SHORT-TERM VEHICLE TRAFFIC FLOW FORECASTING GREY MODEL FOR INTELLIGENT TRANSPORTATION SYSTEM PERFORMANCE

VINCENT BIRUNDU GETANDA

DOCTOR OF PHILOSOPHY

(Electrical Engineering)

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

2024

Short-Term Vehicle Traffic Flow Forecasting Grey Model for Intelligent Transportation System Performance

Vincent Birundu Getanda

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Electrical Engineering of the Jomo Kenyatta University of Agriculture and Technology

2024

DECLARATION

This thesis is my original work and has not been presented for a degree in any other University

Signature……Date……

Vincent Birundu Getanda

This thesis has been submitted for examination with our approval as University Supervisors

Signature……Date……

Dr. Peter Kamita Kihato, PhD

JKUAT, Kenya

Signature……Date……

Prof. Peterson Kinyua Hinga, PhD

JKUAT, Kenya

Signature……Date……

Prof. Hidetoshi Oya, PhD

Tokyo City University, Japan

DEDICATION

I dedicate this work to my parents, brothers, sisters and family with whom I share a happy life and whose efforts have changed my life style.

ACKNOWLEDGEMENT

Am indebted to all my supervisors who through formal consultations have made me shine. I sincerely thank you for your lovely commitment in giving advice, encouragement and constant guidance concerning this Ph.D. research study. I acknowledge the gratitude my family, parents, friends and all others deserve for they have been supporting me materially and financially. Thank you for spending your time and resources to support me in my endeavors. I extend my gratitude to my fellow course mates for their constant encouragement in the consolidation of this thesis. I wish to sincerely thank the entire management of Jomo Kenyatta University of Agriculture and Technology for the opportunity it gave me to pursue my doctorate course at the university. I do acknowledge the financial support by the African Development Bank under Grant MOE/HEST/06/2017-2018. Finally, I thank the Lord my God for making everything possible to me. God be the glory.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF NOMENCLATURES

Symbol	Definition
$\boldsymbol{a}$	Developing coefficient
$\boldsymbol{a_0}$	Fourier coefficient as determined in the Fourier coefficient vector P by the least square method
$\boldsymbol{a_m}$	Fourier coefficient as determined in the Fourier coefficient vector P by the least square method and $m = 1, \cdots, N$ is the mode number if the Fourier series is truncated after N terms
$\boldsymbol{A}$	Data matrix for a univariate GM
$\boldsymbol{A^{-1}}$	The inverse of matrix A
$\boldsymbol{A^T}$	The transpose of matrix A
$\boldsymbol{b}$	Grey control coefficient/grey input/grey action quantity
$\boldsymbol{b_2, \cdots, b_n}$	Multivariate GM parameters for n=number of time series
$\boldsymbol{b_m}$	Fourier coefficient as determined in the Fourier coefficient vector P by the least square method and $m = 1, \cdots, N$ is the mode number if the Fourier series is truncated after N terms
$\boldsymbol{\widehat{B}}$	Parameter vector for the GM
$\boldsymbol{C}$	Initial condition of the traditional GM(1,1)
$\boldsymbol{d}$	De-trended
$\boldsymbol{e}$	Error
$\boldsymbol{m}$	Total number of data points

$\boldsymbol{n}$	Number of time series
$\boldsymbol{N}$	Number of terms after which the Fourier series is truncated
$\boldsymbol{P}$	Fourier coefficient vector
$\boldsymbol{t}$	Time sample
$\boldsymbol{T}$	Period
$\boldsymbol{X^{(0)}}$	Original data series sequence
$\boldsymbol{X^{(1)}}$	AGO data sequence obtained by accumulating $X^{(0)}$
$\boldsymbol{\widehat{X}^{(1)}}$	Predicted value of $X^{(1)}$
$\boldsymbol{\widehat{X}^{(0)}}$	Final predicted/fitted data sequence of $X^{(0)}$ obtained from $\hat{X}^{(1)}$
$\boldsymbol{X}$	Data matrix for a multivariate GM
$\boldsymbol{X^{-1}}$	The inverse of matrix X
$\boldsymbol{X^{T}}$	The transpose of matrix X
$\boldsymbol{y}$	Measured vector for a univariate GM
$\boldsymbol{Y}$	Measured vector for a multivariate GM
$\boldsymbol{Z^{(1)}}$	Background value sequence obtained by mean generation of adjacent neighbors of the AGO
$\boldsymbol{\alpha_1}$	Low order parameter
$\boldsymbol{\alpha_2}$	High order parameter
δ	Percent error

LIST OF APPENDICES

ACRONYMS AND ABBREVIATIONS

AD	Actual Data
AGO	Accumulated Generating Operation
AGO{+}	AGO Operator
ARIMA	Autoregressive Integrated Moving Average
CBD	Central Business District
CEP	Complex Event Processing
CIA	Central Intelligence Agency
CNN-BILSTM	Convolutional Neural Network and Bidirectional Long Short-Term Memory
CPN	Colored Petri Nets
CPU	Central Processing Unit
CTM	Cell Transmission Model
CTMC	Continuous-Time Markov Chain
D1S1	Day 1 Site 1
D1-S2-W	Day 1 Site 2 Westward Direction
D1S4	Day 1 Site 4
D1S5	Day 1 Site 5
D1S7	Day 1 Site 7
D2S2	Day 2 Site 2
D2S5	Day 2 Site 5
D2S6	Day 2 Site 6
D2S7	Day 2 Site 7
D2-S7-E	Day 2 Site 7 Eastward Direction

D3S1	Day 3 Site 1
D3S2	Day 3 Site 2
D3S4	Day 3 Site 4
D3-S4-N	Day 3 Site 4 Northward Direction
D3S7	Day 3 Site 7
DBGM (1, *N*)	Multivariate Grey prediction Model based on Dynamic Background values
DGT	Data Grouping Technique
DP	Data Point
EAGTCN	Ensemble Attention based Graph Time Convolutional Networks
FASTNN	Filter Attention-based Spatiotemporal Neural Network
FDGM(1,1)	Fractional Discrete Grey Model
FGM(1,1)	Fourier residual-corrected Grey prediction Model
FL	Fuzzy Logic
FOBGM(1,N)	Fractional Order optimized OBGM(1,N) model
FSECA	Fourier Series Error Correction Approach
GGM(1,1)	Grouped Grey Model
GM	Grey Model
GM (m,*n*)	Grey Model, m is the order of the differential equation and *n* is the number of variables employed
GPEAe	Grey Prediction Evolution Algorithm based on the even grey model
GPS	Global Positioning System
GST	Grey System Theory
HMM	Hidden Markov Model

IaaS	Infrastructure as a Service
IAGO	Inverse Accumulated Generating Operation
IAGO{+}	IAGO Operator
ICT	Information and Communication Technology
IR	Infrared
ITS	Intelligent Transportation System
KNN	K-nearest neighbor
LSTM	Long Short-Term Memory Network
MAE	Mean Absolute Error
MAPD	Mean Absolute Percentage Deviation
MBV	Modification of Background Value
MBVGGM(1,1)	Modified Background Value Grouped Grey Model
MBVGM(1,1)	Modified Background Value Grey Model
MGO	Mean value Generating Operation
MIC	Modification of Initial Condition
MICGGM(1,1)	Modified Initial Condition Grouped Grey Model
MICGM(1,1)	Modified Initial Condition Grey Model
MITRAM	Microscopic Traffic Simulator
ML	Machine Learning
MOT	Motorcycle
NaN	Not a Number
NG	No Grouping
NGM(1,1)	Non-homogeneous Grey Model
OBGM(1,N)	Optimized Background value, GM(1,N) model

OGM(1,1)	Original Grey Model
PED	Pedestrian
PLC	Programmable Logic Controller
PSO	Particle Swarm Optimization
RFID	Radio Frequency Identification
RMSE	Root Mean Square Error
RMSPE	Root Mean Square Percentage Error
RVSA	Relative Variable Smoothing Approach
SARIMA	Seasonal ARIMA
SG	Strong Grouping
SGGM(1,1,r)	Shale Gas output Grey prediction Model
TSA	Three-Step Approach
UGM(1,1)	Unbiased prediction Grey Model
VEH	Vehicle
VSGFGM(1,*n*)	Variable Smoothed-Grouped data-Fourier series based Multivariate Grey Model
VSGGM(1,*n*)	Variable Smoothed-Grouped data Multivariate Grey Model
WG	Weak Grouping
WGM(1,1)	Wavelet residual-corrected Grey prediction Model

ABSTRACT

Vehicular traffic is continuously increasing around the world and the resulting congestion and pollution is a major concern to transportation specialists and decision makers. As population continues to grow it is a challenge to handle traffic demand, traffic jam, CO_2 emission, global warming and economic loss. Road capacity is not adequate. Roads and highways are unlikely to expand due to cost and dwindling land supply. To manage these issues it is critical to integrate intelligent transportation system (ITS) in the transport management systems. Short-term vehicle traffic flow forecasting by ITS is vital in proactively monitoring a vehicle traffic system. Unfortunately, effective traffic flow forecasting is a key problem of ITS. Therefore, performance improvement of the predictive models which can enhance the forecasting ability of ITS is very crucial. Hence the objective of this study was to develop a short-term vehicle traffic flow forecasting grey model (GM) for ITS performance. Hence, in this thesis the precision of the original GM is improved by three proposed methods namely data grouping technique (DGT), relative variable smoothing approach (RVSA) and a three-step approach (TSA). To further improve the GM's precision these new methods were combined with existing methods such as modification of background value (MBV), modification of initial condition (MIC) and Fourier series error correction approach (FSECA). Consequently, hybrid grey models were established. The accuracy improvement on the conventional grey models were measured by employing measures of model performance, namely root mean square error (RMSE), root mean square percentage error (RMSPE), mean absolute error (MAE) and the mean absolute percentage deviation (MAPD). The evaluation results revealed that the hybrid grey models outperformed the conventional GM in vehicle flow modelling and short-term forecasting. For instance in short-term vehicle traffic flow forecasting the improved models (GGM(1,1) and MBVGGM(1,1)) had good accuracy in the range of 80-90% compared to the corresponding conventional GM(1,1) and MBVGM(1,1) which had reasonable accuracy in the range of 50-80%. On the other hand in validating the DGT in improving the fitting accuracy of the conventional GM(1,3) the accuracy was improved from 60.3270% to 96.9706%. This was great improvement in the conventional GM(1,3)'s fitting accuracy. Further, the results of this research show that the proposed new methods i.e. the DGT, the RVSA and the TSA methods have the potential for improving the prediction accuracy of the conventional GMs. Hence the DGT in hybrid grey models can enhance the short-term forecasting ability of the ITS. A case study based on traffic data collected from Nairobi city, Kenya, was presented and analyzed to show the accuracy improvement in both the univariate (GM(1,1)) and multivariate (GM(1,3)) grey models. For instance from this case study computation of the RMSPE had shown that the fitting accuracy of GM(1,3) was improved from 69.7243% to 99.6281% by the TSA method. Thus an improved multivariate grey model can attain high traffic flow forecasting accuracies compared with an improved univariate grey model. Finally, the performance of the grouping technique based GMs on energy consumption and carbon dioxide emissions, outperformed the conventional GMs. From one of the presented empirical cases the grouping technique based multivariate GGM(1,3) attained an accuracy of 96.9706% against 60.3270% of the conventional GM(1,3). Thus the hybrid grey models developed in this thesis are multidisciplinary. However, in comparison with other state of the art improved GM such as the grey model with cosine term (GM(1,1|cos(ωt))), the performance of the proposed models was below that of the GM(1,1|cos(ωt)). In a recent research GM(1,1|cos(ωt)) had a mean absolute percentage error (MAPE) of 0.1% compared to 0.58% of the original GM(1,1). Therefore, there is need to investigate the performance of the proposed models in this research in comparison with the GM(1,1|cos(ωt)), in the future.

CHAPTER ONE

INTRODUCTION

1.1 Background

Road transport is the major means of mobility in major cities in the world. Cities are the highest populated urban areas in the world because of more births than deaths, availability of facilities, job opportunities or forced off rural land by natural disasters, skyrocketing land prices, to name a few. As population continues to grow, it is a challenge to handle traffic demand, traffic jam, carbon dioxide (CO_2) emission and global warming. Road capacity is not adequate. Roads and highways are unlikely to expand due to cost and dwindling land supply (Toan, 2018). Therefore, better management of traffic demand and capacity is critical to operating the current roadway systems at maximum capacity. Consequently, short-term traffic flow forecasting as an important component of Intelligent Transportation Systems (ITS) can optimize the road transport system operations (Ren et al., 2021).

An inadequate transportation system generates high costs and low customer service levels, which ultimately produces a negative economic impact (Zapata et al., 2013). Low mobility of vehicles, people and goods in urban areas is the greatest obstacle to economic growth and development of developing countries. Traffic congestion, environmental pollution, global warming and low economic growth rates are some of the negative impacts of this low mobility. To manage traffic congestion, ITS have been integrated into the transport management systems to aid the decision-making processes and decongest the roadways (Cheng et al., 2019). Technologies designed to capture and analyze the required real-time information are vital for proper management of a transport system (Zapata et al., 2013). Performance and safety of a transportation system can be improved by use of ITS. Zapata et al. (2013) defines ITS as the interconnection of different information systems aimed to capture, communicate, compute and assist decision making in order to allow smooth flow of vehicles and transportation means. In the past ITS have efficiently and effectively managed up the problem of traffic congestion around the world (Nizar, et al., 2022). Continued use of the current ITS in traffic management systems is likely not going to be efficient with

the rapid trend of urbanization. This is because effective traffic flow forecasting has been regarded as a key problem of ITS (Du et al., 2020; Shah et al., 2022). Thus, the main objective of this research is to develop a short-term vehicle traffic flow forecasting grey model for ITS performance. In particular this research dealt with the low prediction accuracy of the grey model for ITS performance. It is worthy to note that quantification of traffic congestion is a pre-requisite to the implementation of ITS in a road transport system.

1.2 Statement of the Problem

Traffic congestion on motorways is becoming an ever more pressing problem all over the world (Ren et al., 2021; Sun, et al., 2022; Toan, 2018). Road traffic jams continue to remain a major problem in most cities, especially in developing regions resulting in massive delays, increased fuel wastage, environmental pollution, accidents and economic losses (Jain et al., 2012). The most disturbing thing is that this congestion continues to intensify, without any sign of having a limit, thus becoming a nightmare that threatens the quality of urban life. Nairobi is among the most affected cities in the developing regions of the world (Mosoti & Moronge, 2015). As a case in point, Kiiru (2015) states that "there exists very diverse efficiency challenges in the Nairobi public transport provision that continually pose threats to the achievement of efficient transport system. Efficiency is one of the components of a vibrant transport system that cannot be overlooked". Thus, there is need for efficient solutions to this traffic congestion problem (Chama, 2015).

Traffic congestion is caused by many factors such as more people crowding into big cities but more importantly because people are moving more and frequently (Oladimeji et al., 2023). Most of the extra movement is done in cars. Availability of affordable cars has encouraged personal vehicle usage to places of work and recreational trips thus resulting in lesser use of public transport which has led to traffic congestion in urban areas. Additionally, the rapid trend of urbanization in the world, especially in urban areas, has come faster than the current scale of transport infrastructure investments and has also resulted into traffic congestion which is heavy during rush hours (Oladimeji et al., 2023). Therefore, there is need to cope with this increasing

demand for services and stay in control as urbanization causes city traffic to become increasingly dense.

Currently, the lowest capacity and old infrastructure of roads do not support the amount of vehicle flow that cause traffic congestion (Ali et al., 2021). Managing traffic congestion and traffic control calls for a clear understanding of traffic systems and sustainable development has become a priority in the modern society. Therefore, it is crucial to keep mobility safe and under control by leveraging on the available infrastructure and road capacities, yet providing commuters and road users with the most efficient means of travel. Currently, construction of the bypass highways, the Thika Superhighway and the expressway through Kenya's capital, Nairobi, have eased the traffic congestion but there still remains the problem of congestion and it is becoming an ever more pressing problem. Moreover, Shah et al. (2022) points out that despite massive investments in transportation-related infrastructure, traffic congestion remains a societal and public policy problem.

Therefore, as a solution to the problem of traffic congestion, an intelligent system that monitors traffic systems should be in place to keep a check on the efficiency of the transportation systems (Ren et al., 2021). ITS have been proposed as a solution to this issue of traffic congestion (Shah et al., 2022). ITS can play a vital role in monitoring a traffic system but quantification of congestion becomes essential in checking congestion in order to provide a sustainable transportation system (Jain et al, 2017). However, effectively eliminating and relieving traffic congestion is a significant challenge in ITS models (Kołodziej et al., 2022). Moreover, traffic prediction methods for ITS have been faced with numerous problems and challenges. Complex models, such as neural network and combination model, have complex processes and large computation, which are not conducive to the practical application of short-term traffic prediction. Currently, the prediction step size of short-term traffic flow prediction research is mainly single-step prediction and the prediction interval is within 1-15 minutes. The prediction input variables are mostly single variables and the massive traffic data are not effectively used (Chen & Chen, 2019). The use of machine learning (ML) for traffic prediction has the potential to transform ITS. In terms of estimating traffic volume, speed, and trip time, ML models have produced promising results.

However, data accessibility, model scalability and model interpretability are only a few of the issues that still need to be resolved (Manikandan et al., 2023). The focus of prediction research is to optimize the existing models, especially to improve the accuracy and applicability of prediction by combining models (Chen & Chen, 2019). Therefore, studies should concentrate on creating models that can handle traffic prediction model difficulties and boost the precision and efficiency of traffic forecast for ITS (Manikandan et al., 2023). This has motivated this research in dealing with the problem of limited precision of the conventional grey model, GM(1,*n*) in short-term forecasting of vehicle traffic flow using an improved GM(m,*n*). The accuracy improved GM(m,*n*) can enhance the performance of ITS in quantifying traffic congestion on roads and highways for the purpose of proactive vehicle traffic flow control.

1.3 Research Objectives

1.3.1 Main Objective

The main objective of this study is to develop a short-term vehicle traffic flow forecasting grey model for intelligent transportation system performance.

1.3.2 Specific Objectives

In order to accomplish the development of the short-term vehicle traffic flow forecasting grey model, the following specific objectives are addressed in this thesis.

1) To assess the effect of data grouping technique on short term vehicle traffic flow forecasting grey model performance.
2) To investigate the effect of univariate and multivariate formulation on accuracy of grey models on short term vehicle traffic flow forecasting.
3) To assess grouping technique based grey model on energy consumption and carbon dioxide emissions.

1.4 Research Questions

1) What is the effect of data grouping technique on short term vehicle traffic flow forecasting grey model performance?
2) What is the effect of univariate and multivariate formulation on accuracy of grey models on short term vehicle traffic flow forecasting?
3) What is the performance of the grouping technique based grey model on energy consumption and carbon dioxide emissions?

1.5 Scope and Limitations of the Study

1.5.1 Scope

a) Geographical Scope

The study is majorly on the road transport system and in this research the Nairobi CBD was considered as the case study region. Figure 1.1 shows the study site (Google, n.d.). The major road network covered include University Way, Kenyatta Avenue, Haile Sellassie Avenue, Moi Avenue and Uhuru Highway (see section 3.5.4). The training and test data were collected from the intersections of these road network to capture 5-minute pattern of traffic for the city (Sarraj, 2018). The traffic data collected included time series vehicle flow, pedestrian and motorcycle data passing a point of study. The time series data for pedestrian and motorcycle were used and considered as relative variables in the multivariate grey model. In addition, data was sourced from past research contexts as follows. Vehicle traffic flow and vehicular CO_2 emission data were sourced from the national highway route 11 of Tokushima city, Japan. These was sourced from a past research study which was conducted by the author of this thesis (see Table 3.3). Moreover, data from Özceylan, (2016) consists of CO_2 emission *(mt)* as the dependent or output variable and energy consumption *(mtoe)* and number of motor vehicles *(10^6)* as relative or input variables. These data were as tabulated in Table 4.80. The energy consumption data were obtained from two sources; firstly, Kenya's total electricity consumption, expressed in kilowatt-hours (kWh), for the period from the year 2000 to 2019. The data were retrieved from the Central Intelligence Agency (CIA) World Factbook (www.indexmundicom/g/g.aspx?c=

ke&v=81) on December 29, 2019 and tabulated as in Table 4.13. Secondly, data sourced from Cheng et al., (2020) consists of clean energy (10,000 tons of standard coal) as the output variable and economic scale, GDP (CNY 0.1Billion) and population size (10,000 people) as the input variables. These data are as tabulated in Table 4.82.

b) Theoretical Scope

This research is based on the grey system theory and it focuses on the basic grey model (Deng 1989). The basic univariate grey model considered in this study is the single variable first order grey model, GM(1,1). Additionally, the multivariate grey model, GM(1,*n*), is considered. Both the univariate and multivariate grey models are formulated based on the proposed data grouping techniques. The developed models are then assessed on short term vehicle traffic flow forecasting.

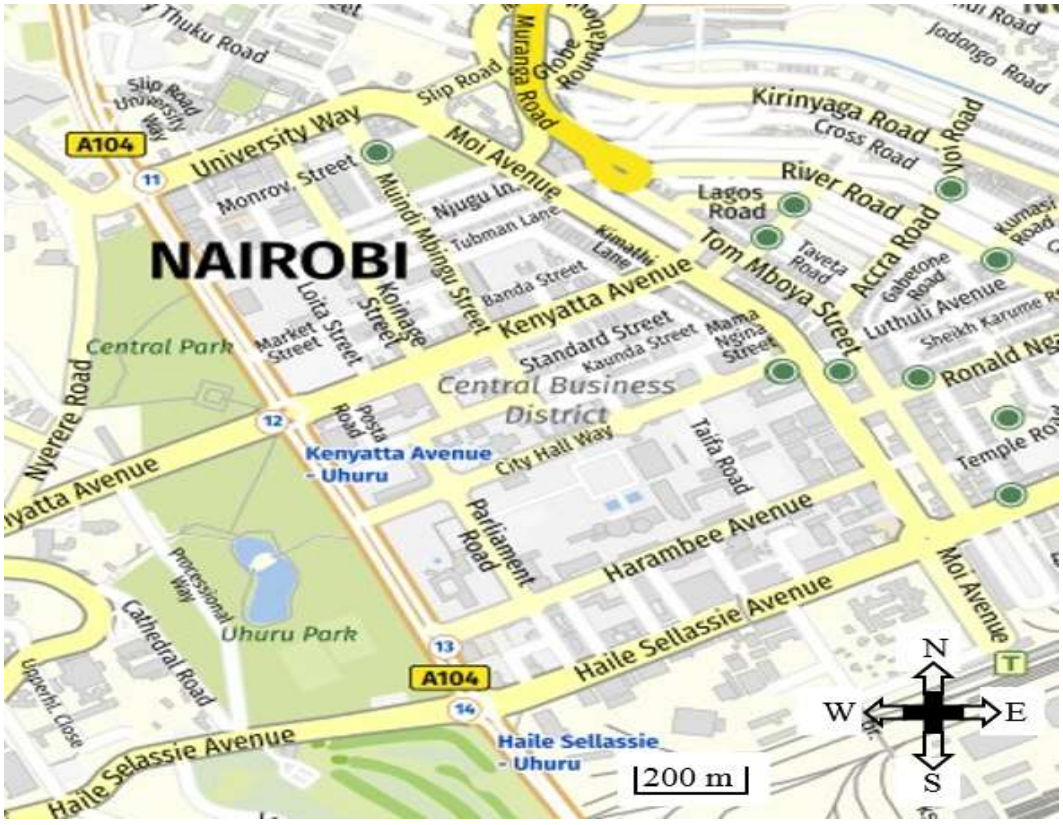

Figure 1.1: Study Site

Source: Google (n.d.)

1.5.2 Limitation

Time constraint was part of the limitations of this study as the author was not fully engaged in the research. This was because of the researcher working as a lecturer at the time of doing this research. However, this constraint was handled and managed through short vacations from the work place and the research was completed successfully at the end. On the same constraint of time the data collection clerks were students and they might have faced the challenges of being present all through to collect the traffic flow data in three consecutive days. However, this was an issue with a few of the clerks and the traffic flow data which were collected were enough representative of the Nairobi CBD traffic transportation network. The time constraint impacted negatively on the duration of the research as this research took five years to be accomplished instead of three years. For similar studies in the future any researcher need plan in time and get a study leave from the place of work and carry out data collection activity during holidays if students are the only available data collecting clerks.

