-arid climate. Saskatchewan has a higher than national land disposal rate of about 86.8% (Statistics Canada, 2010), probably due to the availability of affordable land. One of the major concerns of landfill use is the generation of Greenhouse Gases (GHGs) such as methane (CH$_4$) and carbon dioxide (CO$_2$) from the anaerobic decomposition of buried organics. In Canada, GHG emissions from landfills were 19 mega tonnes (or 550kg per capita), making up more
than 90% of total Canadian GHG emissions within waste management sector (Environment Canada, 2012).

1.3.1 First-order-decay gas generation

Landfill gas generation and emissions are of great importance to landfill operators and regulatory agencies, not only because of the contributions of GHGs, but also because of the explosive risk of methane gas. First order decay (FOD) gas generation models were developed and continue to be used by researchers around the world to quantify gas generation during the landfill lifespan due to their user-friendliness and comprehensive sets of default values. Among these FOD models, LandGEM and Scholl Canyon developed by US Environmental Protection Agency (LandGEM, 2005), Afvalzorg developed by a European private operator (Scharff and Jacobs, 2006), IPCC developed by the Intergovernmental Panel on Climate Change organization (IPCC, 2006), and EPER developed by Geman European Pollutant Emission Register are widely studied and reported (Scharff and Jacobs, 2006; Faour et al., 2007; Machado et al., 2009; Bella et al., 2011; Amini et al., 2012; Govindan and Agamuthu, 2014; Mou et al., 2015a).

However, questions regarding the reliability and applicability of default values and suggested model parameters have been raised over the past decades (Faour et al., 2007; Govindan and Agamuthu, 2014; Mou et al., 2015a, 2015b), specifically for landfills subjected to unconventional climatic conditions (Ishii and Furuichi, 2013).

1.3.2 Accuracy and reliability of the gas model

Bella et al. (2011) conducted a field study with direct measurements using a flux chamber at a municipal landfill in Italy, and found that LandGEM slightly overestimated methane generation. Govindan and Agamuthu (2014) noted that the IPCC model with
default k and degradable organic carbon (DOC) values overestimated methane generation in a Malaysian sanitary landfill. However, Govindan and Agamuthu (2014) found that modeling errors were significantly reduced (by 69 and 81% from 2 different approaches) with the use of site specific parameters. Similarly, Mou et al. (2015a) found that the default parameters used by LandGEM, Afvalzorg and IPCC models failed to represent low-organic waste materials and overestimated methane generation at four landfills in Denmark.

Inconsistencies in the literature on the use of FOD model default parameters are not uncommon. Faour et al. (2007) stated that LandGEM model results fitted reasonably well to actual gas data at three wet landfills in the US with a 1.5 year lag phase. On the contrary, Amini et al. (2012, 2013) and Wang et al. (2013) reported that LandGEM generally underestimated methane generation in a number of US landfills. Similar observations were made by Thompson et al. (2009) with Canadian landfills. Mou et al. (2015a) conducted a comprehensive review on 9 gas modeling studies using over 50 landfills in Europe, North America and Asia, and concluded that the ratios between modeled CH₄ generation estimations and field generation and emission measurement results varied from 0.7 - 3.3.

1.3.3 The parameters k and L₀

Care should be taken when interpreting and comparing these landfill gas modeling studies, as different model inputs, quantification methods, and assumptions were adopted in different studies. However, all of the above studies have stressed the importance of the use of site specific parameters in gas modeling. Table 1-1 summarizes results from different studies on FOD landfill gas model parameters. The decay rate (k)
for cold, semi-arid landfills derived from empirical models are typically lower, ranging from 0.006 to 0.023 year$^{-1}$, whereas the $k$ value for warmer climates ranges from 0.04 to 0.21 year$^{-1}$. The decay rate is sensitive to the climatic conditions where the landfill is located (Machado et al., 2009, Thompson et al., 2009, Ishii and Furuichi, 2013), and $k$ values are lower for landfills located in dry, cold climates (Thompson et al., 2009; Bruce et al., 2018). Sufficient moisture is essential for the growth of the microbial community, while a lack of moisture delays the waste decomposition process (Barlaz et al., 1992; Fourie and Morris, 2004). Unlike the decay rate ($k$), the methane generation potential ($L_0$), or the degradable organic carbon (DOC), only depends on the type and composition of disposed waste (Thompson et al., 2009). $L_0$ values reported from the literature using field studies and experimental measurements varied from 8 to 214.4 m$^3$/Mg (Table 1-1).
Table 1-1: First order decay (FOD) landfill gas model parameters from landfills with different climatic conditions (Vu et al., 2017)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Landfill location</th>
<th>Models</th>
<th>$k$ (year$^{-1}$)</th>
<th>$L_w$ (m$^3$/Mg)</th>
<th>$k$ determined from</th>
<th>$L_w$ determined from</th>
<th>Findings/Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faour et al. (2007)</td>
<td>CSWMC, SSWMC, and Landfill A, US</td>
<td>LandGEM</td>
<td>0.11±0.21</td>
<td>87±115</td>
<td>Regression analysis</td>
<td>Regression analysis</td>
<td>LandGEM results fitted well to the actual data. L$_w$ were lower than those normally presented in the literature for developing tropical countries. The models slightly underestimated methane production.</td>
</tr>
<tr>
<td>Machado et al. (2009)</td>
<td>Salvador, Brazil</td>
<td>FOA USEPA, 1996 and IPCC</td>
<td>0.2±0.21</td>
<td>63.8-66.6</td>
<td>Curve fitting</td>
<td>Curve fitting</td>
<td></td>
</tr>
<tr>
<td>Bella et al. (2011)</td>
<td>Palermo, Italy</td>
<td>LandGEM, Elucic</td>
<td>0.04</td>
<td>100</td>
<td>Used default values</td>
<td>Used default values</td>
<td>LandGEM underestimated methane generation.</td>
</tr>
<tr>
<td>Amini et al. (2013)</td>
<td>Florida, US</td>
<td>LandGEM</td>
<td>0.040-0.090</td>
<td>74-140</td>
<td>Curve fitting and linear regression</td>
<td>Waste composition</td>
<td>Overestimated CH$_4$ generation at all k and DOC values. Estimated error was reduce by 69% and 81% at calculated k and DOC values</td>
</tr>
<tr>
<td>Amini et al. (2012)</td>
<td>Florida, US</td>
<td>LandGEM</td>
<td>0.040-0.130</td>
<td>56-77</td>
<td>Curve fitting and linear regression</td>
<td>Waste composition</td>
<td></td>
</tr>
<tr>
<td>Govindan and Agrawal (2014)</td>
<td>Malaysia</td>
<td>IPCC</td>
<td>0.06; 0.09</td>
<td>0.08; 0.12 (DOC)</td>
<td>Precipitation-based empirical model</td>
<td>Waste composition</td>
<td></td>
</tr>
<tr>
<td>Cold, wet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang et al. (2015, 2012)</td>
<td>US</td>
<td>LandGEM</td>
<td>0.08±0.12</td>
<td>55±100</td>
<td>k was the value determined by minimizing the RSS determined from measured CH$_4$ collection</td>
<td>L$_w$ values given in the legend</td>
<td>Optimal k was higher than the default AP-42</td>
</tr>
<tr>
<td>Tolat et al. (2010)</td>
<td>US</td>
<td>LandGEM</td>
<td>0.06</td>
<td>48</td>
<td>Used defined waste samples</td>
<td>Optimal k was slightly higher than the default AP-42</td>
<td></td>
</tr>
<tr>
<td>Thompson et al. (2009)</td>
<td>Ontario, Canada, Quebec, Canada</td>
<td>EPER, TNO, Belgium, LandGEM, Scholl Canyon</td>
<td>0.037</td>
<td>137</td>
<td>Precipitation-based empirical model</td>
<td>Waste composition</td>
<td>All the models underestimated methane generation except for LandGEM</td>
</tr>
<tr>
<td>Ishii and Furusawa (2013)</td>
<td>Hokkaido, Japan</td>
<td>FOD</td>
<td>0.062 (food) 0.05 (paper)</td>
<td>214.4 (food) 126.7 (paper)</td>
<td>Curve-fitted line using FOD</td>
<td>Aged defined waste samples</td>
<td>k of food waste was smaller than that of the IPCC due to the lower temperature</td>
</tr>
<tr>
<td>Mou et al. (2015a, b)</td>
<td>Denmark</td>
<td>Advanzorg, IPCC, LandGEM</td>
<td>0.013-0.19</td>
<td>3-107</td>
<td>Experimental work</td>
<td>Waste composition</td>
<td>All three models overestimated at all k and L$_w$ values. Advanzorg fitted well at specific k and L$_w$ values for the landfills</td>
</tr>
<tr>
<td>Cold, semi-arid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thompson et al. (2009)</td>
<td>Alberta, Canada</td>
<td>EPER, TNO, Belgium, LandGEM, Scholl Canyon</td>
<td>0.023</td>
<td>152</td>
<td>Precipitation-based empirical model</td>
<td>Waste composition</td>
<td>All the models overestimated methane generation except for the LandGEM</td>
</tr>
<tr>
<td>Thompson et al. (2009)</td>
<td>Regina, Saskatchewan</td>
<td>N/A</td>
<td>0.023</td>
<td>N/A</td>
<td>Precipitation-based empirical model</td>
<td>Waste composition</td>
<td>N/A</td>
</tr>
<tr>
<td>Environment Canada (2005)</td>
<td>Regina, Saskatchewan</td>
<td>N/A</td>
<td>0.006-0.011</td>
<td>132</td>
<td>Precipitation-based empirical model</td>
<td>Waste composition</td>
<td>L$_w$ values derived from the most recent version of the report</td>
</tr>
</tbody>
</table>
Generally, the default set of k values for each landfill gas model are chosen based on the climatic condition (such as precipitation and temperature conditions), while the L₀ values or DOC values are selected based only on the waste sources or waste characteristics. For example, IPCC adopted four distinct climate groups (dry temperate, wet temperate, dry tropical, and wet tropical) with different sets of k values. In many reported studies, different approaches have been proposed to determine site-specific k and L₀ or DOC values, such as the use of multi-variable precipitation-based empirical equations, direct field measurements using tracer dispersion and a flux chamber, theoretical estimates based on waste chemical composition, and curve fitting to the field data (Machado et al., 2009; Amini et al., 2012; Govindan and Agamuthu, 2014).

The decay rates reported in warmer and wetter climates appeared higher in the literature (Table 1-1), suggesting a possible correlation between climatic conditions and waste decay rates in FOD models. However, the literature on landfill gas modeling in cold, semi-arid climates is very limited, and most studies focused on methane oxidation of cover soils (Bogner et al. 1997; Zeiss, 2006). Chapter three reports k, L₀, and DOC parameters by curve fitting using three different FOD models (LandGEM, Afvalzorg Simple and IPCC) with actual gas data in two cold semi-arid landfills in Saskatchewan, Canada. Most landfill gas studies in the literature measured gas generation and emission rates using discrete sampling intervals, and reported gas data as an average value for a given period of time. The present work, however, utilizes continuous daily and monthly recovery gas rates for curve fitting.

1.4 Waste Generation and Collection Route Design
Municipal solid waste collection is one of the most expensive components in modern Waste Management Systems (WMS) as more than 40% of the WMS budget is devoted to the collection and transport of waste materials in various places around the world (Chalkias and Lasaridi, 2009; Jaunich et al., 2016; Sanjeevi and Shahabudeen, 2016). Optimization of waste collection helps to reduce truck travel distance, collection time, fuel consumption, as well as air emission and studies on the optimization of waste collection with regards to the efficiency and effectiveness of WMS are commonly reported. For example, there were a number of recent studies on optimal WMS truck routes (Sanjeevi and Shahabudeen, 2016; Abdelli et al., 2016; Khan and Samadder, 2016; Erfani et al., 2017; Vu et al., 2018a; Vu et al., 2018b), WMS collection cost savings (Zsigraiova et al., 2013; Khan and Samadder, 2014; Abdelli et al., 2016; Boskovic et al., 2016; Nguyen et al., 2017), as well as WMS collection vehicle fuel consumption and emissions (Larsen et al., 2009; Zsigraiova et al., 2013; Abdelli et al., 2016; Edwards et al., 2016; Jaunich et al., 2016). However, most of these studies focus on cross-sectional analysis of a single-stream collection system at a given time and fail to address the continuous and interconnected nature of waste generation trends and waste collection route optimization. Changes in waste composition from the implementation of source-separation campaigns and/or landfill bans may impact waste compaction characteristics. This will ultimately impact the optimal route design due to the spatial constraints of collection trucks.

1.4.1 Residential waste generation rate and composition

Mass-based waste generation rates as well as characteristics such as waste type, waste composition, and waste density have been often used to calculate volume-based
waste generation rates which is employed in geospatial studies to estimate the number of truck trips needed for a given WMS (Zamorano et al., 2009; Abdelli et al., 2016). However, fewer studies have explicitly looked at how compositional changes in different waste streams affect collection truck route optimization. In the United States of America (USA), paper and related paper products accounted for the largest portion of municipal waste stream during the period from 1960 to 2015, representing an average of 31.47% of the total (USEPA, 2016a). However, municipal waste composition in USA, has changed considerably over time. For example, the amount of paper decreased from 30.8% in 2008 (USEPA, 2011) to 25.9% in 2015 (USEPA, 2018). Since this composition continues to change over time, most cross-sectional analyses on route optimization without considering the temporal waste compositional changes are difficult to justify from a theoretical standpoint. Waste density characteristics depend on waste composition and the specific weight of individual waste components. In the USA, the density of recyclables in trucks was estimated to be between 130 and 295 kg/m$^3$ (Jaunich et al., 2016; Maimoun et al., 2016). However, the density of garbage in trucks in the USA is considerably higher. It was estimated to be between 305 and 535 kg/m$^3$ (Jaunich et al., 2016; USEPA, 2016b).

Inaccurate and imprecise waste generation rate projections have been found to likely lead to inadequate or excessive collection facilities (Leao et al., 2001; Rimaityte et al., 2012; Abbasi and Hanandeh, 2016). Along with using recent waste composition data for more accurate waste densities, advanced numerical methods can also be used to estimate waste quantity. Recently, machine learning methods such as artificial neutral network (ANN) have been successfully used to forecast waste generation rates (Noori et
al., 2010; Abdoli et al., 2012; Rimaityte et al., 2012; Shahabi et al., 2012; Abbasi and Hanandeh, 2016; Azadi and Karimi-Jashni, 2016; Kannangara et al., 2018).

1.4.2 Single and multiple stream residential waste collection

In most developed countries, municipal waste is often separated at the source by generators (Nguyen and Wilson, 2010; Edwards et al., 2016; Jaunich et al., 2016; Maimoun et al., 2016). Municipalities provide curbside door-to-door waste collection service using either separated or combined collection vehicles for different waste streams. Unlike most passenger vehicles on road, a waste truck makes frequent stops and consumes more fuel than a comparable size vehicle due to its unique driving pattern (Nguyen and Wilson, 2010). Fuel consumptions and air emissions from door-to-door curbside waste collection services have been explored and quantified in a number of studies. Nguyen and Wilson (2010) studied the fuel consumption of two types of collection trucks and reported that a dual-compartment truck consumes more fuel than a single-compartment truck. They reported that the single-stream collection truck spent considerably shorter time (6.7 to 8.74s) at each stop than the multiple-stream collection truck (21.6s to 29.2s) (Nguyen and Wilson, 2010; Maimoun et al., 2016; Edwards et al., 2016). Edwards et al. (2016) compared 4 different collection systems in Australia, including three single-stream systems (recyclables, garbage, and organics), and a multiple-stream system. They found that collecting organics separately resulted in an increase of truck fuel consumption by 1.38% to 57.59%.

Because of the heterogeneous nature of recyclables, organics, and garbage, waste densities inside the single-stream and multiple-stream collection trucks can be quite different. Field evidences suggest waste composition can be sensitive to fuel use and
collection cost (Nguyen and Wilson, 2010; Maimoun et al., 2016; Edwards et al., 2016). As such, it is important to include recyclables and garbage composition in collection studies. None of the published studies considered the temporal changes of waste generation and composition on collection route design. Analysis in the present study suggests that the changes in waste (recyclables and garbage) generation rate and composition affect the number of trips, travel distance, and waste collection time. As such, a systematic approach combining both the waste generation rate and composition with collection route modelling is warranted.
2. Time-lagged effects on ANN Municipal Yard Waste Prediction

2.1 Objectives of Chapter Two

Chapter two aimed to develop accurate yard waste ANN models with lag effects using weekly social, climatic, and economic inputs. The objectives this part are to (i) conduct correlation analysis and identify the key independent variables for the proposed time series ANN municipal yard waste prediction models; (ii) systematically examine the effects of lag on the model’s performance using a set of statistical metrics; (iii) optimize the structure of the selected yard waste prediction models with a different number of neurons in the hidden layer. Unlike other ANN waste studies, this study integrates time lags to improve the performance of an ANN model. This novel modeling approach allows the use of weekly climatic variables such as ambient temperature and wind speed on yard waste modeling and is more advantageous than conventional ANN approaches from a theoretical point of view.

2.2 Study area and background

Austin, US, is selected as the study area to highlight the potential benefits of the time-lag modeling approach. Austin is the capital city of Texas, the 13th most populous city in the US (Figure 2-1). The population of Austin was about 926,000 in 2016 (City of Austin, 2018c), with a land area of 844.21 km². Weekly data from October 2004 to April 2018 on waste quantity, as well as climatic, and socio-economic data are collected. During the 14-year study period, the average weekly temperature, humidity, maximum wind speed, and precipitation in Austin was 21.34°C, 63.59%, 34.74 km/h, and 17.09mm, respectively (WU, 2018).
Weekly waste and recycling collection services are provided to the residents of Austin. Single, duplex, and triplex homes receive another curbside yard trimmings (yard waste) service on a weekly basis (City of Austin, 2018b). The average solid waste generation rate was 3,965.8 tonnes per week. Mixed garbage, recyclables, yard waste, and other waste composition categories respectively accounted for 57.40%, 21.18%, 11.22%, and 10.19% of the total residential waste. The garbage is sent to the Texas Disposal Systems landfill and the recyclables are transported to nearby material recovery facilities. Yard waste and other bio-solids are transferred to the Hornsby Bend Biosolids Management plant for treatment (City of Austin, 2018a). Significant increase in yard waste generation has been observed in Austin recently. For example, the collected yard waste went up 76.5% from 17,713.8 tonnes in 2005 to 31,263.1 tonnes in 2017 (City of Austin, 2018a).
Figure 2-1: Location of the study area (After National Geographic et al., 2011, WU, 2018, and City of Austin, 2018c, Vu et al., 2019)
2.3 Methodology

The methodology of Chapter two is presented in Figure 2-2. Four climatic variables (ambient temperature (Temp), humidity, maximum wind speed (Wind), and precipitation (Precip)) and four socio-economic variables (population (Pop), Dow Jones Industrial Average (DJIA), NASDAQ Composite (Nasdaq), and West Texas Intermediate Oil Price (Oil Price)), as well as the yard waste data were collected during the study period. Details on variables selection and data sources are provided in the following sections. The complete data set was analyzed using Data Analysis Tools (MS Excel 2013) to identify the correlations between the target variable (weekly municipal yard waste generation rate) and the 8 selected parameters with respect to lag times. Possible correlations between the variables were also determined to minimize multicollinearity among the inputs. Specifically, closely related sets of parameters with correlation coefficients higher than 0.95 (Abdoli et al., 2011) were eliminated.

The independent and significant parameters in relation to the target variable was used to develop ANN models with different lag times. The models were constructed using Matlab (version 2017b). The performance of the models are then evaluated using different statistical indices including mean square error (MSE), mean absolute percentage error (MAPE), and the correlation coefficient (R Value). Finally, the model structures of the better models are optimized by considering the different number of neurons in the hidden layer.
Figure 2-2: Methodology flow chart for Chapter two (Vu et al., 2019)
2.3.1 Data collection and analysis

2.3.1.1 Waste data sets

Waste data used in this study was collected from the City of Austin open data portal (City of Austin, 2018a). The complete waste data set included the mass of daily collected yard waste for the period from October 2004 to April 2018. Yard waste was mainly collected on weekdays and rarely collected on weekends (only 92 days of weekend collection, or about 0.15% of the 14-year study period). As such, the complete daily waste data set is converted to a weekly waste data set for the model development as this provides the most accurate data given the variability that of weekend pickups may introduce throughout the study period. More importantly, it smooths out the differences in daily collection mass that are more related to individual collection routes and neighbourhoods as opposed to the prediction of the citywide collection of yard waste over a decade and a half, which was the focus of this study. There was only one missing data in the entire data set on the 14th of January, 2007, because of adverse weather condition. The mass of yard waste data on the 14th of January, 2007 was estimated using an average of weekly yard waste from 13 years (October 2004 to April 2018). Pre-treated data - interquartile range (IQR) was applied in Kannangara et al. (2018) study to identify and eliminate outliers from datasets. IQR was also applied in the preliminary trials in the current study, and there were 41 data points slightly higher than the upper boundary. A closer look at these data points, however, suggest the identified data is acceptable, as residents prune branches in early spring. To preserve the completeness of the data set, data normalization and pre-processing are not conducted in this study.
2.3.1.2 Climatic parameters

Average weekly temperature, humidity, precipitation, and maximum wind speed over the study period were collected from the Weather Underground (WU, 2018). Ambient temperature is considered as a major factor in many waste generation studies (Gómez et al., 2009; Abdoli et al., 2011; Abdoli et al., 2012; Azadi and Karimi-Jashni, 2016). Precipitation and humidity also affected waste generation rates (Abdoli et al., 2011; Johnson, 2017), as municipal wastes are exposed to weather and their weights are inevitably influenced by moisture content (Han et al., 2018).