1.6 Justification of the Study

Vehicular traffic is continuously increasing around the world. The resulting congestion and pollution is a major concern to transportation specialists and decision makers. As population continues to grow, it is a challenge to handle traffic demand, traffic jam, CO_2 emission, global warming and economic loss. Road capacity is not adequate. Roads and highways are unlikely to expand due to cost and dwindling land supply. Inadequate transportation system generates high costs and low customer service levels, which ultimately produces a negative economic impact (Zapata et al., 2013). Low mobility of vehicles, people and goods in urban areas is the greatest obstacle to economic growth and development. Therefore, there is need for efficient solutions to these problems (Khalil et al., 2010).

Traffic congestion is a severe problem in most cities and therefore it is time to shift from manual mode or fixed timer mode to an automated traffic system with decision making capabilities. Most of the present-day traffic light signaling systems are fixed time based. This renders inefficiency especially if one lane is operational than the

others. Sometimes higher traffic density at one side of a junction demands longer green time as compared to standard allotted time. To optimize such a problem an intelligent traffic control system is necessary (Patil et al., 2019). Further, Lu et al., (2020) asserts that to alleviate traffic congestion problems ITS has become the most popular and effective feasible solution.

Traffic jams can cause severe damages and impact negatively on the lifes of people. Therefore, performance and safety of a transportation system is important and can be improved through anticipation of future traffic flow events. Predicting jams before they happen and taking actions in advance can reduce the impact or duration of the forecasted congestion conditions. In order to control a traffic system in a proactive manner, ITS must have a predictive capability in short-term traffic flow forecasting (Liu et al., 2022). Short-term traffic flow forecasting is a fundamental function in ITS (Lu et al., 2020). Accurate prediction of vehicle flow in a road transport system is of critical importance for efficient traffic control and management (Qin & Zhang, 2022; Wang et al., 2020; Young & Liu, 2015). In addition, information can be disseminated to other parties and road users. For instance, driver information systems, navigation systems and vehicle positioning systems can aid in route selection and guidance to travelers in order to reach the desired destinations in time (Shah et al., 2022); Zhang, 2020). In Singapore, traffic planners receive data from global positioning system (GPS) devices and sensors embedded in the roadway. The information gets analyzed with algorithms to predict future traffic jam conditions and take pre-emptive actions to reduce the impact or duration of the conditions. Thus, a proactive, real-time traffic control system that anticipates future traffic conditions has a wide range of applications in vehicle navigation devices, vehicle routing, and congestion management. Indeed, there is need for the current research which is focused on improving the prediction accuracy of the conventional grey models for the performance of ITS.

Conventional methods of traffic management such as use of police officers at road junctions to direct vehicle movements is problematic in nature. A police officer can get tired or sick and it is not always that they are in control/alert (Hakkert & Gitelman, 2005). A policeman can be biased and cannot predict future traffic flow conditions. It is also difficult for a traffic police officer to monitor the whole scenario round the

clock. For instance, a police officer can only help when the weather permits. Police officers controlling traffic, at different intersections of a road, network cannot share, compare, take judgement and coordinate traffic conditions in real-time like an ITS can do. Generally, police resources are always limited (Hakkert & Gitelman, 2005). But traffic light signal automation can work for 24 hours without break and information can be disseminated to other road users unlike the police officer. Therefore, it is better to emulate the judgment of a traffic police officer on duty by ITS. ITS can resolve many traffic issues, such as traffic congestion, in order to provide an efficient traffic management scheme (Sayed et al., 2023).

However, continued use of the current ITS in traffic management systems is likely not going to be efficient with the rapid trend of urbanization. This is because effective traffic flow forecasting has been regarded as a key problem of ITS (Du et al., 2020; Shah et al., 2022). More accurate vehicle traffic flow prediction can improve the efficiency of ITS (Bharti et al., 2023; Jiang & Liu, 2023). Therefore, performance improvement of the predictive models which can find applications in enhancing the forecasting ability of ITS is very crucial in this error of rapid urbanization (Bharti et al., 2023). The enhancement of existing tools of traffic transportation management is necessary in order to cope with the problem of traffic congestion (Cantarella & Fiori, 2021). This thesis improves the precision of the GM(m,*n*) which has been applied, in the past, in short-term vehicle flow forecasting (Shen, 2022). The improved GM(m,*n*) can find application in enhancing ITS which play a vital role in traffic flow management.

1.7 Significance of the Study

In dealing with the problem of modelling and short-term forecasting of vehicle traffic flow for the purpose of proactive vehicle flow control, accuracy improved grey models (GM(m,*n*)) are proposed and developed in this thesis. Thus, the novelty and importance of this research study is as outlined below.

1) GM(1,*n*)'S Accuracy Improvement and Its Scope Extension

In this research the new idea is to improve the prediction accuracy of the conventional GM(1,*n*), by use of various proposed techniques. A data grouping technique (DGT) and a relative variable smoothing approach (RVSA) are two among the proposed methods to improve the accuracy of the conventional GM(1,*n*). The technique of grouping time series data is introduced into GM(1,*n*), and this improves the GM(1,*n*)'s accuracy. Thus a hybrid model is established and referred to as Grouped data GM(m,*n*) denoted as GGM(1,*n*). The RVSA is vital in smoothing the relative variables of a multivariate grey model for improving its accuracy. Therefore, the accuracy deficiency of the conventional GM(1,*n*) is overcome and the proposed hybrid grey models have improved fitting and short-term forecasting precision. Consequently, this has extended the application scope of the GM(1,*n*) in time series modelling and forecasting. Therefore, the improved GM(1,*n*) can be used to enhance the performance of ITS in short-term traffic flow forecasting.

2) Adherence to the New Information Prior Using Principle for Strong Adaptability

The conventional GM(m,*n*) has the shortcoming of modelling and predicting all data and ignoring new information. Thus, it cannot accurately reflect the characteristics of the current situation of a time series system. In this research this shortcoming is overcome by the proposed DGT which accommodates the concept of the "new information prior using" principle. The process of data grouping involves dropping of an old data point and adding of a new data point and this is in accordance to the new information prior using principle (see section 3.1.2a). A new data value can be collected from the environment being monitored and an older data value get deleted so that a newer model is established. Thus, a series of new predicted values will appear accordingly. Hence this guarantees a strong adaptability of the system under control.

3) Proactive Vehicle Traffic Flow Control

The proposed improved GM(1,*n*) can enhance the performance of ITS in proactive vehicle traffic flow control. By accurately forecasting traffic congestions and taking

actions in advance the impact or duration of the forecasted congestion conditions can be reduced. This way, it becomes possible to prevent a predicted disaster, if any, before it actually occurs and to impose controls in a timely manner.

4) Enhancing the Decision and Policy Making Processes

Proper incorporation and implementation of the proposed GGM(m,*n*) model in ITS can generate accurate predicted information about a traffic system which can enhance the decision and policy making processes to mitigate traffic congestion **(**Saki et al., 2020; Wang, 2017). Accurate forecasts can assist planners create road networks that can accommodate both present and future traffic demands, individuals plan their trips and traffic managers make knowledgeable judgements (Manikandan et al., 2023). Further, enhancing ITS by accurate predictive models is a way of improving the benefits provided by such smart transportation system. Some examples of ITS's application are the optimization of routes and street lighting.

CHAPTER TWO

LITERATURE REVIEW

In this chapter the efforts to evolve the road transport sector are outlined and examined. In addition, traffic flow forecasting methods are discussed and compared to point out the research gaps. The GST is explained together with the structure of the GM(m,*n*) and its optimization least square method. The meaning of grey forecasting is explained and the error causal factors in GM(1,1)'s structure are pointed out. Moreover, methods which have been used in the past to improve accuracy of the conventional GM(m,*n*) have been discussed. Finally, the research gaps have been identified and explained.

2.1 Evolving Road Transport Systems

In an effort to improve and evolve road transport systems a lot of literatures have been written (Chitere & Kibua, 2004; Daniel, 2016; Ngichabe, 2016). All these literatures were concerned to improve safety, determine causes of accidents and to report road accidents and as much as they were explored quantization of traffic flow was not taken into account. It is worthy to note that in order to improve road safety and reduce the rate at which accidents occur on our roads determination of traffic congestion is vital. Steg (2007) argues that behaviour changes of individual car users may help to achieve sustainable transportation. The researcher suggested that people may adopt more energy-efficient driving styles, change their car use, change the time of travel to avoid traffic jams, visit other destinations to reduce travel distance, suppress certain car trips, or travel with other modes of transport (such as public transport, cycling, walking or carpooling), replace their car by an energy efficient car or dispose of their car, and finally people may move residence, or look for another job location to reduce travel needs and distances. This is a psychological perspective aimed at providing sustainable transportation and it cannot be overlooked. However, combination of this perspective together with technological solutions can significantly reduce the negative impacts of road transport. Hence, in order to evolve the road transport sector, development of predictive models for smart transportation systems is equally vital. Nasim (2015) noted that the deployment of ITS is limited in the real-world application because of several challenges associated with its architectural design. This necessitated the study on how

to design a highly flexible and deployable architecture for ITS, which can utilize the recent technologies such as cloud computing and the publish/subscribe communication model. Therefore, Nasim (2015) proposed utilization of advances in Information and Communication Technology (ICT) in ITS to maximize the capacity of existing transportation systems without building new infrastructure. The use of an Infrastructure as a Service (IaaS) model to host large-scale ITS applications in the cloud, which reduces infrastructure cost, improves management flexibility and also ensures better resource utilization was put forth. This was a vital step towards improving the performance of an ITS but quantification of traffic congestion was not considered in the proposed ICT. So, it is important to also enhance the ITS from the short-term forecasting point of view as proposed in this thesis.

On the other hand, implementation and management of traffic control systems in the past has mainly been by programmable logic controller (PLC) and microcontroller-based systems (Liu & Chen, 2009; Udoakah & Okure, 2017). The PLC based methods offer a wide area of expansion which is essential for dynamic traffic control systems. In many countries urban traffic control is focused on mainly sequencing the traffic lights at a junction (Toroman & Mujcic, 2018; Udoakah & Okure, 2017). Many scholars have researched on the concept of unforeseen event handling by PLCs. For instance, using a PLC a traffic control system for emergence vehicle control was designed by (Amir et al., 2017). The transmission of the emergence signal was Radio Frequency Identification (RFID) based. PLC based traffic control systems have been designed and implemented (Muhammad, 2011; Udoakah & Okure, 2017). Udoakah and Okure (2017) designed a density-based traffic light control system with a microcontroller and infrared (IR) sensors. Of curiosity in this area of research is how extensively the PLCs have been utilized to control traffic lights and hence traffic flow in a proactive manner. The aforementioned literatures did not study on proactive traffic flow control. Controlling traffic flow in a proactive manner is a new area to be researched, especially in Kenya. In the current research, incorporation of the improved grey models into a traffic flow control system is recommended for future research. This way vehicle traffic flow can be controlled in a proactive manner.

2.2 Predicting/Forecasting Methods

Forecasting methods can be deterministic, probabilistic or stochastic methods and have been used in weather predictions (Barrera et al., 2005; Del et al., 2015).

Deterministic calculations are made with discrete values. A deterministic model is a mathematical model in which outcomes are precisely determined through known relationships among states and events, without random variation. In such models given a particular input, will always produce the same output, with the underlying machine always passing through the same sequence of states. A deterministic mathematical model is meant to yield a single solution describing the outcome of some "experiment" given appropriate inputs and all data is known beforehand. Deterministic algorithms are by far the most studied and familiar kind of algorithm, as well as one of the most practical, since they can be run on real machines efficiently. The deterministic method has been used in traffic flow prediction as in the following literature. Boyarshinov and Vavilin (2021) analyzed and determined the time dependences of the traffic flow intensity continuously throughout a week. An approach of using the smoothing procedure, i.e. the moving average method, was proposed to define deterministic and stochastic components of the traffic flow intensity. Statistical indicators of the distributions of the intensities and the random components isolated from them were determined. The estimates of the correspondence of the obtained curves to the normal law of probability distribution were also carried out. It was shown that on workdays and weekends, the deterministic components of the traffic flow intensity were similar in shape and location of maximum and minimum values. Therefore, the deterministic method outperformed the stochastic method in traffic flow predictions. From the work of Boyarshinov and Vavilin (2021) it is clear that the deterministic and stochastic methods were not trained and tested. Moreover, the methods were not subjected to short-term traffic flow forecasting. Therefore, even though the deterministic method's performance was good, it is doubtful if it can be effective in predicting traffic flows and in controlling the operation of traffic lights, as the researchers claimed. In this thesis the proposed methods are trained and tested for short-term traffic flow modelling and forecasting.

In probabilistic forecasting element of chance is involved. Forecasts incorporate the stochastic variability of the input variables, combining all the parameters according to their defined probability distributions to generate probabilistic cumulative distribution curves. A probabilistic model is meant to give a distribution of possible outcomes (i.e. it describes all outcomes and gives some measure of how likely each is to occur). Probabilistic methods use stochastic parameters such as a Monte Carlo simulation. Weather forecasting is a common example of probabilistic analysis. Generally, probabilistic forecasting has been applied in some areas as follows. Leverger et al. (2021) proposed forecasting (in a probabilistic way) the whole next season of a time series, rather than forecasting the given time series stepwise. The proposed framework in Leverger et al. (2021) was implemented combining several machine learning techniques. This framework was evaluated using a wide range of real seasonal time series. As demonstrated by their experiences, the proposed framework outperformed competing approaches by achieving lower forecasting errors. Even though their approach outperformed other methods, the proposed frame work has a limitation. The frame work is not applicable on anticipatory traffic flow control (Abdulhai et al., 2002). Forecasting of the whole next season of a time series is not applicable to some situations where stepwise forecasts are required. For instance, in ITS stepwise forecasting is very crucial for timely traffic flow control. In this thesis the proposed grey forecasting models can forecast a traffic flow system in a stepwise manner. Lu et al. (2020) presented an effective approach to forecasting short-term traffic flow based on multi-regime modelling and ensemble learning. Multiple regimes of traffic flow were identified using a probabilistic approach. Each regime characterized a pattern that described a homogeneous traffic condition during the study time period. The identified regimes were then used as the representative features for the forecast modelling. Regime identification was based on probabilistic modelling. And the hidden Markov model (HMM) being one of the most powerful algorithms in probabilistic, was thus used. The experimental results had shown that the identified regimes were able to well explain the different traffic phases, and played an important role in forecasting. Furthermore, the developed forecasting model outperformed four typical models in terms of root mean square error (RMSE) and mean absolute percentage error (MAPE) on three traffic flow measures. Even though the performance of the developed model

was good a shortcoming of the model can be identified. In identifying the separate regimes, it can be deduced that the author sub-divided the traffic flow into unrelated traffic phases-the continuity of the data set in time series is lost. Thus, the method deprives the homogeneity of the whole traffic data set. Supposing the author identified overlapping regimes, perhaps the performance of the developed model could have been much better. This thesis proposes a DGT which sub-divides the traffic data set into overlapping groups and superimposes the group forecast results to obtain final forecasted result. This method maintains the continuity and homogeneity of the traffic data set used.

The adjective “stochastic” implies the presence of a random variable; e.g. a stochastic process is one where the system incorporates an element of randomness as opposed to a deterministic system. A stochastic model is a tool for estimating probability distributions of potential outcomes by allowing for random variation in one or more inputs over time. The random variation is usually based on fluctuations observed in historical data for a selected period. A stochastic process is a random process evolving with time. More specifically, in probability theory, a stochastic process is a time sequence representing the evolution of some system represented by a variable whose change is subject to a random variation. Stochastic traffic flow modelling has been an area of interest by some researchers as is evident in the following two cases. Joubari et al. (2022) developed a stochastic mobility model for urban environments. The model was designed to reflect vehicular activities in urban environments based on vehicular information collected using vehicular communications. The behavior of vehicles along multi-lane roads and intersections was modelled as a stochastic process using queuing theory. Particularly, the queue system was analyzed as a continuous-time Markov chain (CTMC) and by calculating the steady-state probabilities, different performance measures were derived and analyzed under various scenarios. To validate the model, the forecasts were compared with a queue model and realistic traces. The results could show that the model was capable of reproducing the realistic behavior of traffic in urban roads. The obtained estimates were then used to design an actuated traffic light and a vehicle speed adaptor. From the simulation results, it was clear that using the proposed traffic forecasting model helps reduce vehicles idling and travel times. Jabari and Liu (2012) proposed a new stochastic model of traffic flow that addressed issues

of negative traffic densities and mean dynamics that are inconsistent with the original deterministic dynamics. They had shown that their construction implicitly ensures non-negativity of traffic densities and that the fluid limit of the stochastic model was consistent with cell transmission model (CTM) based deterministic dynamics. Although the expected results in Jabari and Liu (2012) and Joubari et al. (2022) were obtained, one can object the clumsy procedure of the proposed model in Jabari and Liu (2012). This thesis proposes the grey model in traffic flow modelling and forecasting. The grey model consists of a short and simple model structure.

2.3 Traffic Flow Forecasting Models

A variety of mathematical models have been explored in the past to predict traffic flow e.g. regression, neural network, historical average algorithms and time series analysis (Kumar, 2017). As mentioned by Kumar (2017) the use of autoregressive integrated moving average (ARIMA) or seasonal ARIMA (SARIMA) in traffic flow prediction requires huge flow data to develop a model and therefore cannot be used in cases where data is limited. For instance, on real time traffic flow forecasting the ARIMA and SARIMA models would require long time to collect huge data for training and eventually forecasting. This will delay traffic flow forecasts. That is to say instant traffic forecasts may not be possible. In addition, the ARIMA model procedure and structure are lengthy and complex. To overcome such limitations there is need for an alternative method which has a simple modelling structure and requires few data to predict future traffic flow on the roadways. In this research the proposed improved GM(m,*n*) has a simple straight forward modelling structure and requires at least four data points for traffic volume prediction. It implies that the GM(m,*n*) can forecast traffic flow at short intervals of time unlike the ARIMA and SARIMA models. Additionally, this will ensure traffic flow monitoring at every short interval of time and hence smooth traffic flow. Conventional statistical methods such as Kalman filtering, ARIMA and SARIMA have been applied to forecast short-term traffic flow based on past data (Kumar & Vanajakshi, 2015; Ma et al., 2017). These statistical techniques can achieve reasonable prediction accuracy, but may not capture the dynamics and nonlinearities existing in traffic flow. In addition, because of the lengthy and complexity of its structure, the ARIMA model is time consuming (Ma et al., 2017).

In addressing these issues, introduced in this research, is the concept of data grouping in grey modelling for predicting short-term traffic flow. The DGT can capture the dynamics and nonlinearities in time series traffic flow because it adheres to the 'new information prior using' principle. Yaping et al. (2015), compared the performance of three prediction models with real-life data in Beijing. The models included ARIMA, neutral network, and nonparametric regression. The nonparametric regression outperformed the other two models. Furthermore, the nonparametric regression model experiences superior accuracy and is portable. However, the K-nearest neighbor (KNN) (a typical nonparametric regression method) model requires more complex process to identify the ''neighbors'' (Rong et al., 2015). Zaki et al. (2020) discussed two different approaches for traffic congestion prediction. They used NeuroFuzzy approach with 11% prediction error and Hidden Markov Model (HMM) with a prediction error of 10%. This confirmed the suitability of HMM for traffic congestion prediction. This is because HMM is a stochastic method and traffic is stochastic in nature. HMM utilizes statistics such as mean speed and standard deviation to predict traffic conditions. However, it is rarely used in traffic prediction although traffic is a stochastic process (Zaki et al., 2016). Smith and Demetsky (1997) proposed four short-term traffic volume forecasting models for the freeway traffic flow forecasting problem at two sites on Northern Virginia's Capital Beltway. The models were based on Historical Average, ARIMA Time-Series, Back-Propagation Neural Network, and Nearest Neighbor Nonparametric Regression models. They tested the forecasting accuracy for each model and the result revealed that the Nonparametric Regression model significantly outperformed the other three models. Moreover, the Nonparametric Regression model was easy to implement, and proved to be portable, performing well at two distinct sites (Smith & Demetsky, 1997).

Due to the dynamics of traffic conditions and complexity of traffic networks, it is difficult to obtain satisfactory traffic flow prediction results with less computation cost (Yao et al., 2022). Thus Yao et al. (2022), proposed a novel deep learning traffic flow forecasting framework, termed as ensemble attention-based graph time convolutional networks (EAGTCN), for accurate traffic flow forecasting. Results of this work show that the forecasting accuracy of the EAGTCN is superior to existing models especially in the long-term predicting situation. Zhou et al. (2022), proposed a filter attention-

based spatiotemporal neural network (FASTNN) for traffic flow forecasting. The experimental results show that the FASTNN has better prediction performance than various baselines and variant models. Zhuang and Cao (2022) proposed combination of convolutional neural network and bidirectional long short-term memory (CNN-BILSTM) model for forecasting short-term traffic flow and the prediction precision of the model was found to be greater than that of the comparison model (support vector regression and gated recurring unit models). The aforementioned deep learning approaches have a limitation which is put forth by Farsi (2020) and Kashyap et al. (2022), who asserts that deep learning architectures have a limitation in that they require large amounts of historical data for training. According to Kashyap et al. (2022), large amounts of historical data can cause over-fitting of the model due to high fluctuations in traffic flow over a small-time interval. In regard to this disadvantage of the deep learning architectures a grey model is proposed, in this thesis, which can model a system with as low as 4 data points.

Nevertheless, the grey model has been utilized in traffic flow forecasting (Duan et al., 2017; Shen, 2022). Duan et al. (2017) proposed four new models of structural parameters and component parameters, inertia nonhomogeneous discrete gray models (referred to as INDGM), and analyzed the important properties of the model. This model examined the construction of the inertia nonhomogeneous discrete gray model from the mechanical properties of data (such as distance, acceleration, force combination, and decomposition), explaining the classic NDGM modelling mechanism. Finally, traffic-flow data was analyzed and the relationship between the inertia model and the traffic-flow state was studied. An optimal INDGM was selected, and better traffic-flow prediction results were achieved at a simulation accuracy and prediction accuracy of up to 0.0248 and 0.0273, respectively. Shen (2022) proposed a seasonal gray Fourier model based on the complex Simpson formula for short-term traffic flow forecasting. A seasonal GM(1, 1) model was used to optimize the background values first, and then the prediction results were adjusted using the Fourier series method. The new model was applied to the prediction of traffic flow and the numerical results indicated that the new model's performance was significantly better than those of the traditional GM(1, 1) model and the seasonal GM(1, 1) model. Although better traffic flow results were obtained by the proposed models of Duan et

al. (2017) and Shen (2022), close examination of the constructional structure and procedure of the proposed models is seen to be lengthy and complex. These features make the proposed models to be tedious so as to cause lengthy delays in the computational process. And thus, this thesis proposes an improved grey model which is not only simple in the constructional structure but also short in its computational form.

Now that the grey model is proposed in this thesis for traffic flow modelling and short-term forecasting it is important to outline and explain its constructional structure as well as the historical attempts to improve its precision by other researchers. Therefore, the next section highlights the grey system and its application in grey modelling.

2.4 Grey System Theory (GST) and Modelling

The GST was introduced in 1982 (Deng, 1989; Liu et al., 2015). In all human endeavors almost, all systems are grey systems (systems lacking information). The term 'grey' means poor, incomplete, uncertain (Deng, 1989; Javanmardi, & Liu, 2019). The theory finds the law governing a grey system through processing a raw data to establish its corresponding mathematical model (Li et al., 2018). The GST is interdisciplinary and stands the test of time (Deng, 1989). The system has been applied in various fields such as ecology, economy, environment, engineering transportation etc. to analyze, estimate, forecast and model the various systems considered to be grey systems (Deng, 1989; Li & Lin, 2014; Li et al., 2021; Liu, 2012; Lu M., 2015; Nguyen et al., 2020; Wang et al., 2018; Zeng et al., 2019).

The GST provides theory, techniques, notions and ideas for resolving (analyzing) grey systems, for example to build a differential model-so called grey model (GM)-by using at least 4 data to model a system (Javanmardi & Liu, 2019).

The system extracts a governing relationship of a system and covers various areas such as grey relational space, grey generating space, grey forecasting, grey decision making, grey control, grey mathematics and grey theory (Deng, 1989; Li & Lin, 2014; Lu M., 2015; Shen et al., 2016; Slavek et al., 2015).

The Grey Model (GM) uses a cumulative model in order to create differential equations and it requires a small number of data (Jiang et al., 2014). The GM is abbreviated as "GM(m,*n*)" where "m" symbolizes the degree of differential equations whereas "*n*" represents the number of variables under consideration. The basic original grey model is the single variable (univariate) first order differential equation prediction model denoted as GM(1,1) and many variants of this model such as the GM(2,1) and the GM(1,*n*) have been developed. The GM(2,1) and GM(1,*n*) are referred to as the second order and multivariate grey models respectively.

2.4.1 Grey Generating Techniques

The grey generating techniques are important in turning the disorderly raw data to a regular series for the benefit of grey modelling (Jiang et al., 2014). Suppose the raw data series is presented as (Caleb et al., 2022; Jiang et al., 2022):

$$X^{(0)} = \left\{ x_{(1)}^{(0)}, x_{(2)}^{(0)}, x_{(3)}^{(0)}, \cdots, x_{(m)}^{(0)} \right\} \tag{2.1}$$

where m is the total number of data points. Accumulating this series by (Jiang et al., 2022; Nguyen et al., 2020**):**

$$X_{(r)}^{(1)} \hat{=} \left\{ \sum_{i=1}^{r} x_{(i)}^{(0)} \right\}, \quad r = 1,2,\cdots,m \tag{2.2}$$

Results to:

$$X^{(1)} \hat{=} \left\{ x_{(1)}^{(1)}, x_{(2)}^{(1)}, \cdots, x_{(m)}^{(1)} \right\} \tag{2.3}$$

This generating operation can be denoted as:

$$AGOX^{(0)} = X^{(1)} \tag{2.4}$$

Which is the Accumulated Generating Operation (AGO) and is vital in grey modelling (Jiang et al., 2022). Thus the Inverse Accumulated Generating Operation (IAGO) is given as:

$$IAGOX^{(1)} = X^{(0)} \tag{2.5}$$

So that IAGO (inverse AGO) restores the original series.

Another grey generating concept is the Mean value Generating Operation (MGO) given as (Jiang et al., 2022):

$$Z_{(k)}^{(1)} = 0.5\left(X_{(k)}^{(1)} + X_{(k-1)}^{(1)}\right), \quad k = 1,2,\cdots,m \tag{2.6}$$

This is mean generation of adjoining neighbors of the AGO (Nguyen et al., 2020).

By AGO a non-negative, smooth, discrete function can be transformed into a series, extended according to an approximate exponential law which is called the grey exponential law and by which a reform to establish a suitable foundation in building a differential model is said to be completed.

2.4.2 Grey Modelling

In GST a dynamic model called grey differential model (GM) was developed. To do this Deng (1985) inferred that a stochastic process whose amplitudes vary with time is referred to as a grey process, the grey modelling is based on the generating series rather than on the raw one, the grey derivative and grey differential equation are defined and proposed in order to build a GM and to build a GM, only a few data (as few as 4) are needed to distinguish it.

As mentioned earlier in grey modelling GM(m,*n*) stands for an m-order differential equation grey model with *n* variables under consideration (Moonchai & Rakpuang, 2015). The basic original grey model is the single variable first order differential equation prediction model denoted as GM(1,1) and many variants of this model such as the GM(2,1) and the GM(1,*n*) have been developed (Zeng et al., 2018). The GM(2,1) and GM(1,*n*) are referred to as the second order and multivariate grey models respectively.

In this concept of grey modelling a single variable first order differential equation prediction model based on GST is given as (Caleb et al., 2022; Guo-Dong et al., 2006; Tien, 2009; Luo et al., 2013):

$$\frac{d}{dt}X^{1}(t)+aX^{1}(t)=b \tag{2.7}$$

where $X^{1}(t)$ is a background grey value at time t, a and b are the coefficients in GST terms, said to be the developing coefficient and the grey input respectively (Guo-Dong et al., 2006). These coefficients are obtained by the least square method (see section 2.4.3).

The time response equation of (2.7) is deduced as:

$$\hat{X}_{(r+1)}^{(1)} \hat{=} \left(X_{(1)}^{(0)} - \frac{b}{a} \right) e^{-ar} + \frac{b}{a}, \quad r = 0,1,2,\cdots,m-1 \tag{2.8}$$

where $\hat{X}_{(r+1)}^{(1)}$ is the predicted value of the data sequence (2.3) and from (2.1) and (2.4) the equation (Luo et al., 2013; Zhang et al., 2017):

$$X_{(k)}^{(0)} + a\, Z_{(k)}^{(1)} = b\,,\ k = 1,2,\cdots,m \tag{2.9}$$

is a grey differential model, called GM(1,l) as it includes only one variable, $X^{(0)}$, where:

$$Z_{(k)}^{(1)} = 0.5\left(X_{(k)}^{(1)} + X_{(k-1)}^{(1)} \right),\ \ k = 1,2,\cdots,m \tag{2.10}$$

and $x_{(k)}^{(0)}$ is a grey derivative which maximizes the information density for a given series to be modelled.

Note that $x_{(1)}^{(0)}$ of (2.8) is the initial condition of the model which causes prediction error (Jiang et al., 2014; Jong & Liu, 2014). Also, the constructional formula for (2.10) is among the factors which produce simulation errors (Jiang et al., 2014; Madhi & Mohamed, 2022). And $Z_{(k)}^{(1)}$ in (2.10) is the background value of GM(1,1).

To recover the original series the IAGO is applied (Jiang et al, 2022):

$$\hat{x}^{(0)}_{(r+1)} = \hat{x}^{(1)}_{(r+1)} - \hat{x}^{(1)}_{(r)}, \quad \hat{x}^{(0)}_{(1)} = \hat{x}^{(1)}_{(1)}\ ,\ r = 1,2,\cdots,m-1 \tag{2.11}$$

Note that in grey modelling according to Deng (1989) the "weight" equals to 0.5 and GM(m,*n*) stands for *m*-th order differential equation of *n* variables.

2.4.3 The Least Square Method

The optimized solution of the parameters a and b of (2.8) is derived by the least square method. This method gives a way to find the best estimate of these parameters. It assume**s** that the errors (i.e. the differences from true values) are random and unbiased. Given observations $(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)$ (N is the number of observations), the error E associated to saying:

$$y_n = a x_n + b \tag{2.12}$$

can be defined by:

$$E(a,b) = \sum_{n=1}^{N} \left(y_n - \left(a x_n + b\right)\right)^2 \tag{2.13}$$

The goal is to find values of a and b that minimize the error. This requires us to find the values of (a,b) such that:

$$\frac{\partial E(a,b)}{\partial a} = 0, \quad \frac{\partial E(a,b)}{\partial b} = 0 \tag{2.14}$$

Eventually, it can be shown that:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \left[A^T A\right]^{-1} A^T y \tag{2.15}$$

where A is the data matrix and y is the measured vector, i.e. $y = (y_1, y_N)^T$ (see references (Deng, 1989; Guo-Dong et al., 2006; Luo et al., 2013; Zou & Wu, 2012) for details).

2.4.4 Grey Forecasting

The concept of grey forecasting is by extrapolating the modelled series of (2.8) as:

$$\hat{X}^{(1)}_{(m+1)}, \hat{X}^{(1)}_{(m+2)}, \hat{X}^{(1)}_{(m+3)}, \tag{2.16}$$

Hence by short-term forecast vehicle volume in the near future is determined at predetermined time intervals. Suppose t is the instant at which data is collected, then to forecast two points into the future means estimating the traffic volume at $t+1$ and at $t+2$ which corresponds to $m+1$ and $m+2$ of grey forecasting, see (2.16).

2.5 Grey Model, GM(1,*n*), Accuracy Improvement

Since in this thesis an improved GM(m,*n*) is proposed it is important to review the literature behind GM(m,*n*))'s prediction accuracy improvement. Many scholars have researched on improving the accuracy of the conventional GM(1,1) by various approaches. Some have improved the accuracy by optimizing the initial condition, background value etc. For instance, aiming at the problem of determining original value of Non-homogeneous Grey Model NGM(1,1) Cai (2010) analyzed model-constructing mechanism of the non-homogeneous grey GM(1,1) model and unveiled the cause of the problem. And in order to minimize the quadratic sum of its fitting error, a new initial value determining approach was proposed and the calculating formula for determining initial value was deduced. The grey model proposed by Cai (2010) was then constructed based on optimal initial value determining method and had the characteristic of high precision as well as high adaptability. Mahdi and Mohamed (2017b), aimed to improve GM(1,1)'s prediction accuracy by improving the initial condition in the response function of the model. They optimized the initial condition by a method of minimizing the error summation of the square. The results had shown that the modified GM(1,1) model gives a better prediction performance compared with traditional GM(1,1) (Mahdi & Mohamed, 2017b). Again, Mahdi and Mohamed (2017a), in their study improved prediction accuracy of GM(1,1) model through an optimization of the background value. They reconstructed the background value to fit the accumulated sequence. The modified model performed better than the traditional grey model GM(1,1). Moreover, the modified GM(1,1) model achieved the

objective of minimizing the forecast errors and had high accurate forecasting power (Mahdi & Mohamed, 2017a). Moreover, Mahdi and Mohamed (2022), in their recent research improved GM(1,1)'s performance by combination of optimized initial and background value using the medium of the time series data. At this point it is critical to note that combination of optimized initial and background value is data discriminative as it assumes the rest of the data in the series, considering only the medium. This shortcoming of this approach is overcome in this research by the proposed DGT which groups data in a non-discriminatory manner and it is based on the new information prior using principle (Javanmardi & Liu, 2019). Khuman et al. (2013), proposed a novel way of improving the overall relative accuracy of the new-information grey model and the metabolic grey model and also improving the filter value accuracy, by incorporating a weight sequence that is populated by a genetic algorithm to minimize the error of the simulated values. The least square parameters $(-a)$ and b, were then scaled by the values contained in the weight sequence, until a satisfactory result was obtained. They proved that the metabolic model in this instance is better suited for forecasting and predicting (Khuman et al., 2013). Changjun et al. (2011), proposed an improved error GM (1,1) model and used this model to predict cultivated land in Yiyang. The results show that the improved error GM (1,1) model has high prediction accuracy and better simulation results. In addition, the improved grey error model is convenient and reliable (Changjun et al., 2011). In order to further improve the prediction accuracy of the multivariate grey prediction model, Lao, et al. (2021), established a novel multivariate grey prediction model based on dynamic background values (abbreviated as DBGM (1, *N*) model) and used the whale optimization algorithm to solve the optimal parameters of the model. The results show that the prediction and fitting accuracies of the DBGM (1, *N*) had been greatly improved compared with those of the GM (1, *N*) model. Luo, et al. (2013), improved the precision of the non-equidistant GM(1,1) by firstly, generating the reciprocal of the original time series data, secondly by adopting the reciprocal AGO process and finally, by optimzing the background values of the improved unequal interval GRM(1,1). The GRM(1,1) was found to have good performance. However, this method is tedious and complex in its structure form compared to the proposed improved GM(m,*n*) of this thesis. Additionally, Luo et al. (2013) did not modify the initial condtion which is

considered as a causal to prediction error. In this theisis the proposed model only data grouping is adopted and modification of the raw data is not part of the data grouping process. So that the proposed improved GM(m,*n*) is more adaptable to the raw data as opposed to other methods (such as the ones discussed above) which require data modification before grey modelling. Moreover, from the above literatue review it is evident that the methods used in improving GM(m,*n*)'s accuracy are lengthy, complex, data discriminatory, time consuming and involve data pre-processing which poses much modification to the raw data. Complexity of any system modelling method increases uncertinity in the performance of the sysstem (Javanmardi & Liu, 2019). To overcome these shortcomings a simple improved GM(m,*n*) is proposed in this thesis. The proposed model's procedure and construction structure is simple and hence less computational time. The DGT proposed in improving the GM(m,*n*)'s forecasting perfomance does not modify the raw data like in the case of Luo, et al. (2013), in which accumulated generating operation of the reciprocal number is used. Instead, the DGT groups the raw data directly without any data pre-processing procedure. It is worthy to note that data pre-processing, some times, has issues to do with multicollinearity. Moreover, this DGT adheres to the 'new information prior using" principle. This principle takes the advantage of new/current information of a time series in system modelling (Javanmardi & Liu, 2019). Note that none of the above outlined accuracy improving methods adheres to this important principle in improving the performnce of the conventional GM(m,*n*).

To overcome the deficiencies of the existing unbiased prediction grey model (abbreviated as UGM(1,1)), Zeng et al. (2020) proposed a new shale gas output grey prediction model and abbreviated it as SGGM(1,1,r). They modelled the SGGM(1,1,r) by combining new initial conditions and reprocessing of the original data using a grey average weakening buffer operator. Then the new model's accumulative order was optimized by fraction accumulation generation operation. Moreover, the new model was designed based on the "new information priority" principle and thus its time response function was derived by using the latest value as the initial value. Comparing the experimental results of the new model with that of the existing model, the former was significantly better than the latter. However, recognition of the latest value as the initial value discriminates the old values and this violates one of the six basic axioms

of the GST (Javanmardi & Liu, 2019). The Axiom states that “Difference” connotes the existence of information-each piece of information must carry some kind of “difference” which is known as principle of informational differences. Thus, all values in the original data sequence have information about the trend of the system and need to be considered in grey modelling.

Chen et al. (2020) proposed an improved GM (1,7) prediction model to predict the volume fraction of 7 dissolved gases in transformers. In this study the original data sequence of the traditional GM (1, 7) prediction model, was transformed so that the original data sequence after the transformation had better exponential properties, to meet the model's requirements for sequence smoothness. In order to greatly improve the prediction accuracy of the model, the construction method of background value optimization was also introduced. The average residual of the improved GM (1,7) model was found to be smaller compared to those of the traditional GM (1,1) and GM (1,7) models. Thus, the superiority of the improved model was proved. In this raw data transformation and background value modification approach a limitation is noticed. The initial condition in the improved GM (1,7) is not modified and yet it is one of the error causal factors in grey modelling (Jiang et al., 2014). To overcome this limitation this thesis proposed a DGT which keeps modifying the initial condition from one formed group to the other in accordance to the new information prior using principle.

In Li et al. (2023) the mechanism defects, parameter defects, and structural defects of the GM(1,N) model were compensated by adding linear correction term and grey action, and the OGM(1,N) model was established. Then, the background value coefficients of the OGM(1,N) model were optimized using the particle swarm optimization (PSO) algorithm, and the OBGM(1,N) model was established. Then, by introducing the fractional order idea, the PSO algorithm was used to optimize the cumulative order of the OBGM(1,N) and the order of the OBGM(1,N) model was extended from an integer field to a real number field to establish FOBGM(1,N) model. From the result analysis the prediction accuracy of the FOBGM model reached 99.996%. This verified the effectiveness of the model improvement by Li et al. (2023). Li et al. (2023) concluded this study by quoting that the model still has room for improvement, such as from the perspectives of initial value improvement, residual

correction, and metabolism, as a subject of further research. Therefore, this thesis extends and proposes improvement of the GM(1,N) from the initial value improvement and residual correction perspectives.

Ohene-Akoto et al. (2023) enhanced the prediction accuracy of the traditional grey system model using the PSO algorithm. The enhancement was mainly by the use of PSO to predict an optimum initial condition value based on the input dataset to improve the prediction accuracy of the original grey model that uses the first data of the input dataset as its initial condition. The optimum initial condition chosen becomes the initial condition value for the response function of the whitenization equation of the original grey system model. The accuracy of the proposed model was tested using a generated monotonic increasing and decreasing data. The proposed PSO optimized initial condition (PSOOIC) had the best performance as compared to the traditional grey system model and a model that also improved the initial condition. The results showed that the enhanced model using the PSO to choose an optimum initial condition is a better modification of the grey system model.

In dealing with the problem of prediction accuracy of stochastic volatility series Xiao et al. (2021) proposed a method to optimize the GM(1,1) from the perspective of residual error. In this study, a new fitting method which combines the wavelet function basis and the least square method to fit the residual data of the GM(1,1) was used. The residual prediction function was constructed by using the fitting method. Then, the prediction function of the GM(1,1) was modified by the residual prediction function. As a result a wavelet residual-corrected grey prediction model (WGM(1,1)) was established. The WGM(1,1) was compared with the GM(1,1) and a Fourier residual corrected grey prediction model (FGM(1,1)). From the experimental analysis, the WGM(1,1) had better simulation effect and higher prediction accuracy than the GM(1,1) and FGM(1,1).

2.6 Research Gaps

From the analysis of the relevant literature, it was found that scholars from different spheres have improved and optimized the GM(1,*n*) model from different perspectives in different research fields and obtained relatively better prediction accuracy compared

with the traditional GM(1,*n*) model. Some of the methods and approaches used to improve the existing GM(1,*n*) had deficiencies and some of the researchers (e.g. Li et al. (2023)) gave recommendations for further improvement of the GM(1,*n*)'s prediction accuracy. From the literature the following research gaps were identified.

2.6.1 GM(m,*n*)'s Accuracy Improvement

The proposed accuracy improvement models in Cai (2010), Mahdi and Mohamed (2017, 2022), Lao et al. (2021), Luo et al. (2013) and Xiao et al. (2021) are inconsistent with the GST. According to the principles of minimum information, new information priority and difference information, most of these existing grey forecast models and their improvement are inconsistent with the GST (Tu et al., 2023). Therefore, a novel grey model consistent with the GST is necessary. Thus, to improve the accuracy of the grey model, a DGT which is consistent with the GST is proposed in this thesis. The DGT obeys both the "difference information" and "new information priority" principles. Moreover, the DGT has a simple and straight forward procedure and model structure. Further, Li et al. (2023) recommended GM(1,*n*)'s accuracy improvement from three perspectives, namely initial value improvement, residual correction and metabolism. In respect to this recommendation this thesis improves GM(1,*n*)'s accuracy by employing the DGT and the FSECA. As mentioned earlier the DGT modifies the initial condition whereas the FSECA corrects the prediction error. Additionally, the RVSA is proposed for improving GM(1,*n*)'s prediction accuracy in vehicle flow prediction.

2.6.2 Short-Term Forecasting

These efforts (Changjun et al., 2011; Khuman et al., 2013; Lao et al., 2021; Li et al., 2023; Luo, et al., 2013; Mahdi & Mohamed, 2017a, 2017b, 2022; Ohene-Akoto et al. 2023; Twumasi et al., 2021; Xiao et al., 2021) to improve the precision of the original GM(1,1) had only concentrated on methodologies to improve the GM(1,*n*)'s prediction accuracy. The researchers did not apply the improved GM(1,*n*) in any short-term forecasting. Thus, this research not only introduces new approaches, which include data grouping and relative variable smoothing, for improving the precision of the original grey models but also extended the application of the improved models in

short-term forecasting of vehicle traffic flow. Moreover, this area of short-term traffic flow forecasting is receiving much attention nowadays (Shah et al., 2022).

2.6.3 Relative Factors in the Multivariate Grey Model

In many mathematical models (Changjun et al., 2011; Kumar, 2017; Kumar & Vanajakshi, 2015; Liu et al., 2014; Ma et al., 2017; Shen et al., 2016; Shen, 2022; Smith & Demetsky, 1997; Wu et al., 2013; Xu & Dang, 2015; Yao et al., 2022; Yaping et al., 2015; Zaki et al., 2016, 2020; Zhou et al., 2022; Zhuang & Cao, 2022) which have been used to predict traffic flow the researchers have not considered other factors such as pedestrian and motorcycle which affect traffic flow. In such studies the researchers have overlooked important factors that affect the performance of a traffic system. In other words, the researchers have employed univariate models to predict traffic flow. Variables that affect the system of interest are external and is important that they are determined and taken into account in system modelling (Caleb et al., 2022). The pedestrian mode as a relative variable is an important component of urban networks and greatly affects the performance of sidewalks and crosswalks, as well as the entire network traffic operations by interacting with other traffic modes (automobile, bicycle, transit) (Zheng et al., 2016). In this research these relative factors are considered in a multivariate grey model, GM(1,n), where n stands for the number of factors/variables affecting the system. In particular considered is the pedestrian and motorcycle factors in modelling and forecasting of vehicle traffic flow.

CHAPTER THREE

METHODOLOGY

In addressing city issues of traffic congestion an approach of improved grey model for short-term vehicle flow forecasting (estimating traffic flow conditions in the near future) is proposed in this research. Therefore, in this chapter, to facilitate the assessment of the effect of DGT on short term traffic flow forecasting grey model performance both the univariate (GM(1,1)) and multivariate (GM(1,*n*)) grey models are defined and their modelling structure provided. Then the least square method is discussed as the optimizing algorithm for determining the optimal parameters of the GM(m,*n*). Next formulation of the proposed accuracy improving methods are discussed. The chapter also presents methods and materials on how to analyze the performance of the DGT. Evaluation of accuracy improvement by error indicators is explained in addition to highlighting on how to validate the reliability and applicability of the improved grey models. Finally, research data sources, collection methods and procedures are discussed.

3.1 Assessment of the Effect of Data Grouping Technique on Short Term Vehicle Traffic Flow Forecasting Grey Model Performance

3.1.1 Formulation of the Basic Conventional Grey Model and Its Variants

The grey model as introduced by Deng (1989) is a widely employed prediction model and in this thesis modelling and forecasting of vehicle traffic flow was based on first-order *1*-variable and *n*-variable grey models denoted by GM(1,1) and GM(1,*n*) respectively.

I. The Basic Grey Model (GM(1,1)

The basic grey model i.e. the first-order one-variable (univariate) grey model (GM(1,1)), utilizes grey generating techniques to construct a differential equation. These generating techniques include the accumulated generating operation (AGO), inverse accumulated generating operation (IAGO) and mean value generating operation (MGO) which are summarize as follows (Deng, 1989; Xie et al., 2013).

a) AGO Formation

In grey modelling the grey generating techniques develops a systematic series from a raw data series. In this thesis the raw data series is presented as (Lu M., 2015; Wang et al., 2018):

$$X^{(0)} = \left\{ x_{(1)}^{(0)}, x_{(2)}^{(0)}, x_{(3)}^{(0)}, \cdots, x_{(m)}^{(0)} \right\} \tag{3.1}$$

where m is the total number of data points. Accumulation of the series in (3.1) by:

$$X_{(r)}^{(1)} \hat{=} \left\{ \sum_{i=1}^{r} x_{(i)}^{(0)} \right\}, \quad r = 1,2,\cdots,m \tag{3.2}$$

results to:

$$X^{(1)} \hat{=} \left\{ x_{(1)}^{(1)}, x_{(2)}^{(1)}, \cdots, x_{(m)}^{(1)} \right\} \tag{3.3}$$

which is the Accumulated Generating Operation (AGO) (Li et al., 2021). The AGO is a pre-processing technique which reduces the randomness of the actual data to improve data regularity and smoothness (Jiang et al., 2014). This generating operation can be denoted as:

$$AGO\{X^{(0)}\} = X^{(1)} \tag{3.4}$$

and $AGO\{\cdot\}$ denotes an operator.

b) MGO Formation

From adjacent AGO neighbors a Mean value Generating Operation (MGO) can be obtained as (Deng, 1989; Li & Lin, 2014; Lu M., 2015):

$$z_{(k)}^{(1)} = 0.5\left(x_{(k)}^{(1)} + x_{(k-1)}^{(1)} \right), \quad k = 1,2,\cdots,m \tag{3.5}$$

where $Z_{(k)}^{(1)}$ is the background value.

Thus, the mean generation of adjacent neighbors of the AGO yields the background value $Z_{(k)}^{(1)}$ and according to the existing result the "weight" equals to 0.5 (Deng, 1989).

c) The Grey Generated Model

The grey modelling algorithm has a unique characteristic of first order differential equation given as (Chang et al., 2013; Iqelan, 2017; Lu M., 2015):

$$\frac{d}{dt}X^{1}(t)+aX^{1}(t)=b \tag{3.6}$$

where $X^{1}(t)$ is a background grey value at time t, a and b are the developing coefficient and grey input respectively (Deng, 1989; Tien, 2009; Wang et al., 2016, 2018). The parameters a and b are optimized and estimated by the least squares method (see section 3.1.1 part IV) (Lu M., 2015). Eq. (3.6) is the 1st-order grey differential equation denoted by GM(1,1) (Tien, 2009) where the first 1 stands for the first-order derivative of AGO data sequence of (3.3) and the second 1 stands for only 1 time series data used in the grey differential equation.

Based on initial condition:

$$X^{1}(0)=x_{(1)}^{(0)}=x_{(1)}^{(1)} \tag{3.7}$$

the solution of (3.6) is deduced as:

$$X^{1}(t)=\left(x_{(1)}^{(0)}-\frac{b}{a}\right)e^{-at}+\frac{b}{a} \tag{3.8}$$

Therefore, the time response function of (3.6) is deduced as (Chang et al., 2013):

$$\hat{x}_{(r+1)}^{(1)}\hat{=}\left(x_{(1)}^{(0)}-\frac{b}{a}\right)e^{-ar}+\frac{b}{a},\qquad r=0,1,2,\cdots,m-1 \tag{3.9}$$

where $\hat{x}_{(r+1)}^{(1)}$ is the prediction of the AGO and $x_{(1)}^{(0)}$ of (3.9) is the initial condition of the model which causes prediction error (Jiang et al., 2014; Jong & Liu, 2014).