Unlike other waste studies, maximum wind speed data was used in the proposed municipal yard waste prediction model. The effects of wind damages on crop plants and fruit trees (Jim and Liu, 1997; Cataldo et al., 2013), as well as on urban and forest trees (Ciftci et al., 2014; Gardiner et al., 2016) are well-documented in the literature. The leaves and small branches from the gust damaged shrubs and trees are likely to be a source of municipal yard waste. Climatic data such as wind speed is also adopted in Kontokosta et al.’s study (2018) to develop daily and weekly predictive models of building-level waste and recycling generation in New York City.

2.3.1.3 Socio-economic parameters

Weekly data on population, Dow Jones Industrial Average, NASDAQ Composite index, and West Texas Intermediate Oil Price during the study period were selected as socio-economic variables. Population is one of the most reported parameters on waste generation (Daskalopoulos et al., 1998; Gidarakos et al., 2006; Intharathirat et al., 2015; Oribe-Garcia et al., 2015; Younes et al., 2015; Azadi and Karimi-Jashni, 2016;
Urbanization from population growth is expected to increase the quantity of municipal yard waste.

The quantity of municipal yard waste collected also depends on the frequency and extent of grass cutting and tree pruning, which in turn depends on the economy. GDP and consumer expenditures have been commonly reported as factors on waste generation (Daskalopoulos et al., 1998; Bruce et al., 2016; Wang et al., 2016). DJIA and Nasdaq indices as well as Oil Price fluctuate with time and both are responsive to GDP and consumer expenditures (Edelstein and Kilian, 2009; Kilian and Vigfusson, 2012; Huang et al., 2016), and therefore are also selected.

2.3.1.4 Data variability and correlation analysis

The standard deviation (STDEV) as well as the coefficient of variation (ratio of STDEV divided by mean) of the variables are computed to assess the relative variability of the parameters, as well as the bias and the distribution of the data sets. Unlike other waste generation studies (Abdoli et al., 2011, Azadi and Karimi-Jashni, 2016, and Kannangara et al., 2018), the correlation coefficients between the variables are considered at different time lags to investigate the time lag’s effects on the accuracy and precision of the yard waste models. This is the first waste generation study to explicitly consider lag effects in the development of time series ANN models.

2.3.2 Lag time and ANN model development and optimization

Grass from yard work is the major constituent of the municipal yard waste in the US, accounting for more than 50% of yard waste by mass (Ragsdale et al., 1992; State of Oregon, 2016). Well-maintained lawns and gardens in a municipality require frequent
cutting and trimming, and hence the steady production of yard waste. Grass leaf growth time ranges from 5 to 65 days depending on the species and climatic conditions (Durand et al., 1999). The lag time considered in the proposed yard waste models conservatively ranged from 1 to 16 weeks.

In this study, a nonlinear autoregressive with external input (NARX) model is constructed in Matlab (version 2017b) using time series inputs. Correlation analysis ensures that the inputs are independent and statistically significant. The target variable is the weekly municipal yard waste by mass. There are 600 data sets for training and 105 data sets for testing, corresponding to an 85:15 training to testing ratio. This 85:15 training to testing ratio was also used by Abbasi and Hanandeh (2016) as well as Kannangara et al. (2018).

A total of 3 layers (input, hidden, output) are used to build each model, with 10 neurons in the hidden layer (default setting). Levenberg-Marquardt algorithm was adopted in all the models for data training. A total of 128 scenarios (8 models including climatic, socio-economic, and hybrid groups, each with 16 lag times) were created and assessed using various performance indicators. Early stopping is applied in each NARX model to avoid overfitting (i.e. overtraining of the model). For each scenario, over 30 trials are conducted with randomly assigned inputs from the dataset to study the model precision and to determine the case with the minimum MSE.

To determine the optimized model structure, the number of neurons of the better models were identified using their MSE. The defaulted number of neurons within the hidden layer is 10, however, slightly smaller optimal number of neurons (8~9) have been
reported (Abbasi and Hanandeh, 2016; Azadi and Karimi-Jashni, 2016). To address the case-specific nature of the proposed non-linear yard waste model, a wider range of number of neurons from 7 to 13 was selected to determine the optimized model structure. When number of neurons in hidden layers changes, weight of each neuron will be changed as well as the output of the model will be changed. Trials with different number of neurons will help to fine the optimal model structure with the lowest error.

2.3.3 Models performance assessment

MSE measures how close the predicted values are to observed values in numerical models. A MSE value closer to zero represents better model performance. MAPE and R Value were reported in various waste generation studies (Rimaityte et al., 2012; Shahabi et al., 2012; Shamshiry et al., 2014; Younes et al., 2015; Abbasi and Hanandeh, 2016; Azadi and Karimi-Jashni, 2016; Ghinea et al., 2016; Younes et al., 2016; Kumar and Samadder, 2017; Kannangara et al., 2018), and are used in the present work for comparison purposes. The drawbacks and strengths of these performance metrics in the evaluation of ANN modeling were discussed elsewhere (Azadi and Karimi-Jashni, 2016) and is not repeated here. MSE, MAPE, and R Value are defined as:

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (Y_m^i - Y_a^i)^2
\]  \hspace{1cm} (2 - 1)

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y_m^i - Y_a^i}{Y_a^i} \right| \times 100
\]  \hspace{1cm} (2 - 2)

\[
R = \frac{\sum_{i=1}^{n} (Y_m^i - \bar{Y}_m) \times (Y_a^i - \bar{Y}_a)}{\sqrt{\sum_{i=1}^{n} (Y_m^i - \bar{Y}_m)^2} \times \sqrt{\sum_{i=1}^{n} (Y_a^i - \bar{Y}_a)^2}}
\]  \hspace{1cm} (2 - 3)
Where:

- \( n \): Number of data points
- \( Y_a \): Actual mass of yard waste
- \( Y_m \): Predicted mass of yard waste
- \( \bar{Y}_a \): The mean actual waste mass
- \( \bar{Y}_m \): The mean predicted waste mass

### 2.4 Results and discussion

#### 2.4.1 Data analysis and screening

Table 2-1 shows the characteristics of the weekly data sets. Large variations of the yard waste quantity are observed during the study period, with a coefficient of variation of 0.6. Periodicity of the data is observed, and is consistent with the local growing seasons. The highest yard waste quantities are observed in the spring, when frequent cutting and trimming of vegetation is performed. The average weekly yard waste was 445.24 tonnes during the study period, or approximately 0.55kg/capita per week. The median is lower than the mean, suggesting the target data is skewed. The mean temperature, humidity, maximum wind speed, and precipitation were 21.34°C, 63.59%, 34.74 km/h, and 17.09 mm, respectively (Table 2-1). Among all variables studied, the precipitation fluctuated most during the study period, with a coefficient of variation above one (1.71). This is possibly due to a few extreme rainfall events, as the set is heavily skewed.

The mean population of Austin during the 14-year period was about 808,550. The weekly stock indices and Oil Price varied considerably, with coefficients of variation
ranging from 0.28 to 0.43 The Nasdaq varied the most among the four social-economic factors, probably due to the volatile nature of the high-growth technology sectors, of which the Nasdaq is more exposed than the DJIA.
Table 2-1: Weekly variable statistics (October 2004 - April 2018) (Vu et al., 2019)

<table>
<thead>
<tr>
<th>Variable Type</th>
<th>Variable</th>
<th>Min</th>
<th>Mean</th>
<th>Median</th>
<th>Max</th>
<th>STDEV</th>
<th>STDEV/Mean</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waste Variable</td>
<td>Yard Waste (tonnes)</td>
<td>39.69</td>
<td>445.24</td>
<td>360.22</td>
<td>1748.29</td>
<td>268.74</td>
<td>0.60</td>
<td>City of Austin</td>
</tr>
<tr>
<td>Climatic Variables</td>
<td>Temperature (°C)</td>
<td>2.43</td>
<td>21.34</td>
<td>21.71</td>
<td>33.43</td>
<td>7.40</td>
<td>0.35</td>
<td>Weather Underground</td>
</tr>
<tr>
<td></td>
<td>Humidity (%)</td>
<td>29.43</td>
<td>63.59</td>
<td>63.43</td>
<td>92.71</td>
<td>8.81</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Max. Wind Speed (km/h)</td>
<td>20.71</td>
<td>34.74</td>
<td>34.71</td>
<td>124.50</td>
<td>6.24</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Precipitation (mm)</td>
<td>0</td>
<td>17.09</td>
<td>3.04</td>
<td>192.28</td>
<td>29.29</td>
<td>1.71</td>
<td></td>
</tr>
<tr>
<td>Social-economic Variables</td>
<td>Population (people)</td>
<td>691,110</td>
<td>808,547</td>
<td>801,731</td>
<td>955,827</td>
<td>79,316</td>
<td>0.10</td>
<td>City of Austin</td>
</tr>
<tr>
<td></td>
<td>Nasdaq (USD)</td>
<td>1,239.85</td>
<td>3,371.34</td>
<td>2,725.16</td>
<td>7,560.81</td>
<td>1,451.73</td>
<td>0.43</td>
<td>Investing.com</td>
</tr>
<tr>
<td></td>
<td>DJIA (USD)</td>
<td>6,026.94</td>
<td>14,015.92</td>
<td>12,820.60</td>
<td>26,616.71</td>
<td>3,957.73</td>
<td>0.28</td>
<td>Investing.com</td>
</tr>
<tr>
<td></td>
<td>Oil Price (USD)</td>
<td>28.14</td>
<td>73.07</td>
<td>70.80</td>
<td>142.52</td>
<td>22.76</td>
<td>0.31</td>
<td>EIA</td>
</tr>
</tbody>
</table>

Notes:

The correlation coefficients between all nine variables with no lag time are first compared, as shown in Table 2-2. The first column shows the correlation coefficients of the climatic and socio-economic variables on the quantity of yard waste. Humidity, precipitation, and Oil Price appear to have minimal impacts on yard waste generation, with coefficients less than |0.1|. The maximum wind speed shows the highest coefficient of +0.322. The temperature has a moderate negative relationship with the mass of yard waste, with a coefficient of -0.225. Given that the majority of US municipal yard waste is comprised of grass (Ragsdale et al., 1992), higher ambient temperatures promote the evaporation of moisture in yard waste and reduces the waste's weight. This may not be as applicable to other waste types. For example, studies by Abdoli et al. (2011) and Azadi and Karimi-Jashni (2016) examining mixed solid waste generation in Iran found that temperature had a positive correlation with waste generation, possibly due to the changes of Iranian household consumption patterns on warmer days (Abdoli et al., 2011).

The correlation coefficients of Nasdaq and DJIA indices on weekly yard waste are very similar, ranging from 0.306 to 0.309. A closer look at the coefficients between these indices confirms that they are also highly correlated to each other, with a coefficient of +0.985, above the 0.95 cutoff value (Abdoli et al., 2011). It is found that population in the study area is related surprisingly well to the DJIA and Nasdaq indices, with coefficients larger than +0.85. This is probably due to a vibrant advanced manufacturing and technology sector located within the greater Austin area. The results suggested multicollinearity exists among these variables.
Table 2-2: Correlations of the input variables and weekly municipal yard waste from October 2004 to April 2018 (without the effect of lag time) (Vu et al., 2019)

<table>
<thead>
<tr>
<th></th>
<th>Yard Waste</th>
<th>Average Temperature</th>
<th>Average Humidity</th>
<th>Max. Wind Speed</th>
<th>Precip.</th>
<th>Pop.</th>
<th>Nasdaq</th>
<th>DJIA</th>
<th>Oil Price</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Waste variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yard Waste</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Climatic variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Temperature</td>
<td>-0.225</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Humidity</td>
<td>0.027</td>
<td>-0.077</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max. Wind Speed</td>
<td>0.322</td>
<td>0.002</td>
<td>-0.120</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precipitation</td>
<td>-0.047</td>
<td>0.001</td>
<td>0.485</td>
<td>0.078</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Socio-economic variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>0.294</td>
<td>0.016</td>
<td>0.147</td>
<td>-0.046</td>
<td>0.048</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasdaq</td>
<td>0.306</td>
<td>-0.014</td>
<td>0.197</td>
<td>-0.085</td>
<td>0.064</td>
<td>0.917</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DJIA</td>
<td>0.309</td>
<td>-0.027</td>
<td>0.194</td>
<td>-0.081</td>
<td>0.065</td>
<td>0.870</td>
<td>0.985</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Oil Price</td>
<td>-0.077</td>
<td>0.140</td>
<td>-0.217</td>
<td>0.140</td>
<td>-0.059</td>
<td>-0.174</td>
<td>-0.290</td>
<td>-0.225</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Figure 2-3 confirms that lag exists between the parameters and the target variable. In general, more lag effects are observed in the climatic variables than the socio-economic variables. Nasdaq, DJIA, and population are positively correlated (about +0.3) with yard waste generation. However, lag effect is not observed for these parameters with the target variable, as the slope of the curves is almost constant. The DJIA and Nasdaq curves overlap with each other, suggesting the lack of independency of these two parameters. To avoid multicollinearity, the DJIA variable is discarded and only the Nasdaq variable is used in the ANN model development because of its higher data variability, similar to the target yard waste variable’s high data variability (Table 2-1). Population is less related to these economic indices and is included in the final yard model.

Temperature is strongly correlated to the target variable, with coefficients ranged from -0.22 to -0.57 (Figure 2-3). Temperature has the strongest influences on weekly municipal yard waste generation at lags between 6 and 13 weeks, with correlation coefficients > |0.5|. A strong lag effect is also observed for the maximum wind speed variable. Unlike other parameters, both positive and negative coefficients are observed (Figure 2-3). Higher wind speed may damage trees and leaves and yield more municipal yard waste in the short term, however the total amount of yard waste susceptible to wind damages is finite. Due to these postulated causations, both temperature and wind speed are selected for use in the proposed models.

Humidity and precipitation curves in Figure 2-3 are similar with each other, and this result is consistent to the findings from correlation analysis (Table 2-2). Although both show effects of lag, the coefficients over 16 lag weeks are generally small, between
+0.1 to -0.1 (Figure 2-3). As such, these two climatic variables were not further considered. Similarly, Oil Price has a relatively low correlation (< |0.2|) at any lag time (1-16 weeks) with yard waste, and is not included in the proposed model.
Figure 2-3: Correlation coefficient between the input variables and the yard waste variable at different lag times (Vu et al., 2019)
A total of eight models are developed using the significant and independent variables. These include two climatic models (Temp, Temp-Wind), one socio-economic model (Nasdaq), and five hybrid models (Temp-Nasdaq, Temp-Pop, Temp-Wind-Nasdaq, Temp-Wind-Pop, and Nasdaq-Pop-Temp). A detailed description of the models are provided in Table 2-3.
**Table 2-3: The proposed weekly municipal yard waste prediction models (Vu et al., 2019)**

<table>
<thead>
<tr>
<th>Type of Models</th>
<th>Explanation</th>
<th>Model Acronym</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Climatic Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Weekly yard waste prediction model using temperature as an input</td>
<td>In this model, the input contains the weekly average temperature over the period from October 2004 to April 2018 in the city of Austin. The temperature variable was selected as an input of the model due to it having the highest correlation with yard waste in comparison to the other climatic variables.</td>
<td>Temp</td>
</tr>
<tr>
<td>(2) Weekly yard waste prediction model using temperature and wind as the inputs</td>
<td>This model was developed based on the Temp model but the maximum speed of wind was also added as a second input variable. Wind was selected as an additional input of the model due to it having the second highest correlation with yard waste in comparison with the other climatic variables.</td>
<td>Temp-Wind</td>
</tr>
<tr>
<td><strong>Socio-economic Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Weekly yard waste prediction model using Nasdaq Composite stock market index as an input</td>
<td>In this model, the input contains the weekly Nasdaq Composite stock market index over the period from October 2004 to April 2018 in the city of Austin. The Nasdaq Composite stock market index variable was selected as an input of the model due to it having the highest correlation with yard waste in comparison to the other socio-economic variables.</td>
<td>Nasdaq</td>
</tr>
<tr>
<td><strong>Hybrid Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Weekly yard waste prediction model using temperature and the Nasdaq Composite stock market index as the inputs</td>
<td>This model was developed based on the Temp model but the Nasdaq Composite variable was also added as a second input. The Nasdaq Composite stock market index was selected as an additional input of the model due to it having the highest correlation with yard waste in comparison to the other socio-economic variables.</td>
<td>Temp-Nasdaq</td>
</tr>
<tr>
<td>(5) Weekly yard waste prediction model using temperature and population as the inputs</td>
<td>This model was developed based on the Temp model but the population was also added as a second input. Population was selected as an additional input of the model due to it having the second highest correlation with yard waste in comparison to the other socio-economic variables after eliminating the Dow Jones Industrial Average variable.</td>
<td>Temp-Pop</td>
</tr>
<tr>
<td>(6) Weekly yard waste prediction model using temperature, maximum wind speed, and NASDAQ Composite stock market as the inputs</td>
<td>This model was developed based on the Temp-Wind model but the Nasdaq Composite stock market index was also added as a third input. The Nasdaq Composite stock market index was selected as an additional input of the model due to it having the highest correlation with yard waste in comparison to the other socio-economic variables.</td>
<td>Temp-Wind-Nasdaq</td>
</tr>
<tr>
<td>(7) Weekly yard waste prediction model using maximum wind, temperature, and population as the inputs</td>
<td>This model was developed based on the Temp-Wind model but the population was also added as an input. Population was selected as an additional input of the model due to it having the second highest correlation with yard waste in comparison with other socio-economic variables after eliminating the Dow Jones Industrial Average variable.</td>
<td>Temp-Wind-Pop</td>
</tr>
<tr>
<td>(8) Weekly yard waste prediction model using Nasdaq Composite stock market index, population and temperature as the inputs</td>
<td>This model was developed based on the Nasdaq model but the population and temperature were also added as inputs. The temperature and population were selected as additional inputs of the model due to their correlation ranking of Temp being the highest ranked climatic variable and population being the second highest ranking socio-economic variable after eliminating the Dow Jones Industrial Average variable.</td>
<td>Nasdaq-Pop-Temp</td>
</tr>
</tbody>
</table>
2.4.2 Optimal lag times and data overfitting

In this study, time series ANN weekly yard waste prediction models are evaluated using 128 scenarios. The results are tabulated in Table 2-4. With the exception of the Temp model, the optimal lag time of all models ranged from 1 to 5 weeks. Temperature has the strongest correlation (Figure 2-3) in this study and the parameter is used in 7 out of the 8 models. The optimal lag time of the Temp model corresponds well to the temperature variable curve in Figure 2-3. All models performed better in the training stage, with the MSE at the training stage 1.8 to 2.8 times lower than that of the testing stage, and data overfitting is not observed. Models with the lowest MSE in the testing stage (Temp and Temp-Pop) models are selected for model structure optimization, as further discussed in Section 3.4.

2.4.3 Performance of models at the testing stage

MSE is scale dependent, and as such only the MAPE and R Values at the testing stages are used to evaluate the relative performance of the models. Models in the climatic group (Temp and Temp-Wind) performed well, with an average MAPE of 20.19% and an average R Value of 0.84 at the testing stage. The R Values of these two ANN yard waste prediction models compared favorably to other ANN municipal solid waste prediction models, in which R Values ranged from 0.76-0.92 (Abbasi and Hanandeh, 2016; Azadi and Karimi-Jashni, 2016; Kannangara et al., 2018). The performance of the two climatic yard waste models is better than Kannangara et al.’s (2018) paper waste prediction model with an R Value of 0.59.
The performance of the five hybrid models are acceptable, with an average R Value slightly higher than 0.8. Larger variations of the MAPE (19.97%-23.79%) are observed among the models, making the precision of the hybrid group lower than that of the climatic group (Table 2-4).

The Nasdaq model is the only model from the socio-economic group. The Nasdaq model has the highest MAPE of 25.24% among all 8 models and an R Value of 0.75 at testing. Socio-economic factors are traditionally used in development of prediction models for mixed municipal waste. The results from this study, however, highlight the importance of climatic data on the accuracy of yard waste prediction models. Of the three hybrid models with 3 inputs, the Nasdaq-Pop-Temp model had the poorest estimates with an MAPE of 23.79%. A higher correlation coefficient is observed between Nasdaq and population (Table 2-2) and the collinearity of the parameters may contribute to a higher MAPE value at the testing stage.
Table 2-4: Comparison of each model’s performance at their optimal lag time (Vu et al., 2019)

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Lag time (weeks)</th>
<th>MSE</th>
<th>MAPE</th>
<th>R value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Train</td>
<td>Test</td>
<td></td>
</tr>
<tr>
<td><strong>Climatic Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp</td>
<td>11</td>
<td>10,882</td>
<td>30,158</td>
<td>19.48</td>
</tr>
<tr>
<td>Temp-Wind</td>
<td>3</td>
<td>13,107</td>
<td>30,946</td>
<td>20.82</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>7.00</td>
<td>11,995</td>
<td>30,552</td>
<td>20.15</td>
</tr>
<tr>
<td><strong>Socio-economic Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasdaq</td>
<td>1</td>
<td>18,187</td>
<td>38,728</td>
<td>23.99</td>
</tr>
<tr>
<td><strong>Hybrid Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temp-Nasdaq</td>
<td>3</td>
<td>14,077</td>
<td>30,378</td>
<td>20.98</td>
</tr>
<tr>
<td>Temp-Pop</td>
<td>1</td>
<td>15,864</td>
<td>28,003</td>
<td>22.21</td>
</tr>
<tr>
<td>Temp-Wind-Nasdaq</td>
<td>5</td>
<td>11,888</td>
<td>31,976</td>
<td>21.22</td>
</tr>
<tr>
<td>Temp-Wind-Pop</td>
<td>2</td>
<td>15,366</td>
<td>31,422</td>
<td>21.37</td>
</tr>
<tr>
<td>Nasdaq-Pop-Temp</td>
<td>2</td>
<td>15,331</td>
<td>33,982</td>
<td>21.77</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>2.60</td>
<td>14,505</td>
<td>31,152</td>
<td>21.51</td>
</tr>
</tbody>
</table>
To better illustrate the lag effects of the inputs on the proposed yard waste models, results from the single-input-variable models (Temp and Nasdaq) are compared side by side (Figure 2-4). From Figure 2-4a, it can be seen that in both the training and testing stages, R Values of the Temp model are consistently higher than that of the Nasdaq model at any lag time. R Value of the Temp (train) model peaked at 10 weeks, with an R Value of 0.93. The peak shifted to the right by one week with a lag time of 11 weeks using the Temp (test) model data set. R Value of the Nasdaq (train) model appears to be stable at around 0.85, with the exceptions of two peaks at 12 and 16 weeks. The peaks from the Nasdaq (test) set are however located at 11 and 16. The testing curve for the Nasdaq model has a lower correlation than any of the testing Temp models at any lag time.