And from (3.1) and (3.3) the following equation can be obtained (Li & Lin, 2014):

$$x_{(k)}^{(0)} + a\, z_{(k)}^{(1)} = b\,,\ k = 1,2,\cdots,m, \tag{3.10}$$

This is a grey differential model, called 'Grey Model (first-order, single-variable)', GM(1,1). where $x_{(k)}^{(0)}$ is a grey derivative (Iqelan, 2017; Lu M., 2015; Slavek et al., 2015).

Now the data sequence of the predicted values arising from (3.9) is:

$$\hat{X}^{(1)} \hat{=} \left\{\hat{x}_{(1)}^{(1)}, \hat{x}_{(2)}^{(1)}, \cdots, \hat{x}_{(m)}^{(1)}\right\} \tag{3.11}$$

d) IAGO Formation

Now, based on (3.11), a systematical sequence of the original series is obtained by retrogressing through Inverse Accumulated Generating Operation (IAGO) given by (Chang et al., 2013; Shen et al., 2016):

$$\hat{x}_{(r+1)}^{(0)} = \hat{x}_{(r+1)}^{(1)} - \hat{x}_{(r)}^{(1)}, \qquad \hat{x}_{(1)}^{(0)} = \hat{x}_{(1)}^{(1)} \tag{3.12}$$

where $r = 1,2,\cdots,m-1$.

Then the simulative data sequence $\hat{X}^{(0)}$ of original data sequence $X^{(0)}$ of (3.1) is generated as:

$$\hat{X}^{(0)} \hat{=} \left\{\hat{x}_{(1)}^{(0)}, \hat{x}_{(2)}^{(0)}, \cdots, \hat{x}_{(m)}^{(0)}\right\} \tag{3.13}$$

where the data sequence $\hat{X}^{(0)}$ is the final predicted/fitted data sequence of $X^{(0)}$ of (3.1).

In other words, by an inverse accumulated generating operation (IAGO) it is shown that:

$$IAGO\{X^{(1)}\} = X^{(0)} \tag{3.14}$$

Notice that IAGO (inverse AGO) re-models the original series, and $IAGO\{\cdot\}$ means IAGO operator

e) Estimating Model Parameters

The ordinary least square method best estimates the parameter values a and b which are calculated as (Chang et al., 2013; Lu M., 2015):

$$\begin{bmatrix} a \\ b \end{bmatrix} = \left[A^T A\right]^{-1} A^T y, \tag{3.15}$$

where A is the data matrix and y is the measured vector and are given as:

$$A = \begin{bmatrix} -Z^{(1)}_{(2)} & 1 \\ -Z^{(1)}_{(3)} & 1 \\ \vdots & \vdots \\ -Z^{(1)}_{(m)} & 1 \end{bmatrix} \text{ and } y = \begin{bmatrix} X^{(0)}_{(2)} \\ X^{(0)}_{(3)} \\ \vdots \\ X^{(0)}_{(m)} \end{bmatrix} \tag{3.16}$$

II. The Second-Order Grey Model, (GM(2,1)

For the sake of understanding the grey model in detail the second-order grey model, GM(2,1), as a variant of the basic grey model is discussed here. Now, based on the basic grey model described above the conventional second-order one-variable grey model, (GM(2,1)), can be constructed.

The operators AGO and IAGO are used to generate the sequences $X^{(1)}$ and $X^{(-1)}$ from the raw data sequence $X^{(0)}$. The sequences $X^{(0)}$ and $X^{(1)}$ are as in (3.1) and (3.3) respectively. Now, the sequence $X^{(-1)}$ is generated by IAGO as:

$$X^{(-1)} \hat{=} \left\{X^{(-1)}_{(1)}, X^{(-1)}_{(2)}, \cdots, X^{(-1)}_{(m)}\right\} \tag{3.17}$$

where:

$$X^{(-1)}_{(r)} \hat{=} \left\{X^{(0)}_{(r)} - X^{(0)}_{(r-1)}\right\}, \; r = 1,2,\cdots,m \tag{3.18}$$

Besides, the background sequence (i.e. the sequence of the mean of adjacent AGO series data) is constructed:

$$Z^{(1)} \hat{=} \left\{ z_{(2)}^{(1)}, z_{(3)}^{(1)}, \cdots, z_{(m)}^{(1)} \right\} \tag{3.19}$$

where the mean generation of adjoining AGO neighbors is:

$$z_{(r)}^{(1)} = \frac{1}{2}\left(x_{(r)}^{(1)} + x_{(r-1)}^{(1)} \right),\ r = 2,3,\cdots,m \tag{3.20}$$

This is the mean value generating operation (MGO) and according to the existing result (Deng, 1989), the "weight" equals to 0.5, as mentioned earlier.

Then the grey differential equation model is constructed:

$$\alpha_1 x_{(r)}^{(-1)} + x_{(r)}^{(0)} + \alpha_2 z_{(r)}^{(1)} = b \tag{3.21}$$

where α_1 and α_2 are the low order parameter and high order parameter, respectively, b denotes the control variable and $Z^{(1)}$ is the background sequence as mentioned earlier.

These parameters are best estimated by the least square error method. From which the vector parameter p is given by (Xie et al., 2013):

$$p = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ b \end{bmatrix} = \left[A^T A \right]^{-1} A^T y \tag{3.22}$$

where A is the data matrix and y is the measured vector (Deng, 1989; Xie et al., 2013) obtained as:

$$A=\begin{bmatrix} -X_{(2)}^{(-1)} & -Z_{(2)}^{(1)} & 1 \\ -X_{(3)}^{(-1)} & -Z_{(3)}^{(1)} & 1 \\ \vdots & \vdots & \vdots \\ -X_{(m)}^{(-1)} & -Z_{(m)}^{(1)} & 1 \end{bmatrix} \text{ and } \quad y=\begin{bmatrix} X_{(2)}^{(0)} \\ X_{(3)}^{(0)} \\ \vdots \\ X_{(m)}^{(0)} \end{bmatrix} \tag{3.23}$$

and m is the total number of time series data.

Next, the grey reflection equation (or grey whitenization equation) is constructed and solved to obtain a time response equation. Finally, the data sequence $X^{(0)}$ of (3.1) can be forecasted.

Additionally, in this thesis a multivariable grey model, discussed in the next section, is considered.

III. The Multivariable Grey Model, GM(1,*n*)

The multivariable grey model is another variant of the basic grey model denoted as GM(1,*n*), where 1 stands for first order and *n* stands for the number of time series used in the grey differential equation (Li & Wu, 2021). The constructional structure of the conventional multivariable grey model is as follows (Cheng et al., 2020; Moonchai & Rakpuang).

The dependent variable time series sequence is presented by $x_1^{(0)}(k)$ and the relative variables' time series sequences are presented by $x_i^{(0)}(k)$ $(k = 1,2,3,4,\cdots,m\ ,\ i = 2,3,4,\cdots,n)$, where m is the total number of time series data and n is the number of time series.

The conventional GM(1,*n*)'s first order accumulated generating operation (1-AGO) is (Li & Wu, 2021):

$$x_i^{(1)}(k)=\sum_{j=1}^{(k)} x_i^{(0)}(j)\ k = 1,2,3,4,\cdots,m \text{ and } i = 1,2,3,4,\cdots,n \tag{3.24}$$

and its differential equation is given by:

$$x_1^{(0)}(k) + az_1^{(1)}(k) = b_2x_2^{(1)} + b_3x_3^{(1)}(k) + \cdots + b_nx_n^{(1)}(k) \quad (3.25)$$

where a, b_2 to b_n are parameters determined by the least square method and $z_1^{(1)}$ is the background value obtained as:

$$z_i^{(1)}(k) = 0.5\left(x_i^{(1)}(k) + x_i^{(1)}(k-1)\right), k = 2,3,4,\cdots,m, i = 1,2,3,4,\cdots,n \quad (3.26)$$

The whitening equation is given as:

$$\frac{d}{dt}x_1^{(1)}(t) + ax_1^{(1)}(k) = b_2x_2^{(1)}(k) + b_3x_3^{(1)}(k) + \cdots + b_nX_n^{(1)}(k) \quad (3.27)$$

and its solution is:

$$x_1^{(1)}(k) = e^{-at}\left(x_1^{(1)}(1) + t\sum_{i=2}^{n} b_i\, x_i^{(1)}(1) + \sum_{i=2}^{n}\int b_i\, x_i^{(1)}(t)e^{at}dt\right) \quad (3.28)$$

When the time sequence $x_i^{(1)}(k)$, $(i = 2,3,4,\cdots,n)$ has few changes in the time interval $[t_k, t_{k+1}]$, the time response equation of GM(1,*n*)'s whitening equation approximates to:

$$\hat{x}_1^{(1)}(k+1) = e^{-at}\left(x_1^{(1)}(1) - \frac{1}{a}\sum_{i=2}^{n} b_i x_i^{(1)}(k+1)\right) + \frac{1}{a}\sum_{i=2}^{n} b_i x_i^{(1)}(k+1) \quad (3.29)$$

where a vector parameter $\hat{B}$, data matrix X and measured vector Y are given by:

$$\hat{B} = \begin{bmatrix} a \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = (X'X)^{-1}X'Y \quad (3.30)$$

$$X = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) & \ldots & x_n^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) & \ldots & x_n^{(1)}(3) \\ \vdots & \vdots & \ldots & \vdots \\ -z_1^{(1)}(m) & x_2^{(1)}(m) & \ldots & x_n^{(1)}(m) \end{bmatrix}, \quad Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(m) \end{bmatrix} \quad (3.31)$$

where m is the total number of time series data and n is the number of time series, as mentioned earlier.

Then the forecasted value of the original sequence is:

$$\hat{x}_1^{(0)}(k) = \hat{x}_1^{(1)}(k) - \hat{x}_1^{(1)}(k-1)\ ,\ k = 2,3,4,\cdots,m,\ \hat{x}_1^{(0)}(1) = \hat{x}_1^{(1)}(1) \qquad (3.32)$$

If time series sequence $x_i^{(1)}(k)$, $i = 1,2,3,4,\cdots,n$, has great changes in time interval $[t_{k,}t_{k+1}]$, the time response function of GM(1,1) modelled by the series is substituted into $x_i^{(1)}(t)$ $(i = 2,3,4,\cdots,n)$ of $x_1^{(1)}(t) = e^{-at}\left(x_1^{(1)}(1) + t\sum_{i=2}^{n} b_i\, x_i^{(1)}(1) + \sum_{i=2}^{n}\int b_i\, x_i^{(1)}(t)e^{at}dt\right)$ to get the analytical expression for time response equation of GM(1,*n*).

At this point it is worthy to make clear the deficiency of the GM(1,*n*). According to (Wei & Dang, 2016) the trend of the relative variables is a causal to the low accuracy of the GM(1,*n*). Moreover, in the study of Li et al. (2023), it was concluded that GM(1,*n*) still has room for improvement, such as from the perspectives of initial value improvement, residual correction, and metabolism, as a subject of further research. Now in this research the adopted DGT modifies the initial condition whereas the RVSA smooths the relative variables. Additionally, the FSECA performs residual correction.

IV. Method of Least Squares

The parameters $(-a)$ and b of the grey differential equation model are referred to as the *development coefficient* and the *grey action quantity* respectively and are determined by the method of least squares. The parameter $(-a)$ describes the development states of $X^{(1)}$ and $X^{(0)}$. It represents the variance and dynamic nature of the system. In other words, it reflects the variation in the data. A small variance indicates that the data points tend to be very close to the mean and to each other. Parameter b is referred to as the grey action quantity value, and is obtained from the background values (behavioral sequence).

The grey model parameter values a and b are such that the best fit result minimizes the sum of squared errors or residuals which are the differences between the observed or experimental value and corresponding fitted value given in the model. Thus, the

ordinary least square method best estimates these parameters which are calculated as (Caleb et al., 2022; Chang et al., 2013; Lu M., 2015; Nguyen et al., 2020):

$$\begin{bmatrix} a \\ b \end{bmatrix} = \left[A^T A\right]^{-1} A^T y \text{, which is applicable in GM(1,1)}$$

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ b \end{bmatrix} = \left[A^T A\right]^{-1} A^T y \text{, which is applicable in GM(2,1)}$$

$$\begin{bmatrix} a \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = (X'X)^{-1}X'Y \text{, which is applicable in GM(1,}n\text{)} \tag{3.33}$$

where a and b are as defined in (3.6), α_1 and α_2 are as defined in (3.21), A is the data matrix and y is the measured vector (Deng, 1989; Guo-Dong et al., 2006; Luo et al., 2013; Zou & Wu, 2012), X and Y are the data matrix and measured vector for a multivariable grey model as defined in (3.31). Whereas a, b_2 to b_n are the multivariate grey model parameters as defined in (3.25).

And as long as the matrix A has a full rank and the inverse of $A^T A$ exists, then it is possible to compute "good" values of the parameters a and b that minimizes the sum of squared errors.

3.1.2 Formulation of the Grey Models Based on Accuracy Improving Methods

In this thesis six methods for improving GM's accuracy in traffic flow fitting and short-term forecasting are proposed. To improve the conventional grey models' precision this thesis introduces three new methods, namely a data grouping technique (DGT), a relative variable smoothing approach (RVSA) and a three-step approach (TSA). The DGT has the consequence of a lot of calculations but with the aid of MATLAB code the computation is done easily and faster. The RVSA is used as a data pre-processing tool in improving the precision of the multivariate grey model. TSA method was proposed to improve the accuracy of the conventional multivariate grey model

(GM(1,*n*)). The other three methods employed in this thesis were existing approaches, namely modification of background value (MBV), modification of initial condition (MIC) and Fourier series error correction approach (FSECA). These methods were used to improve prediction accuracy of the conventional grey models through data grouping. The TSA entails combining three approaches i.e. combining the DGT, RVSA and FSECA. Moreover, these methods are compared in improving the accuracy of the conventional grey models. The methods are explained and formulated in the following sub-sections of this thesis.

a) Data Grouping Technique (DGT)

The concept of DGT can be explained from a curve point of view. When a curve is divided into several small portions, the portions assume a straight-line shape. In other words the trends in a single portion tend to relate linearly with each other. This is the concept of "smoothing" and it is the "hidden inherent characteristic" in this concept of DGT. So if a set of time series data is grouped into overlapping sub-groups (small portions) and forecasted separately and the results of forecasts superimposed to obtain overall result of prediction, it is definite that the prediction accuracy can be improved. This can be referred to as "Group Smoothing" and can be used all across data analysis. The general idea is to group data points that are expected to have similar characteristics and fit a simple model. It is almost similar to bin smoothing.

Data grouping has been employed in other time series prediction methods. But in this research the grouping is very unique in the sense that it is based on the overlapping concept (Farsi, 2020) whereas in other cases grouping is based on data type, data intervals, data similarity etc. In the process of data grouping the groups overlap and depending on the intensity of overlapping the strong grouping and weak grouping techniques are formulated. In addition, the DGT is based on the "new information prior using" principle (see part *III* below). The 'new information prior using' principle prioritizes new information in a given data series and is useful in reducing the inherent prediction errors in the conventional GM (Mahdi and Mohamed, 2017b).

According to the GST the minimum number of data points to estimate and model a system is four (4). Therefore, in this research the minimum number of group data

points is four. Consequently, data grouping in 5s, 6s, 7s, and so on is considerable. Moreover, the more the groups formed the more the accuracy improvement as explained in section 4.1.3 of this thesis.

Three categories of data grouping techniques, namely no grouping or one group (NG or OG), weak grouping (WG), and strong grouping (SG) are considered in this thesis. The no/one group technique correspond to the original GM (m,*n*) whereas the weak grouping and strong grouping techniques correspond to the proposed Grouped GM (m,*n*) (GGM (m,*n*)). These data grouping techniques are explained below and their performance analysis is done in section 4.1.3. Moreover, part IV outlines a mathematical form of the DGT in 4s.

(I) No Grouping (NG) Method

In this method whole data sequence to be predicted forms a single group. For example, for ten elements in a data sequence $m = 10$, where m is the total number of elements. This is illustrated in Figure 3.1. The original GM(1,1) is based on this grouping method in which case there is no data overlapping. This method is also referred to as the "one group" or "non-grouping" (abbreviated as NG) method.

Figure 3.1: No Grouping (NG)

(II) Weak Grouping (WG) Technique

Considering a data set of 10 sample points 3 groups of 4s are formed (see Figure 3.2). In Figure 3.2, shown is the first group containing elements 1 to 4 and the third group consists of elements 7 to 10. The second group contains elements 4 to 7. Note that the three groups only overlap at data points 4 and 7. If the data is grouped in 5s, two groups are obtained in which case the first group is that of elements 1 through 5, and the second group has elements 5 through 9. Thus, one data is left out and this is the

disadvantage of this technique. If the data is grouped in 6s, one group of elements 1 through 6 is formed but the second group is not full. Therefore, based on this grouping technique, the formed groups do not overlap so much as compared to the strong grouping technique and thus the simulation errors do not cancel out. Consequently, the prediction accuracy is low, as it is evident from section 4.1.3d. In this thesis this technique is referred to as weak grouping. Notice that in this case a maximum of three groups are formed if no data is left out.

Figure 3.2: Weak Grouping (WG)

(III) Strong Grouping (SG) Technique

This is the most accurate DGT as the groups overlap so much and hence the errors tend to reduce to a minimum. For any particular data set the number of groups is given by:

$$N = m - [k - 1] \tag{3.34}$$

where N is the number of groups, m is the total number of data used, and k is the number of group data points. This technique is illustrated in Figure 3.3. With 10 sample points, 7 groups of 4s are formed. In addition to the three groups of the WG technique, four more groups are formed.

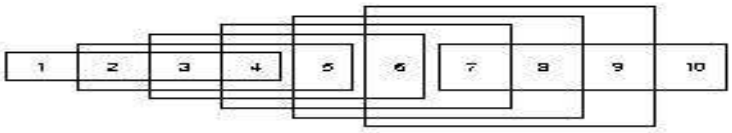

Figure 3.3: Strong Grouping (SG)

The first group includes elements 1 to 4 and the last group is composed of elements 7 to 10. Note that at data points 2 and 9 the overlap is once, at data points 3 and 8 the overlap is twice and at data points 4, 5, 6 and 7 the overlap is thrice. Note that a lot of overlaps are involved in this DGT and consequently many groups are formed. Now this is referred to as strong grouping (SG). With many groups and overlaps the simulation errors tend to reduce to a minimum, making this technique to be more accurate compared with the weak grouping technique. See section 4.1.3d. If now the data is grouped in 5s, six groups are obtained in which case the first group is that of elements 1 through 5, and the sixth group has elements 6 through 10. Here no data is left out and this is the advantage of this technique. If grouping is in 6s, five groups are formed. Thus, the more the data points per group the less the number of groups formed. The SG technique has an advantage of being based on the "new information prior using" principle, which is discussed in section 3.1.2b.

As a matter of illustration, the traffic flow data of Table 3.1 is employed to show the concept of SG in 4s and this is the data grouping method adopted throughout this thesis in improving the precision of the original grey model. The Table 3.1 data are from 6:00 am to 8:00 am of the day which were recorded at every data point (DP). The vehicle flow data was from the national highway route 11 of Tokushima city, Japan.

For example, the first 22 data points given in Table 3.1 are grouped into 19 groups of 4s based on (3.34) and Figure. 3.3. These groups are formed and tabulated as in Table 3.2 where DP stands for data point and AD stands for actual data. Notice that the process of data grouping involves dropping of an old data point from a group and adding of a new data point to that group until all data set elements are grouped.

Table 3.1: Traffic Flow Data

Data Point	Time of Day	Traffic Volume
1	6:00	0
2	6:05	14
3	6:10	35
4	6:15	54
5	6:20	55
6	6:25	95
7	6:30	83
8	6:35	89
9	6:40	98
10	6:45	134
11	6:50	103
12	6:55	173
13	7:00	110
14	7:05	167
15	7:10	160
16	7:15	150
17	7:20	210
18	7:25	200
19	7:30	172
20	7:35	149
21	7:40	154
22	7:45	140
23	7:50	157
24	7:55	146
25	8:00	145

Note: Data by author; A-1-8 Traffic Congestion and CO_2 Emission Analysis via MITRAM in Tokushima City. *Proceedings of the IEICE Engineering Sciences Society/NOLTA Society* Conference. January 3rd, 2016.

(IV) Data Grouping Technique in a Mathematical Form

The process of data grouping can be easily presented in a general mathematical form. For instance, the process of strongly grouping data in 4s can be summarized as follows. Any time series data can be grouped to form $m-3$ groups. The first group is presented as:

$$x_i^{(0)}(k), k = 1, \cdots, 4, \quad i = 1,2, \cdots n \tag{3.35}$$

where n is number of time series under consideration.

And in accordance with the “new information prior using” principle (Mahdi & Mohamed, 2017b) the second group is presented as:

Table 3.2: Actual Data (AD) Grouping by SG in 4s

DP	AD	GROUPS 1- 19																	
		G1	G 2	G 3	G 4	G 5	G6	G7	G8	G9	G 10	G 11	G 12	G 13	G 14	G 15	G 16	G 17	G18	G19
1	0	0																		
2	14	14	14																	
3	35	35	35	35																
4	54	54	54	54	54															
5	55		55	55	55	55														
6	95			95	95	95	95													
7	83				83	83	83	83												
8	89					89	89	89	89											
9	98						98	98	98	98										
10	134							134	134	134	134									
11	103								103	103	103	103								
12	173									173	173	173	173							
13	110										110	110	110	110						
14	167											167	167	167	167					
15	160												160	160	160	160				
16	150													150	150	150	150			
17	210														210	210	210	210		
18	200															200	200	200	200	
19	172																172	172	172	172
20	149																	149	149	149
21	154																		154	154
22	140																			140

$$x_i^{(0)}(k), k = 2, \cdots, 5, \quad i = 1,2, \cdots n \tag{3.36}$$

This grouping process continues until the last group is presented as:

$$x_i^{(0)}(k), k = m - 3, \cdots, m, \quad i = 1,2, \cdots n \tag{3.37}$$

where m is the total number of time series data.

Therefore, the SG in 4s (i.e. four data points per group) keeps dropping an old data and adding a new data in the grouping process. This is the point at which the SG technique has the advantage of adhering to the 'new information prior using' principle. Thus the procedure of grouped grey modelling entails repeated application of the OGM(1,1) on each group of data and averaging the predicted values at points of overlap. In other words, the final combined prediction is obtained by superimposing those group predictions. Thus, grouped grey modelling inherently modifies the initial condition (MIC) of (3.9) and this promotes accuracy improvement.

In this thesis the original GM(1,1) is denoted as OGM(1,1). Introduction of the DGT into the OGM(1,1) results to a hybrid grey model referred to as Grouped GM(1,1) and denoted as GGM(1,1).

b) New Information Prior Using Principle

In (Mahdi & Mohamed, 2017b) both initial and final conditions of GM(1,1)'s time response function were improved in order to express this principle. In the proposed SG technique, the new information priority principle is well and repeatedly demonstrated in each formed group. For example, the initial condition of the first group is data point 1, that of the second group is data point 2, that of the third group is data point 3, and so on. This process of dropping an old data point and adding a new data point keeps modifying the initial condition. Thus, the SG technique takes advantage of the new pieces of information in the original data (Mahdi & Mohamed, 2017b; Jiang et al., 2022).

c) Modification of the Background Value (MBV)

The computation of the parameters a and b of (3.9) depends on the original data sequence of (3.1) and the background value of (3.20). In GST, the value of the background value coefficient has a significant impact on the performance of the grey forecasting models (Zeng & Li, 2018). However, for most existing grey models, the background value coefficient is fixed at 0.5 to simplify the modelling process (Li et al., 2021; Zeng & Li, 2018). The constructional formula for (3.20) is among the factors which produce simulation errors (Jiang et al., 2014). Thus remodelling (3.20) will improve the precision of the model and in this thesis, remodelling is by integrating both sides of (3.6) from $k-1$ to k to obtain (Mahdi & Mohamed, 2017a):

$$\int_{k-1}^{k} \frac{dx_{(t)}^{(1)}}{dt} dt + a \int_{k-1}^{k} x_{(t)}^{(1)}\, dt = \int_{k-1}^{k} b dt$$

$$x_{(k)}^{(1)} - x_{(k-1)}^{(1)} + a \int_{k-1}^{k} x_{(t)}^{(1)} dt = b\, x_{(k)}^{(0)} + a \int_{k-1}^{k} x_{(t)}^{(1)}\, dt = b \tag{3.38}$$

Compare (3.10) and (3.38) and notice that the parameters a and b are now estimated by using:

$$Z_{(k)}^{(1)} = \int_{k-1}^{k} x_{(t)}^{(1)} dt \tag{3.39}$$

as the background value. Thus a and b are now more adaptive to the whitenization equation (Mahdi & Mohamed, 2017a).