The MSE of the Temp model was lower than the Nasdaq model in both the training and testing stages at all lag times (Figure 2-4b). The MSE of the Temp (train) model went down 55.4% from 15,820 to 8,769 when the lag time increased from 1 to 10 weeks. The training curve of Nasdaq model is quite constant at about 17,800, particularly from lag of 1 to 11 weeks. There are however two troughs in the MSE of the Nasdaq (train) model located at 12 and 16 weeks. Considerable larger errors are observed in the testing stage for both models. The trough is shifted to 11 weeks for Temp model at testing stage. There are multiple troughs in the Nasdaq (test), at lags of 1, 11, and 16 weeks. An inverse relationship between R Value and MSE is observed (Figure 2-4a and 2-4b), and the performance indicators generally agree well with each other. The results also suggest that the lag time predicted by the training and testing sets agree within ± 1 week.
Figure 2-4: Example of the two single variable model performances at different lag times: a) using R Value; b) using MSE (Vu et al., 2019)
2.4.4 The optimal weekly yard waste ANN prediction models

In order to obtain the optimal ANN model structure, the two best yard waste models (Temp and Temp-Pop) with different number in neurons of the hidden layer are investigated using the MSE metric. At the training stage, there are negligible differences between the numbers of neurons on model performance for both models (Figure 2-5). The MSE of the models is however much more sensitive to the number of neurons at the testing stage. The optimized number of neurons are 10 and 11 for Temp and Temp-Pop models, respectively. Larger MSE differences between training and testing stages are observed in the single-variable Temp model. For example, the MSE at testing is 4 times higher than the MSE at training using 8 neurons in the hidden layer. The MSE differences between training and testing stages for Temp-Pop model generally differ by a factor of 2.

Performance indices of the Temp and Temp-Pop models at their respective optimal number of neurons are tabulated in Table 2-5. Temp-Pop model with a 2-11-1 structure provides the best yard waste estimates of the two models (Temp and Temp-Pop), with the MSE at testing of 27,425 for Temp-Pop. The optimization of the Temp-Pop model structure is found beneficial at both the training and testing stages. For example, the MSE of the Temp-Pop model at both stages after optimization was reduced by 3.2% for the training stage and 2.1% for the testing stage (Table 2-4 and Table 2-5). The MAPE and R Value (18.72% and 0.84, respectively) of the optimized Temp-Pop model at testing compares favorably to other published models (Abbasi and Hanandeh, 2016; Azadi and Karimi-Jashni, 2016; Kannangara et al., 2018).
Figure 2-5: A comparison of the two superior model’s (Temp and Temp-Pop) performance using MSE with a different number of neurons (Vu et al., 2019)
Table 2-5: Performance of the optimized yard waste models (Vu et al., 2019)

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Structure</th>
<th>MSE</th>
<th>MAPE</th>
<th>R Value</th>
<th>Lag Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Train/Test</td>
<td>Train/Test</td>
<td>Train/Test</td>
<td></td>
</tr>
<tr>
<td>Temp</td>
<td>1-10-1</td>
<td>10.882/30,158</td>
<td>19.48/21.46</td>
<td>0.92/0.85</td>
<td>11</td>
</tr>
<tr>
<td>Temp-Pop</td>
<td>2-11-1</td>
<td>15,363/27,425</td>
<td>21.46/18.72</td>
<td>0.88/0.84</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 2-6 graphically compares the predicted and observed yard waste using the optimized Temp-Pop model at both the training and testing stages. Figure 2-6a shows a training period from weeks 10/1/2006 to 9/30/2007. The Temp-Pop model appears to capture the observed yard waste data well, including the scattering of data during the December holiday and the single generation peak in early spring. Although the position of the peak at week 3/25/2007 is relatively accurate, the magnitude of the peak prediction is off by 17%. Figure 2-6b shows the optimized model results from a testing period from weeks 4/2/2017 to 4/1/2018. Compared to the training stage, the prediction curve in Figure 2-6b appears smoother or less responsive to the actual field data. The location of the peak at early spring is again well captured. The model is however less useful at the prediction of outlier weeks (see, for example, data on 8/27/2017, 9/3/2017, 1/7/2018, 2/11/2018, and 3/11/2018). The model tends to slightly overestimate the actual yard waste data, particularly from weeks 6/18/2017 to 8/20/2017. However, given the inherent difficulties associated with waste forecasting, the proposed time-lag ANN modeling approach is promising.
Figure 2-6: Performance of the optimized Temp-Pop ANN model (a) Training stage from weeks 10/1/2006 to 9/30/2007; (b) Testing stage from weeks 4/2/2017 to 4/01/2018 (Vu et al., 2019)
In this Chapter, ANN time series with external inputs was used to predict weekly yard waste generation in Austin, Texas. Examination of how the lag time affect on yard waste prediction models was implemented. The optimal lag time of all models was found ranging from 1 to 5, except for the Temp model. The climatic group had more accurate and precise results in comparison with those of either the socio-economic or hybrid group. The limitation of this study is to require continuous data points which are not always available in any city. This study can be applied for predict E-waste prediction, forecast landfill gas generation and leachate generation in the future.
3. Parameter Interrelationships in a Dual Phase Collection Model

GIS-based studies solely focused on either phase (pre-collection phase or collection phase) have provided us with many valuable insights on search and optimization techniques. However, they have failed to deliver a comprehensive solution to the complete waste collection system. Due to the interconnectivities of the waste collection point distribution and the route design (Khan and Samadder, 2016; Erfani et al., 2017), studies considering both stages are more appropriate from a practical standpoint. Unlike industrialized countries utilizing permanent structures (or transfer stations) for temporary waste storage, many Asian countries relied on TCPs located at the side of streets. The size, number and spatial distribution of these TCPs are constantly evolving, responding elastically to the needs of the community and the constraints of infrastructural needs. Because of a higher degree of freedom in TCPs assignment, the use of GIS tools on waste management is particularly advantageous in developing countries utilizing dual-phase collection system.

3.1 Objectives of Chapter Three

Unlike other GIS based waste studies, this Chapter focuses on the interrelationships between the parameters in a dual-phase collection system by explicitly considering scenarios with different numbers of TCPs and source-to-TCP distances. The objectives of are: (i) to develop a simple GIS-based approach to minimize system costs considering both pre-collection and collection stages in a Vietnamese city, and (ii) to examine the interrelationships between the number of TCPs and the maximum source-to-TCP distance affects the waste system’s cost. In the present study, both of the number and distribution of TCPs are model variables and are determined by the optimization
process. TCPs in this study referring to the temporary collection points located on the side of streets and back alleys of the neighborhood where handcarts gather to meet a compactor truck. These TCPs are not physical buildings and are often mobile. TCPs in this study should not be confused with permanent waste facilities such as a waste transfer station in industrial countries.

3.2 Study area and existing waste collection system

Hai Phong is the third largest city in Vietnam with a population of 1,837,000 and an area of 1,530 km² (Hai Phong Information and Communication Center, 2009). The study area was located in the Hai Phong’s city centre including 5 wards with a population of 28,450, and covering an area of 1.41 km² (URENCO Hai Phong, 2016). The waste generation rate in the study area was about 16,200 tonnes per year (URENCO Hai Phong, 2016). The existing waste collection system of the study area consists of both pre-collection and collection stages. In the first stage, waste within the study area is collected daily from each source, including households, small shops, and institutions, by handcarts with an average capacity of 0.5m³. A typical pre-collector worker using a handcart in the study area is responsible for a stretch of road with an average length of 400 m per shift. When the handcarts are full, they are moved and emptied at the 10 TCPs located on the side of streets (represented by the circles in Figure 3-1). Five truck routes (including both afternoon and evening shifts) are used in the collection phase to transport the materials daily from this mixed residential, commercial, and institutional neighborhood TCPs to the final disposal sites. The capacity of a single truck used for the collection phase was approximately 10m³. The daily total travel distance of the trucks is
about 119.8 km, and the total collection and transportation time for the operation is about 8 hours per day (URENCO Hai Phong, 2016).
Figure 3-1: Map of the study area (After OpenStreetMap)
3.3 Methodology

In typical waste collection phase studies, the total system cost decreases with the decrease in the total number of bins or TCPs (Karadimas and Loumos, 2008). Given the flexibilities of the TCP assignments in the study area and the interconnectivities of the model parameters with the total system costs, a simple model using Network Analysis in ESRI ArcGIS (Version 10.3) was proposed. The proposed model consisted of three components: (i) a location-allocation problem for the distributions of TCPs using different maximum source-to-TCP distances ($D_{\text{max}}$); (ii) a vehicle route problem for handcarts in the pre-collection phase; and (iii) a vehicle route problem for trucks in the collection phase. Figure 3-2 presents a methodological flowchart of how these three components interact within this study. Using this proposed model, different scenarios were studied to explore the interrelationships of the model parameters. For each scenario, assignment of TCPs and the corresponding handcart routes were optimized. Truck route problems were then solved to minimize the system costs. Finally, comparisons among the different scenarios with respect to system costs were discussed.

A total of 30 scenarios were considered in this study (Figure 3-2). For each maximum distance set ($D_{\text{max}} = 300 \text{ m}, 400 \text{ m}, \text{ and } 500 \text{ m}$), a total of ten scenarios were considered with different numbers of TCPs (TCP = 7, 8, 10, 11, 13, 15, 16, 19, 22, and 25). The wide range of TCPs (7-25) was selected by applying factors to the benchmark values to cover any reasonable network uncertainties of the system. Section 2.2 will provide further details to explain how the number of TCPs was calculated and $D_{\text{max}}$ set was selected.
Figure 3-2: Diagram of Location-Allocation and VRP solutions using GIS (Vu et al., 2018b)
3.3.1 Geodatabase development and field data

Field data such as the road network, number, location, and type (household, commercial house, and institution) of waste generation source, waste generation rates, and candidate waste TCPs were used to develop the Geodatabase (Figure 3-2). Roads in the study area were divided into 4 categories, namely secondary roads (road width between 14 m to 21 m and with daily traffic flow between 20,000 to 30,000 vehicles), tertiary roads (road width between 7 m to 14 m and with daily traffic flow between 10,000 to 20,000 vehicles), residential roads (road width between 2.75 m to 6.5 m; with daily traffic flow less than 10,000 vehicles), and alleys (road width between 1.5 m to 2.75 m and with low traffic flow) (Ministry of Construction, 2007). The alleys accounted for about 30.67% of the total length of road network in the study area, while the other three road categories made up around 69.33% (URENCO Hai Phong, 2016). Trucks cannot access any alleys due to the large truck width compared to the small alley width, and only handcarts can access the alleys. Two road network from Open Street Map were developed: one for handcarts and one for trucks. The complete handcart road network consists of all of the above roads, however, alleys were almost exclusively used by handcart collectors in the pre-collection phase to transport materials from sources to the TCPs in practice while truck road network only contained the first three types of roads with the width larger than 2.75 m. The study area has a high population density (20,177 people/km²) (URENCO Hai Phong, 2016). Residential houses were built along main streets and alleys. A typical unit in the main street consisted of two to four floors of which the first floor was used for commercial purposes and the up floors for living while the unit in alleys was just for living purposes only. The average distance between two
doors of two units in the study area was about 4 to 5 m (URENCO Hai Phong, 2016) and there was no space available for placing garbage bins along streets or alleys. A waste generation site was assumed to be located at the front door of each unit. People stored waste in their own bins placed in their units and deposited the waste into a handcart when it passed through their area every day.

The waste generation sources were identified using ArcGIS (ArcCatalog, ArcMap, version 10.3.1) with World Imagery Basemap. There were 5,620 waste generation sites in the study area. The locations of TCPs and other geographical features were determined by a handheld GPS tool (Garmin eTrax10). All the input parameters were collected from government reports and verified by a field study in the summer of 2017.

3.3.2 Location-allocation of TCPs

Maximize coverage and minimize facility tools of GIS were used to determine the optimal TCP locations with respect to service coverage. The maximize coverage tool helps to identify TCPs that covered the most waste generation sites, whereas the minimize facility tool helps to identify the least amount of TCPs (Esri, 2015). The candidate TCPs were selected at locations with sufficient space to accommodate the temporary storage of wastes assuming an average waste density of 380kg/m³. All candidate TCPs were located at secondary, tertiary, and residential roads to facilitate rapid waste transfer. A total of 25 candidate TCPs were inputted to the model, of which 10 were existing TCPs and 15 were additional sites. The benchmark number of TCPs was estimated by the following equation.

\[
N_{cp} = \frac{H \cdot P \cdot W_k + I \cdot W_l}{D_w \cdot V \cdot N_{ha} \cdot f}
\]  

(3-1)
Where:

\(N_{tcp}\): Number of TCPs needed on a given day, \(0 < N_{tcp}\): integer \(\leq N_{ctcp} = 25\) (\(N_{ctcp}\): Number of candidate TCP)

\(H\): Number of households and commercial units to be served in the study area (\(H = 5417\) units)

\(I\): Number of institutions to be served in the study area (\(I = 203\) units)

\(P\): Average number of people in a household/commercial unit (\(P = 5.25\) people/unit)

\(W_h\): Daily average household/commercial waste generation rate (\(W_h = 1.2\) kg/capita∙day)

\(W_i\): Daily average institution waste generation rate (\(W_i = 34.3\) kg/institution∙day)

\(D_w\): Average density of waste in a handcart (\(D_w = 380\) kg/m\(^3\))

\(V\): Volume of a handcart (\(V = 0.5\) m\(^3\))

\(N_{ha}\): Average number of handcarts served in a TCP (\(N_{ha} = 14\))

\(f\): Average handcart collection frequency (\(f = 1\) day\(^{-1}\)).

The benchmark TCP number from equation 3-1 is varied (\(\pm 20\%\), \(\pm 40\%\) and \(\pm 60\%\)) to develop six additional scenarios for analyses using the maximize facilities tool. A range of \(\pm 20\%\) was commonly used in many environmental studies (Arribas et al., 2010; Ficklin et al., 2009; Rasouli et al., 2014; Tabari and Talaee, 2014), however, a larger range was used in this study to further explore the interrelationships between the
variables. In addition, the solutions from the GIS minimize facility tool (TCP=8, TCP=11, TCP=15) were also adopted as special scenarios for comparison purposes.

Optimal source-to-temporary collection point distances appear to be site-specific, ranging from 75 m to 600 m (Alvarez et al., 2008; Boskovic and Jovicic, 2015; Gallardo et al., 2015; Kanchanabhan et al., 2010; Khan and Samadder, 2016; Valeo et al., 1998; Vijay et al. 2005). The lower range distances reported in the literature were unlikely to be appropriate at the study site as the number of potential TCPs were limited due to physical constraints. A 400 m maximum distance ($D_{\text{max}}$) was selected as the benchmark to better reflect the local conditions, as a typical handcart worker in Hong Bang district is responsible for 400 m per shift. Variations of ±25% were applied to the benchmark distance, thus $D_{\text{max}}$ of 300 m, 400 m and 500 m was selected for use in the location-allocation of TCPs.

### 3.3.3 Handcart route optimization

The Vehicle Route Problem (VRP) module from the GIS Network Analysis program was used to optimize handcart routes for all 30 scenarios in the pre-collection phase. Inputs needed for the handcart VRPs consisted of the handcart road network data set, waste generation data, TCP characteristics, and route class (Figure 3-2). Inside the route class, information such as handcart start depot location, handcart end depot location, handcart capacity, maximum number of waste generation sources which can be served by one handcart, maximum total collection time for a handcart route per day, maximum travel time, and maximum travel distance of a handcart were included.

The network for handcarts consisted of the entire road network, including narrower alleys (width < 2.75m) which are not accessible to trucks. Restrictions on U-turns, one-
way lanes, and vehicle height were not imposed for handcarts as they often travel on sidewalks on one-way roads, make U-turns indiscriminately, and face no practical height restrictions (full height < 1.5 m). Unlike trucks, the average speed of handcarts was modified to walking speed (3km/hr).

The capacity of a handcart, maximum number of waste generation sources which can be served by one handcart (MWS_H), maximum total collection time for a handcart route per day and maximum travel distance of a handcart were measured from a field study in the summer of 2017. Capacity of a handcart ranged from 0.35 m$^3$ to 1.0 m$^3$ (URENCO Hai Phong, 2016), with an average capacity of 0.5 m$^3$. Maximum number of waste generation sources which can be served by one handcart (MWS_H) was calculated based on handcart capacity and average waste generation rate at a generation site. The optimized TCPs from the location-allocation solution, in all scenarios, were also assumed to be the depots where the handcarts would be stored overnight.

Waste generation data such as the volume of waste, pick up time, and time windows at each source were inputted as factors. The average pick up time at sources was measured from field data. A time window of 8 working hours per day was assumed for handcarts. Time windows were then chosen to be from 12:30 to 20:30. Parameters and constraints for handcart route optimization are summarized in Table 3-1.

### 3.3.4 Truck route problem and constraints

In addition to the handcart VRP for the pre-collection phase, a VRP was also applied to the collection phase to optimize truck routes and schedules for waste transportation between TCPs and the Trang Cat landfill (Figure 3-1). Geodatabase for the truck VRP included study area map, truck road network data set, waste TCP characteristics, and
route class which included the start depot location, end depot location, truck capacity, maximum travel distance, maximum total time, maximum travel time, and maximum TCP count. A road network data set was built using the road map from OpenStreetMap. The network for trucks included only roads that are accessible to trucks (>2.75 m). Input data consisted of waste volume, pickup time and time windows at each optimized TCP obtained from the location-allocation solution (Figure 3-2).

The volume of waste (in trucks) at TCP$i$ was calculated using equation 3-2.

$$V_i = \frac{(H_i \cdot G_{hi} + I_i \cdot G_{ii}) \cdot k_c}{f}$$

(3-2)

Where:

$V_i$: Volume of waste (in trucks) at the temporary collection point i (L)

$H_i$: Number of household and commercial houses served at temporary collection point i

$I_i$: Number of institutions served at temporary collection point i

$G_{hi}$: Average household/commercial waste generation rate ($G_{hi}= 16$ L/unit·day)

$G_{ii}$: Average institution waste generation rate ($G_{ii}= 90.4$ L/institution·day)

$k_c$: Waste compaction ratio in trucks ($k_c= 2.5$)

$f$: Collection frequency ($f= 1$ day$^{-1}$)

Pickup time of a truck at a TCP ($T_p$) was calculated by the equation below:

$$T_p = N_{hi} \cdot t + T_w$$

(3-3)

Where:

$N_{hi}$: Number of handcarts served at temporary collection point i
t: Average time for a handcart to unload waste to truck (t = 1 minute)

T_w: Average idling time for a truck at temporary collection point (T_w = 10 minutes)

Parameters and constraints for truck route optimization were given in Table 3-1. Optimization of truck routes was implemented for all 30 scenarios above.
Table 3-1 Model parameters and constrains (Vu et al., 2018b)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameters</th>
<th>Description</th>
<th>Unit</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Handcart</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS_H</td>
<td>Start Depot Name</td>
<td>The handcart starts at the temporary collection point</td>
<td>Temporary collection point</td>
<td></td>
</tr>
<tr>
<td>DE_H</td>
<td>End Depot Name</td>
<td>The handcart returns to the temporary collection point at the end of the route</td>
<td>Temporary collection point</td>
<td></td>
</tr>
<tr>
<td>CH_H</td>
<td>Handcart capacity</td>
<td>The handcart can load a maximum of 0.5 m³ of waste</td>
<td>m³</td>
<td>0.5</td>
</tr>
<tr>
<td>MWS_H</td>
<td>Maximum number of waste generation sources which can be served by one handcart</td>
<td>The handcart can fully load waste at maximum of 31 sources</td>
<td>Source</td>
<td>31</td>
</tr>
<tr>
<td>TP_H</td>
<td>Pickup time</td>
<td>Time the handcart spent to pick up waste at a household/institution</td>
<td>minute</td>
<td>0.75-10.00</td>
</tr>
<tr>
<td>WGR_H</td>
<td>Waste generation rate</td>
<td>Waste generation rate at a source</td>
<td>(L/source·day)</td>
<td>16.00-356.06</td>
</tr>
<tr>
<td>TW_H</td>
<td>Time windows</td>
<td>Started pick up time and Ended pick up time at a final source in a route</td>
<td></td>
<td>12:30:00 to 20:30:00</td>
</tr>
<tr>
<td>TMT_H</td>
<td>Maximum total time</td>
<td>Workers cannot work 4 hours per shift</td>
<td>minute</td>
<td>240</td>
</tr>
<tr>
<td>TMTR_H</td>
<td>Maximum travel time</td>
<td>The road with lots of traffic and dust, thus the worker should not spend more than 2 hours walking on the road</td>
<td>minute</td>
<td>120</td>
</tr>
<tr>
<td>DMT_H</td>
<td>Maximum travel distance</td>
<td>The road with lots of traffic and dust, thus the worker should not spend more than 1.5 km walking on the road</td>
<td>km</td>
<td>1.5</td>
</tr>
<tr>
<td><strong>Truck</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DS_T</td>
<td>Start Depot Name</td>
<td>The truck starts at the landfill</td>
<td>Trang Cat landfill</td>
<td></td>
</tr>
<tr>
<td>DE_T</td>
<td>End Depot Name</td>
<td>The truck returns to the landfill at the end of the route</td>
<td>Trang Cat landfill</td>
<td></td>
</tr>
<tr>
<td>CH_T</td>
<td>Truck capacity</td>
<td>The truck can load a maximum of 10 m³ of waste</td>
<td>m³</td>
<td>10</td>
</tr>
<tr>
<td>MCS_T</td>
<td>Max collection source count</td>
<td>The truck can pick up waste at maximum of 5 temporary collection points</td>
<td>point</td>
<td>5</td>
</tr>
<tr>
<td>TP_T</td>
<td>Pickup time</td>
<td>Time the handcart spent to pick up waste at a household/institution</td>
<td>minute</td>
<td>Depends on volume of waste at each selected temporary collection point</td>
</tr>
<tr>
<td>QW_T</td>
<td>Quantity of waste</td>
<td>Quantity of waste that the truck picks up at each temporary collection point</td>
<td>L</td>
<td>Depends on volume of waste at each selected temporary collection point</td>
</tr>
<tr>
<td>TW_T</td>
<td>Time windows</td>
<td>Started pick up time and Ended pick up time at a final source in a route</td>
<td></td>
<td>12:30:00 to 22:30:00</td>
</tr>
<tr>
<td>TMT_T</td>
<td>Maximum total time</td>
<td>Workers cannot work 5 hours per shift</td>
<td>minute</td>
<td>300</td>
</tr>
<tr>
<td>TMTR_T</td>
<td>Maximum travel time</td>
<td>The truck should not spend more than two hours (120 minutes) driving on the streets</td>
<td>minute</td>
<td>120</td>
</tr>
<tr>
<td>DMT_T</td>
<td>Maximum travel distance</td>
<td>Due to operation and maintenance reasons the truck should not travel more than 30 km</td>
<td>km</td>
<td>30</td>
</tr>
</tbody>
</table>
3.3.5 Comparison among different scenarios using cost estimates

The capital, operation, and maintenance costs for the pre-collection and collection system were used as performance indicators. The topography of the study area is relatively flat, therefore fuel and operational costs were assumed to be linearly related to the travel distance. The 2014 repair and maintenance rates were used in this study (Ministry of Construction, 2014) and the travel costs between the study area and the landfill (Figure 3-1) was included. Non-compounding interest rates were assumed, and they were expressed daily (for equivalent uniform daily cost), monthly (labour and fuel costs), or yearly (repair and maintenance). Total daily waste collection cost \( C_t \) for each scenario was based on the following equations:

\[
C_t = C_h + C_{tr}
\]  \hspace{1cm} (3-4)

Where:

\( C_h \): Daily capital, operation, and maintenance costs for handcarts (USD/day)

\( C_{tr} \): Daily capital, operation, and maintenance costs for trucks (USD/day)

\[
C_h = \left[ (N_{tha} \cdot P_h) + (PW_{h-rep} \cdot N_{tha} \cdot P_h \cdot n_h) + (30.4 \cdot PW_{h-lab} \cdot W \cdot S_h) \right] \cdot CRF_h
\]  \hspace{1cm} (3-5)

Where:

\( N_{tha} \): Total number of handcarts used in a day

\( P_h \): Capital cost of a 0.5 m\(^3\) handcart (US$132/handcart)

\( n_h \): Standard repair and maintenance rate of a handcart: 15% per year (Ministry of Construction, 2014)
W: Total work hours of the handcart workers within the study area (total travel time plus the total pickup time) (hours/day)

$S_h$: Average salary and benefits of a handcart operator (US$2.79/hour)

$PW_{h\text{-rep}}$: series present worth factor for handcart repair and maintenance: 1.690 ($i=12\%$, $n=2$ year). Average numbers of day in a month: 30.4.

$PW_{h\text{-lab}}$: series present worth factor for handcart labour: 21.243 ($i=1\%$, $n=24$ months)

$CRF_h$: Capital recovery factor for handcarts: 0.00154 ($i=0.0329\%$, $n=730$ days)

\[
C_{tr} = \left[ 2 \cdot P_{tr} + \left( 2 \cdot PW_{tr\text{-rep}} \cdot P_{tr} \cdot n_{tr} \right) + \left( 30.4 \cdot PW_{tr\text{-gas}} \cdot D \cdot F \right) + \left( 30.4 \cdot PW_{tr\text{-lab}} \cdot D \cdot S_{tr} \right) \right] \cdot CRF_{tr} \tag{6}
\]

Where:

$P_{tr}$: Capital cost of a 10 m$^3$ truck (US$43,264/truck)

$n_{tr}$: Standard repair and maintenance rate of a truck: 15% per year (Ministry of Construction, 2014)

$D$: Total travel distance of all trucks (km/day)

$F$: Typical fuel cost (US$0.84/km)

$S_{tr}$: Total salary for the truck drivers (US$3.13/km)

$PW_{tr\text{-rep}}$: series present worth factor for truck repair and maintenance: 4.111 ($i=12\%/year$, $n=6$ year)

$PW_{tr\text{-gas}}$: series present worth factor for truck fuel: 51.150 ($i=1\%/month$, $n=72$ months). Average numbers of day in a month: 30.4.

$PW_{tr\text{-lab}}$: series present worth factor for truck labour: 51.150 ($i=1\%/month$, $n=72$ months). Average numbers of day in a month: 30.4.

$CRF_h$: Capital recovery factor for trucks: 0.00154 ($i=0.0329\%$, $n=2,191$ days)
3.4 Results and discussion

3.4.1 Waste temporary collection points for different scenarios

Selected location-allocation results using the maximize coverage and minimize facility tools are shown on Figure 3-3. A total of 25 candidate TCPs were used for the optimization, as shown on Figure 3-3a. Figure 3-3b shows the minimize facility results for TCP=15 scenario with $D_{\text{max}}= 300$ m, while Figure 3-3c, 3-3d, and 3-3e showed the maximize coverage results for TCP=7 and a $D_{\text{max}}$ of 300m, 400m, and 500 m, respectively. Figure 3-3f shows the maximize coverage result for TCP= 25 scenario with $D_{\text{max}}= 300$ m. Comparison between Figure 3-3b and 3-3c showed the optimized TCP solution either relocated (“C21” at the southwest corner on Figure 3-3b moved to “C4” on Figure 3-3c), or merged together (“C9,” “C12,” “C15”, and “C16” at the northwest corner on Figure 3-3b combined into a mega “C16” on Figure 3-3c) as TCPs decreases under the same $D_{\text{max}}$. On the other hand, fixing the number of TCPs while changing $D_{\text{max}}$ also changed the spatial distributions of optimized TCPs. Figures 3-3c, 3-3d, and 3-3e showed noticeable differences in the TCP distribution. For example, only 3 out of 7 (“C4,” “C8,” and “C18”) were not affected by the variations (±25%) in $D_{\text{max}}$. Since generation sites were not uniformly distributed across the study area, changes in TCP locations led to changes in the volume of waste collected at each TCP.
Figure 3-3: Results of Location-Allocation: a) candidate temporary collection points; b) minimize facility function with $D_{\text{max}}$ of 300 m; c) maximize coverage function with the number of temporary collection points equal to 7 and $D_{\text{max}}$ of 300 m, d) 400 m, and e) 500 m; f) maximize coverage with number of temporary collection points of 25 and $D_{\text{max}}$ of 300 m (Vu et al., 2018b)

Figure 3-4 shows the service coverage with respect to the number of TCP for all 30 scenarios. It appeared that both the $D_{\text{max}}$ and number of TCPs have a positive impact.
on the percentage of coverage. $D_{\text{max}} = 500$ m had the highest coverage (99.5-99.9%), followed by $D_{\text{max}} = 400$ m (92.9-96.8%), and $D_{\text{max}} = 300$ m (69.4-84.4%). This could be explained by the larger handcart service area associated with $D_{\text{max}}$ of 500 m. However, due to the increased overlapping of the coverage area and more of the coverage area falling outside of the study area with a higher $D_{\text{max}}$, coverage improved more significantly from $D_{\text{max}} = 300$ m to $D_{\text{max}} = 400$ m, than from $D_{\text{max}} = 400$ m to $D_{\text{max}} = 500$ m.

The percentage of coverage increases with the number of TCPs until reaching TCP= 15, TCP= 11, and TCP= 8 for the curves of $D_{\text{max}} = 300$ m, 400 m, and 500 m, respectively. The results were consistent with the minimize facility solution obtained from different $D_{\text{max}}$ values, suggesting their usefulness in optimization. The percentage of coverage remained constant after reaching the optimal TCP numbers due to the spatial constraints on the assignment of candidate TCPs. For example, there was a large gap (more than 600m) between the TCPs “C9” and “C10” (Figure 3-3a) and no additional candidate TCP was allowed because of the existing infrastructures and the limited physical space in this area. As such, some generation sources were not covered, even for the scenarios with $D_{\text{max}} = 500$ m (99.9% coverage). The results suggest that the coverage is more sensitive to the spatial distribution of candidate TCPs than the number of TCPs or $D_{\text{max}}$. This finding regarding the importance of candidate TCPs highlights one of the potential benefits of the GIS minimize facility solution.
Figure 3-4: Percentage of coverage with respect to the number of sources and the maximum distances to be covered by handcarts (Vu et al., 2018b)
3.4.2 VRP Solution for handcarts

The outputs of the VRP solution for handcarts represented the shortest travel distances between TCPs and the sources. It was found that the total handcart travel distance dropped with an increasing number of TCP (Table 3-2). Since TCPs are the final destination for handcart operators, more TCPs reduce the average handcart travel distance.

The total handcarts travel distance associated with the TCP distribution of $D_{\text{max}} = 500$ m seemed to be the lowest with an average total handcart travel distance of 160.1 km, followed by $D_{\text{max}} = 300$ m (160.3 km) and $D_{\text{max}} = 400$ m (160.9 km) as shown in Table 3-2. The highest percentage of coverage associated with $D_{\text{max}} = 500$ m (Figure 3-4) may lead to shorter handcart travel distances. It is, however, unclear why the scenario with $D_{\text{max}} = 400$ m had the highest average handcart travel distances. A closer look at the $D_{\text{max}} = 400$ m data reveal that the travel distances vary more when compared to the other cases (with standard deviations of 9.54, compared to 8.65 and 9.05 for $D_{\text{max}} = 300$ m and $D_{\text{max}} = 500$ m respectively). The average handcart travel distances among all 211 routes were 724.4 m, 728.1 m, and 725.4 m with $D_{\text{max}} = 500$ m, $D_{\text{max}} = 400$ m, and $D_{\text{max}} = 300$ m). These results suggest that the handcart travel distances were more sensitive to the TCP spatial distribution associated with $D_{\text{max}} = 400$ m. The average handcart travel distances in the current study were approximately 3.8 times lower than the average tricycle travel distances in Kanchanabhan et al.’s study in India (2010). This perhaps, because the population density of the current study was higher than that of the Kanchanabhan et al.’s study (2010).
The numbers of handcarts were found to be insensitive to the spatial distribution TCP, with an average value of 221 handcarts for different $D_{\text{max}}$ considered (Table 3-2). The required number of handcart increased slightly with the number of TCPs. It should be noted that the handcarts were used daily for a single shift, as the collection trucks visited the TCPs only once a day.
Table 3-2: Results of VRP for trucks and handcarts (Vu et al., 2018b)

<table>
<thead>
<tr>
<th>Number of specified TCPs</th>
<th>Total handcart travel distance (km)</th>
<th>Number of handcarts</th>
<th>Total truck travel distance (km)</th>
<th>Number of truck</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D300 m</td>
<td>D400 m</td>
<td>D500 m</td>
<td>D300 m</td>
</tr>
<tr>
<td>7</td>
<td>178.0</td>
<td>179.0</td>
<td>178.4</td>
<td>220</td>
</tr>
<tr>
<td>8</td>
<td>171.3</td>
<td>173.1</td>
<td>173.7</td>
<td>221</td>
</tr>
<tr>
<td>10</td>
<td>163.0</td>
<td>167.3</td>
<td>161.8</td>
<td>222</td>
</tr>
<tr>
<td>11</td>
<td>161.0</td>
<td>163.9</td>
<td>158.3</td>
<td>222</td>
</tr>
<tr>
<td>13</td>
<td>157.9</td>
<td>156.9</td>
<td>158.2</td>
<td>220</td>
</tr>
<tr>
<td>15</td>
<td>156.9</td>
<td>154.5</td>
<td>157.2</td>
<td>222</td>
</tr>
<tr>
<td>16</td>
<td>155.2</td>
<td>156.1</td>
<td>156.0</td>
<td>222</td>
</tr>
<tr>
<td>19</td>
<td>154.6</td>
<td>153.1</td>
<td>153.3</td>
<td>222</td>
</tr>
<tr>
<td>22</td>
<td>152.5</td>
<td>151.8</td>
<td>152.0</td>
<td>224</td>
</tr>
<tr>
<td>25</td>
<td>152.0</td>
<td>151.3</td>
<td>151.7</td>
<td>225</td>
</tr>
<tr>
<td>Mean</td>
<td>160.3</td>
<td>160.9</td>
<td>160.1</td>
<td>222</td>
</tr>
<tr>
<td>STDEV</td>
<td>8.65</td>
<td>9.54</td>
<td>9.05</td>
<td>1.56</td>
</tr>
</tbody>
</table>
3.4.3 VRP Solution for the trucks

The VRP solution for trucks was developed. Table 3-2 shows the results of the total travel distances of trucks. In all cases, two trucks are found as adequate for the study area (Figure 3-1) which is identical to the number of trucks of the existing system.

The total truck travel distances generally increased with the number of TCPs, as shown in Table 3-2. Average total travel distances of trucks for different $D_{\text{max}}$ (300 m, 400 m and 500 m) were similar, ranged from 104.1 km to 104.2 km. Unlike the handcart cases, a constant decreasing trend of travel distance with TCP numbers was not observed. Instead, the truck travel distances fluctuated with the number of TCPs. For example, the travel distances for $D_{\text{max}}= 400$ m at TCP= 13 (105.6 km) was higher than TCP= 11 and TCP= 15 (103.5 and 103.3 km, respectively). It was found that truck capacity ultimately influenced the truck travel distances and the overall efficiency of the collection system.

Table 3-3 reports the TCP servicing sequence for each route and the changes of truck waste volume with respect to different TCPs for the benchmark case ($D_{\text{max}}= 400$ m). Smaller trucks (10,000 L) were used in the study area, and the waste mass balance from the simulations indicated that the truck capacity was not fully utilized on all routes. The number of TCPs served per route was not consistent, ranged from 2 to 5. The truck volume utilization rates shown in Table 3-3 ranged from 42.8% to 99.9% in different scenarios and different truck routes. However, the overall percentage of truck space utilization averaged for each of the 5 trucks was similar for each of the three scenarios (84.0%). The results suggest that larger trucks or having a varied truck size may improve the collection efficiency by increasing the average truck space utilization. However, these two suggestions may not be compatible with the existing the system in the study area and
may leave it more vulnerable to overcapacity with any future increases in waste
generation rates.
Table 3-3: Pickup waste volume balance of the five truck routes in selected scenarios ($D_{max} = 400\text{ m}, \ TCP = 11, 13, \text{ and } 15$) (Vu et al., 2018b)

<table>
<thead>
<tr>
<th>TCP#</th>
<th>Pickup waste (L)</th>
<th>TCP#</th>
<th>Pickup waste (L)</th>
<th>TCP#</th>
<th>Pickup waste (L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Route 1 (truck 10,000 L)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1,659.32</td>
<td>6</td>
<td>5,101.99</td>
<td>12</td>
<td>2,729.46</td>
</tr>
<tr>
<td>10</td>
<td>2,617.88</td>
<td>4</td>
<td>3,211.27</td>
<td>16</td>
<td>1,404.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>1,350.40</td>
<td>15</td>
<td>591.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>1,037.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>4,075.16</td>
</tr>
<tr>
<td>Total</td>
<td>4,277.20</td>
<td>Total</td>
<td>9,663.66</td>
<td>Total</td>
<td>9,838.45</td>
</tr>
<tr>
<td>Route 2 (truck 10,000 L)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>6,102.28</td>
<td>12</td>
<td>3,681.42</td>
<td>4</td>
<td>3,211.27</td>
</tr>
<tr>
<td>1</td>
<td>2,675.07</td>
<td>11</td>
<td>2,094.38</td>
<td>2</td>
<td>2,303.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>4,075.16</td>
<td>1</td>
<td>1,350.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2,916.61</td>
</tr>
<tr>
<td>Total</td>
<td>8,393.68</td>
<td>Total</td>
<td>9,850.97</td>
<td>Total</td>
<td>9,782.15</td>
</tr>
<tr>
<td>Route 3 (truck 10,000 L)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4,209.67</td>
<td>2</td>
<td>2,297.47</td>
<td>10</td>
<td>1,659.32</td>
</tr>
<tr>
<td>8</td>
<td>2,094.38</td>
<td>3</td>
<td>2,910.21</td>
<td>13</td>
<td>2,617.88</td>
</tr>
<tr>
<td>9</td>
<td>3,681.42</td>
<td>5</td>
<td>3,979.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9,985.48</td>
<td>Total</td>
<td>9,186.95</td>
<td>Total</td>
<td>4,277.20</td>
</tr>
<tr>
<td>Route 4 (truck 10,000 L)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5,563.80</td>
<td>10</td>
<td>1,659.32</td>
<td>8</td>
<td>5,563.80</td>
</tr>
<tr>
<td>6</td>
<td>3,463.23</td>
<td>13</td>
<td>2,617.88</td>
<td>9</td>
<td>3,463.23</td>
</tr>
<tr>
<td>Total</td>
<td>9,027.04</td>
<td>Total</td>
<td>4,277.20</td>
<td>Total</td>
<td>9,027.04</td>
</tr>
<tr>
<td>Route 5 (truck 10,000 L)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4,075.16</td>
<td>8</td>
<td>5,563.80</td>
<td>5</td>
<td>3,979.27</td>
</tr>
<tr>
<td>3</td>
<td>5,863.59</td>
<td>9</td>
<td>3,463.23</td>
<td>6</td>
<td>5,101.99</td>
</tr>
<tr>
<td>Total</td>
<td>9,938.76</td>
<td>Total</td>
<td>9,027.04</td>
<td>Total</td>
<td>9,081.26</td>
</tr>
</tbody>
</table>
3.3.4 Effects of the number and spatial distribution of TCPs on handcart and truck costs

The capital, operating, and maintenance costs of handcarts and trucks are shown in Figure 5-5a. In both cases, the costs were sensitive to both the number and spatial distribution of TCPs. The handcart cost curves generally decrease with the numbers of TCPs as less travel distance was required for each pick up service (Table 3-2). The handcart cost curves also slightly increased from TCP= 19 to TCP= 25, probably due to the fact that more handcarts were used in scenarios with higher TCP numbers (Table 3-2) and a larger number of handcarts increases the total cost estimates (“Ntha” in equation 3-5). Compared to the truck costs, the handcart costs were less sensitive to the TCP spatial distribution and the handcart cost curves for different Dmax were similar, especially for TCP>13.

The truck cost curves generally increased with the number of TCPs due to the fact that total truck travel distances increased when the number of TCPs increased. On average, the Dmax= 400 m case had highest truck collection costs ($1,048.8/day) due to it having a slightly higher truck travel distance (104.2km, see Table 3-2). There exists also local truck cost peaks at TCP= 13 (Dmax= 400 m) and TCP= 15 (Dmax= 500 m) corresponded well to the truck travel distance peaks for Dmax= 400 m (TCP= 13, Table 3-2), and Dmax= 500 m (TCP= 15, Table 3-2). Local truck cost peak is not observed for Dmax= 300 m.
Figure 3-5: Comparison of (a) handcart and truck collection costs; (b) system collection cost among different scenarios (Vu et al., 2018b)
3.4.5 Total capital, operating, and maintenance costs

Similar to handcart and truck travel distances, total collection costs change with the number and spatial distribution of TCPs (Figure 3-5b). It is interesting to note that the collection costs associated with the special scenarios using the minimize facility module ($D_{\text{max}}= 300 \text{ m and TCP}= 15$; $D_{\text{max}}=400 \text{ m and TCP}= 11$; and $D_{\text{max}}= 500 \text{ m and TCP}= 8$) did not result in lower total collection cost. The lowest costs were identified at TCP= 13 ($D_{\text{max}}= 300 \text{ m}$), TCP=16 ($D_{\text{max}}= 400 \text{ m}$), and TCP= 11($D_{\text{max}}= 500 \text{ m}$). The results suggested that minimize facility module alone may not be able to minimize the system costs in dual phase collection system with flexible TCP assignments.

The average costs were similar between different $D_{\text{max}}$ scenarios after optimization. The scenarios associated with $D_{\text{max}}= 500 \text{ m}$ had the lowest minimum daily cost at USD $1,040.6, and the scenarios associated with $D_{\text{max}}= 300 \text{ m}$ had the highest minimum daily cost at USD $1041.5. Among the 30 scenarios optimized in this study, the case associated with $D_{\text{max}}= 500 \text{ m and TCP}= 11$ appeared to be the most economical given the constraints and the field conditions.

According to these optimization results, a total of 11 TCPs were found appropriate for the Hai Phong region (area= 1.41 km$^2$, population= 28,450 people, waste generation rate= 1.2kg/capita-day). Thus, on average, about 2,586 people within an area of 0.11 km$^2$ were served by each proposed TCP. Compared to the minimize facility solution, the recommended solution from the dual phase model reduces waste collection costs by about USD $9.57/day or USD $0.23/tonne. About 7.8 TCPs/km$^2$ was obtained in this study, slightly greater than that of Kanchanabhan et al.’s study (2010), with a reported value of 6.5 TCPs/km$^2$.  

80
Table 3-4 highlighted the potential benefits of the optimized scenario (TCP= 11 and $D_{\text{max}}= 500$ m) over the existing system. The optimized truck routes reduced the truck travel distances by 9.63% to 19.61% (Table 3-4) which also reduce the associated vehicle emission. The average total truck travel distances fell by 13.76%. The finding is consistent with other waste collection studies in Vietnam. For example, Nguyen et al. (2017) reported a 11.3% reduction in travel distance in Ha Giang city and Thanh et al.’s (2009) reported a 19% reduction in Can Tho city.