Use of (3.39) with a few steps of integration as in (Liu et al., 2014) yields:

$$\boldsymbol{Z}_{(k)}^{(1)} = x_{(k)}^{(1)} + \frac{x_{(k)}^{(0)}}{\ln x_{(k)}^{(0)} - \ln x_{(k-1)}^{(0)}} - \frac{[x_{(k)}^{(0)}]\text{\^{}}2}{x_{(k)}^{(0)} - x_{(k-1)}^{(0)}}, k = 2,3,\cdots.n \tag{3.40}$$

Then this MBV and the DGT method are combined to improve the OGM(1,1)'s accuracy in fitting and short-term forecasting. Thus a modified background value GM(1,1) named as MBVGM(1,1) is established. Further, combination of the MBV and the DGT in improving the OGM(1,1)'s accuracy establishes a Modified Background Value Grouped GM(1,1), MBVGGM(1,1).

d) Modification of the Initial Condition (MIC)

The initial condition of (3.9) is a cause to precision error of the OGM(1,1) (Jong & Liu, 2014). Thus, modification of this initial condition to enhance the model's accuracy is important and as in (Madhi & Mohamed, 2017c) MIC is accomplished as follows. From (3.12) the restored value of the original sequence can as well be given by:

$$\hat{\boldsymbol{X}}_{(r)}^{(0)} = \hat{\boldsymbol{X}}_{(r)}^{(1)} - \hat{\boldsymbol{X}}_{(r-1)}^{(1)}, r = 2,3\cdots, m \tag{3.41}$$

Substituting (3.9) in (3.41) results to:

$$\hat{x}_{(r)}^{(0)} = (1 - e^{a})\left(x_{(1)}^{(0)} - \frac{b}{a}\right) * e^{-ar} * e^{a} \tag{3.42}$$

Let C, that is the initial condition of the traditional GM(1,1), be expressed as:

$$C = \left(x_{(1)}^{(0)} - \frac{b}{a}\right) * e^{a} \tag{3.43}$$

So that (3.42) becomes:

$$\hat{x}_{(r)}^{(0)} = C * (1 - e^{a}) * e^{-ar} \tag{3.44}$$

Therefore, the discrete form of (3.9) results to:

$$\hat{x}_{(r)}^{(1)} \triangleq C * e^{-ar} + \frac{b}{a}, \tag{3.45}$$

By applying IAGO on $\hat{x}_{(r)}^{(1)}$, the restored (predicted) value of the raw data is given as follows:

$$\hat{x}_{(r)}^{(0)} = \hat{x}_{(r)}^{(1)} - \hat{x}_{(r-1)}^{(1)}$$

$$= C * (1 - e^{a}) * e^{-ar}$$

$$= C * (e^{-ar} - e^{a} * e^{-ar})$$

$$= C * \left(e^{-ar} - e^{-a(r-1)}\right),\ r = 2,3,..m \qquad (3.46)$$

Optimization of the Initial Condition

Minimizing the sum of the squared error from the predicted values can optimize the initial condition. A function of C is constructed as follows (Madhi & Mohamed, 2017c):

$$f(C) = \sum_{r=2}^{m} \left[\hat{x}_{(r)}^{(0)} - x_{(r)}^{(0)}\right] \text{^}2 \qquad (3.47)$$

Substituting (3.46) into (3.47) yields:

$$f(C) = \sum_{r=2}^{m} \left[C * \left(e^{-ar} - e^{-a(r-1)}\right) - x_{(r)}^{(0)}\right] \text{^}2 \qquad (3.48)$$

Now, (3.48) is differentiated with respect to C and let the derivative be equal to zero. Then as in (Madhi & Mohamed, 2017c) the optimized C is given as:

$$C = \frac{\sum_{r=2}^{m}\left(e^{-ar} - e^{-a(r-1)}\right) * x_{(r)}^{(0)}}{\sum_{r=2}^{m}\left(e^{-ar} - e^{-a(r-1)}\right) \text{^}2} \qquad (3.49)$$

Therefore, the process of prediction by the modified initial condition grey model is outlined in four steps which include:

a. Calculation of the background value from the AGO sequence of (3.3) by (3.20),
b. Computation of the developing coefficient and grey input parameters by (3.33),
c. Computation of the optimized value of C by (3.49) and
d. Computation of the restored (predicted) values by (3.46).

When MIC is used to improve OGM(1,1)'s accuracy a Modified Initial Condition GM(1,1) is developed and referred to as MICGM(1,1). And combination of DGT and MIC in improving OGM(1,1)'s accuracy results to a hybrid model called Modified Initial Condition Grouped Grey Model (MICGGM(1,1)).

e) Relative Variable Smoothing Approach (RVSA)

According to (Wei & Dang, 2016) the trend of the relative variables is a causal to the low accuracy of the GM(1,*n*). Thus this thesis proposes to adopt the procedure of the conventional grey model GM(1,1) (He & Tao, 2014) as a data pre-processing tool and smooth the relative variables. Thus a GM(1,1) fitted data series is generated for each relative variable and the new smoothed data series are used, instead of the observed data series, in GM(1,*n*) modelling. Note that GM(1,1) is a special case of the GM(1,*n*) in which *n*=1 and its constructional structure is as outlined in section 3.1.1 part I. Now, introducing RVSA and DGT in OGM(1,1) results to a Variable Smoothed-Grouped data multivariate Grey Model denoted as VSGGM(1,*n*). This approach is employed in section 4.3.4 to improve the precision of the conventional GM(1,3).

f) Fourier Series Error Correction Approach (FSECA)

In this section, based on Fourier series theory, the use of the Fourier series in time series fitting error correction is discussed. A Fourier series takes a signal and decomposes it into a sum of sines and cosines at different frequencies and is a very good way of approximating functions in a finite range, using only the first few modes (i.e. truncating the sum over m after some low value $m = N$) (Hu, YC., 2021). Generally, Fourier components get smaller as the mode number m increases and if the Fourier series is truncated after N terms, an error E_N that measures how much the truncated Fourier series differs from the original function, can defined. That is to say if:

$$f_N(t) = \frac{a_0}{2} + \sum_{m=1}^{N} \left[a_m \cos\left(\frac{2\pi m}{T} t\right) + b_m \sin\left(\frac{2\pi m}{T} t\right) \right] \tag{3.50}$$

the error can be defined as:

$$E_N = \int_{-\frac{T}{2}}^{\frac{T}{2}} dt|f(t) - f_N(t)|^2 \geq 0 \tag{3.51}$$

where $f_{(t)}$ is the signal waveform, $a_0/2$ is a constant term representing functions that are entirely above the x-axis, T is the period, a_m and b_m are Fourier coefficients to be determined by the least square method (Hu, YC., 2021). Furthermore, the Fourier coefficients are designed to minimize the square of the error from the actual function. The Fourier series is optimal in the least-squares sense because the Fourier coefficients which minimize E_N for some given N, are exactly the coefficients that can be obtained by solving the full Fourier problem.

Now the Fourier analysis allows to isolate (filter) certain frequency ranges such as those of noise and can compensate the random error of a system (Niu et al., 2014). This property makes it useful in modifying the residual to effectively improve the prediction accuracy of the conventional GM(1,*n*). Thus, modification of the error based on Fourier series is achieved as follows. The relative error $e(k)$, can be computed as:

$$e(k) = x^{(0)}(k-1) - \hat{x}^{(0)}(k-1); \quad k = 2, 3, \cdots, m; \quad e(1) = 0 \tag{3.52}$$

where $x^{(0)}$ is the original dependent time series and $\hat{x}^{(0)}$ is the restored value of this series by the conventional GM(1,*n*) as obtained before error correction.

This results to a relative error series presented as:

$$e = [e(1), e(2), e(3), \cdots, e(m)] \tag{3.53}$$

By Fourier series and truncating the sum over m after some low value $m = N$ the modified random error sequence can be approximated as (Hsu, 2009):

$$\hat{e}(k) = \frac{a_0}{2} + \sum_{m=1}^{N} \left[a_m \cos\left(\frac{2\pi m}{T} k\right) + b_m \sin\left(\frac{2\pi m}{T} k\right) \right] \tag{3.54}$$

where a_0, a_m and b_m are Fourier coefficients, $N = \{(m-1)/2\} - 1, k = 2, 3, \cdots, m$, and $T = m - 1$.

For convenience and to compute the Fourier coefficients a_0, a_m and b_m based on the original relative error sequence, e, (3.54) is rewritten in matrix form as:

$$e = AP \tag{3.55}$$

where A is a $2N+1$ by $m-1$ matrix obtained as:

$$A=\begin{bmatrix} \frac{1}{2} & \cos\left(\frac{2\pi.2}{T}\right) & \sin\left(\frac{2\pi.2}{T}\right) & \cos\left(\frac{2.2\pi.2}{T}\right) & \sin\left(\frac{2.2\pi.2}{T}\right) & \cdots & \cos\left(\frac{N..2\pi.2}{T}\right) & \sin\left(\frac{N.2\pi.2}{T}\right) \\ \frac{1}{2} & \cos\left(\frac{2\pi.3}{T}\right) & \sin\left(\frac{2\pi.3}{T}\right) & \cos\left(\frac{2.2\pi.3}{T}\right) & \sin\left(\frac{2.2\pi.3}{T}\right) & \cdots & \cos\left(\frac{N.2\pi.3}{T}\right) & \sin\left(\frac{N.2\pi.3}{T}\right) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ \frac{1}{2} & \cos\left(\frac{2\pi.m}{T}\right) & \sin\left(\frac{2\pi.m}{T}\right) & \cos\left(\frac{2.2\pi.m}{T}\right) & \sin\left(\frac{2.2\pi.m}{T}\right) & \cdots & \cos\left(\frac{N.2\pi.m}{T}\right) & \sin\left(\frac{N.2\pi.m}{T}\right) \end{bmatrix} \tag{3.56}$$

and P is the Fourier coefficient vector given as:

$$P = [a_0\,, a_1, b_1, \cdots, a_N, b_N]^T \tag{3.57}$$

And from the least square method (Nguyen et al., 2020):

$$P = [A^T A]^{-1} A^T e^T \tag{3.58}$$

Then the computed values of the Fourier coefficients are to be substituted in (3.54) and the final restored value of the dependent variable is obtained as:

$$(k) = \hat{x}^{(0)}(k) + \hat{e}(k) \tag{3.59}$$

This error correction method is employed in section 4.3.4 to improve the precision of the conventional GM(1,*n*) in vehicle traffic flow prediction. Incorporation of FSECA in VSGGM(1,*n*) yields a Variable Smoothed-Grouped data-Fourier series based multivariate Grey Model denoted as VSGFGM(1,*n*).

g) Three-Step Approach (TSA)

The conventional GM(1,*n*) has a complex structure compared to the GM(1,1) and its low accuracy has been attributed to its defects in modelling mechanism, parameter

estimation and model structure (Cheng et al., 2020; Li et al., 2023; Zeng et al., 2016). According to (Lao et al., 2021; Moonchai & Rakpuang, 2015), the prediction accuracy of the conventional GM(1,*n*) is limited due to some factors such as the smoothness of the raw data, its background value computation and its model equation. Additionally, Li et al. (2023) in their study recommended residual correction to further improve GM(1,*n*)'s accuracy. In this thesis the prediction accuracy of a conventional first order three variable grey model denoted as GM(1,3), is improved by smoothing its two relative variables in addition to adopting the DGT and FSECA. It is simply the combination of the three approaches which have been explained above.

This TSA consists of smoothing the relative variables by a univariate grey model, GM(1,1), adoption of a DGT in the GM(,*n*) and correction of the prediction error (residual) based on FSECA. The univariate grey model, GM(1,1), is proposed as the data pre-processing tool for smoothing the relative variables. The FSECA filters noise components in the prediction error. This approach is simply the combination of the RVSA, DGT and FSECA methods, in that order, for improving the precision of the conventional GM(m,*n*). The TSA yields the two improved hybrid grey models as follows.

i) VSGGM(1,*n*) Improved Model

Firstly, by combining the RVSA and DGT a Variable Smoothed-Grouped data multivariate Grey Model denoted as VSGGM(1,*n*) is established.

ii) VSGFGM(1,*n*) Improved Model

Secondly, introducing FSECA in VSGGM(1,*n*) results to a Variable Smoothed-Grouped data-Fourier series based multivariate Grey Model denoted as VSGFGM(1,*n*).

h) Relative Factors; Motorcycle and Pedestrian in the GM(1,*n*)

Various factors affects vehicle traffic flow and can increase the intensity of traffic congestion. The pedestrian mode as relative factor is an important component of urban networks and greatly affects the performance of sidewalks and crosswalks, as well as

the entire network traffic operations by interacting with other traffic modes (automobile, bicycle, transit) (Zheng et al., 2016). On the other hand, the effects of motorcycle on vehicle traffic operations cannot be neglected. Traffic mixed with motorcycle is more hazardous under interruption caused by motorcycle. Such aspects affect vehicle traffic flow in one way or the other depending on whether an intersection is signalized or not (Zheng et al., 2016).

Such factors are taken into account by the multivariable grey model GM(1,n), as relative variables. Therefore, these factors are considered in predicting vehicle traffic flow in this study. These factors along with the main factor of vehicle traffic flow are the n variables in the GM(1,n). For instance, considering three variables ($n = 3$), say the main variable (vehicle traffic flow) and two relative variable (i.e. motorcycle and pedestrian), then these three variables can be defined in GM(1,3) as follows (Shen et al., 2019):

$X_{(1)}^{(0)}(t)$ is the time series of traffic flow (number of vehicles),

$X_{(2)}^{(0)}(t)$ is the time series of motorcycle (number of motorcycles) and

$X_{(3)}^{(0)}(t)$ is the time series of pedestrian (number of people),

where t is an order of the time series and can be minute, hour, month or year.

In this thesis the motorcycle and pedestrian are considered as the relative variables affecting vehicle traffic flow. See section 4.3.4.

i) Short-Term Traffic Flow Forecasting

Traffic flow forecasting is a crucial technology for building ITS and has gained more and more attention with the rapid development and deployment of the ITS. Traffic flow prediction means to predict the distribution of traffic flow in the near future. The prediction is based on historical and real-time traffic flow data and can include short-term, medium-term and long-term traffic flow prediction (Sun et al., 2022). In this

thesis the focus was on short-term vehicle flow forecast and the concept of grey forecasting was used.

The concept of grey forecasting is by extrapolating the modelled series of (3.9) as:

$$\hat{x}^{(1)}_{(m+1)}, \hat{x}^{(1)}_{(m+2)}, \hat{x}^{(1)}_{(m+3)}, \tag{3.60}$$

Hence by short-term forecast vehicle volume in the near future was determined at predetermined time intervals. In general, short-term traffic flow forecasting refers to the case in which the time span between t and $t+1$ does not exceed 15 minutes (or even is smaller than 5 minutes). Where t is the instant at which data is recorded and $t+1$ is the immediate instant of data collection. So, to forecast two points ahead means estimating the traffic volume at $t+1$ and at $t+2$ which corresponds to $m+1$ and $m+2$ of grey forecasting, see (3.60).

3.1.3 Evaluating the Performance of the Improved GM (m,*n*)

Measuring and judging accuracy improvement was of paramount importance in this research. Measures of model performance, namely Root Mean Square Error (RMSE), Root Mean Square Percentage Error (RMSPE), Mean Absolute Error (MAE), Mean Absolute Percentage Deviation (MAPD) and the Percent Error (δ) were adopted to evaluate the accuracy improvement of the grey models. These error indicators were computed by (3.61) to (3.65) (Chai & Draxler, 2014; Guo et al., 2015; Liu & Cocea, 2017; Lotfalipour et al., 2013; Willmott & Matsuura, 2005; Zhang et al., 2015):

$$RMSE \triangleq \sqrt{\left(\frac{1}{m}\sum_{i=1}^{m}\left(x^{(0)}_{(i)} - \hat{x}^{(0)}_{(i)}\right)^2\right)} \tag{3.61}$$

$$RMSPE \triangleq \sqrt{\left(\frac{1}{m}\sum_{i=1}^{m}\left(\frac{x^{(0)}_{(i)} - \hat{x}^{(0)}_{(i)}}{x^{(0)}_{(i)}}\right)^2\right)} \times 100 \tag{3.62}$$

$$MAE \triangleq \frac{1}{m}\sum_{i=1}^{m}\left|\hat{x}_{(i)}^{(0)} - x_{(i)}^{(0)}\right| \tag{3.63}$$

$$MAPD \triangleq \frac{\sum_{i=1}^{m}\left|\hat{x}_{(i)}^{(0)} - x_{(i)}^{(0)}\right|}{\sum_{i=1}^{m}\left|x_{(i)}^{(0)}\right|} \times 100 \tag{3.64}$$

$$\delta \triangleq \left|\frac{\hat{x}_{(i)}^{(0)} - x_{(i)}^{(0)}}{x_{(i)}^{(0)}}\right| \times 100 \tag{3.65}$$

where $x_{(i)}^{(0)}$ is the original data, and $\hat{x}_{(i)}^{(0)}$ is it's forecasting, $n = m$ is the total number of data, (see (3.1) and (3.13)). In addition accuracy improvement was judged by comparing the improved grey models with other best models available (e.g. MBV GM(1,1) of (Mahdi & Mohamed, 2017a) and MIC GM(1,1) of (Mahdi & Mohamed, 2017b) in terms of accuracy. In other words, comparison of the DGT with MBV and MIC in improving the precision of the original grey models was evaluated.

3.1.4 Analyzing the Performance of the Data Grouping Techniques

To improve the prediction accuracy of the conventional GM (1,1) the DGT, MBV and MIC techniques are adopted. For the DGT various techniques are proposed and graphically analyzed and their performance is evaluated in detail. In particular, the strong and weak grouping techniques are discussed to show that prediction accuracy continues to improve as the number of data groups increases. Therefore, detailed performance analysis of the proposed strong grouping (SG) and weak grouping (WG) techniques is presented and consequently the usefulness of the proposed technique (i.e. the DGT) is precisely discussed and shown. In particular, the connection between data points per group and the number of groups is evaluated. It is shown that the more the number of groups, the more the overlaps and the more the data is smoothed to increase prediction accuracy. In order to achieve this performance analysis two types of data (from Table 3.3), namely vehicle volume and CO_2 emissions, up to 25 data points are employed. Both vehicle volume and CO_2 emissions data were grouped in 4s and 5s. Generally, these four grouping techniques and the one group technique are compared

to determine the most accurate technique for improving the prediction accuracy of the conventional GM(1,1).

3.2 Investigating the Effect of Univariate and Multivariate Formulation on Accuracy of Grey Models on Short Term Vehicle Traffic Flow Forecasting

The conventional univariate and multivariate grey models are compared to improved univariate and multivariate grey models. This comparison establishes the effect of the relative factors in vehicle flow modelling. The univariate model is used to fit the main variable data set without considering the relative variables. On the other hand, the multivariate model is used to fit the same main variable data set with consideration of the relative variables. The main variable in this thesis is the vehicle flow time series data whereas the relative variables are pedestrian and motorcycle time series data. The three variables are from traffic flow data collected from the Nairobi CBD. Since $n = 3$ the conventional multivariate grey model considered in this thesis is GM(1,3). On the other hand the conventional univariate grey model is GM(1,1). These conventional grey models are improved by the DGT to develop hybrid grey models named as univariate Grouped Grey Model (GGM(1,1) and multivariate Grouped grey Model (GGM(1,3). Then these four models are compared in terms of their prediction accuracy to determine the effect of the relative variables on vehicle traffic flow forecasting.

3.3 Assessment of Grouping Technique Based Grey Model on Energy Consumption and Carbon Dioxide Emissions

In this research the proposed improved grey models were subjected to a variety of time series data in order to assess their performance. The improved grey models were employed in fitting and forecasting two types of time series data, namely CO_2 emission and electricity consumption. The detailed modelling of these time series data by the improved models is presented in Chapter 4, sections 4.2 and 4.4. Additionally, in this research the existing MBV and MIC methods have been compared with the proposed DGT in improving the precision of the conventional grey models.

The GM uses a cumulative model in order to create differential equations and it requires a small number of data. As mentioned earlier, in section 3.1.1 of this thesis,

the grey forecasting model is abbreviated as "GM(m,*n*)" and "m" symbolizes the degree of differential equations whereas "*n*" represents the number of variables under consideration. In this thesis a conventional first order three variable grey model denoted as OGM(1,3) (Original GM(1,3)) was considered. This model is applicable in modelling systems that involve one dependent variable and *n*-1 relative factors (Moonchai & Rakpuang, 2015). According to Moonchai and Rakpuang (2015) the prediction accuracy of the conventional GM(1,*n*) is limited due to some factors such as the smoothness of the raw data, its background value computation and its model equation.

Now in this thesis to improve the fitting accuracy of the OGM(1,*n*) the DGT is introduced into OGM(1,*n*). Thus a Grouped GM(1,*n*) denoted as GGM(1,*n*) is developed. The DGT has been discussed in section 3.1.2 of this thesis and it plays a great role in smoothing the raw data and thereby improving the smoothing and fitting accuracy of the OGM(1,*n*). The accuracies of the OGM(1,*n*) and those of the GGM(1,*n*) are compared for the purpose of validating the DGT in accuracy improvement. Additionally, these two models are subjected to two data scenarios (CO_2 emission and clean energy consumption) for the purpose of further assessing the DGT in boosting the accuracy of the OGM(1,*n*).

3.4 Research Data

As mentioned earlier a variety of time series data were used in this research for the purpose of validating the proposed new concept of data grouping in improving the precision of the original grey models. These data included vehicle traffic flow, vehicular CO_2 emission, electricity consumption data, energy consumption, economic scale and population size. Therefore, several empirical examples were given in this thesis based on different data scenarios sourced from different origins. This is to prove that the improved grey models are portable and multidisciplinary applicable.

3.4.1 Data Source

For a vehicle traffic system there are three cyclical variations: Traffic flow characteristics varies as hourly pattern, daily pattern, monthly and yearly pattern

(Sarraj, 2018). Of particular interest in this study is the hourly vehicle traffic pattern. For the hourly pattern, vehicle traffic congestion data for the whole day is essential (Sarraj, 2018). Of particular interest in this study was minute time traffic pattern data (i.e. traffic data collected and recorded in minute time segments). In this study vehicle traffic data was sourced from various sources as categorized below.

Table 3.3: 6:00 Am to 8:00 Am Vehicle Flow and Vehicular CO_2 Emission Data

DP	Time of Day	Traffic Flow	CO_2 [g]
1	6:00	0	0.0
2	6:05	14	13.62
3	6:10	35	56.06
4	6:15	54	67.88
5	6:20	55	83.08
6	6:25	95	96.13
7	6:30	83	84.74
8	6:35	89	181.12
9	6:40	98	128.72
10	6:45	134	132.58
11	6:50	103	249.10
12	6:55	173	311.93
13	7:00	110	300.76
14	7:05	167	271.66
15	7:10	160	331.52
16	7:15	150	250.40
17	7:20	210	275.48
18	7:25	200	311.09
19	7:30	172	321.01
20	7:35	149	258.03
21	7:40	154	338.53
22	7:45	140	378.92
23	7:50	157	174.47
24	7:55	146	330.33
25	8:00	145	339.23

Note: Data by author; A-1-8 Traffic Congestion and CO_2 Emission Analysis via MITRAM in Tokushima City. *Proceedings of the IEICE Engineering Sciences Society/NOLTA Society* Conference. January 3rd, 2016.

I. Data Sourced from Past Research Contexts

a) Vehicle Flow Data from the National Highway Route 11 of Tokushima City

The Table 3.3 data were from 6:00 am to 8:00 am of the day which were as recorded

at every data point (DP). The vehicle flow data was from the national highway route 11 of Tokushima city, Japan. The data were simulated in a microscopic traffic simulator (MITRAM) and output data were recorded (See Table 3.3). For more details about MITRAM see (Ishikawa, et al., 2005; Mori et al., 2013).

b) Vehicular CO_2 Emission Data

The vehicular CO_2 emission time series data employed in this research consists of data output as simulated by MITRAM and recorded in Table 3.3. The vehicle flow data from the national highway route 11 of Tokushima city (see (a) above) were simulated in MITRAM, and the simulator could output assumed CO_2 emission data based on vehicle traffic density.

The data from (Özceylan, 2016) consists of CO_2 emission *(mt)* as the dependent or output variable and energy consumption *(mtoe)* and number of motor vehicles *(10^6)* as relative or input variables. These data were as tabulated in Table 4.80.

c) Electricity/Energy Consumption Data

The electricity consumption data used in this study consists of Kenya's total electricity consumption, expressed in kilowatt-hours (kWh), for the period from the year 2000 to 2019 except for the year 2015 whose value was approximated by computing the mean of the adjacent data. The data were retrieved from the Central Intelligence Agency (CIA) World Factbook (www.indexmundi.com/g/g.aspx?c=ke&v=81) on December 29, 2019 and tabulated as in Table 4.13.

Data sourced from (Cheng et al., 2020) consists of clean energy (10,000 tons of standard coal) as the output variable and economic scale, GDP (CNY 0.1Billion) and population size (10,000 people) as the input variables. These data are as tabulated in Table 4.82.

II. Traffic Data Collected from Nairobi CBD

For the case study, vehicle, pedestrian and motorcycle traffic data were collected from real-time traffic system information, from Nairobi CBD, the study site (see section 3.5.4). Normally it is impractical to study a whole population and thus the need for sampling. Moreover, sampling is necessary for economic reasons. Traffic counts are normally not taken on a holiday nor on the day before or after a holiday. Monday mornings and Friday evenings generally show high volumes (Regehr et al., 2015; Sampson, 2017). Usually, days counted include Tuesdays, Wednesdays and Thursdays. These are days that are likely to have the same pattern of vehicle, pedestrian and motorcycle volumes within a week and these days are representative enough for one week. Mondays and Fridays have different volume patterns (Regehr et al., 2015) and higher volumes are expected on weekends. Fridays and Saturdays are usually different and have higher volumes due to traditional activities like "after works", recreational and due to the night life (Sampson, 2017). However, these days cause a bias on the volume of traffic (e.g. pedestrians) in comparison to the remaining days of the week; consequently it is better to count in other days rather than these (Miranda et al., 2011). Therefore, data were collected at intervals of 5 minutes for a duration of 15 hours per day for three consecutive days, i.e. from Tuesday to Thursday. Examples of these data were as tabulated in Tables 1 to 21 of Appendix IV.

3.4.2 Training and Test Data Sets

In time series modelling it is common to use a portion of the available data for fitting and the rest of the data for testing a model (Liu & Cocea, 2017). Ajiboye et al., (2015), asserts that a model trained with the largest size of training sets is the most accurate and it consistently delivers much better and stable results. Therefore, in this study the data for modelling were divided into two sub-sets, namely training data set and testing data set (see Figure 4.1 and Figure 4.30). To achieve good performance in modelling the training data set was made large compared to the testing data set (Ajiboye et al., 2015). The goal is to develop a trained (fitted) model for well generalization to new, unknown data. Then the fitted model is evaluated using "new" data from the held-out dataset (validation dataset) to estimate the model's accuracy in classifying the new

data. The reason for sub dividing the original data set is to reduce the risk of issues such as over fitting and, therefore, the data in the validation dataset is not used to train the models. Thus, in this thesis employed is the classical hold out method.

3.5 Traffic Data Collection and Sampling

Traffic data were collected from the study region (i.e. the Nairobi CBD). This region is shown in Figure 1.1. For the purpose of this research study the number of vehicles, motorcycles and pedestrians, passing a given point, were counted at intervals of 5 minutes. It was ensured that traffic flow data during important periods of a day were well captured in the counting-by-counting data from 6:00 am to 21:00 pm. These are the morning peak, off peak and evening peak traffic volumes (Sampson, 2017). Thus, day time traffic data were of paramount importance in this research.

3.5.1 Types of Traffic Counts

This research is tailored towards the most typical count i.e. the volume count. Due to complexity and time taken other counts based on vehicle classification, occupancy, and truck percentage are not considered. Traffic counts involve counting vehicles passing a point for varying intervals of time and can range from 365 days per year, 24 hours per day, to short term intervals of fifteen minutes (Aldrin, 1998). The various types of traffic counts include directional traffic, lane traffic and pedestrian counts. In this research vehicle traffic, motorcycle traffic and pedestrian traffic were counted (Ling et al., 2013; Yong & Xiuchun, 2013; Zheng et al., 2016). Pedestrian counts are counts of the number of people walking through the area being studied and are typically taken only for 12 hours (day time) say between 8:00 A.M. and 6:00 P.M (Miranda et al., 2011; Obiri et al., 2021). This period, accounts for most of the pedestrian volume for a whole day. Therefore, in this research interest was not only in vehicle traffic counts but also in motorcycle and pedestrian counts because the flow of motorcycles and the human factor are among the factors that affect vehicle traffic flow.

3.5.2 Method of Traffic Count

Traffic counting is categorized into two, namely manual counts and automatic counts (Bas et al., 2007; Obiri et al., 2021). The difference between these two methods can be deduced from the discussions of the respective methods below.

a) Manual/Observation Counts

The most common method of collecting traffic volume data is the manual method which involves assigning a group of people to record the number of vehicles passing, on a pre-determined location, using tally marks in inventories (Leduc, 2008). Raw data from those inventories is then organized for compilation and analysis. This method of data collection is usually expensive in terms of manpower, however it is necessary in most cases where vehicles are to be classified with a number of movements recorded separately, such as at intersections, also in case where automatic methods cannot be used due to lack of infrastructure, necessary authorization etc. This method is also appropriate for short periods of counting time (e.g. 5, 15 minutes) (Schneider et al., 2009).

b) Automatic/Machine Counts

These are counts by machines that can record passing vehicles automatically, hence known as "automatic traffic counts". This method is employed in cases where manual count method is not feasible. Various instruments are available for automatic count, which have their own merits and demerits. Some of the widely used instruments are pneumatic tubes, inductive loops, weigh-in-motion Sensor, micro-millimeter wave Radar detectors and video camera (Bas et al., 2007).

Because the manual/observation traffic count method is the most common and it is necessary in most cases where automatic methods cannot be used, it was selected for collecting data for this study. Moreover, for economic reasons (from equipment requirement point of view) the manual count method was adopted and in the direct manual method tally sheets were used due to the fact that they are rather cheap (Miranda et al., 2011). Further, in this method subjective bias is eliminated,

information obtained relates to what is currently happening and it is independent of respondents' willingness to respond.

At road intersection sites, the traffic on each arm were counted and recorded separately for each direction of movement. Traffic-counting teams (counting clerks) were set up to carry out the counting of vehicles, motorcycles and pedestrians at the various predetermined locations throughout the road networks at set interval of time.

3.5.3 Traffic Data Sampling

As it is impractical to study a whole population and to reduce the cost and workload in traffic counting, sampling of the counting locations from the study site was necessary. Sampling is a method that allows researchers to infer information about a population from a subset of the population, without having to investigate every individual. Reducing the number of individuals in a study reduces the cost and workload, and may make it easier to obtain high quality information, but this has to be balanced against having a large enough sample size with enough power to detect a true parent population (Claffy et al., 1993). By whatever method a sample is chosen, it is important that the individuals selected are representative of the whole population.

For quantification of vehicle traffic congestion which is obviously heavy during rush hours and at road junctions, then the best representative count locations were generally at or near road junctions. All junctions qualified to be selected. However, reducing the cost and workload in traffic counting was necessary to make it easier to obtain high quality representative information. Therefore, traffic data from randomly selected traffic counting locations (at road junctions) of the Nairobi CBD were collected.

There are several different sampling techniques available, and they can be subdivided into two groups: probability sampling and non-probability sampling (Hamed, 2016). In this study random sampling which is a probability method was used. The seven '+' junctions from within the CBD were randomly selected, as long as these junctions experiences heavy traffic jams.

Sample sizes can vary from a fraction of an hour to 24 hours a day, 365 days a year. Generally, peak periods need to be included in all samples. Traffic counts are normally not taken on a holiday nor on the day before or after a holiday. Monday mornings and Friday evenings generally show high volumes (Regehr et al., 2015; Sampson, 2017). Usually, days counted include Tuesdays, Wednesdays and Thursdays. These are days that are likely to have the same pattern of vehicle and pedestrian volumes within a week. Mondays and Fridays have different volume patterns (Regehr et al., 2015) and higher volumes are expected on weekends. Fridays and Saturdays are usually different and have higher volumes due to traditional activities like “after works”, recreational and due to the night life (Sampson, 2017). However, these days cause a bias on the volume of traffic (e.g. pedestrians) in comparison to the remaining days of the week; consequently it is better to count in other days rather than these (Miranda et al., 2011).

Therefore, based on random sampling, data were collected from the following seven Nairobi CBD road junctions. See Figure 3.4 for the exact location of these sites.

1. Haile Selassie Roundabout,
2. Kenyatta Avenue Uhuru Highway Roundabout,
3. University Way Uhuru Highway Roundabout,
4. Kenyatta Avenue-Moi Avenue-Mondlane Street Junction,
5. Moi Avenue-Slip Road Junction,
6. City Hall Way-Wabera Street T-Roundabout and
7. Haile Selassie Avenue-Moi Avenue Roundabout.

3.5.4 Selection of Counting Sites

For obvious reasons traffic cannot be possibly counted on every street in any given area. Therefore, it is necessary to collect traffic count samples at a variety of locations. Counting sites thought to best represent typical urban traffic congestion conditions were selected. The specific locations for counting sites were determined keeping in

mind the following facts before deciding on the counting site (National Highway Authority, 2017; Roads and Highways Department, 2001; Roads Department, 2004);

- The road section should have uniform geometric characteristics and be near junctions,
- Location should be on a horizontal (flat) and geometrically straight road section. So that enumerators have good vision of traffic approaching from all directions. The site should not be located on bends or at places where trees/buildings obscure vision.
- Section of the road to have an uninterrupted traffic flow and free of animal traffic and
- Section to meet safety requirements. For instance a lighted location would be of advantage for counts conducted after daylight.

See Figure 3.4 for the exact location of the selected counting sites 1 to 7 in the Nairobi CBD.

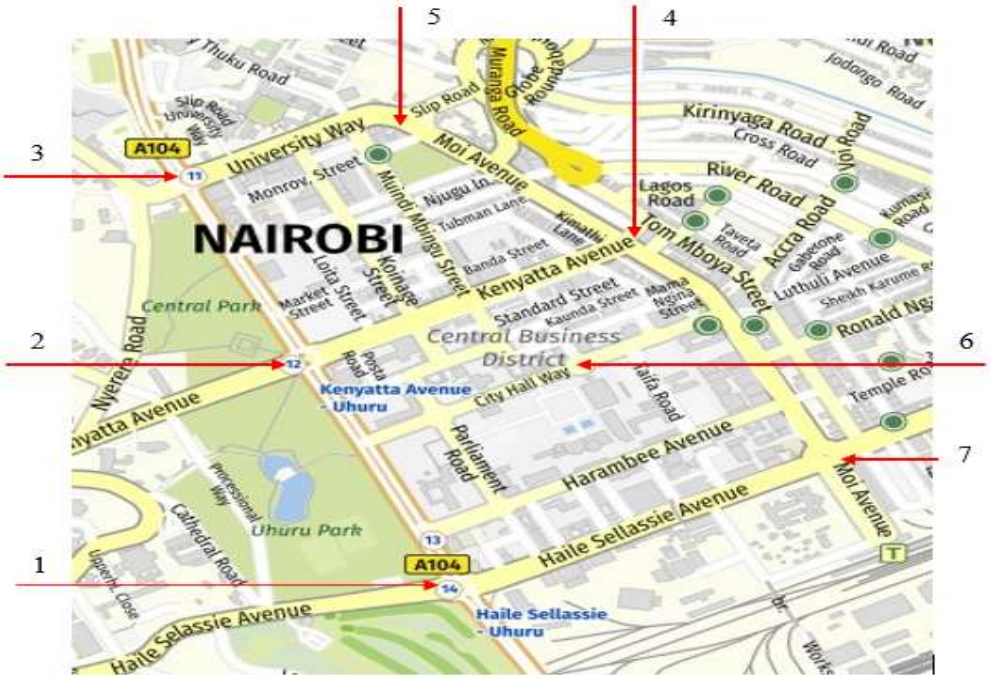

Figure 3.4: Traffic Data Counting Site Locations

Source: Google (n.d.)

3.5.5 Traffic Counting

Every 5-minute interval traffic were counted. For example, at this increment of 5 minutes counting started at 06:00, then the number of vehicles crossing a specific point on the road from 06:00 to 06:05 was recorded. The next count began at 06:06 to 06:10 and so forth. A 24-hour clock approach was used so that no confusion between morning and evening counts. Tally marks were mostly used to represent each vehicle on the 15-hour traffic count. Four directions of traffic flow – northward, southward, eastward and westward directions were considered and recorded separately.

3.5.6 Timing of Counts

The traffic counting was conducted for 15 hours (from 06:00 to 21:00). The traffic counting took place for a period of three consecutive days. Counts were conducted on days for which the traffic flow is typical of an average day of the week (Roads and Highways Department, 2001). Generally, the best days for counting were from Tuesday to Thursday (Regehr et al., 2015). In the timing of counts the following days were avoided; public holidays, Mondays, Fridays and any other days when it is known from local knowledge that traffic flows will be unusual, for instance when there are local religious ceremonies (Roads and Highways Department, 2001). Traffic flow during those days contains more complex spatial-temporal characteristics, with a large range of changes (Chen L. et al., 2023). For instance, it experiences sudden and irregular characteristics. Because of such complexity of vehicle flow, traffic flow prediction during such days is beyond the scope of this thesis.

The traffic counting clerks worked in two shifts. The first shift counted as from 06:00-14:00 and the second shift as from 14:00-21:00.

3.5.7 Staff Requirements

The total number of persons (clerks) required to conduct the count was dependent on the number of selected counting locations. At four-way junction eight enumerators (counting clerks) were required at all times, two clerks counted vehicles, motorcycles and pedestrians in each direction of traffic flow. For example, for the northward

direction, one clerk counted the northbound vehicles and the other counted the northbound motorcycles and pedestrians crossing at that point. A supervisor was also required at all times. The supervisor was responsible for ensuring that the enumerators filled the forms in correctly, collating the completed forms and acting as a relief for the clerks to provide breaks during the shift. Similarly, at a three-way junction six enumerators and a supervisor were required. Table 3.4 gives the number of clerks who did the counting for the 15-hour count, the standard 8-hour shift for each counting team was adopted (National Highway Authority, 2017; Roads and Highways Department, 2001).

Table 3.4: Staff Requirements per Shift

	15 hour count (in 2 shifts)			
Type of Junction	**Number of Junctions**	**Enumerators per Junction**	**Supervisors per Junction**	**Total Staff**
Three-way	4	6	1	28
Four-way	3	8	1	27

Note that this was the total required counting staff per shift and there were two shifts per day for three counting days. Therefore, a total staff of 330 clerks participated in the data collection exercise.

3.5.8 Equipment Requirements

The enumerators and supervisors required watches, reflector jackets among other items. Clipboards were also required for each of the team members with pencils, erasers and sharpeners. The equipment requirements for a team of enumerators and one supervisor were as shown in Table 3.5 (National Highway Authority, 2017; Roads and Highways Department, 2001).

Table 3.5: Equipment Requirements for a Team of Enumerators and One Supervisor

Item	Quantity Per Four-Way Junction	Quantity Per Three-Way Junction
Watch	4	3
Clipboards	9	7
Pencils	9	7
Pencil sharpeners	4	3
Erasers	4	3
Spring files (to store forms	5	4
Jacket reflectors	9	7
Umbrellas	5	4
Led power torches	5	4
Traffic count tally sheets	Enough	Enough
Daily summary sheets	3	3

3.5.9 Conducting the Count

The volume of vehicles, motorcycles and pedestrians were recorded onto a standard form, the Traffic Count Tally Sheet (see Appendix I) (Roads and Highways Department, 2001). Data were recorded in 5-minute's time increments in order that variations in traffic flow over the hour and day could be identified. Data were also recorded in four directions of travel, which included northward, southward, eastward and westward directions. During data collection various activities took place as outlined below.

a) Prior to Starting

The first shift assembled at the station half an hour before the count was due to start. The supervisor issued Traffic Count Tally Sheets (see Appendix I) attached to a clipboard to each enumerator, together with a pencil. The enumerators then filled in details of the count on the top of the sheet, according to the supervisor's instructions. Finally, the enumerator entered the starting time of the count in the left hand "Time" column.

b) Counting

The supervisor directed the enumerators to their assigned sites of the road five minutes prior to the starting time. At the start the supervisor could announce the start of the count and the clerks began to record the number of vehicles, motorcycles and pedestrians passing on that point of the road. A five bar tally was used when and where possible. Record of data was done in the Traffic Count Tally Sheets of Appendix I.

c) Supervision

The supervisor ensured that the enumerators were filling in the tally sheets correctly. He/she also acted as a relief to the enumerators allowing them to have alternate meal breaks.

In addition to these duties the supervisor was responsible for completing a short report on the shifts count. This was done on the Traffic Count Report Sheet (see Appendix II).

d) Weather Report

The Traffic Count Report Sheet of Appendix II included a brief summary of the weather conditions and incidents that may have affected the validity of the count (especially accidents or road closures in the vicinity of the count station). It was the responsibility of the supervisor to complete this report.

Finally, at the end of the shift the supervisor could collate all the sheets in order and keep them in safe custody.

3.5.10 Data Presentation

As mentioned earlier data were collected from the Nairobi CBD for 15 hours of a day. Information on traffic data is not always easily accessible. Thus, the essential part of any data collection process is to analyze and present the data in a format that is easily understandable. In this thesis tables were used to present the collected data. The tables typically show five-minute interval traffic flow pattern. See Tables 1- 21 in Appendix

IV. This gave useful and comprehensive information of the traffic volume data on the Nairobi CBD Road network. In addition, the error and accuracy evaluations of the forecasted results were tabulated in table form throughout the thesis.

CHAPTER FOUR

RESEARCH RESULTS AND DISCUSSION

Analyzed and discussed in this chapter include improvement of the original GM(1,1) by the MBV, MIC and DGT methods, evaluation of the improved grey models, analysis of the proposed DGTs in order to determine the most accurate DGT, validation of the DGTs in improving both GM(1,1) and GM(1,3), case study analysis and finally comparison of the univariate (GM(1,1)) and multivariate (GM(1,*n*)) grey models to determine the influence of the relative factors on their prediction performance.

4.1 Assessment of the Effect of Data Grouping Technique on Short Term Vehicle Traffic Flow Forecasting Grey Model Performance

4.1.1 MBV, MIC and DGT in GM(1,1) Modelling and Short-Term Traffic Flow Forecasting

In this thesis the conventional GM(1,1) is referred to as the Original GM(1,1) and denoted as OGM(1,1). The precision of the OGM(1,1) is improved by, firstly, introducing modification of the background value (MBV) (Liu et al., 2014) in the OGM(1,1) and establishing a Modified Background Value Grey Model which is denoted as MBVGM(1,1). Secondly, a DGT is introduced into OGM(1,1) to establish a Grouped Grey Model abbreviated as GGM(1,1). Overall, by combination of the MBV and DGT in OGM(1,1) a Modified Background Value Grouped Grey Model abbreviated as MBVGGM(1,1) is established. Moreover, from the GST, to improve the precision of the OGM(1,1), emphasis should be based on the "new information prior using" principle (Jong & Liu, 2014). Therefore in this section proposed is combination of the DGT and MIC methods, which adheres to this principle, in improving the performance of the OGM(1,1). Hence, this thesis section institutes a new perspective of amalgamating the DGT and MIC in optimizing the accuracy of the OGM(1,1). The developed optimized grey model is referred to as the single variable first-order Modified Initial Condition Grouped Grey Model and denoted as MICGGM(1,1). Additionally, employing MIC alone to improve OGM(1,1)'s accuracy results to a Modified Initial Condition Grey Model named as MICGM(1,1). So

modelling and short-term forecasting of traffic flow was accomplished by these improved grey models.

Therefore, in this section the thesis introduces a new approach of the DGT and combines it with existing MBV and MIC methods to improve fitting and short-term forecasting accuracy of the OGM(1,1) and the results are as discussed below.

I. Traffic Flow Fitting

a) Data Source

In this section traffic flow data of Table 3.3 was employed. The traffic flow data were from 06:00 to 08:00 of the day. At 5-minute interval this constitutes 25 data points.

b) Training and Test Data Sets

It is common to use a portion of the available data for fitting and the rest of the data for testing a model. Ajiboye et al., (2015), asserts that a model trained with the largest size of training sets is the most accurate and it consistently delivers much better and stable results. Thus, traffic flow data of Table 3.3 is subdivided into training and test data sets, as in Figure 4.1, for estimating and evaluating the proposed grey models, respectively (Liu & Cocea, 2017).

Thus, to analyze the performance of the proposed grey models a numerical example was simulated in MATLAB. The vehicle flow data in Table 3.3 was partitioned into two as in Figure 4.1. The first 22 points (i.e. from 6:00 am to 7:45 am) were used to train the proposed models whereas the last 3 points (i.e. from 7:50 am to 8:00 am) were used as the test data set.

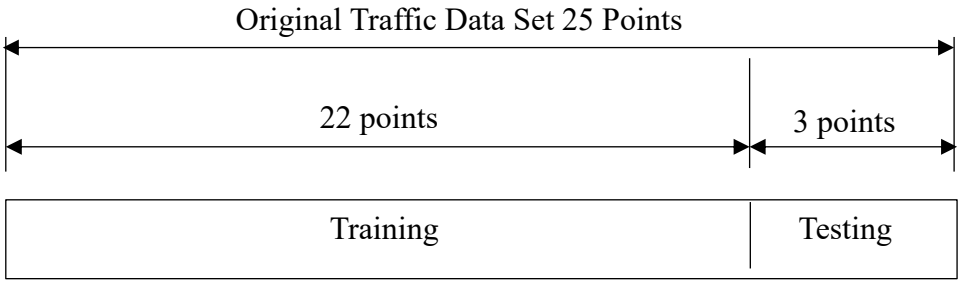

Figure 4.1: Two-Way 25-Point Data Splitting

c) Training the Grey Models

The AGO, MGO and IAGO operations were performed on the training data set by (3.3), (3.19) and (3.13), respectively. Then the models were trained as follows.

For the OGM(1,1) these operations were computed and the resultant data were:

$$X^{(1)} = \{0\ 14\ 49\ 103\ 158\ 253\ 336\ 425\ 523\ 657\ 760\ 933\ 1043\ 1210\ 1370\ 1520\ 1730\ 1930\ 2102\ 2251\ 2405\ 2545\}, \text{this is the AGO.} \quad (4.1)$$

$$Z^{(1)} = 1.0e{+}03 * \{0.0070\ 0.0315\ 0.0760\ 0.1305\ 0.2055\ 0.2945\ 0.3805\ 0.4740\ 0.5900\ 0.7085\ 0.8465\ 0.9880\ 1.1265\ 1.2900\ 1.4450\ 1.6250\ 1.8300\ 2.0160\ 2.1765\ 2.32802.4750\}, \text{this is the MGO} \quad (4.2)$$

$$y = \{14\ 35\ 54\ 55\ 95\ 83\ 89\ 98\ 134\ 103\ 173\ 110\ 167\ 160\ 150\ 210\ 200\ 172\ 149\ 154\ 140\}, \text{this is the measured data.} \quad (4.3)$$

The data matrix A and the measured vector y are obtained as:

$$A = \begin{bmatrix} -\mathbf{Z}^{(1)}_{(2)} & 1 \\ -\mathbf{Z}^{(1)}_{(3)} & 1 \\ \vdots & \vdots \\ -\mathbf{Z}^{(1)}_{(m)} & 1 \end{bmatrix} = 1.0e{+}03 * \begin{bmatrix} -0.0070 & 1 \\ -0.0315 & 1 \\ \vdots & \vdots \\ -2.4750 & 1 \end{bmatrix}$$

$$y = \begin{bmatrix} X_{(2)}^{(0)} \\ X_{(3)}^{(0)} \\ \vdots \\ X_{(m)}^{(0)} \end{bmatrix} = \begin{bmatrix} 14 \\ 35 \\ \vdots \\ 140 \end{bmatrix} \tag{4.4}$$

For the OGM(1,1) all the training data set is used to compute parameters a and b from (3.33) and the obtained values are a = -0.0516 and b= 69.4717. Consequently, the time response function of (3.9) for this model was obtained as:

$$\hat{X}_{(r+1)}^{(1)} \hat{=} 1{,}346.3508\, e^{0.0516r} - 1{,}346.3508,\ r = 0,1,2,\cdots,m-1 \tag{4.5}$$

The computed values from (4.5) were subjected to IAGO and the final fitting values obtained as:

$$\begin{aligned} &\hat{X}^{(0)} \\ &= \{0\ 71.2956\ 75.0718\ 79.0479\ 83.2346\ 87.6431\ 92.2851\ 97.1729\ 102.3196 \\ &\quad 107.7389\ 113.4452\ 119.4538\ 125.7805\ 132.4424\ 139.4572\ 146.8434 \\ &\quad 154.6209\ 162.8103\ 171.4335\ 180.5133\ 190.0741\ 200.1413\} \end{aligned} \tag{4.6}$$

The fitting values of (4.6) is the IAGO and these values were tabulated in Table 4.1 where DP stands for data point. Figure 4.2 shows the plots for the real values, OGM(1,1)'s fitted values and the error curve (the de-trended vehicle volume). The error curve indicates the difference between the simulated and actual values. The actual values are the real values (or raw data).