These savings were attained by minimizing truck route distances and the total number of TCPs visited per route (Figure 3-6). Figures 3-6a, 3-6c, 3-6e, 3-6g, and 3-6i show the existing truck routes 1 to 5 which are organized empirically. TCPs are visited by the trucks multiple times on a given day, especially in the south west corner of the study area. The proposed truck routes are shown on Figures 3-6b, 3-6d, 3-6f, 3-6h, and 3-6j. Multiple collection within a day was mostly eliminated and the route distance and total number of TCPs were greatly reduced.
Table 3-4: Truck travel distances compared between the optimized case ($D_{\text{max}} = 500$, TCP= 11) and the status quo (Vu et al., 2018b)

<table>
<thead>
<tr>
<th>Name of route</th>
<th>Total proposed travel distance of trucks (km)</th>
<th>Existing travel distance of trucks (km)</th>
<th>Reduction of travel distance of trucks (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Route 1</td>
<td>22.1</td>
<td>27.5</td>
<td>19.61%</td>
</tr>
<tr>
<td>Route 2</td>
<td>19.5</td>
<td>22.1</td>
<td>11.88%</td>
</tr>
<tr>
<td>Route 3</td>
<td>20.6</td>
<td>23.8</td>
<td>13.48%</td>
</tr>
<tr>
<td>Route 4</td>
<td>19.2</td>
<td>22.1</td>
<td>13.18%</td>
</tr>
<tr>
<td>Route 5</td>
<td>22.0</td>
<td>24.3</td>
<td>9.63%</td>
</tr>
<tr>
<td>Total</td>
<td>103.3</td>
<td>119.8</td>
<td>13.76%</td>
</tr>
</tbody>
</table>
Figure 3-6: Comparison of the existing truck routes (a, c, e, g, i) and the proposed truck routes (b, d, f, h j) for the optimized scenario ($D_{\text{max}}= 500$, TCP=11) (Vu et al., 2018b)
In Chapter 3, a simple GIS based waste collection was developed to minimize waste collection cost considering both pre-collection and collection phases. Location Allocation tool was applied to find the optimal number and locations of temporary collection points while Vehicle Routing Problem tool was used to obtain the optimal handcart and truck routes. The maximum distance from a temporary collection point to a source impacted on both pre-collection and collection phases. Scenario with the temporary collection point of 11 was found as the optimal scenario with the lowest collection cost. The study can be applied to waste collection systems in developing countries with narrow alleys, as well as to selection drop-off sites and transfer station sites in developed countries.
4. Model Optimization for Landfills located in Cold Semi-arid Climates

4.1 Objectives of Chapter Four

The objectives of Chapter Four are (i) to estimate and compare methane generation rates in Regina and Saskatoon landfills with the selected FOD models at default $k$ and $L_0$ or DOC values; (ii) to determine optimal sets of FOD model parameters by minimizing the residual sum of squares (RSS) values between the predicted and the actual gas rates; (iii) and to examine the effects of different iterative methods on the curve fitting process using mean percentage error.

4.2 Methodology

4.2.1 The landfills and the waste records

Regina and Saskatoon are both located in the province of Saskatchewan, Canada. Regina is located approximately 160 km from the US border (bordering Montana), while Saskatoon is located more centrally in the province, approximately 250 km northwest of Regina. The city's landfills are about 270 km apart and both are located in a cold semi-arid climate area. The climate in Regina and Saskatoon is classified as “Dfb” using the Koppen-Geiger classification system, implying a cold climate, with no dry season, and a warm summer (Peel et al., 2007). The average temperatures in Regina and Saskatoon are $3.1^\circ$C and $3.3^\circ$C, respectively (Canada Climate Normals, 2016). The annual precipitation in Regina and Saskatoon in the period from 1981 to 2010 are 389.7 mm and 353.7 mm, respectively (Canada Climate Normals, 2016).

The Regina landfill began operating in 1961, and served as a non-hazardous municipal waste landfill for Regina and surrounding areas. The landfill was capped with a 1m compacted clay liner and 0.15 m of topsoil (Conestoga-Rovers & Associates, 2006).
The total buried waste in the study area is approximately 3.1 million tonnes. The input waste mass was estimated using national statistics and local waste records from three separate periods due to changes in disposal practices and regulations. For example, construction and demolition waste was separated from the main disposal area in 1992 (Conestoga-Rovers & Associates, 2006; City of Regina, 1995-2011; Barlishen, 1996), thereby reducing inert materials and altering average methane generation potential of the buried waste. The 16 ha well field is located in the North part of the old disposal area. The Regina gas collection and management system consists of 27 vertical gas wells at a 15m depth, a condensation sump and a flare system. Real-time gas data was collected from July 2008 to December 2014. LFG from Regina and Saskatoon landfills was measured automatically via SCADA (Supervisory Control and Data Acquisition system) on a daily basis by a gas analyzer system. In Regina, about 44% of the collected landfill gas was CH$_4$, 37% was CO$_2$ and 19% was residual gas during the study period, and the average methane generation rate was about 4.17 m$^3$/min. Gas residuals were ignored in the gas modeling, and CH$_4$ and CO$_2$ fractions were adjusted accordingly. A CH$_4$ fraction of 54% was used as input for the Regina landfill.

The Saskatoon landfill has been serving the City of Saskatoon and nearby towns since 1955. The non-hazardous municipal landfill site was lined with a 0.45 m clay liner. From 1955 to 1981, waste disposed in the landfill was estimated based on population records and an average waste generation rate of 1.75 kg/capita, and from 1982 to 1997, waste disposed in the landfill was calculated based on a generation rate of 1.7 kg/capita. From 1998 to 2010, the mass of disposed waste was taken from landfill waste records (City of Saskatoon, 2012 and 2014). The total waste placed within the study area is
estimated at 5.3 million tonnes. The study area of the Saskatoon landfill includes 29 vertical gas wells, with an area of about 27 ha (Comcor Environmental Limited, 2010). The study area was capped in 2010 and monthly gas data was collected from 2014 to 2015 for 2 years. About 57% of the landfill gas was CH\(_4\), 43% was CO\(_2\) and the average methane generation rate was about 6.24 m\(^3\)/min.

### 4.2.2 FOD models and default model parameters

The three FOD models selected in the present study are (i) LandGEM developed by US Environmental Protection Agency (LandGEM, 2005), (ii) Afvalzorg Simple Model developed by a European private operator (Scharff and Jacobs, 2006), and (iii) IPCC model developed by the Intergovernmental Panel on Climate Change organization (IPCC, 2006). These FOD models are used extensively by engineering practitioners and were selected to facilitate result comparisons.

Landfill gas generation from these models were converted to gas collection with an efficiency of 80%. The LFG collection efficiencies of 80% was used in Thompson et al.’s study for 35 landfills in Canada (Thomson et al., 2009). The collection efficiency ranging from 67-91%, 62-86%, and 55-78% were reported at k of 0.02, 0.04 and 0.07, respectively by Barlaz et al. (2009). Bruce et al. (2017) also used a collection efficiency of 70-80% in their study of the multiple Western Canadian landfills.

LandGEM decay rates are set to a default value of 0.02 yr\(^{-1}\) for Clean Air Act (CAA) arid area (with annual rainfall less than 635 mm), 0.05 yr\(^{-1}\) for CAA conventional landfills, and 0.7 yr\(^{-1}\) for wet landfills. The Afvalzorg model default decay rates ranged from 0.02 yr\(^{-1}\) to 0.4 yr\(^{-1}\), depending on the waste type and climate. IPCC default k values vary from 0.01 yr\(^{-1}\) to 0.7 yr\(^{-1}\), depending on waste categories. In cold, dry areas, the
IPCC default $k$ ranged from 0.03 yr$^{-1}$ to 0.08 yr$^{-1}$ (IPCC, 2006). Given the climatic conditions in Regina and Saskatoon, the following default $k$ values were selected in accordance with the models' recommendations: $k_{\text{LandGEM}} = 0.02$ yr$^{-1}$, $k_{\text{Afvalzorg}} = 0.05$ yr$^{-1}$ for MSW, $k_{\text{IPCC}} = 0.06$ yr$^{-1}$ for food waste, $k_{\text{IPCC}} = 0.05$ yr$^{-1}$ for garden waste, $k_{\text{IPCC}} = 0.04$ yr$^{-1}$ for paper, $k_{\text{IPCC}} = 0.05$ yr$^{-1}$ for nappies, $k_{\text{IPCC}} = 0.02$ yr$^{-1}$ for wood, and $k_{\text{IPCC}} = 0.04$ yr$^{-1}$ for textiles.

The default methane generation potential ($L_0$) in LandGEM ranges from 96 m$^3$/Mg for wet landfills, 100 m$^3$/Mg for inventory conventional and arid area landfills, and 170 m$^3$/Mg for CAA conventional landfills and arid area landfills. The default DOC values for the Afvalzorg model range from 0.12 to 0.28 tonne/tonne (wet basis) for MSW (Gronert and Scharff, 2011), and the default DOC values for the IPCC model vary from 0.08 to 0.2 for food waste, 0.18 to 0.22 for garden waste, 0.36 to 0.45 for paper, 0.18 to 0.32 for hygiene nappies, 0.39 to 0.46 for wood and 0.20 to 0.40 for textile (IPCC, 2006). Methane generation potential depends on the type and composition of waste placed in the landfill. In Saskatchewan, paper made up the largest proportion of waste at 32.0%, followed by food waste, yard waste, wood, textiles and hygiene waste at 24.1%, 13.7%, 4.3%, 4.0% and 0.5%, respectively (Sinclair, 2006). The waste composition in 2002 was used as the input for the IPCC model. The following default values were used in this study: $L_0 = 100$ m$^3$/Mg (LandGEM), DOC = 0.19 tonne/tonne (Afvalzorg), DOC$_{\text{IPCC}} = 0.15$ for food waste, DOC$_{\text{IPCC}} = 0.2$ for garden waste, DOC$_{\text{IPCC}} = 0.4$ for paper, DOC$_{\text{IPCC}} = 0.43$ for wood, and DOC$_{\text{IPCC}} = 0.24$ for hygiene nappies and textiles.
Methane correction factor (MCF) and fraction of DOC dissimilated (DOC\textsubscript{f}) were chosen at 1.0 and 0.5, respectively. Fraction of methane (F) was selected at 54% for Regina landfill and 57% for Saskatoon landfill.

4.2.3 **Iterative methods and curve fitting by minimizing RSS values**

Six year of continuous gas data from the Regina landfill was used to examine the effects of three iterative approaches on curve fitting: (i) changing both k and L\textsubscript{0} or DOC; (ii) changing k first and then L\textsubscript{0} or DOC, and (iii) changing L\textsubscript{0} or DOC first and then k. The selected iterative method was then applied to optimize k and L\textsubscript{0} or DOC at the Regina landfill, and the optimal parameters from Regina landfill was used as input for Saskatoon landfill. The Excel Solver function was used to minimize RSS for the FOD models. The maximum number of iterations of 100 was used with a tolerance of 5% and convergence of 10\textsuperscript{-4}. LandGEM and Afvalzorg use single k and L\textsubscript{0} or DOC values, thus additional constraints were not used in the optimization process. On the other hand, constraints on the k and DOC values were imposed to the IPCC optimization to eliminate divergence issues and the return of negative roots. The ranges of possible k and DOC values reported in literature were collected (Ballinger, 2011; Barlaz, 1998; De La Cruz & Barlaz, 2010; Larsen, 2013; Manfredi, 2009; Ministry of the Environment Japan, 2015; US EPA, 2012; Environment Agency of Iceland, 2015) and 10% adjustments were applied to the extreme values reported, in order to establish broader upper and lower boundaries for IPCC optimization constraints. The overall system RSS, mean absolute percentage error, and distribution of errors were examined to select the best iterative method.
4.3 Results and discussion

4.3.1 Methane estimates using default k and L0 (DOC) values

Each FOD model with default inputs yielded gas estimates much higher than the measured data in Saskatchewan (Table 4-2). This finding is consistent with observations made regarding the use of default inputs in other landfills (Bella et al. 2011; Govindan and Agamuthu, 2014; Mou et al., 2015a). The mean percentage error between the estimated and measured values of the 3 models at both sites were substantial (55.09% - 134.92%), even after considering the possible unfavorable effects of non-ideal gas recovery efficiency and elevated fugitive emissions. LandGEM default inputs produced the lowest mean percentage error (55.09-64.28%), followed by Afvalzorg and IPCC models. This is probably due to the smaller default k value (k=0.02yr⁻¹) recommended by LandGEM. IPCC requires additional waste composition data and is more complex than LandGEM and Afvalzorg, however, its sets of default values failed to represent landfills in Saskatchewan. The results suggest that the default values in these FOD models are not applicable to the Regina and Saskatoon landfills, probably due to the dry and cold climate in Saskatchewan. LandGEM was developed by the United States Environmental Protection Agency using predominantly American landfill data. Unlike Regina and Saskatoon, there is actually very little (< 10%) of the United States that has the "Dfb" designation under Koppen-Geiger classification (Peel et al. 2007). Fourie and Morris (2004) noted that the waste degradation process in many South African landfills seemed to be slow to none-existent due to lack of moisture. According to a Japanese field study, landfills located in colder climates may have actual decay rates lower than the minimum default k values (Ishii and Furuichi, 2013). The percentage error values were not constant
with respect to time for all models in both sites. For the Regina landfill, the percentage errors were smallest at the peak in 2011.

It is not clear why the mean percentage errors were consistently lower in Saskatoon than Regina (about 9-21% lower) regardless of the model used. Due to the close proximity of the landfills, the waste composition, disposal practices, landfill regulations and climatic conditions are very similar. One possible explanation is that the Saskatoon gas field is larger, with more buried waste (27 ha and 5.3 million tonnes of waste), and therefore the landfill conditions could be better represented by a set of generic default values. The landfill gas methane content in the Saskatoon landfill was also higher (57%), indicating a mature and favorable anaerobic condition for methanogenic microorganisms, as assumed by all FOD gas models. However, it is important to note that only two years of Saskatoon gas data was available and used in the analysis, and more gas data is needed for a definite conclusion. A study period of 2 to 10 years is suggested for future studies of this nature, similar to Amini et al., (2013); Faour et al., (2007); Machado et al., (2009); and Wang et al., (2013).

Despite the fact that the default inputs grossly overestimate gas quantity, all the models successfully predicted the gas generation peak in 2011 at the Regina landfill except for Afvalzorg model, which had an extended peak from 1998 to 2011. The result suggests that LandGEM and IPCC models could be good gas modeling tools, provided that accurate waste quantity inputs are used. Afvalzorg model appeared to be more sensitive to the changes of waste mass inputs over the study period and the elevated decay rate after optimization ($k_{Afvalzorg}=0.048$). Modeled Saskatoon gas generation peaked
at 2011; however, comparison with field data is not possible due to the limited field data range.
Table 4-1: Measured methane generation data and estimates from FOD models using default parameters (Vu et al., 2017)

<table>
<thead>
<tr>
<th>Landfill</th>
<th>Year</th>
<th>CH₄ estimates (m³/year)</th>
<th>Actual data (m³/year)</th>
<th>Percentage Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LandGEM</td>
<td>Afvalzorg</td>
<td>IPCC</td>
</tr>
<tr>
<td>Regina</td>
<td>2009</td>
<td>3,087,547</td>
<td>4,358,627</td>
<td>4,543,365</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>3,123,760</td>
<td>4,377,418</td>
<td>4,569,032</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>3,155,506</td>
<td>4,386,378</td>
<td>4,585,208</td>
</tr>
<tr>
<td></td>
<td>2012</td>
<td>3,133,491</td>
<td>4,268,631</td>
<td>4,480,789</td>
</tr>
<tr>
<td></td>
<td>2013</td>
<td>3,071,443</td>
<td>4,060,448</td>
<td>4,289,695</td>
</tr>
<tr>
<td></td>
<td>2014</td>
<td>3,010,625</td>
<td>3,862,417</td>
<td>4,107,146</td>
</tr>
<tr>
<td></td>
<td>2015</td>
<td>2,951,010</td>
<td>3,674,045</td>
<td>3,932,741</td>
</tr>
</tbody>
</table>

Mean Percentage Error (%) Regina 64.28 123.76 134.92

| Saskatoon | 2009 | 5,242,509 | 7,672,893 | 7,897,505 | N/A |
|           | 2010 | 5,343,007 | 7,784,237 | 8,035,886 | N/A |
|           | 2011 | 5,458,404 | 7,930,290 | 8,178,615 | N/A |
|           | 2012 | 5,350,321 | 7,543,525 | 7,828,202 | N/A |
|           | 2013 | 5,244,377 | 7,175,623 | 7,493,521 | N/A |
|           | 2014 | 5,140,531 | 6,825,664 | 7,173,837 | 3,304,582 | 55.56 | 106.55 | 117.09 |
|           | 2015 | 5,038,742 | 6,492,772 | 6,868,450 | 3,258,588 | 54.63 | 99.25 | 110.78 |

Mean Percentage Error (%) Saskatoon 55.09 102.90 113.93
4.3.2 Iterative approaches

The k and L₀ or DOC values with the lowest RSS at the Regina landfill are shown in Table 4-2 for all 3 FOD models. Annual gas generation rates were expressed in cubic meter (m³/yr), resulting in high magnitude RSS values. The RSS values using different FOD models and iterative approaches are similar (about 4.36E+11), except for iterative method 1 of the IPCC model (9.56E+11). This is probably due to a higher number of model parameters associated with the IPCC model and a larger degree of freedom associated with method 1). Instead of a single set of k and L₀ or DOC values, IPCC requires waste composition as well as k and DOC values for each of the six waste categories, and therefore more constraints to be satisfied before reaching an optimized solution.

It is found that for all three models considered, method 2 (change k, then L₀ or DOC) generally resulted in smaller k values, and method 3 (change L₀ or DOC, then k) generally resulted in smaller L₀ or DOC values. An inverse relationship exists between k and L₀ or DOC values using modeled results at the Regina site: the higher the k values, the lower the L₀ or DOC values, or vice versa. A similar trend is also observed in other FOD models, for example, Mou et al. (2015 a,b), Tolaymat et al. (2010), and Wang et al. (2013), as shown in Table 1-1.
Table 4-2: Optimal FOD landfill gas model parameters using different iterative approaches at Regina (Vu et al., 2017)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Method 1 (Change k and ( L_0 ) or DOC)</th>
<th>Method 2 (Change k then ( L_0 ) or DOC)</th>
<th>Method 3 (Change ( L_0 ) or DOC then k)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LandGEM</td>
<td>Afvalzorg</td>
<td>IPCC</td>
</tr>
<tr>
<td>( k (\text{year}^{-1}) )</td>
<td>0.01</td>
<td>0.016</td>
<td>0.049</td>
</tr>
<tr>
<td>( L_0 ) or DOC (m(^3)/Mg or Mg/Mg)</td>
<td>100</td>
<td>0.139</td>
<td>0.101</td>
</tr>
</tbody>
</table>
Figure 4-1 graphically compares the calculated methane generation rates using the different iterative methods and FOD models at the Regina landfill. Method 1 tends to overestimate the generation rates, as more points are located above the 45 degree line (Figure 4-1a, 4-1d and 4-1g). On the other hand, modeling errors are more evenly distributed for methods 2 and 3. It appears that iterative method 2 (Figure 4-1b, 4-1e, and 4-1h) provided slightly more consistent results than the other iterative methods among the 3 FOD models, and thus was selected for subsequent discussion.
Figure 4-1: Comparisons of actual and estimated gas data using optimal model parameters at Regina landfill (Vu et al., 2017)
Regina gas estimates using optimized parameters were plotted with respect to time in Figure 4-2. The total area under the curve for a given case depends on the amount of waste (mass inputs) and the characteristics of waste (methane generation potential). The slope of the curve, however, depends on the k value. In all cases, it can be seen that the curves were generally steeper for iterative method 3 because of the higher waste decomposition rates (i.e. higher k values). A distinct gas generation peak in 2011 was predicted by all 3 models, except for Afvalzorg model method 3 (Figure 4-2b), where an extended equal peak gas rates were predicted from 1998 to 2011. This is likely due to (i) the elevated decay rate after optimization ($k_{Afvalzorg} = 0.048$) and (ii) the higher sensitivity of the Afvalzorg model to changes in the waste mass inputs during the study period. Unlike results from other models and iterative methods, the IPCC gas estimates using method 1 (Figure 4-2c) are noticeably higher. This can be explained by a higher RSS value from the optimization process (Table 4-2).
Figure 4-2: Effects of iterative methods on landfill gas generation characterises using (a) LandGEM, (b) Afvalzorg, and (c) IPCC model (Vu et al., 2017)
4.3.3 Optimal FOD gas model parameters

When compared to the default inputs recommended by each FOD model, the optimized parameters using Regina field data are much lower (Table 4-2). Using results from method 2 as an example, the optimal k value was 2 times lower than the LandGEM default ($k_{\text{LandGEM}} = 0.02$) and 5 times lower than the Afvalzorg default ($k_{\text{Afvalzorg}} = 0.05$), while the optimal $L_0$ or DOC values were identical to the model default values. On average, optimal k values for the IPCC model were about 2 times lower than the default values, while the DOC values were about 1.2 times lower than the default values. Table 4-3 tabulated the mean percentage errors between estimated gas data with optimized parameters and the actual data over the study period for both Regina and Saskatoon. The results of the three models seemed to fit well to the actual data with the mean percentage errors ranging from 11.60% to 19.93% for the Regina landfill and about 1.65% to 10.83% for the Saskatoon landfill.

The mean percentage errors for IPCC Method 1 (19.93% for Regina and 10.83% for Saskatoon) were much higher than others. This is probably due to the additional constraints imposed on the optimization process as previously discussed, and as such method 1 is not recommended for complex multi-variable models such as IPCC. The mean percentage errors for methods 2 and 3 were similar. Method 2 had better estimates in Saskatoon, while method 3 had better estimates in Regina. When comparing the results from the two sites, the mean percentage errors were significantly lower in Saskatoon than in Regina.