Now the OGM(1,1) is modified by reconstructing the background value (see section 3.1.2c). Thus parameters a and b are calculated from (3.33) using (3.40) instead of (3.20). This modification resulted to a new background value given as:

$$\begin{aligned} Z^{(1)} = 1.0\text{e+}03\ *\{&0\ 0.0289\ 0.0741\ 0.1304\ 0.2012\ 0.2954\ 0.3800\ 0.4732 \\ &0.5865\ 0.7108\ 0.8391\ 0.9921\ 1.1207\ 1.2906\ 1.4458\ 1.6191 \\ &1.8308\ 2.0182\ 2.1783\ 2.327\ 2.4761\} \end{aligned} \tag{4.7}$$

Thus the modified data matrix A was computed as:

$$A = 1.0\text{e}+03 * \begin{bmatrix} 0 & 1 \\ -0.0289 & 1 \\ \vdots & \vdots \\ -2.4761 & 1 \end{bmatrix} \tag{4.8}$$

With this modification a= -0.0515 and b= 69.6675. Hence the MBVGM(1,1)'s time response function of (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} = 1{,}352.7670\, e^{0.0515r} - 1{,}352.7670, \quad r = 0,1,2,\cdots,m-1 \tag{4.9}$$

The corresponding fitting values are as recorded in Table 4.1 and Figure 4.3 shows the plot of the real, fitted and de-trended values. De-trended values are the errors involved in fitting and forecasting, i.e. the residues.

Table 4.1: Original and Improved Grey Models' Traffic Flow Fitting and Forecasting Values

Raw Data		Grey Model					
DP	Real Value	OGM(1,1)	MBVGM(1,1)	MICGM(1,1)	GGM(1,1)	MBVGGM(1,1)	MICGGM (1,1)
Training		Fitted Values					
1	0	0	0	0	0	0	0
2	14	71.2956	71.4918	69.5654	15.6615	17.6566	15.9281
3	35	75.0718	75.2683	73.2499	34.7612	36.1439	35.1083
4	54	79.0479	79.2444	77.1295	50.1448	50.8296	50.8452
5	55	83.2346	83.4305	81.2147	60.5172	57.5155	60.9350
6	95	87.6431	87.8377	85.5162	88.2873	84.9983	88.8117
7	83	92.2851	92.4777	90.0455	86.2357	84.8559	86.2782
8	89	97.1729	97.3628	94.8147	87.1706	84.9820	87.3272
9	98	102.3196	102.5060	99.8365	102.3888	98.6913	102.5789
10	134	107.7389	107.9209	105.1242	122.6047	116.2246	122.9481
11	103	113.4452	113.6218	110.6921	119.3534	114.3712	119.5043
12	173	119.4538	119.6239	116.5548	153.3755	148.3196	153.5545
13	110	125.7805	125.9430	122.7281	128.6640	126.6200	128.7178
14	167	132.4424	132.5959	129.2283	156.3421	155.6647	156.4133
15	160	139.4572	139.6003	136.0728	158.7244	153.6437	158.9242
16	150	146.8434	146.9747	143.2798	158.8813	153.6201	159.0757
17	210	154.6209	154.7386	150.8685	202.0690	197.7780	202.3187
18	200	162.8103	162.9127	158.8592	200.7975	202.6486	200.9595
19	172	171.4335	171.5185	167.2731	171.8332	175.9304	171.9366
20	149	180.5133	180.5790	176.1326	152.9252	155.8546	153.0167
21	154	190.0741	190.1180	185.4613	148.3633	150.3881	148.3897
22	140	200.1413	200.1610	195.2842	143.2761	144.6307	143.2855
Testing		Short-Term Forecasted Values					
23	157	210.7416	210.7345	205.6273	122.4088	122.1732	122.4716
24	146	221.9034	221.8666	216.5182	130.0374	130.8847	130.0599
25	145	233.6564	233.5866	227.9860	131.0119	131.8423	131.0204

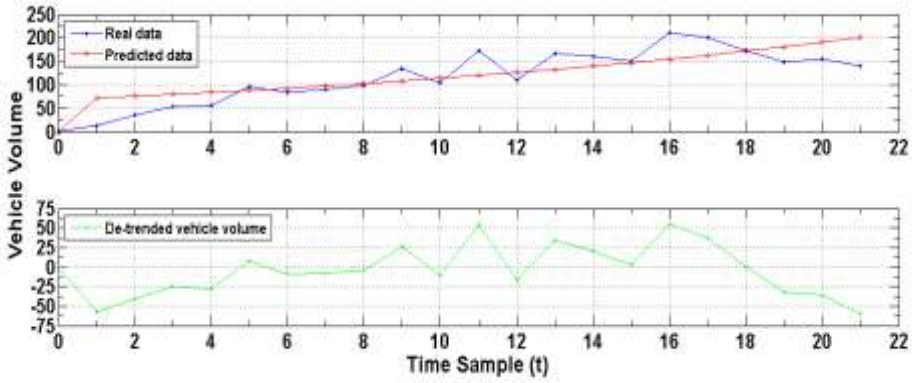

Figure 4.2: Vehicle Flow OGM(1,1) Fitting

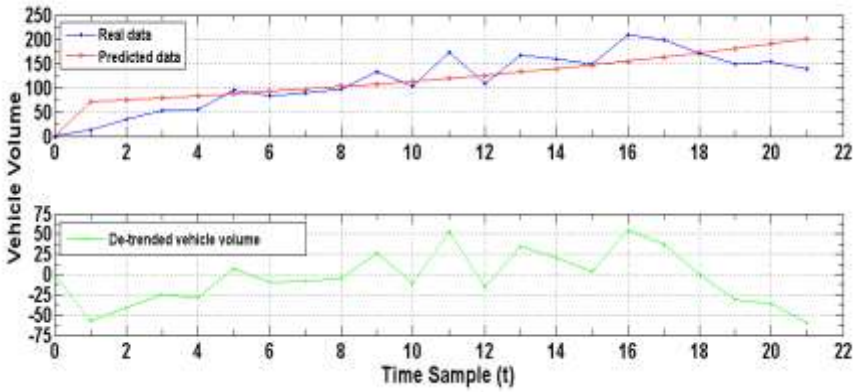

Figure 4.3: Vehicle Flow MBVGM(1,1) Fitting

Next MIC is introduction into OGM(1,1) (see section 3.1.2d) to establish the MICGM(1,1). The AGO series for MICGM(1,1) is the same as that of OGM(1,1). Since the background value is not modified its value also remains the same as that of the OGM(1,1). Consequently, The data matrix A and the measured vector y are also the same as those of the OGM(1,1). So the computed parameter values were a= -0.0516 and b= 69.4717 and the optimized initial condition value computed by (3.49), C= 1.2474e+03. Therefore, the resulting modified initial condition model structure of the grey model from (3.46) is given as:

$$\hat{x}_{(r)}^{(0)} = 1.2474 * e^{3} * \left(e^{0.0516r} - e^{0.0516(r-1)}\right), \quad r = 2,3,..m \tag{4.10}$$

The fitting values computed from (4.10) were tabulated in Table 4.1. Figure 4.4 is an indication of how well the real and fitted values of the MICGM(1,1) fit onto each other and the de-trended values.

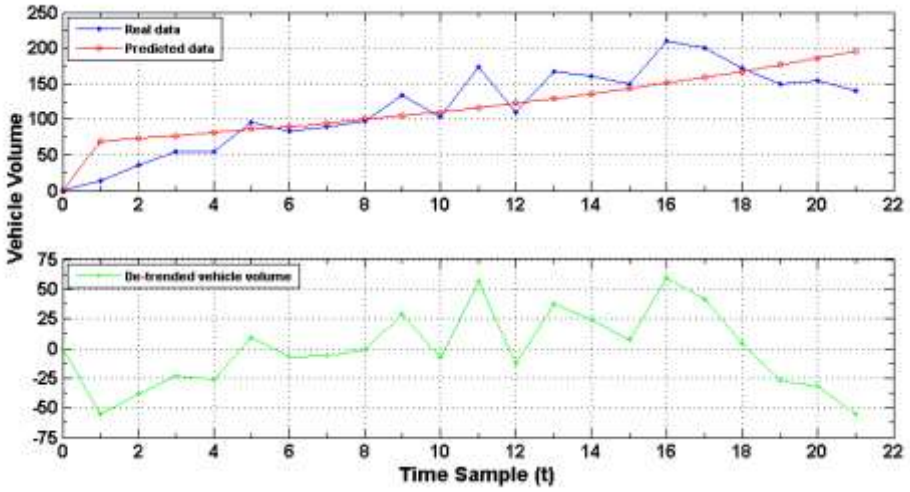

Figure 4.4: Vehicle Flow MICGM(1,1) Fitting

MICGM(1,1)'s fitting accuracies were found to be 77.1957% and 77.9812% from RMSPE and MAPD calculations, respectively, as in Table 4.4.

For the GGM(1,1), the DGT is adopted in the simulation process as follows. Firstly, the first 22 data points shown in Table 3.3 were grouped into 19 groups of 4s based on (3.34) and Figure 4.5. These groups are formed and tabulated as in Table 3.2. Secondly, the OGM(1,1) procedure was applied on each group to obtain fitted data (FD) for each group as tabulated in Table 4.2. Since data is grouped the parameters a and b are many and have unique values for each group and therefore these values have not been provided in this thesis. Moreover, each formed group has a different time response function arising from (3.9). Thus 19 different time response functions are formulated. Also note that in Table 4.2 the FD for groups 6 to 14 are not shown because of wanting to reduce the size of Table 4.2. Thirdly, the overall simulation data sequence is obtained by superimposing the group simulation data at points of overlaps and this overall (final) sequence is as indicated in Table 4.2. For instance, groups 1 and 2 overlap once at data point 2 and, thus, the final FD is obtained as (17.3230+14)/2=15.6615 and groups 1, 2 and 3 overlaps twice at data point 3 resulting to a final fitted value

computed as (30.3630+38.9204+35)/3=34.7611. Also, groups 1, 2, 3 and 4 overlap three times at data point 4 and the final fitted value is obtained as (53.2192+47.3115+46.0486+54)/4=50.1448. This sequence of computations is continued to generate the final simulation sequence (i.e. final FD) which is also indicated in Table 4.1. Figure 4.6 shows the corresponding plots and it can be observed that this improved grey model has good fitting ability.

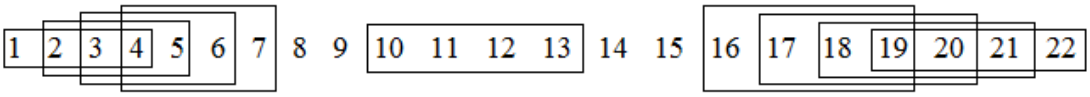

Figure 4.5: Strong Grouping (SG). The 22 Data Points Have Been Grouped in 4s to Form 19 Groups

Now, if the DGT is introduced into MBVGM(1,1) simulation the MBVGGM(1,1) is established. Following similar procedure to that of the GGM(1,1) 19 groups of data are formed. Thus the MBVGGM(1,1) equally has different time response function for each formed group with unique values of parameters $\boldsymbol{a}$ and $\boldsymbol{b}$. Its computed fitting values are tabulated in Table 4.1 and plotted in Figure 4.7.

Table 4.2: Group and Final Fitted Data for GGM(1,1)

Groups 1 - 5 and 15 - 19 fitted data											
AD	**G1**	**G2**	**G3**	**G4**	**G5**	**G 15**	**G 16**	**G 17**	**G 18**	**G 19**	**Final FD**
0	0										0
14	17.3230	14.0000									15.6615
35	30.3630	38.9204	35.0000								34.7612
54	53.2192	47.3115	46.0486	54.0000							50.1448
55		57.5117	64.0354	65.5217	55.0000						60.5172
95			89.0480	76.9918	92.1096						88.2873
83				90.4698	88.9531						86.2357
89					85.9048						87.1706
98											102.3888
134											122.6047
103											119.3534
173											153.3755
110											128.6640
167											156.3421
160						160.0000					158.7244
150						163.6233	150.0000				158.8813
210						185.5501	212.8429	210.0000			202.0690
200						210.4152	193.2895	199.4853	200.0000		200.7975
172							175.5324	172.1672	167.6332	172.0000	171.8332
149								148.5902	158.1018	152.0837	152.9252
154									149.1124	147.6142	148.3633
140										143.2761	143.2761

Finally, in training, the MICGGM(1,1), the DGT and MIC are adopt in the simulation as follows. Firstly, the first 22 data points of Table 3.3 are grouped into 19 groups of 4s based on (3.34) and Figure 4.5. These groups are formed and tabulated as in Table 3.2.

Secondly, MIC is introduced into OGM(1,1) procedure (Wang et al., 2018) and applied on each group to obtain fitted data (FD) for each group as tabulated in Table 4.3. Note that each group has unique values of the parameters a and b which have not been provided in this thesis. Consequently, the corresponding time functions are different. Also note that in Table 4.3 the FD for groups 6 to 14 are not shown because of wanting to reduce the size of Table 4.3.

Thirdly, the overall simulation data sequence is obtained by superimposing the group simulation data at points of overlaps and this overall (final) sequence is as indicated in Table 4.3. For instance, groups 1 and 2 overlap once at data point 2 and, thus, the final FD is obtained as (17.8561+14)/2=15.9281 and groups 1, 2 and 3 overlaps twice at data point 3 resulting to a final fitted value computed as (31.2975+39.0274+35)/3=35.1083. Also, groups 1, 2, 3 and 4 overlap three times at data point 4 and the final fitted value is obtained as (54.8572+47.4416+47.082+54)/4=50.8452. This sequence of computations is continued to generate the final simulation sequence (i.e. final FD) which is also tabulated in Table 4.1.

Lastly, this proposed model, MICGGM(1,1), resulted with simulation values tabulated in Table 4.1 and Figure 4.8 is its plot of the real, fitted and de-trended values.

Table 4.3: Group and Final Fitted Data for MICGGM(1,1)

Groups 1 - 5 and 15 - 19 fitted data											
AD	**G1**	**G2**	**G3**	**G4**	**G5**	**G 15**	**G 16**	**G 17**	**G 18**	**G 19**	**Final FD**
0	0										0
14	17.8561	14									15.9281
35	31.2975	39.0274	35								35.1083
54	54.8572	47.4416	47.082	54							50.8452
55		57.6698	65.4726	65.5975	55						60.9350
95			91.0465	77.0809	92.1196						88.8117
83				90.5745	88.9628						86.2782
89					85.9141						87.3272
98											102.5789
134											122.9481
103											119.5043
173											153.5545
110											128.7178
167											156.4133
160						160					158.9242
150						163.8034	150				159.0757
210						185.7543	212.9778	210			202.3187
200						210.6468	193.412	199.7791	200		200.9595
172							175.6437	172.4208	167.6817	172	171.9366
149								148.809	158.1476	152.0936	153.0167
154									149.1556	147.6238	148.3897
140										143.2855	143.2855

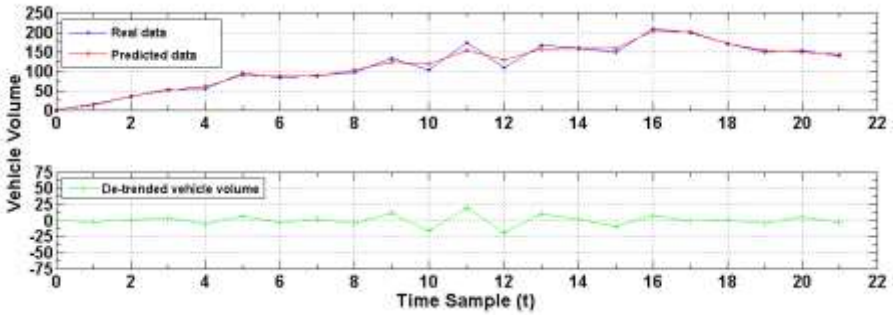

Figure 4.6: Vehicle Flow GGM(1,1) Fitting

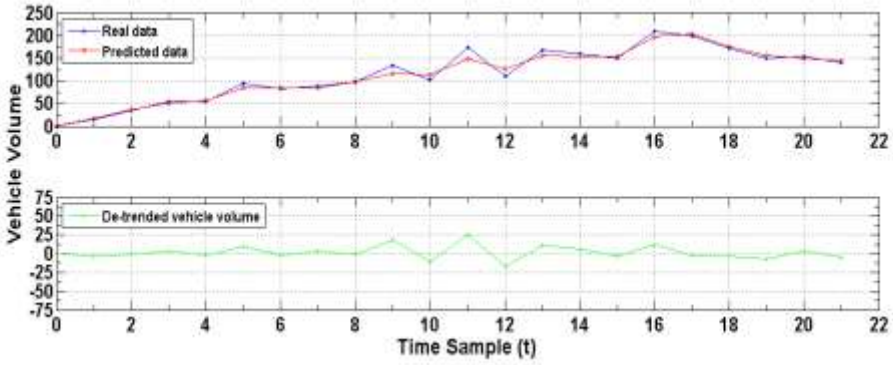

Figure 4.7: Vehicle Flow MBVGGM(1,1) Fitting

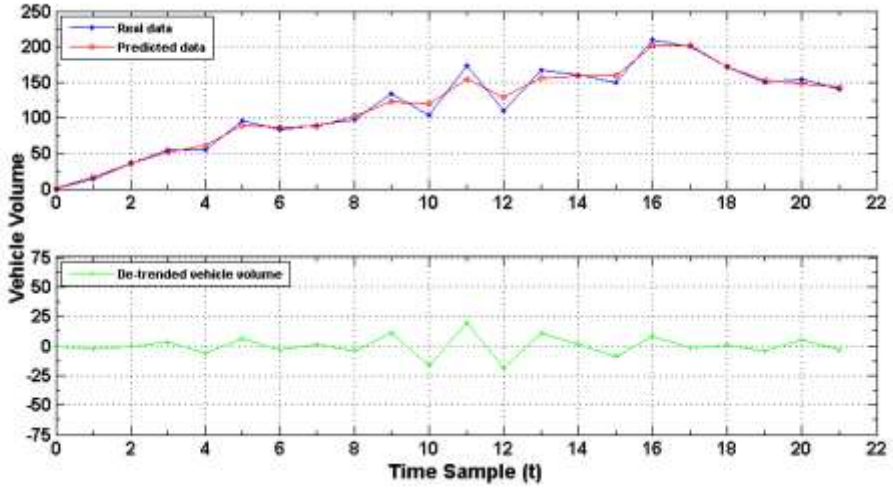

Figure 4.8: Vehicle Flow MICGGM(1,1) Fitting

Close observation of the error curves in Figures 4.2-4.8 reveals that MICGGM(1,1) is the most accurate model in predicting vehicle flow and the OGM(1,1) is the least

accurate. Moreover, from Figures 4.2, 4.3 and 4.4 notice that the de-trended vehicle volume curves look similar meaning that the OGM(1,1), MBVGM(1,1) and MICGM(1,1) have almost the same fitting accuracy. It was expected that the accuracy would be improved after the background and initial conditions are modified. Unfortunately, it was not and this was an indication that MBV and MIC have minimal effect in improving the prediction accuracy of the OGM(1,1). From Figures 4.6, 4.7 and 4.8 the de-trended curves also look similar and are approaching zero level. It means that GGM(1,1), MBVGGM(1,1) and MICGGM(1,1) are more accurate compared to the OGM(1,1), MBVGM(1,1) and MICGM(1,1). Thus the DGT greatly improves the fitting accuracy of the OGM(1,1). Now, it is evident that the proposed DGT is a powerful technique, compared to MBV and MIC, in improving the fitting accuracy of the OGM(1,1). For error evaluation of these grey models see Table 4.4.

II. Testing the Grey Models in Short-Term Forecasting

In this section short-term traffic flow forecasting is carried out, three points into the future, by extrapolating (3.9). Thus, the test data set (i.e. the three points) are forecasted and used to evaluate the performance of the proposed grey models (Liu & Cocea, 2017).

From Table 3.3 the testing data set consists of the last three data points, i.e. 157, 146 and 145. To forecast these three points (4.5), (4.9) and (4.10) are extrapolated three points ahead (see section 4.1.1 part I). Hence short-term traffic flow forecasting. Extrapolation of (4.5), (4.9) and (4.10) gives the forecasts by the OGM(1,1), MBVGM(1,1) and MICGM(1,1) respectively. For GGM(1,1), MBVGGM(1,1) and MICGGM(1,1) the 19 time response functions are also extrapolated to forecast the three points. The forecasted values by the six models are as indicated in Table 4.1. Plotted in Figures 4.9-4.14 are the real, predicted and de-trended vehicle volume curves for the six grey models. The last three-time sample points (i.e. at t=22, t=23 and t=24) are the extrapolated points of focus. The error curves in Figures 4.9, 4.10 and 4.11 shows that the OGM(1,1), MBVGM(1,1) and MICGM(1,1) have almost the same short-term forecasting accuracy. On the other hand from Figures 4.12, 4.13 and 4.14 the GGM(1,1), MBVGGM(1,1) and MICGGM(1,1) have much improved short-term

forecasting accuracy. MBVGGM(1,1) is the most accurate in short-term forecasting whereas OGM(1,1) is the least accurate. This is clear and evident in the next discussion on evaluation of the grey models.