It is also found that all 3 FOD models were capable of estimating gas rates within uncertainties when optimized inputs were used. There were negligible differences
between the performance of these FOD models, however, the Regina and Saskatoon landfills failed to demonstrate the potential benefits of using a more complex FOD model such as the IPCC model. This finding was consistent with implications from Klause et al (2016) and Wang et al. (2015). Klause et al. (2016) noted that reducing the uncertainty of the input parameters of a single-phase model will make the model more accurate than using multiple-phase model. Wang et al. (2015) found that the single-phase model provided lower error in measuring methane collection with a best fit k value.
Table 4-3: Mean percentage error between estimated rates using optimal parameters and actual gas rates (Vu et al., 2017)

<table>
<thead>
<tr>
<th>Landfill</th>
<th>Mean Percentage Error (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LandGEM</td>
<td>Afvalzorg</td>
<td>IPCC</td>
</tr>
<tr>
<td>Regina</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method 1 (Change $k$ and $L_0$ or DOC)</td>
<td>12.05</td>
<td>11.96</td>
<td>19.93</td>
<td></td>
</tr>
<tr>
<td>Method 2 (Change $k$ then $L_0$ or DOC)</td>
<td>12.04</td>
<td>12.04</td>
<td>11.83</td>
<td></td>
</tr>
<tr>
<td>Method 3 (Change $L_0$ or DOC then $k$)</td>
<td>11.89</td>
<td>11.74</td>
<td>11.60</td>
<td></td>
</tr>
<tr>
<td>Saskatoon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method 1 (Change $k$ and $L_0$ or DOC)</td>
<td>1.65</td>
<td>2.90</td>
<td>10.83</td>
<td></td>
</tr>
<tr>
<td>Method 2 (Change $k$ then $L_0$ or DOC)</td>
<td>1.66</td>
<td>1.75</td>
<td>3.52</td>
<td></td>
</tr>
<tr>
<td>Method 3 (Change $L_0$ or DOC then $k$)</td>
<td>3.60</td>
<td>7.93</td>
<td>5.25</td>
<td></td>
</tr>
</tbody>
</table>
4.3.4 Comparison of the selected optimal $k$ and $L_0$/DOC values and literature

Figure 4-3 compares the results from this study to the model parameters reported in the literature. From Figure 4-3a, the optimal $k$ values ($k_{\text{LandGEM}}= 0.010$ yr$^{-1}$, $k_{\text{Afvalzorg}}= 0.010$ yr$^{-1}$, $k_{\text{IPCC}}= 0.020$ yr$^{-1}$) obtained from Saskatchewan landfills matched well with the recommended values from Environment Canada (2015). In general, the optimal $k$ values of the LandGEM, Afvalzorg and IPCC models obtained from this study were lower than those reported in the literature. LandGEM, developed by the US Environmental Protection Agency, is commonly used in North America for preliminary studies on landfill gas estimation. The $k_{\text{LandGEM}}$ obtained in this study was 1.3 - 19 times lower than that of cold wet climate (Table 1-1), and 4 - 21 times lower than that of warmer climate. Similarly, $k_{\text{IPCC}}$ was about 0.65-9.5 times lower than that of cold, wet climate area and 2 - 10.5 times lower than that of warmer climate. The results suggest that the default values in these FOD models fail to represent field conditions at Saskatchewan landfills.

In general, $L_0$ values from this study ($L_{\text{LandGEM}}= 100$ m$^3$/Mg, $L_{\text{Afvalzorg}}= 97$ m$^3$/Mg, $L_{\text{IPCC}} = 91$ m$^3$/Mg) compared well with other published studies (Figure 4-3b). DOCs (Mg/Mg) of IPCC and Afvalzorg were converted to $L_0$ (m$^3$/Mg) based on the equation: $L_0= F \times \text{DOC} \times \text{DOC}_f \times 16/12 \times \text{MCF}$ (IPCC, 2006). $L_0$ is related to the waste type and characteristic, and thus is less sensitive to the climatic conditions of the landfills. $L_0$ values recommended by Thompson et al. (2009) were at least 1.3 - 1.95 times higher than the results from the present study. It should be noted that $L_0$ values reported by Thompson et al. (2009) were calculated based on a higher fraction of degradable organic carbon ($\text{DOC}_f=0.77$), while this paper used default $\text{DOC}_f$ at only 0.50. Buried waste
composition, and thus $L_0$ values, may also be affected by local waste regulations and disposal practices.
Figure 4-3: Reported (a) k and (b) $L_0$ values in Saskatchewan and other studies (Vu et al., 2017)
In this study, curve fitting was implemented to obtain the optimal parameters of the three FOD models: LandGEM, Afvalzorg, and IPCC for landfills located in the cold semi-arid climatic areas. The optimal constant decay rates were at least 2 times lower than the default values of the models. The proposed techniques can also be used to estimate methane for power plant design, or to prepare Greenhouse gas inventory calculations. However, the optimal parameters in this study are site specific and should be carefully checked before apply elsewhere.
5. Effects of Waste Characteristics on Optimal Collection Routes

5.1 Objectives of Chapter Five

In this Chapter, a waste collection model was developed using a Vehicle Routing Problem (VRP) network analysis within a Geographic Information System (GIS), with the waste inputs from a time series Artificial Neural Networks (ANN) nonlinear autoregressive model. The objectives of Chapter 5 are to (i) forecast the generation rates and characteristics of recyclables and garbage in four sub-areas in the study area using a time series ANN; and (ii) examine how waste quantity, composition, collection types, and material density characteristics affect optimized vehicle routes and truck emissions. Unlike other cross-sectional analysis, the novelty of this study is in the quantification of the effects that changes in waste characteristics have on optimal truck routes by combining a machine-learning generation model with a geo-spatial optimization method on a given WMS.

5.2 Selection of target area and waste collection routes

City of Austin is again selected as the study area. A description of the City is provided in Chapter three of this thesis. There are five waste collection zones in the City of Austin. The one-way road density in these five waste collection zones varied from 0.0045 to 0.3268 km per km$^2$. The household density of these five waste collection zones ranged from 296 to 529 households per km$^2$. Both the one-way road density and the household density were considered in order to select the target area for this study. Minimizing one-way road density in a waste collection zone introduces more degrees of freedom to the subsequent VRP optimization of truck routes. A higher density in a given area, on the other hand, will generally lead to a higher amount of waste and more truck
trips. Combining these two considerations, the areas with the lower density of one-way roads and higher density of households are therefore selected in this study for route optimization.

As shown in Figure 5-1, the Monday waste collection zone was selected due to its lowest one-way road density (0.0045 km/km²) and the second highest household density (434 hh/km²). There is a total of 43 sub-collection areas within the Monday zone, and detailed waste mass data required for an ANN time series analysis was not available for all the 43 sub-areas within the target area. As such, four sub-areas located in the southwest of the Monday zone, namely RMBU1, RMBU12, RMBU13, and RMBU14 were selected for ANN time series analysis and VRP modeling. At least 10 years of data is available for these four sub-areas (recyclables data from 2008 to 2018 and garbage data from 2004 to 2018).

Curbside door-to-door collection has been adopted in the target area. There were five collection streams in the target area: yard, large brushes, single stream recycling materials (recyclables), garbage (residual), and pilot organic streams (City of Austin, 2018a) of which recyclables were collected bi-weekly while garbage was collected weekly. Automatic side-load trucks have been used for both the recyclables and garbage streams. Based on the waste data from City of Austin (2018a), the mean mass of recyclables collected from 2008 to 2018 was found about 11,112 to 12,352 kg/2weeks (5,555.1 to 6,176 kg/week) in the target area and the mean mass of garbage collected was about 13,312 to 14,932 kg/week in the period from 2004 to 2018.
Figure 5-1: Study area (After City of Austin, 2018d, 2018e)
5.3 Methodology

The methodology of the study has been illustrated in Figure 5-2. Firstly, the data of weekly collected garbage, bi-weekly collected single stream recyclables, waste composition, and number of households were collected from Austin’s Open Data Portal. Density of waste was calculated based on waste composition and specific weight of each material from a USEPA study (2016b). The next stage is an ANN time-series analysis to forecast future recyclables and garbage generation rates of each sub-collection area in the year 2023. Different scenarios are considered with different recyclables and garbage compositions. The predicted recyclables and garbage generation rates were used to compute expected volume of waste in trucks. The predicted waste volumes from the target areas are inputs into the GIS - Network Analysis tool (ArcGIS - version 10.5.1) to develop optimal truck routes. It is assumed that during the 5-year forecasting period there will be no major changes in the street and road network configuration and the number and location of collection points (household) are similar. A total of 36 scenarios are developed to examine the effects of changing waste characteristics on optimal truck routes with minimal travel distance. Details of each stage of the study are presented in the following sections.
Figure 5-2: Methodology flow chart of Chapter five
5.3.1 Data collection and analysis

5.3.1.1 Waste data sets

Bi-weekly recyclables and weekly garbage data were collected and verified by using waste records from Austin’s Open Data Portal. The recyclables and garbage data sets were divided into two sub-sets: a training set and a testing set with the ratio of 85:15. The ratio was selected based on the literature which was reported to range from 60:40 to 90:10 (Abbasi and Hanandeh, 2016; Azadi and Karimi-Jashni, 2016; Kannangara et al., 2018; Pan et al. 2019; Vu et al., 2019). Standard deviation, minimum, maximum, and the ratio of maximum to minimum was calculated to assess the variation of the recyclables and garbage data sets. Recyclables composition and garbage composition of the study area were collected from the City-Serviced Residential Waste Characterization Study Report (CB&I Environmental & Infrastructure, Inc, 2015). Table 5-1 shows the characteristics of recyclables and garbage data in each of the four sub-study areas. The recyclables and garbage characteristics will be forecasted for different scenarios and discussed in section 5.3.2.
Table 5-1: Waste generation in the study area

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
<th>STDEV</th>
<th>Max/Min</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bi-weekly recyclables (kg/2weeks) (2008-2018)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training stage (Sep 2008-Feb 2017)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMBU11</td>
<td>2.957</td>
<td>12,525</td>
<td>28,177</td>
<td>3,395</td>
<td>9.53</td>
</tr>
<tr>
<td>RMBU12</td>
<td>5.117</td>
<td>11,826</td>
<td>24,680</td>
<td>3,045</td>
<td>4.82</td>
</tr>
<tr>
<td>RMBU13</td>
<td>2.703</td>
<td>10,948</td>
<td>18,316</td>
<td>2,277</td>
<td>6.78</td>
</tr>
<tr>
<td>RMBU14</td>
<td>4.536</td>
<td>11,305</td>
<td>22,734</td>
<td>2,663</td>
<td>5.01</td>
</tr>
<tr>
<td>Testing stage (Feb 2017-Aug 2018)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMBU11</td>
<td>8,845</td>
<td>11,357</td>
<td>16,175</td>
<td>1,547</td>
<td>1.83</td>
</tr>
<tr>
<td>RMBU12</td>
<td>6,078</td>
<td>9,555</td>
<td>11,258</td>
<td>1,168</td>
<td>1.85</td>
</tr>
<tr>
<td>RMBU13</td>
<td>8,210</td>
<td>12,001</td>
<td>28,350</td>
<td>3,708</td>
<td>3.45</td>
</tr>
<tr>
<td>RMBU14</td>
<td>6,985</td>
<td>11,993</td>
<td>22,008</td>
<td>3,194</td>
<td>3.15</td>
</tr>
<tr>
<td><strong>Weekly Garbage (kg/week) (2004-2018)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training stage (Oct 2004-Jul 2016)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMBU11</td>
<td>1,715</td>
<td>14,756</td>
<td>27,823</td>
<td>3,503</td>
<td>16.22</td>
</tr>
<tr>
<td>RMBU12</td>
<td>2,749</td>
<td>13,437</td>
<td>23,242</td>
<td>2,943</td>
<td>8.45</td>
</tr>
<tr>
<td>RMBU13</td>
<td>2,350</td>
<td>13,028</td>
<td>41,159</td>
<td>3,118</td>
<td>17.51</td>
</tr>
<tr>
<td>RMBU14</td>
<td>1,207</td>
<td>13,854</td>
<td>27,724</td>
<td>3,468</td>
<td>22.97</td>
</tr>
<tr>
<td>Testing stage (Aug 2016-Aug 2018)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMBU11</td>
<td>10,904</td>
<td>15,930</td>
<td>23,215</td>
<td>2,242</td>
<td>2.13</td>
</tr>
<tr>
<td>RMBU12</td>
<td>5,978</td>
<td>15,675</td>
<td>23,668</td>
<td>3,065</td>
<td>3.96</td>
</tr>
<tr>
<td>RMBU13</td>
<td>7,965</td>
<td>14,924</td>
<td>24,893</td>
<td>2,904</td>
<td>3.13</td>
</tr>
<tr>
<td>RMBU14</td>
<td>10,278</td>
<td>17,095</td>
<td>24,966</td>
<td>2,666</td>
<td>2.43</td>
</tr>
</tbody>
</table>
5.3.1.2 Road network data sets, target areas and number of stops

The road map of the city of Austin was collected from OpenStreetMap, an open source (OpenStreetMap contributors, 2018). Road network data set such as names of roads details on one-way, two-ways roads, road restriction, vehicle speed limitation, and types of roads was created using this road map. Roads in the study area are mainly residential roads with vehicle speed limitation of 30 miles/h (Texas Department of Public Safety, 2017), or about 48 km/h.

Boundaries of the waste collection areas was taken from Austin’s Open Data Portal created in 2015 and updated in May 2018 (City of Austin, 2018e). Locations and number of households (service stops for recyclables and garbage) were also collected from Austin’s Open Data Portal created in 2017 from the city of Austin (City of Austin, 2018f). There were 1024, 869, 855, and 964 households (service stops) in the areas RMBU11, 12, 13, 14, respectively.

5.3.2 ANN waste generation forecast

ANN time series - Nonlinear Autoregressive was used to forecast recyclables and garbage in 2023. The approach was used to overcome the lack of independent variables for the future forecast (Abbasi and Hanandeh, 2016). The ANN model contained three layers: an input layer, a single hidden layer with 10 neurons, and an output layer. From the preliminary trials, a lag time of 3 months was applied for recyclables and a 12 months lag was applied for garbage, similar to the methodology reported by Vu et al. (2019). Bi-weekly recyclables and weekly garbage were forecasted in the future for a medium term of five years ahead, until 2023. The medium term generation forecast was applied for numerous waste studies such as Navarro-Esbri et al. (2002), Abbasi et al. (2014), and
Abbasi and Hanandeh (2016). The mean bi-weekly recyclables and weekly garbage over the five years will be used for the GIS - Network Analysis - VRPs models.

5.3.3 Models performance assessment

The mean squared error (MSE), mean absolute percentage error (MAPE), and R value have been reported in a range of waste generation studies (Rimaityte et al., 2012; Shahabi et al., 2012; Abbasi and Hanandeh, 2016; Azadi and Karimi-Jashni, 2016; Kannangara et al., 2018). Similar to other performance indices, each index has both drawbacks and benefits when it comes to ANN modeling assessment. A treatment on the performance indices is reported by Azadi and Karimi-Jashni (2016) and is not repeated here. MATLAB (version 2017a) was used for the ANN simulations. MAPE is widely reported by other waste forecast studies and thus is adopted in this study to facilitate comparison. Equations to calculate MAPE are provided below:

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{Y_{ai} - Y_{mi}}{Y_{ai}} \right) \times 100 \quad (5 - 1)
\]

Where:

n: Number of data points

\(Y_{ai}\): Actual mass of yard waste

\(Y_{mi}\): Predicted mass of yard waste

\(\overline{Y}_{a}\): The mean actual waste mass

\(\overline{Y}_{m}\): The mean predicted waste mass
5.3.4 Waste characteristics forecast

Three scenarios of recyclables composition in 2023 were assumed for the target area and include: scenario R1 - recyclables composition was similar with that of the year 2014; scenario R2 - the percentage of the largest component (paper) of recyclables decreased 15% compared to that of the year 2014; scenario R3 - the percentage of the paper of recyclables increased 15% compared to that of the year 2014. The remaining components of the recyclables will be adjusted accordingly to have 100% of total composition. The 15% change in percentage of paper of the recyclables in 2023 was assumed based on the overall national recycling trend in the USA, which the percentage of paper changed 15.9% during the period of 8 years from 2008 to 2014 as mentioned in Section 1.2. A ±15% change in percentage of the major component of waste such as paper and organic was also assumed by Slagstad and Brattebø (2013) to examine how household waste composition influenced on WMS using life cycle analysis. The composition of garbage in 2023 is estimated using a similar approach.

Three scenarios of garbage composition in 2023 were assumed for the target area and include: scenario G1 - garbage composition was similar with that of the year 2014; scenario G2 - the percentage of the largest component (organics) of garbage decreased 15% compared to that of the year 2014; scenario G3 - the percentage of the organics inside the garbage stream increased 15% compared to that of the year 2014. Waste density was computed for the six scenarios using the specific density of each waste component. The waste density equation is as follows:

\[ \rho = \frac{K}{\sum_{i=1}^{n} \frac{r_i}{\rho_i}} \]  

(5-2)
Where:

\( \rho \): Density of recyclables or garbage (kg/m\(^3\)) in trucks

\( K \): compaction ratio of trucks, \( K \) ranged from 2.2 to 2.5 (Gomes et al., 2008; Vaccari et al., 2013; Vu et al., 2018b), in this study \( K \) was selected at 2.5 for recyclables trucks. Because side load garbage trucks can compact waste up to 700 lb per cubic yard or 415 kg/m\(^3\) (Dover Corporation, 2017), the compaction ratio for garbage trucks in this study was selected at 3.

\( P_i \): Percentage by mass of material \( i \) of recyclables or garbage (%)

\( \rho_i \): Density of material \( i \) of recyclables or garbage (kg/m\(^3\))

Waste composition and density of the six scenarios are presented in Table 5-2.
Table 5-2: Waste composition in the study area

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scenario R1</td>
</tr>
<tr>
<td><strong>Recyclables</strong></td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>69.6%</td>
</tr>
<tr>
<td>Plastics</td>
<td>13.1%</td>
</tr>
<tr>
<td>Metals</td>
<td>5.8%</td>
</tr>
<tr>
<td>Glass</td>
<td>8.7%</td>
</tr>
<tr>
<td>Organics</td>
<td>2.8%</td>
</tr>
<tr>
<td>Density of recyclables (kg/m$^3$)</td>
<td>193.1</td>
</tr>
<tr>
<td><strong>Garbage</strong></td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>20.2%</td>
</tr>
<tr>
<td>Plastics</td>
<td>7.8%</td>
</tr>
<tr>
<td>Metals</td>
<td>4.5%</td>
</tr>
<tr>
<td>Glass</td>
<td>0.0%</td>
</tr>
<tr>
<td><strong>Organics</strong></td>
<td>63.7%</td>
</tr>
<tr>
<td>Other materials</td>
<td>3.9%</td>
</tr>
<tr>
<td>Density of garbage (kg/m$^3$)</td>
<td>340.9</td>
</tr>
</tbody>
</table>

Note:

1: CB&I Environmental & Infrastructure, Inc, 2015
5.3.5 GIS - Network Analysis - VRP models

5.3.5.1 Model development

A total of 9 scenarios were developed for different waste streams and collection types in order to assess the effects of waste composition and mass of waste on the optimal truck routes at each sub-collection area. These scenarios included: three scenarios of composition for recyclables trucks (R1, R2, R3), three scenarios of composition for garbage trucks (G1, G2, G3), and three scenarios for dual-compartment trucks (M1, M2, M3) in accordance with the three sets of different material compositions. Therefore, there were a total of 36 scenarios for four sub-collection areas.

Model inputs of the GIS Network Analysis VRPs (ArcGIS - version 10.5.1) included network data sets, time spent to pick up waste at each stop, and volume of recyclables or garbage at each stop. The volume of recyclables or garbage was estimated from their mass and composition. In addition to these parameters, other factors such as time windows (8:00-20:00), capacity of truck (20 m$^3$), maximum travel distance of a truck (120 km), maximum travel time (3 hrs), maximum collection time (3 hrs), and maximum number of stops (number of stops that a truck trip can serve) were used as inputs of the models. Research has shown that pick-up times at each stop for garbage trucks ranged from 6.70 to 8.74 seconds (Nguyen and Wilson, 2010; Edwards et al., 2016; Maimoun et al., 2016). The pick-up time for garbage trucks was selected at 8.5 second (0.142 minute) per household. A slightly longer pick-up time for recyclables trucks of 9 seconds (0.15 minute) per household was selected. The pick-up time was selected according to Maimoun et al.’s study (2016). The pick-up time for dual-compartment trucks at each household was considerably higher than that of garbage
trucks or recyclables trucks. It varied from 21.6 to 29.3 seconds (Nguyen and Wilson, 2010; Maimoun et al., 2016), depending on the spatial distribution and density of households (Nguyen and Wilson, 2010). Table 3 presents the major inputs of the VRP models.

Speed of trucks within the collection area range from 7 km/h to 15 km/h (Edwards et al., 2016; Maimoun et al., 2016), depending on the density of households in the collection area (Edwards et al., 2016). In this study, the speed of truck in the collection area was conservatively selected at 10 km/h. Volume of waste of each scenario was calculated based on the waste generation rate obtained from ANN models and waste density from each scenario using equation below:

\[ V_i = \frac{M_i}{\rho_{wi}} \]  

\( V_i \): Volume of recyclables or garbage (L/hh-week) of scenario i  
\( M_i \): Recyclables or garbage generation rate at each stop (kg/hh-week) of scenario i  
\( \rho_{wi} \): Density of recyclables or garbage (kg/1000 L) of scenario i
### Table 5-3: Major inputs to the VRP solution

<table>
<thead>
<tr>
<th>Sub-area</th>
<th>Number of households</th>
<th>Volume of waste at each stop (L/hh)</th>
<th>Number of households that a 20 m³ truck can serve in a trip</th>
<th>Number of 20 m³ truck trips served each sub-area</th>
<th>Pick up time at each stop (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recyclables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMBU11</td>
<td>1024</td>
<td>58.39</td>
<td>68.06</td>
<td>48.44</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>343</td>
<td>294</td>
<td>413</td>
<td>4</td>
</tr>
<tr>
<td>RMBU12</td>
<td>869</td>
<td>52.09</td>
<td>60.72</td>
<td>43.21</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>384</td>
<td>329</td>
<td>463</td>
<td>3</td>
</tr>
<tr>
<td>RMBU13</td>
<td>855</td>
<td>68.37</td>
<td>79.70</td>
<td>56.72</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>293</td>
<td>251</td>
<td>353</td>
<td>3</td>
</tr>
<tr>
<td>RMBU14</td>
<td>964</td>
<td>61.86</td>
<td>72.10</td>
<td>51.31</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>323</td>
<td>277</td>
<td>390</td>
<td>4</td>
</tr>
<tr>
<td><strong>Garbage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMBU11</td>
<td>1024</td>
<td>28.86</td>
<td>33.19</td>
<td>24.67</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>693</td>
<td>603</td>
<td>811</td>
<td>2</td>
</tr>
<tr>
<td>RMBU12</td>
<td>869</td>
<td>33.11</td>
<td>38.08</td>
<td>28.31</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>604</td>
<td>525</td>
<td>707</td>
<td>2</td>
</tr>
<tr>
<td>RMBU13</td>
<td>855</td>
<td>35.13</td>
<td>40.40</td>
<td>30.03</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>569</td>
<td>495</td>
<td>666</td>
<td>2</td>
</tr>
<tr>
<td>RMBU14</td>
<td>964</td>
<td>35.00</td>
<td>40.25</td>
<td>29.92</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>571</td>
<td>497</td>
<td>668</td>
<td>2</td>
</tr>
<tr>
<td><strong>Recyclables and garbage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMBU11</td>
<td>1024</td>
<td>58.05</td>
<td>67.22</td>
<td>48.89</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>345</td>
<td>298</td>
<td>409</td>
<td>4</td>
</tr>
<tr>
<td>RMBU12</td>
<td>869</td>
<td>59.16</td>
<td>68.44</td>
<td>49.91</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>338</td>
<td>292</td>
<td>401</td>
<td>3</td>
</tr>
<tr>
<td>RMBU13</td>
<td>855</td>
<td>69.31</td>
<td>80.25</td>
<td>58.39</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>289</td>
<td>249</td>
<td>343</td>
<td>4</td>
</tr>
<tr>
<td>RMBU14</td>
<td>964</td>
<td>65.93</td>
<td>76.31</td>
<td>55.58</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>303</td>
<td>262</td>
<td>360</td>
<td>4</td>
</tr>
</tbody>
</table>
5.3.5.2 Model result comparison

In this Chapter, the GIS-VRP model was applied to identify the shortest route and to determine the optimal solution. Truck travel distances, truck travel time, and air emissions were used to evaluate different scenarios in order to examine the effects of waste characteristics on the optimized truck routes. According to Maimoun et al.’s study (2013) on emissions from USA waste collection vehicles, air emissions (E) was a function of travel distances, speed of truck, and idle time:

\[ E = (D_c \cdot E_c + D_{oc} \cdot E_{oc} + T_i \cdot E_i) \]  \hspace{1cm} (5-4)

Where:

- \( D_c \): Total truck travel distances within the collection area
- \( E_c \): Emission rate (g/km) when trucks travel in the collection area, \( E_c = 16.13 \text{ g/km} \) for NO\(_x\) (average value calculated based on Chen et al., 2007; Maimoun et al., 2013; Sandhu et al., 2015) and 2358.81 g/km for CO\(_2\) (average value calculated based on Maimoun et al., 2013; Sandhu et al., 2015)
- \( D_{oc} \): Total truck travel distances outside the collection area
- \( E_{oc} \): Emission rate (g/km) when trucks travel outside the collection area \( E_{oc} = 5.30 \text{ g/km} \) for NO\(_x\) (average value calculated based on Chen et al., 2007; Maimoun et al., 2013; Sandhu et al., 2015) and 1111.37 g/km for CO\(_2\) (average value calculated based on Maimoun et al., 2013; Sandhu et al., 2015)
- \( T_i \): Total idle time trucks spent to pick up waste at households and unload at landfills and material recovery facilities
E_i: Emission rate (g/h) for truck idle time, E_i = 58.77 g/h for NO_x (average value calculated based on Chen et al., 2007; USEPA, 2008; Khan et al., 2009; Maimoun et al., 2015) and 5610 g/h for CO_2 (average value calculated based on Khan et al., 2006; Khan et al., 2009; Maimoun et al., 2013)

5.4 Results and discussion

5.4.1 Future waste forecast and performance of ANN models at the testing stage

The mean values of recyclables between the training stage and testing stages of the input data are quite similar, as shown in Table 5-1. On the other hand, the mean values of garbage are higher in the testing stage than the training stage, probably due to an increasing generation trend and a longer garbage modeling period (2004-2018). More data variations in the garbage data than the recyclable set are also observed, with higher standard deviations in both training and testing stages. The computed errors of ANN models in the testing stage for recyclables and garbage predictions are shown in Figures 5-3a and b. In general, most MAPE of the waste prediction models were lower than 15% and model performances were comparable to other published ANN waste studies (Azadi and Karimi-Jashni 2016, Kannangara et al., 2018). Sub-area RMBU11 on the west side of the study area had the lowest MAPE at 10.92% and 11.76% for both recyclables and garbage respectively compared to the three other sub-areas. This can be partly explained by the ratio of Max and Min values of the input data in the testing stage were the lowest at 1.83 and 2.13 respectively compared to those of the three other sub-areas (Table 5-1). It appears that the time-series ANN modeling results are sensitive to the extreme values (max and min) in the data sets.
Figures 5-3c and 5-3d present future recyclables and garbage generation rates for four sub-areas from RMBU11 to RMBU14. From the figure, it can be seen that the sub-area RMBU13 located at the north had the highest rates for both recyclables and garbage at 13.20 kg/hh·2weeks and 11.97 kg/hh·week, respectively. An upward generation trend is predicted by the ANN model at RMBU13, as the sub-area has a moderate mean number of recyclables (10,948 kg/2weeks) and garbage (13,028 kg/2weeks) during the respective training periods, as shown in Table 5-1. Sub-area RMBU12 had the lowest rate for recyclables at 10.06 kg/hh·2weeks while RMBU11 had the lowest rate for garbage at 9.84 kg/hh·week. Compared to RMBU13, decreasing trends are predicted by ANN in these two sub-areas based on the historical waste records.
Figure 5-3: Results of ANN: a) MAPE of recyclables; b) MAPE of garbage; c) Future recyclables generation rate; d) Future garbage generation rate
5.4.3 GIS - Network Analysis - VRP solutions

5.4.3.1 Effects of waste mass and waste composition on optimized GIS truck travel distances

Figure 5-4 demonstrates travel distances of recyclables trucks (Scenarios R1, R2, and R3), garbage trucks (Scenarios G1, G2, and G3), and dual-compartment trucks (Scenarios M1, M2, and M3) for four sub-collection areas RMBU11 to RMBU14 in the target study area. Truck travel distances are generally higher in RMBU11 and RMBU 14 in all cases, probably due to the higher numbers of collection points (1024 households for RMBU11 and 964 households for RMBU14, as shown in Table 5-3).

Recyclables truck travel distances of scenario R2 (Figure 5-4a) were found more sensitive than those of scenarios R1 and R3. Three out of four sub-areas (RMBU11, RMBU13, RMBU14) had recyclables truck travel distances actually increase for scenario R2, the scenario with the paper composition reduced 15% compared to the baseline scenario (R1). Only sub-area RMBU12 had recyclables truck travel distances drastically changed for scenario R3, the scenario with the paper composition increased 15% compared to scenario R1. The density of recyclables in trucks of the scenario R2 (165.6 kg/m³) was lower than that of scenario R3 (232.7 kg/m³) (Table 5-2) or recyclables of scenario R2 were likely to occupy more space in trucks as compared to scenario R3. Comparing to scenario R1, the truck travel distance of scenario R2 of sub-area RMBU11 increased the most among the four sub-areas, from 90.87 km to 117.66 km, or 29.48%. Contrary to the other sub-areas, sub-area RMBU12 had the recyclable truck travel distance of scenario R2 slightly decreasing from 79.23 km to 77.81 km compared to scenario R1 despite scenario R2 having the same recyclables composition for all sub-
areas. The difference in the recyclables generation rate of each sub-area (Figure 5-3c) and the number of households (Table 5-3) also contributed to differences in its truck travel distances.

From Figure 5-4b, it can be seen that garbage truck travel distances seemed constant among scenarios G1, G2, and G3 when the garbage composition changed. This is probably due to the fact that garbage has a noticeably higher density inside the compactor truck, with densities ranging from 296.4 to 398.7 kg/m$^3$ (Table 5-2), almost doubling the densities of recyclables. A closer look at the required garbage truck trips in the four subareas are constant despite of the scenarios (Table 5-3). The truck travel distances of the dual-compartment trucks are much higher, ranging from 155.42 km to 233.85 km (Figure 5-4c).
Figure 5-4: Travel distances: a) of recyclables trucks; b) of garbage trucks; c) of dual-compartment trucks in four sub-areas
Figure 5-4 suggests that recyclable truck routes are more sensitive to waste composition than those of the garbage. To demonstrate visually the optimal truck route changes when the waste composition changes, an example of recyclables truck routes in sub-area RMBU14 was provided. Figure 5-5a, 5-5c, and 5-5e demonstrate the three truck routes of the baseline scenario (R1) while Figure 5-5b, 5-5d, 5-5f, and 5-5g present the four truck routes of scenario R2. Compared to those of scenario R1, fewer number of stops were served in route 1, route 2, and route 3 of scenario R2. As such, scenario R2 needed an additional route (route 4, shown in Figure 5-5g) to serve all the households in this sub-area. Route 3 of scenario R2 (Figure 5-5f) was much shorter than that of scenario R1 (Figure 5-5e), signifying that there may likely be excess space in the shorter route 3 truck, but is needed because all the others have been already filled. This is likely because the volume of recyclables of scenario R2 is larger compared to that of scenario R1 due to changes in recyclables composition and therefore the density of recyclables. Changes in the volume of waste generated at each stop led to changes in the number of stops that a 20 m³ recyclables truck can serve in a trip in a given area (Table 5-3). Thus, truck travel distances are greatly affected by changes in waste composition.
Figure 5-5: Example of recycling material truck routes in 2023: a, c, e) Routes using waste composition in 2014; b, d, f, g) Routes using waste composition in 2023 (Maps used were from © OpenStreetMap contributors, 2018)
5.4.3.2 Comparison of total truck travel distances and collection time

Comparison of total truck travel distances and time of different scenarios in the study area was showed in Figure 5-6. Overall, the total recyclables truck travel distance of scenario R2 grew from 331.0 km to 396.8 km or 19.89%, while the total recyclables truck travel distance of scenario R3 slightly went down from 331.0 km to 311.0 km or 6.04% compared to that of the baseline scenario (R1). More paper in recyclables in R2 led to less truck travel distances because the overall density of the recyclables was denser and therefore truck space was utilized more efficiently. On the other hands, less paper and more plastics in recyclables made truck travel distances rise. A similar trend was observed for total dual-compartment truck travel distances. The total dual-compartment truck travel distance changed from 697.6 km to 791.7 km (+13.48%) for the M2 route and from 697.6 km to 670.9 km (-3.83%) for the M3 route when recyclables and garbage composition changed in comparison to the baseline scenario of M1. Garbage truck travel distances for difference scenario is surprising constant. Results suggest the optimized truck travel distance using GIS is very sensitive to the recyclables and garbage composition, at least for the target areas considered in this study.

Figure 5-6a shows that the total dual-compartment truck travel distances (Scenarios M1, M2, and M3) were lower than those of both separated recyclables and garbage truck travel distances when added together (Scenarios R1+G1, R2+G2, R3+G3). The dual-compartment truck scenarios M1, M2, and M3 had shorter total travel distances amounting to 697.6 km, 791.7 km, and 670.9 km, respectively, when compared to the single-compartment truck scenarios R1+G1, R2+G2, and R3+G3 which had longer total travel distances amounting to 818.2 km, 882.6 km, and 798.7 km, respectively.
Compared to the separated recyclables and garbage trucks, the use of dual-compartment trucks can save 10.3% to 16.0% of total travel distances. In addition, less truck trips are required in each sub-area (Table 5-3).

However, in terms of total collection time, separated recyclables and garbage trucks were likely to be more efficient (Figure 5-6b). Collecting recyclables and garbage by separated trucks helped save 8.1 hours to 10.3 hours per week, or 15.7% to 19.8% of collection time per week. Collecting recyclables and garbage by separate trucks versus dual-compartment trucks was also found to save collection time according to some studies (Nguyen and Wilson. 2010; Maimoun et al. 2016). This is perhaps due to the pick-up times of the dual-compartment trucks at each stop (27 sec) being 3 times higher than that of either the garbage truck or the recyclables truck (9 sec), as shown in Table 5-3. Differences between scenarios are less obvious when results are expressed in collection time. Results suggest that both truck travel distance and collection time should be considered when perform GIS-based route design and optimization.
Figure 5-6: Comparison among different scenarios in term of a) Total truck travel distance; b) Total collection time
5.4.3.3 Air emissions from waste collection trucks at the study area

Air emissions parameters CO\textsubscript{2} and NO\textsubscript{x} from the waste collection trucks are shown in Figure 5-7. Recyclables trucks generated the lowest emissions, generally less than 590,000 g CO\textsubscript{2} per week and 3,500 g NO\textsubscript{x} per week. This is due to travel distances of recyclables trucks being the shortest (56.9 km - 117.7 km) when compared to garbage trucks and dual-compartment trucks (Figure 5-4). Dual-compartment trucks produced slightly less air emissions than both recycling trucks and garbage trucks did (Figure 5-7) due to a lower travel distance, despite its total collection time was higher than that of both recycling trucks and garbage trucks collection time (Figure 5-6b). It should be noted that both travel distance and idle time are included in the calculation of air emissions (eq. 5-4). Air emissions from garbage trucks were stable as garbage trucks had minor changes in total travel distances and collection times in each scenario. Unlike CO\textsubscript{2} emissions, NO\textsubscript{x} emissions are more sensitive to the truck idle time. Emissions from traveling are considerably higher in this study, as such, NO\textsubscript{x} emissions are found less sensitive between scenarios than that of CO\textsubscript{2}. 
Figure 5-7: CO₂ and NOₓ emissions from the truck in different scenarios
In Chapter 5, a combination of ANN time series and GIS -Network Analysis model was developed to examine how waste characteristics and type of waste collection affected on optimal truck waste routes. The results suggested that the optimal recyclables truck can change -6.04% to 29.48% when changing in recyclables composition. The dual-compartment truck model can save 10.3% to 16.0% of total travel distance when compared to using separated garbage and recyclables trucks. The study results can be applied for solid waste master plan preparation and waste truck route designs in the future.
6. Conclusion from the studies

The work reported in this thesis is significant and of great importance to researchers working on numerical modeling of non-hazardous WMS. Key conclusions from the four studies are summarized in the following four sections.

6.1 Effects of time lag on ANN Yard Waste Modeling

Factors contributing to municipal yard waste generation are time sensitive and the effects of a lag time must be explicitly considered in an accurate and theoretically sound prediction model. In Chapter two, correlation analysis was conducted to identify various climatic and socio-economic variables. Eight time series ANN models were developed, and a total of 128 scenarios were considered with respect to optimal lag time. The optimal lag time of all of the models, except for the Temp model ranged from 1 to 5 weeks. It was found that the models of the climatic group seemed to have more accurate and precise results in comparison with those of either the socio-economic or hybrid group. An inverse relationship between R Value and MSE was generally observed.

The Temp and Temp-Pop models were found as the two better weekly yard waste generation models with low MAPEs of 21.46% and 18.72% respectively for the testing stage. The lag time predicted by the training and testing sets agreed within ± 1 week in this study. At the training stage, there were negligible differences between the numbers of neurons on model performance. The MSE of the models was however much more sensitive to the number of neurons at the testing stage. The optimal ANN model was the Temp-Pop model with a 2-11-1 structure. The results from this study have highlighted the importance of climatic variables and their lag effects on the performance of yard waste prediction models.
The work reported in Chapter two is the first study using ANN modeling to carry out weekly municipal yard waste prediction testing using optimized lag times. This method proved effective as it reduced MSE by a substantial amount, for example it was reduced by 55.4% for the Temp model (train) as lag time was adjusted from 1 week to 10 weeks. Error was further reduced on the models with optimal structures. This proposed method can be valuable for refining yard waste models in any municipality, possibly leading to more effective waste system design and operations. This study approach can be applied for prediction of E-waste in the future, though continuous time series data is required.

6.2 Dual phase Municipal Solid Waste Collection

Chapter three applied a GIS (location-allocation and vehicle route problem) to obtain the best number of TCPs and their spatial distributions to satisfy conditions in both the pre-collection phase and the collection phase for the minimization of collection costs. A total of 30 scenarios were considered in the study area in 5 wards in Hong Bang district, Hai Phong city, Vietnam. The handcart and truck collection costs were then added together and the resulting system costs were compared among scenarios to reveal the interrelationships of the parameters.

It was found that the $D_{\text{max}}$ greatly affected both pre-collection and collection phases which also affected total system costs. The maximum distance from a source to a TCP of 500 m was deemed to be appropriate with the 1.41 km² study area. The percentage coverage of the different scenarios with $D_{\text{max}}= 500$ m was the highest with a coverage of over 99.9%. The number and spatial distribution of candidate TCPs were found critical to the optimization process using location-allocation.
GIS minimize facility solution performed well in the location of the TCPs in the study area, however they were found ineffective in identifying the scenario with the lowest system cost. Handcart travel distances, but not the total number of handcarts, were sensitive to the TCP spatial distribution. Travel handcart collection costs fell when the number of TCPs increased from 7 to 19 and then slightly rose when the number of TCPs rose from 19 to 25. In contrast, truck collection costs continued to climb when the number of TCPs increased. The truck capacity played a significant role in the truck travel distances and the overall efficiency of the collection system. There was a fluctuation of total collection costs among different scenarios with different number of TCPs.

The scenario with 11 TCPs and the maximum distance of 500 m was found to be the most economical option for the study area with a population density of 20,177/km². Each of the TCP served about 2,586 people in an area of 0.11 km². The recommended scenario helped to reduce truck travel distances by 13.76% compared to status quo.

This study approach can be extended to other systems, or incorporate real-time data to the optimization process.

6.3 Landfill Gas Model Optimization

LandGEM, Afvalzorg, and IPCC models were analyzed in Chapter four to determine the optimal FOD model parameters for two municipal solid waste landfills located in a cold semi-arid climate. Continuous daily (Regina) and monthly (Saskatoon) field gas recovery measurements were used for curve fitting by minimizing the residual sum of squares. Different iterative methods were examined and compared. These FOD models with default inputs yielded gas estimates substantially higher than the measured data in all cases. An inverse relationship was observed between k and the $L_0$ or DOC
values. The mean percentage errors between the estimated and measured values of the 3 models in both sites were between 55% - 135%, suggesting the defaults are not applicable to Saskatchewan landfills. The mean percentage errors were found to be consistently lower in Saskatoon than Regina despite the similarities of the landfills. It is found that the LandGEM and the IPCC models provide accurate prediction on gas generation peak at the Regina landfill.

The results of the three models, using optimal parameters, seemed to fit well to the actual data with mean percentage errors ranging from 11.60% to 19.93% the Regina landfill and about 1.65% to 10.83% for the Saskatoon landfill. In general, the optimal k values obtained from this study (k_{LandGEM}= 0.010 yr^{-1}, k_{Afvalzorg}= 0.010 yr^{-1}, k_{IPCC}= 0.020 yr^{-1}) were lower than those reported from the literature, whereas the L_0 or DOC values (L_{LandGEM}= 100 m^3/Mg, L_{Afvalzorg}= 97 m^3/Mg, L_{IPCC}= 91 m^3/Mg) were consistent with reported values. There were negligible differences between the performance of these FOD models, however, the Regina and Saskatoon landfills failed to demonstrate the potential benefits of using a more complex FOD model such as the IPCC model.

The RSS values using different FOD models and iterative approaches were similar, except for iterative method 1 of the IPCC model. Method 1 (fit k and L_0 or DOC simultaneously) in general tended to overestimate the gas generation rates. It appears that iterative method 2 (change k first, then L_0 or DOC) provided slightly more consistent results, at least using Regina data from the present study.

Climate change has an effect on weather patterns, and environmental factors which may affect landfill gas generation. Understanding how methane generation rates
differ at landfills located in different climates is important and relevant to today’s discussions on climate change.

The results are significant to the research community, particularly in many greenhouse gas inventory studies. However, the optimal parameters reported are site specific and must be checked when apply elsewhere.

6.4 Waste Characteristics and Collection Route Optimization

ANN time series and GIS - Network Analysis - VRPs models were combined to examine how waste composition and the mass of waste affects truck routes as well as air emissions from the trucks. Four Monday routes were selected as the target areas in the City of Austin. An ANN time series was applied to forecast recyclables and garbage generation rates in the year 2023. The ANN model showed better results when waste input data had fewer extreme values for both recyclables and garbage. The MAPE ranged from 10.92% to 14.83% for recyclables and from 11.76% to 16.51% for garbage.

Composition and compaction density of both garbage and recyclables were explored in the models. A total of 36 scenarios were studied. Optimal truck routes were found by using GIS - Network Analysis - VRPs. The results showed the recyclables truck travel distance of scenarios R2 or R3 changed for all four sub-areas when their recyclables composition changed. The largest increase in recyclables truck travel distance outside the collection area was observed in scenario R2 for sub-area RMBU11, at 29.48% compared to baseline scenario R1. The total recyclables truck travel distance of scenario R2 grew 19.89% while the figure for scenario R3 slightly went down 6.04% compared to that of scenario R1. Travel distances of garbage trucks were on the other hand less sensitive to the changes in compositions in each sub-collection area.
Results suggest waste composition and therefore waste density are important in GIS route optimization studies and these variables must be explicitly considered. In this study, a higher percentage of paper in recyclables led to less truck travel distance within the WMS, and the dual-compartment truck model can save 10.3% to 16.0% of total travel distance when compared to using separated garbage and recyclables trucks. However, separated garbage and recyclables trucks were likely to be more efficient in terms of reduced total collection time, at least using data from the study area. Dual-compartment trucks produced lower emissions than both recycling trucks and garbage trucks in the simulations. This is heavily due to the savings in distance used by dual-compartment trucks. This distance savings even offsets the increased emissions from higher dual-compartment truck idling time ($T_i$, equation 5-4) as the engine is at a lower power setting than when travelling greater distances ($D_c$ and $D_{oc}$, equation 5-4). This leads to a tradeoff between truck distance and emissions versus total collection time.

This study has important implications for WMS planners as they need to look very carefully at their municipalities’ waste composition. Waste compositions change with time due to changes in waste generators’ behaviours, waste polices, as well as attitudes towards recycling. ANN time-series model proves to be a promising tool for municipal waste prediction. The integration of ANN prediction model with GIS optimization reported in this study reveals the interrelationships between waste composition and GIS optimized routes and allows WMS managers to better response to the changes in waste composition.


Canada Climate Normals, 2016. http://climate.weather.gc.ca/climate_normals/results_1981_2010_e.html?searchType=stnName&txtStationName=Regina&searchMethod=contains&txtCentralLatMin=0&txtCentralLatSec=0&txtCentralLongMin=0&txtCentralLongSec=0&stnID=3002&dispBack=0 (accessed 15.05.2016).


160


Appendices

A1 – Chapter Two

The equation of Nonlinear Autoregressive with External Input (NARX) model is provided below:

Predict series $y(t)$ given $d$ past values of $y(t)$ and another series $x(t)$.

$y(t) = f(x(t-1), \ldots, x(t-d), y(t-1), \ldots, y(t-d))$ (A1-1)

(Source: MathWorks Inc., 2017)
Results of ANN-NARX

Table A1-1: R values of different scenarios

<table>
<thead>
<tr>
<th>Lag time (week)</th>
<th>Temp (Train)</th>
<th>Temp (Test)</th>
<th>Nasdaq (Train)</th>
<th>Nasdaq (Test)</th>
<th>Temp-Wind (Train)</th>
<th>Temp-Wind (Test)</th>
<th>Temp-Nasdaq (Train)</th>
<th>Temp-Nasdaq (Test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.87</td>
<td>0.81</td>
<td>0.85</td>
<td>0.75</td>
<td>0.88</td>
<td>0.80</td>
<td>0.86</td>
<td>0.80</td>
</tr>
<tr>
<td>2</td>
<td>0.88</td>
<td>0.80</td>
<td>0.85</td>
<td>0.76</td>
<td>0.88</td>
<td>0.82</td>
<td>0.88</td>
<td>0.80</td>
</tr>
<tr>
<td>3</td>
<td>0.89</td>
<td>0.81</td>
<td>0.85</td>
<td>0.74</td>
<td>0.89</td>
<td>0.82</td>
<td>0.89</td>
<td>0.82</td>
</tr>
<tr>
<td>4</td>
<td>0.90</td>
<td>0.79</td>
<td>0.85</td>
<td>0.71</td>
<td>0.90</td>
<td>0.81</td>
<td>0.90</td>
<td>0.77</td>
</tr>
<tr>
<td>5</td>
<td>0.90</td>
<td>0.80</td>
<td>0.86</td>
<td>0.69</td>
<td>0.90</td>
<td>0.81</td>
<td>0.90</td>
<td>0.81</td>
</tr>
<tr>
<td>6</td>
<td>0.91</td>
<td>0.81</td>
<td>0.86</td>
<td>0.73</td>
<td>0.90</td>
<td>0.82</td>
<td>0.91</td>
<td>0.80</td>
</tr>
<tr>
<td>7</td>
<td>0.91</td>
<td>0.78</td>
<td>0.86</td>
<td>0.71</td>
<td>0.91</td>
<td>0.80</td>
<td>0.91</td>
<td>0.77</td>
</tr>
<tr>
<td>8</td>
<td>0.90</td>
<td>0.82</td>
<td>0.86</td>
<td>0.72</td>
<td>0.91</td>
<td>0.82</td>
<td>0.91</td>
<td>0.82</td>
</tr>
<tr>
<td>9</td>
<td>0.92</td>
<td>0.81</td>
<td>0.85</td>
<td>0.72</td>
<td>0.93</td>
<td>0.77</td>
<td>0.90</td>
<td>0.78</td>
</tr>
<tr>
<td>10</td>
<td>0.93</td>
<td>0.82</td>
<td>0.86</td>
<td>0.73</td>
<td>0.93</td>
<td>0.78</td>
<td>0.91</td>
<td>0.77</td>
</tr>
<tr>
<td>11</td>
<td>0.92</td>
<td>0.85</td>
<td>0.86</td>
<td>0.77</td>
<td>0.92</td>
<td>0.81</td>
<td>0.90</td>
<td>0.78</td>
</tr>
<tr>
<td>12</td>
<td>0.92</td>
<td>0.82</td>
<td>0.88</td>
<td>0.71</td>
<td>0.91</td>
<td>0.85</td>
<td>0.92</td>
<td>0.77</td>
</tr>
<tr>
<td>13</td>
<td>0.92</td>
<td>0.80</td>
<td>0.86</td>
<td>0.72</td>
<td>0.93</td>
<td>0.83</td>
<td>0.91</td>
<td>0.79</td>
</tr>
<tr>
<td>14</td>
<td>0.93</td>
<td>0.79</td>
<td>0.86</td>
<td>0.76</td>
<td>0.91</td>
<td>0.82</td>
<td>0.91</td>
<td>0.80</td>
</tr>
<tr>
<td>15</td>
<td>0.91</td>
<td>0.78</td>
<td>0.86</td>
<td>0.74</td>
<td>0.93</td>
<td>0.80</td>
<td>0.92</td>
<td>0.79</td>
</tr>
<tr>
<td>16</td>
<td>0.92</td>
<td>0.82</td>
<td>0.89</td>
<td>0.77</td>
<td>0.91</td>
<td>0.80</td>
<td>0.93</td>
<td>0.78</td>
</tr>
</tbody>
</table>
Table A1-1: R value of different scenarios (continue)

<table>
<thead>
<tr>
<th>Lag time (week)</th>
<th>Temp-Pop (Train)</th>
<th>Temp-Pop (Test)</th>
<th>Temp-Wind-Nasdaq (Train)</th>
<th>Temp-Wind-Nasdaq (Test)</th>
<th>Temp-Wind-Pop (Train)</th>
<th>Temp-Wind-Pop (Test)</th>
<th>Nasdaq-Pop (Train)</th>
<th>Nasdaq-Pop (Test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.87</td>
<td>0.83</td>
<td>0.87</td>
<td>0.80</td>
<td>0.87</td>
<td>0.80</td>
<td>0.87</td>
<td>0.77</td>
</tr>
<tr>
<td>2</td>
<td>0.88</td>
<td>0.81</td>
<td>0.88</td>
<td>0.77</td>
<td>0.87</td>
<td>0.80</td>
<td>0.87</td>
<td>0.79</td>
</tr>
<tr>
<td>3</td>
<td>0.89</td>
<td>0.80</td>
<td>0.88</td>
<td>0.78</td>
<td>0.88</td>
<td>0.79</td>
<td>0.88</td>
<td>0.76</td>
</tr>
<tr>
<td>4</td>
<td>0.90</td>
<td>0.78</td>
<td>0.90</td>
<td>0.76</td>
<td>0.89</td>
<td>0.80</td>
<td>0.89</td>
<td>0.78</td>
</tr>
<tr>
<td>5</td>
<td>0.91</td>
<td>0.84</td>
<td>0.91</td>
<td>0.81</td>
<td>0.91</td>
<td>0.78</td>
<td>0.90</td>
<td>0.79</td>
</tr>
<tr>
<td>6</td>
<td>0.90</td>
<td>0.80</td>
<td>0.90</td>
<td>0.80</td>
<td>0.90</td>
<td>0.78</td>
<td>0.89</td>
<td>0.79</td>
</tr>
<tr>
<td>7</td>
<td>0.92</td>
<td>0.75</td>
<td>0.91</td>
<td>0.77</td>
<td>0.90</td>
<td>0.80</td>
<td>0.90</td>
<td>0.80</td>
</tr>
<tr>
<td>8</td>
<td>0.92</td>
<td>0.80</td>
<td>0.92</td>
<td>0.77</td>
<td>0.92</td>
<td>0.78</td>
<td>0.90</td>
<td>0.79</td>
</tr>
<tr>
<td>9</td>
<td>0.90</td>
<td>0.78</td>
<td>0.91</td>
<td>0.78</td>
<td>0.91</td>
<td>0.79</td>
<td>0.91</td>
<td>0.78</td>
</tr>
<tr>
<td>10</td>
<td>0.91</td>
<td>0.81</td>
<td>0.91</td>
<td>0.81</td>
<td>0.92</td>
<td>0.79</td>
<td>0.91</td>
<td>0.81</td>
</tr>
<tr>
<td>11</td>
<td>0.93</td>
<td>0.80</td>
<td>0.90</td>
<td>0.78</td>
<td>0.92</td>
<td>0.77</td>
<td>0.91</td>
<td>0.77</td>
</tr>
<tr>
<td>12</td>
<td>0.91</td>
<td>0.80</td>
<td>0.91</td>
<td>0.82</td>
<td>0.92</td>
<td>0.81</td>
<td>0.91</td>
<td>0.79</td>
</tr>
<tr>
<td>13</td>
<td>0.92</td>
<td>0.80</td>
<td>0.91</td>
<td>0.81</td>
<td>0.91</td>
<td>0.80</td>
<td>0.89</td>
<td>0.81</td>
</tr>
<tr>
<td>14</td>
<td>0.91</td>
<td>0.79</td>
<td>0.92</td>
<td>0.78</td>
<td>0.92</td>
<td>0.79</td>
<td>0.89</td>
<td>0.81</td>
</tr>
<tr>
<td>15</td>
<td>0.92</td>
<td>0.84</td>
<td>0.92</td>
<td>0.77</td>
<td>0.91</td>
<td>0.78</td>
<td>0.90</td>
<td>0.79</td>
</tr>
<tr>
<td>16</td>
<td>0.93</td>
<td>0.80</td>
<td>0.92</td>
<td>0.81</td>
<td>0.91</td>
<td>0.79</td>
<td>0.90</td>
<td>0.80</td>
</tr>
</tbody>
</table>
Table A1-2: Mean Square Error (Training stage)

<table>
<thead>
<tr>
<th>Lag time (week)</th>
<th>Temp</th>
<th>Nasdaq</th>
<th>Temp-Wind</th>
<th>Temp-Nasdaq</th>
<th>Temp-Pop</th>
<th>Temp-Wind-Nasdaq</th>
<th>Temp-Wind-Pop</th>
<th>Nasdaq-Pop-Temp</th>
</tr>
</thead>
</table>
Table A1-3: Mean Square Error (Testing stage)

<table>
<thead>
<tr>
<th>Lag time (week)</th>
<th>Temp</th>
<th>Nasdaq</th>
<th>Temp-Wind</th>
<th>Temp-Nasdaq</th>
<th>Temp-Pop</th>
<th>Temp-Wind-Nasdaq</th>
<th>Temp-Wind-Pop</th>
<th>Nasdaq-Pop-Temp</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>33,085</td>
<td>38,728</td>
<td>36,048</td>
<td>33,730</td>
<td>28,003</td>
<td>32,376</td>
<td>32,177</td>
<td>38,063</td>
</tr>
<tr>
<td>2</td>
<td>35,608</td>
<td>40,600</td>
<td>31,313</td>
<td>32,228</td>
<td>31,307</td>
<td>36,400</td>
<td>31,422</td>
<td>33,982</td>
</tr>
<tr>
<td>3</td>
<td>34,640</td>
<td>41,008</td>
<td>30,946</td>
<td>30,378</td>
<td>34,116</td>
<td>37,396</td>
<td>35,641</td>
<td>37,416</td>
</tr>
<tr>
<td>4</td>
<td>35,956</td>
<td>47,250</td>
<td>35,852</td>
<td>38,403</td>
<td>36,887</td>
<td>40,678</td>
<td>35,842</td>
<td>36,540</td>
</tr>
<tr>
<td>5</td>
<td>35,462</td>
<td>47,462</td>
<td>37,244</td>
<td>37,422</td>
<td>28,829</td>
<td>31,976</td>
<td>36,158</td>
<td>35,617</td>
</tr>
<tr>
<td>6</td>
<td>36,553</td>
<td>43,990</td>
<td>34,292</td>
<td>35,041</td>
<td>37,181</td>
<td>35,515</td>
<td>36,362</td>
<td>37,032</td>
</tr>
<tr>
<td>7</td>
<td>41,904</td>
<td>46,511</td>
<td>43,126</td>
<td>42,841</td>
<td>39,789</td>
<td>37,160</td>
<td>35,753</td>
<td>37,068</td>
</tr>
<tr>
<td>8</td>
<td>32,483</td>
<td>46,281</td>
<td>34,659</td>
<td>31,347</td>
<td>37,393</td>
<td>41,293</td>
<td>37,383</td>
<td>38,696</td>
</tr>
<tr>
<td>9</td>
<td>41,364</td>
<td>46,842</td>
<td>41,881</td>
<td>37,238</td>
<td>36,289</td>
<td>39,508</td>
<td>35,910</td>
<td>37,579</td>
</tr>
<tr>
<td>10</td>
<td>36,060</td>
<td>57,597</td>
<td>44,707</td>
<td>41,115</td>
<td>32,374</td>
<td>41,784</td>
<td>36,356</td>
<td>36,228</td>
</tr>
<tr>
<td>11</td>
<td>30,158</td>
<td>39,900</td>
<td>37,868</td>
<td>38,456</td>
<td>39,181</td>
<td>37,850</td>
<td>38,848</td>
<td>39,900</td>
</tr>
<tr>
<td>12</td>
<td>35,241</td>
<td>47,882</td>
<td>31,649</td>
<td>40,719</td>
<td>36,843</td>
<td>38,840</td>
<td>35,125</td>
<td>38,335</td>
</tr>
<tr>
<td>13</td>
<td>35,989</td>
<td>48,375</td>
<td>39,655</td>
<td>37,864</td>
<td>36,060</td>
<td>36,686</td>
<td>35,599</td>
<td>34,058</td>
</tr>
<tr>
<td>14</td>
<td>38,762</td>
<td>44,437</td>
<td>34,833</td>
<td>36,274</td>
<td>38,358</td>
<td>39,464</td>
<td>39,316</td>
<td>34,621</td>
</tr>
<tr>
<td>15</td>
<td>41,861</td>
<td>45,323</td>
<td>36,650</td>
<td>38,619</td>
<td>34,374</td>
<td>41,101</td>
<td>39,296</td>
<td>38,580</td>
</tr>
<tr>
<td>16</td>
<td>41,672</td>
<td>41,862</td>
<td>38,324</td>
<td>42,346</td>
<td>36,168</td>
<td>39,032</td>
<td>39,573</td>
<td>38,510</td>
</tr>
</tbody>
</table>
Figure A1-1: Comparison of the optimal model results (Temp-Pop) in training stage (from October 2004 to March 2016) and in testing stage (from April 2016 to April 2018) with the observation data from
A2 – Chapter Three

Results from the Maximize Coverage and the Minimum Facility functions with different $D_{\text{max}}$ and CP values.

Figure A2-1 Result of Maximize Coverage with $D_{\text{max}}=300$, CP= 7

Figure A2-2 Result of Maximize Coverage with $D_{\text{max}}=300$, CP= 8

Figure A2-3 Result of Maximize Coverage with $D_{\text{max}}=300$, CP= 10

Figure A2-4 Result of Maximize Coverage with $D_{\text{max}}=300$, CP= 11

Figure A2-5 Result of Maximize Coverage with $D_{\text{max}}=300$, CP= 13

Figure A2-6 Result of Minimize Facility with $D_{\text{max}}=300$
Figure A2-7 Result of Maximize Coverage with $D_{\text{max}} = 300$, $CP= 16$

Figure A2-8 Result of Maximize Coverage with $D_{\text{max}} = 300$, $CP= 19$

Figure A2-9 Result of Maximize Coverage with $D_{\text{max}} = 300$, $CP= 22$

Figure A2-10 Result of Maximize Coverage with $D_{\text{max}} = 300$, $CP= 25$

Figure A2-11 Result of Maximize Coverage with $D_{\text{max}} = 400$, $CP= 7$

Figure A2-12 Result of Maximize Coverage with $D_{\text{max}} = 400$, $CP= 8$

Figure A2-13 Result of Maximize Coverage with $D_{\text{max}} = 400$, $CP= 10$

Figure A2-14 Result of Minimize Facility with $D_{\text{max}} = 400$
Figure A2-15 Result of Maximize Coverage with $D_{\text{max}} = 400$, CP= 13

Figure A2-16 Result of Maximize Coverage with $D_{\text{max}} = 400$, CP= 15

Figure A2-17 Result of Maximize Coverage with $D_{\text{max}} = 400$, CP= 16

Figure A2-18 Result of Maximize Coverage with $D_{\text{max}} = 400$, CP= 19

Figure A2-19 Result of Maximize Coverage with $D_{\text{max}} = 400$, CP= 22

Figure A2-20 Result of Maximize Coverage with $D_{\text{max}} = 400$, CP= 25

Figure A2-21 Result of Maximize Coverage with $D_{\text{max}} = 500$, CP= 7

Figure A2-22 Result of Minimize Facility with $D_{\text{max}} = 500$
Figure A2-23 Result of Maximize Coverage with $D_{\text{max}} = 500$, CP= 10

Figure A2-24 Result of Maximize Coverage with $D_{\text{max}} = 500$, CP= 11

Figure A2-25 Result of Maximize Coverage with $D_{\text{max}} = 500$, CP= 13

Figure A2-26 Result of Maximize Coverage with $D_{\text{max}} = 500$, CP= 15

Figure A2-27 Result of Maximize Coverage with $D_{\text{max}} = 500$, CP= 16

Figure A2-28 Result of Maximize Coverage with $D_{\text{max}} = 500$, CP= 19

Figure A2-29 Result of Maximize Coverage with $D_{\text{max}} = 500$, CP= 22

Figure A2-30 Result of Maximize Coverage with $D_{\text{max}} = 500$, CP= 25
Governor Equations of the three models:

**LandGEM’s formula:**

\[
Q_{CH_4} = \sum_{i=1}^{n} \sum_{j=0.1}^{1} kL_j \left( \frac{M_i}{10} \right) e^{-kt_i}
\]  

(A3-1)

Where:

- \(Q_{CH_4}\) = annual methane generation in the year of the calculation (m³/year)
- \(i\) = 1-year time increment
- \(t_{ij}\) = age of the \(j^{th}\) section of waste mass \(M_i\) accepted in the \(i^{th}\) year (decimal years, e.g., 3.2 years)
- \(n\) = (year of the calculation) - (initial year of waste acceptance)
- \(j\) = 0.1 - year time increment
- \(k\) = methane generation rate (year⁻¹)
- \(Lo\) = potential methane generation capacity (m³/Mg)

**Afvalzorg and IPCC’s formulas:**

\[
CH_4 \text{ gen } t = DDOCm \text{ decomp } t \cdot F \cdot \frac{16}{12}
\]

(A3-2)

\[
DDOCm \text{ decomp } t = DDOCm a t \cdot (1 - e^{-k})
\]

(A3-3)

\[
DDOCm a t = DDOCm d t + (DDOCm a t - 1 \cdot e^{-k})
\]

(A3-4)

\[
DDOCm = \text{waste mass} \times \text{DOC} \times \text{DOCf} \times \text{MCF}
\]

(A3-5)

Where:

- \(CH_4 \text{ gen } t\): Methane generated in year \(t\)
- \(F\): Fraction of methane (v/v) in generated landfill gas
- \(16/12\): Conversion factor of kg \(CH_4\) from kg C (molecular weight ratio: 16/12)
DDOCmdecomp t: DDOCm decomposed in the landfill in year t

DDOCma t: Mass of decomposable DOC accumulated in landfill at the end of year t

DDOCmd t: Mass of decomposable DOC deposited in landfill in the year t

DDOCma t-1: Mass of decomposable DOC accumulated in landfill at the end of year t-1

k: reaction rate constant (years⁻¹)

DDOCm: Mass of decomposable DOC

DOC: Degradable organic carbon

DOC_f: Fraction of DOC that can decompose

MCF: Correction factor for aerobic decomposition in the year of deposition
Nonlinear Autoregressive (NAR): Predict series $y(t)$ given $d$ past values of $y(t)$. The equation is given below:

$$y(t) = f(y(t-1), \ldots, y(t-d))$$  \hspace{1cm} (A4-1)

(Source: MathWorks Inc., 2017)
Figure A4-1: Comparison of a garbage NAR model results in training stage (from October 2004 to March 2016) and in testing stage (from April 2016 to April 2018) with the observation data from in sub-area RMBU11.
Figure A4-2: Recycling material truck routes in sub-area RMBU12 in 2023: a1-a3) Routes using waste composition in 2014 (scenario R1); b1-b2) Routes using waste composition in 2023 (scenario R3)
Figure A4-3: Recycling material truck routes in sub-area RMBU11 in 2023: a1-a3) Routes using waste composition in 2014 (scenario R1); b1-b4) Routes using waste composition in 2023 (scenario R2)
Figure A4-4: Dual chamber truck routes in sub-area RMBU11 in 2023: a1-a3) Routes using waste composition in 2014 (scenario M1); b1-b4) Routes using waste composition in 2023 (scenario M2)
Figure A4-5: Dual chamber truck routes in sub-area RMBU13 in 2023: a1-a3) Routes using waste composition in 2014 (scenario M1); b1-b4) Routes using waste composition in 2023 (scenario M2)
Figure A4-6: Dual chamber truck routes in sub-area RMBU14 in 2023: a1-a3) Routes using waste composition in 2023 (scenario M3); b1-b4) Routes using waste composition in 2014 (scenario M1)