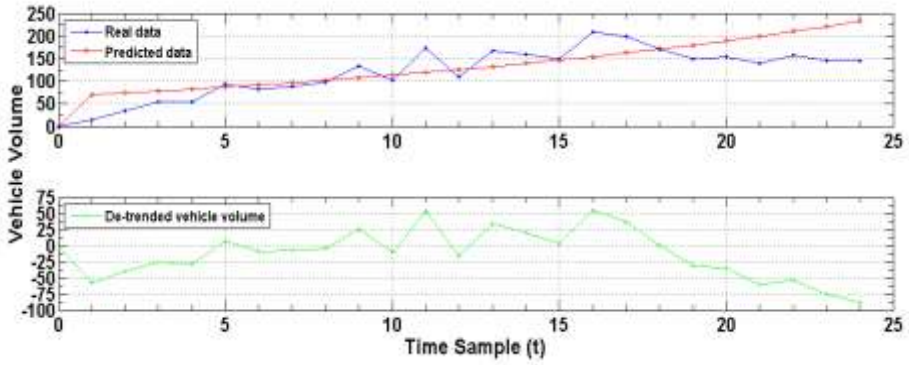

Figure 4.9: Short-Term Vehicle Flow Forecast by OGM(1,1)

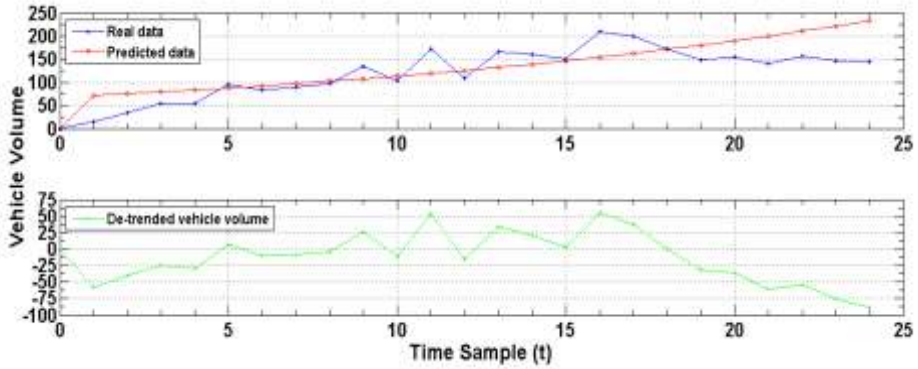

Figure 4.10: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

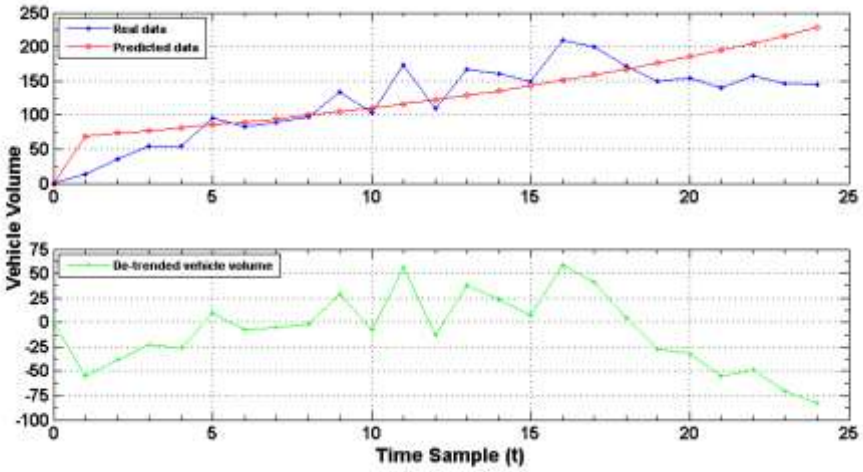

Figure 4.11: Short-Term Vehicle Flow Forecast by MICGM(1,1)

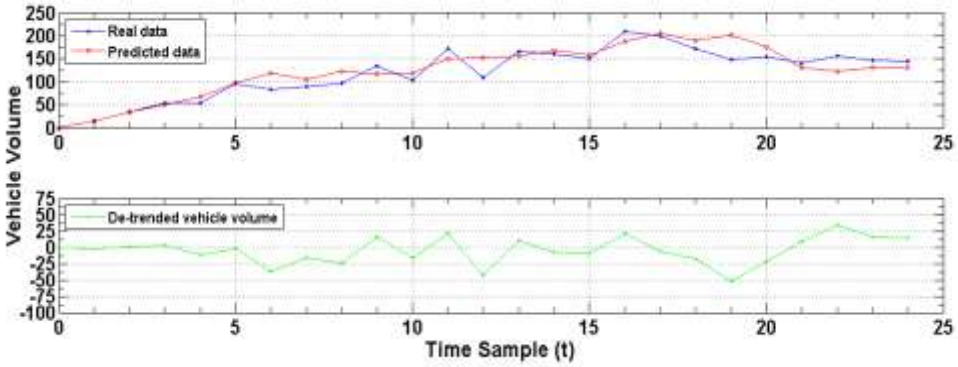

Figure 4.12: Short-Term Vehicle Flow Forecast by GGM(1,1)

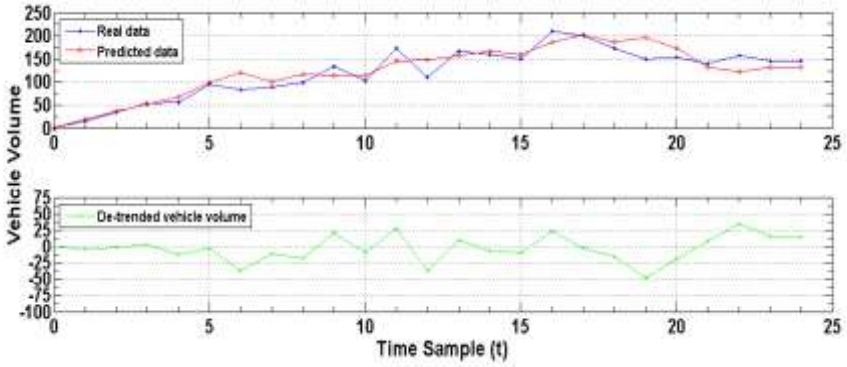

Figure 4.13: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

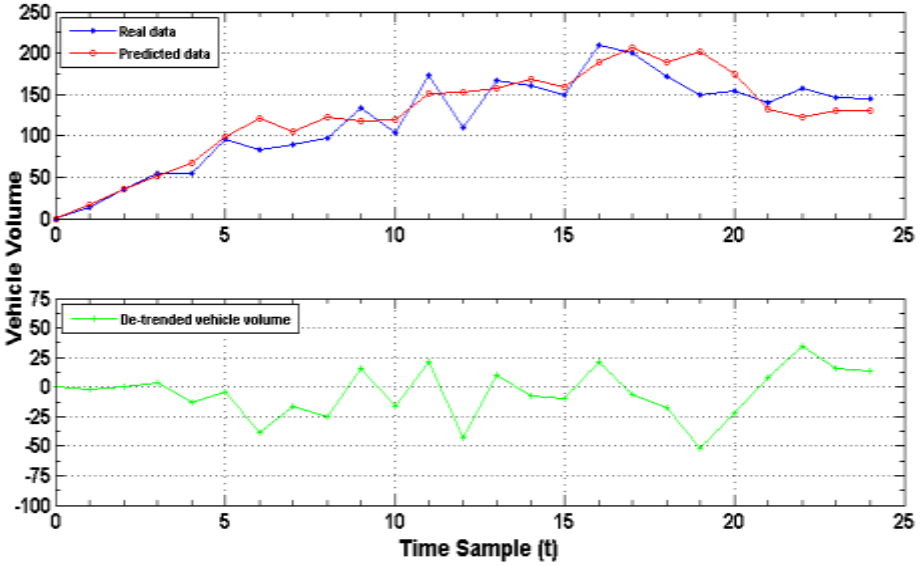

Figure 4.14: Short-Term Vehicle Flow Forecast by MICGGM(1,1)

4.1.2 Evaluation of the Improved GM (1,1) in Traffic Flow Fitting and Forecasting

The fitting and forecasting values for the six grey models were as tabulated in Table 4.1 and the errors involved are indicated in Tables 4.4 and 4.5. The OGM(1,1) and the other five improved grey models are evaluated based on their fitting and forecasting errors. This Error evaluation is done as follows. Using the error indicators discussed in section 3.1.3 and MATLAB the fitting and short-term forecasting errors were calculated and recorded in Tables 4.4 and 4.5 respectively. The corresponding

accuracies were as well recorded in Tables 4.4 and 4.5. From Table 4.4 notice that the OGM(1,1), MBVGM(1,1) and MICGM(1,1) had lower fitting accuracies compared to GGM(1,1), MBVGGM(1,1) and MICGGM(1,1). MICGGM(1,1) was the most accurate and its mean absolute percentage deviation (MAPD) was 5.3013% which translates to an accuracy of 94.6987%. Therefore, DGT is a powerful method in improving the fitting accuracy of the OGM(1,1) as compared to the MBV and MIC methods (see Table 4.4).

In short-term forecasting it is clear that the MBVGGM(1,1) is the most accurate model (see Table 4.5). Consider the MAPD for all the models; the accuracies are 51.2720%, 51.2974%, 54.8814%, 85.5933%, 85.9152% and 85.6143% for the OGM(1,1), MBVGM(1,1), MICGM(1,1), GGM(1,1), MBVGGM(1,1) and MICGGM(1,1) respectively. Thus MBV slightly improves the short-term forecasting accuracy of the GGM(1,1). It means that the MGO by (3.20), which is the background value, is among the factors causing forecasting errors in the OGM(1,1). This is evident from Table 4.5, that after modification of the background value the established MBVGM(1,1) is more accurate in short-term forecasting compared to the OGM(1,1). However, in fitting the OGM(1,1) is slightly more accurate compared to the MBVGM(1,1) (see Table 4.4). Overall, the DGT is a powerful technique for improving the accuracy of the OGM(1,1).

Table 4.4: Vehicle Flow Fitting Error and Accuracy Evaluation

Error/ Accuracy Indicator	Grey Model					
	OGM (1,1)	MBVGM (1,1)	MICGM (1,1)	GGM (1,1)	MBVGGM (1,1)	MICGGM (1,1)
Error						
RMSE	31.9387	31.9513	31.7867	8.4546	9.3349	8.4172
RMSPE	22.1441	22.1060	22.8043	5.6729	6.7574	5.6318
MAE	25.6790	25.7037	25.4717	6.1829	6.9414	6.1327
MAPD	22.1980	22.2193	22.0188	5.3448	6.0004	5.3013
Accuracy						
100-RMSPE	78.8559	78.8940	77.1957	94.3271	93.2426	94.3682
100-MAPD	78.8020	77.7807	77.9812	94.6552	93.9996	94.6987

Table 4.5: Short-Term Vehicle Flow Forecast Error and Accuracy Evaluation

Error/ Accuracy Indicator	Grey Model					
	OGM (1,1)	MBVGM (1,1)	MICGM (1,1)	GGM (1,1)	MBVGGM (1,1)	MICGGM (1,1)
Error						
RMSE	74.1832	74.1411	68.8576	23.4309	23.1984	23.3932
RMSPE	48.8740	48.8469	45.3326	16.1111	15.9802	16.0849
MAE	72.7671	72.7292	67.3772	21.5140	21.0333	21.4827
MAPD	48.7280	48.7026	45.1186	14.4067	14.0848	14.3857
Accuracy						
100-RMSPE	51.1260	51.1531	54.6674	83.8889	84.0198	83.9151
100-MAPD	51.2720	51.2974	54.8814	85.5933	85.9152	85.6143

The criteria of MAPD and RMSPE (Lotfalipour et al., 2013; Zhang et al., 2015) are as tabulated in Table 4.6. Notice that the fitting errors of GGM(1,1), MBVGGM(1,1) and MICGGM(1,1) from Table 4.4 are less than 10%. Thus from Table 4.6 it is clear that the fitting accuracies of GGM(1,1), MBVGGM(1,1) and MICGGM(1,1) are high. Considering RMSPE and MAPD their short-term forecast errors ranges between 10 to 20% as seen in Table 4.5. This is good short-term forecasting accuracy as seen in Table 4.6).

Table 4.6: Criteria for MAPD and RMSPE

MAPD and RMSPE	Forecasting power
Less than 10%	High accuracy
10 to 20%	Good
20 to 50%	Reasonable
More than 50%	Inaccurate

In this section the conventional grey model, OGM(1,1), has been modified to establish improved grey models denoted by MBVGM(1,1), MICGM(1,1), GGM(1,1), MBVGGM(1,1) and MICGGM(1,1). In essence the OGM(1,1)'s fitting and forecasting accuracies have been improved and generally, the GGM(1,1), MBVGGM(1,1) and MICGGM(1,1) have high fitting accuracy and good short-term forecasting accuracy. Thus DGT, MBV and MIC methods have the ability to improve the fitting and forecasting accuracy of the OGM(1,1). However, the DGT has better performance in improving OGM(1,1)'s accuracy compared to the MBV and MIC.

4.1.3 Data Grouping Techniques' Performance Analysis

In the earlier sections of this thesis the prediction accuracy of the conventional GM (1,1) was improved by adopting the DGT, MBV and MIC techniques. For the DGT various data grouping techniques were proposed and in this section these data grouping techniques are graphically analyzed and their performance is evaluated in detail. In particular, the strong and weak grouping techniques are discussed and it is shown that prediction accuracy continues to improve as the number of data groups increases. Therefore, the novelty of this section is that detailed performance analysis of the proposed strong grouping (SG) and weak grouping (WG) techniques is presented and consequently the usefulness of the proposed technique (i.e. the DGT) is precisely discussed and shown. In particular, the connection between data points per group and the number of groups is evaluated. It is shown that the more the number of groups, the more the overlaps and the more the data is smoothed to increase prediction accuracy.

For analyzing the data grouping techniques two types of data were employed, namely vehicle traffic flow and CO_2 emission, as discussed in sub-sections 4.1.3b and 4.1.3c. The vehicle flow and CO_2 emissions data of Table 3.3 was simulated in MATLAB by OGM(1,1) and GGM(1,1). From Table 3.3 the total number of data points is 25.

a) The GGM(1,1) Prediction Method

The GGM(1,1) was established from the conventional GM(1,1) based on the data grouping concept. This DGT involves a lot of calculations. However, with a MATLAB code the calculations are made easier and faster. The forecasting model was deduced and presented by (3.9) and rewritten here as:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} \left(x^{(0)}_{(1)} - \frac{b}{a} \right) e^{-ar} + \frac{b}{a},\ r = 0,1,2,\cdots,m-1 \tag{4.11}$$

Generally, five (5) data grouping techniques are adopted. They include the no grouping (NG) (all data form one group), WG in 4s, WG in 5s, SG in 4s and SG in 5s. And it is confirmed that strongly grouping the training data in 4s is the most accurate method. This work is further evaluated by the performed error evaluation in section 4.1.3d and the tabulated errors of Tables 4.7 and 4.8.

b) Vehicle Flow Forecast

First, considering the NG method and employing traffic volume data of Table 3.3, the OGM(1,1) $\hat{x}^{(1)}_{(r+1)} = 2075.209e^{0.0378r} - 2075.2090$ was obtained, where $a = -0.0378$ and $b = 78.4429$. The 2-h (25 data-points) forecast data set $\hat{X}^{(0)}$ was obtained as:

$$\hat{X}^{(0)} \triangleq \{0.0 \;\; 79.9449 \;\; 83.0259 \;\; 86.2256 \;\; 89.5486 \;\; 92.9997 \;\; 96.5838 \;\; 100.3060 \;\; 104.1717$$
$$108.1863 \;\; 112.3557 \;\; 116.6857 \;\; 121.1827 \;\; 125.8529 \;\; 130.7031 \;\; 135.7402 \;\; 140.9715$$
$$146.4043 \;\; 152.0466 \;\; 157.9063 \;\; 163.9918 \;\; 170.3118 \;\; 176.8754 \;\; 183.6920 \;\; 190.7712 \}$$

(4.12)

The residue time series set $X^{(d)} \triangleq X^{(0)} - \hat{X}^{(0)}$ was obtained as:

$$X^{(d)} \triangleq \{0.0 \;\; -65.9449 \;\; -48.0259 \;\; -32.2256 \;\; -34.5486 \;\; 2.0003 \;\; -13.5838 \;\; -11.3060 \;\; -6.1717$$
$$25.8137 \;\; -9.3557 \;\; 56.3143 \;\; -11.1827 \;\; 41.1471 \;\; 29.2969 \;\; 14.2598 \;\; 69.0285$$
$$53.5957 \;\; 19.9534 \;\; -8.9063 \;\; -9.9918 \;\; -30.3118 \;\; -19.8754 \;\; -37.6920 \;\; -45.7712\}$$

(4.13)

where d stands for de-trended. This is actually the error involved in forecasting, i.e. the residue.

Second, the SG method's forecasts are computed by (3.9). However, the parameters a and b differ from one group to the other. Hence, strongly grouping data in 4s yields 22 data groups (see section 4.1.3d) whose forecast, data sequence $\hat{X}^{(0)}$, is:

$$\hat{X}^{(0)} \triangleq \{0.0\ \ 15.6615\ \ 34.7612\ \ 50.1448\ \ 60.5172\ \ 88.2873\ \ 86.2357\ \ 87.1706\ \ 102.3888$$
$$122.6047\ \ 119.3534\ \ 153.3755\ \ 128.6640\ \ 156.3421\ \ 158.7244\ \ 158.8813\ \ 202.0690$$
$$200.7975\ \ 171.8332\ \ 151.9439\ \ 149.8770\ \ 144.5920\ \ 151.6495\ \ 149.9049\ \ 143.2892\}$$

(4.14)

In this case, the residue data sequence $\hat{X}^{(d)}$ is as follows:

$$X^{(d)} \triangleq \{0.0\ \ -1.6615\ \ 0.2388\ \ 3.8552\ \ -5.5172\ \ 6.7127\ \ -3.2357\ \ 1.8294\ \ -4.3888$$
$$11.3953\ \ -16.3534\ \ 19.6245\ \ -18.6640\ \ 10.6579\ \ 1.2756\ \ -8.8813\ \ 7.9310$$
$$-0.7975\ \ 0.1668\ \ -2.9439\ \ 4.1230\ \ -4.5920\ \ 5.3505\ \ -3.9049\ \ 1.7108\}$$

(4.15)

Figure 4.15 shows the vehicle flow forecasting for the NG, SG and WG methods and their corresponding residues. The residue data is obtained by subtracting the predicted data from the original data. This difference is the error (residue). The residue gives a comparison between original and fitted curves. If the residue is zero it shows 100% prediction accuracy. Otherwise, the accuracy is low. Observing the de-trending (residue) curves in Figure 4.15 shows that the data grouping concept improves GM(1,1)'s precision (see Table 4.7). The SG in 4s (see Figure 4.15) is the most accurate data grouping method because many data groups (22 groups in this case) are formed. Forecasts for SG in 5s (forming 21 data groups) and WG cases are as shown in the same Figure 4.15. Similarly, for the WG method, grouping data in 4s is more accurate than grouping data in 5s (see Figure 4.15). However, strong grouping technique's improvement on the conventional GM(1,1)'s precision is greater than the weak grouping technique's improvement. For the SG in 5s case and the WG cases the data sequence $\hat{X}^{(0)}$ and residue sequence $X^{(d)}$ are not given in this thesis but are easily obtainable through the same methods.

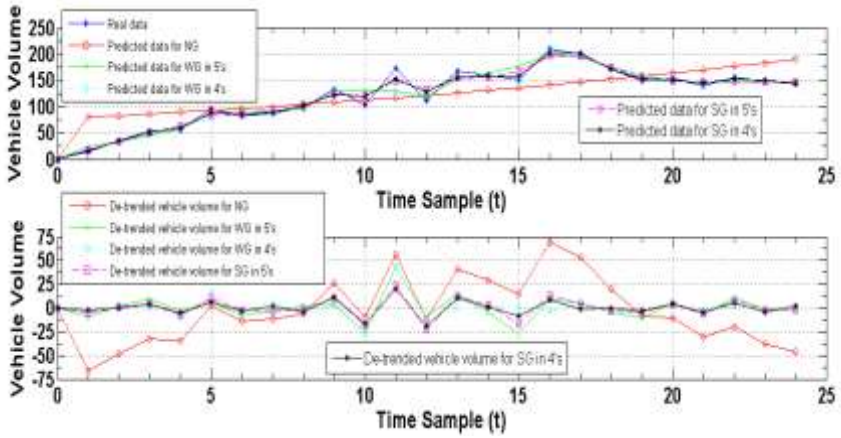

Figure 4.15: Vehicle Flow Forecasting

c) CO_2 Emission Forecast

The CO_2 emissions prediction is based on the model presented by (4.11) and the parameters $\boldsymbol{a}$ and b differ from group to group. However, the application of the AGO and IAGO procedures are alike. The CO_2 emission forecasts are as shown in Figure 4.16. Figure 4.16 also shows the residues for CO_2 emission forecasts. For NG of data all the 25 data points form 1 group (refer to sections 4.1.3d(ii) and 4.1.3d(iii) for more information on data grouping techniques). Comparison of the NG CO_2 emission residues and those of the SG (in Figure 4.16) again indicates that the concept of grouping data plays a great role in improving the precision of the original GM(1,1). Further, from the de-trended curves, it can be observed that strongly grouping data in 4s is more accurate than in 5s. This is because grouping data in 4s yields more data groups (in this case 22 groups) compared to grouping data in 5s (in which case 21 groups are formed). With the WG method (see Figure 4.16) similar situation of accuracy improvement is evident. CO_2 emission forecasting error evaluation in Table 4.8 shows that the concept of DGT in improving the precision of the OGM(1,1) is valid and reliable.

Vehicle volume and CO_2 emissions have been forecasted for two hours and therefore, in Figures 4.15 and 4.16, time sample (t=0) corresponds to 6:00 am and time sample (t=24) corresponds to 8:00 am of the day.

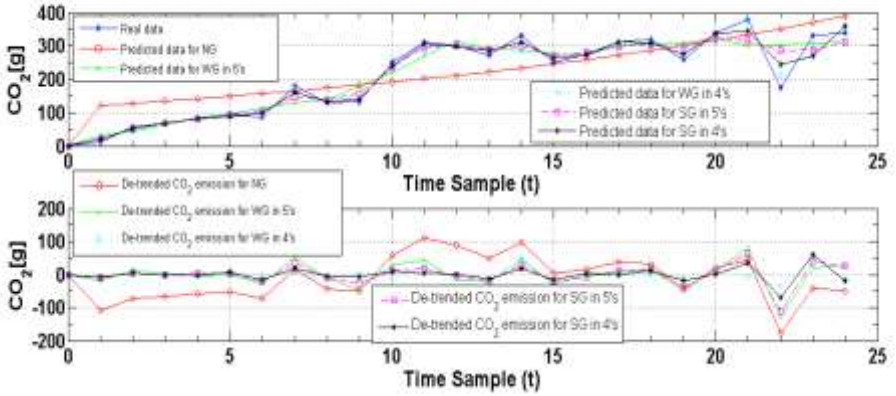

Figure 4.16: CO_2 Emissions Forecasting

d) Error and Performance Evaluation

Discussed here is error evaluation as computed by the various error indicators and the performance analysis of the strong grouping (SG) and weak grouping (WG) techniques.

i) Error Evaluation

The root mean square error (RMSE), root mean square percentage error (RMSPE), mean absolute error (MAE), and the mean absolute percentage deviation (MAPD) error indicators are adopted in analyzing the accuracy of the forecasting. These indicators are computed as given in section 3.1.3.

The computed errors show that GGM(1,1) is of high accuracy than the conventional GM(1,1). Moreover, grouping in 4s is the most accurate DGT (see Tables 4.7 and 4.8). In Tables 4.7 and 4.8 RMSPE and MAPD errors were computed in percentage. The number in curly brackets in Tables 4.7-4.9 indicate the number of data groups formed. The criteria of MAPD and RMSPE are as tabulated before in Table 4.6 (Lotfalipour et al., 2013; Zhang et al., 2015).

Considering the indicator MAPD, the accuracies of conventional GM(1,1) and improved GM(1,1) (GGM(1,1)) in forecasting the traffic parameters are as tabulated in Table 4.9. The accuracy improvement by the data grouping concept is clearly

evident from Table 4.9. Thus, the proposed GGM(1,1) is effective and useful in analyzing, estimating, forecasting and modelling grey systems.

ii) Strong Grouping (SG) Technique Performance Analysis

From the Strong grouping (SG) technique formula (see section 3.1.2a):

$$M = m - [k - 1] \tag{4.16}$$

Table 4.7: Vehicle Flow Forecast Error Evaluation

Error Indicator	Grouping Method				
	In GM (1,1)	In GGM (1,1)			
	NG	WG in		SG in	
	All data {1}	4s {8}	5s {6}	4s {22}	5s {21}
RMSE	34.1121	11.4549	13.2773	8.0205	10.2526
RMSPE	24.0135	8.1548	9.5584	5.3415	7.0417
MAE	27.8521	6.5161	9.1310	5.8325	7.7149
MAPD	23.2644	5.4428	7.6270	4.8718	6.4441

Table 4.8: CO_2 Emissions Forecast Error Evaluation

Error Indicator	Grouping Method				
	In GM (1,1)	In GGM (1,1)			
	NG	WG in		SG in	
	all data {1}	4s {8}	5s {6}	4s {22}	5s {21}
RMSE	67.8673	19.8205	39.7873	22.0666	32.0497
RMSPE	20.4764	7.9574	13.5263	8.1039	11.0460
MAE	56.3588	14.2036	28.1121	13.8517	21.7322
MAPD	26.6528	6.7171	13.2946	6.5506	10.2774

Table 4.9: GM(1,1) and GGM(1,1) Prediction Accuracy

Accuracy Indicator	Traffic Parameter	Grouping Method				
		In GM (1,1)	In GGM (1,1)			
		NG	WG in		SG in	
		all data {1}	4s {8}	5s {6}	4s {22}	5s {21}
100-MAPD	Vehicle volume	76.7356	94.5572	92.3730	95.1282	93.5559
	CO_2 emission	73.3472	93.2829	86.7054	93.4494	89.7226

where $M = N$ is the number of groups, $m = n$ is the total number of data, and k is the number of data points per group; the SG method's performance is graphically presented and it is shown that the more the number of groups the more the prediction accuracy is improved. Figures 4.17 and 4.18 shows the graph of number of groups versus the prediction accuracy by the various error indicators. The graphs are self-explanatory and the blue vertical line indicates the accuracy levels of the OGM(1,1), whose accuracy is improved in this study. With OGM(1,1) all data form one group and therefore, $M = 1,\ k = m = 25$. Observe that the prediction accuracy of the OGM(1,1) without data grouping is low.

Now, when the concept of data grouping is introduced in OGM(1,1), there is significant improvement in the prediction accuracy. This is evident from Figures 4.17 and 4.18. Data was grouped in 4s, 5s, 6s,$\cdots$, to 25s which corresponds to 22, 21, 20, $\cdots$, to 1 group (s), respectively, and the accuracies were computed by error indicators discussed in section 3.1.3. In Figures 4.17 and 4.18 the accuracy levels of GGM(1,1) are as indicated. Notice that the prediction accuracy continues to improve as the number of groups increases from 1 to 22. Consider the MAPD error indicator curve, and notice a very neat trend in accuracy improvement.

From both graphs observe that data grouping in 4s is the most accurate. It involves so many data groups, overlaps and few data points per group. The graphs present the connection between the data points per group and the corresponding number of groups generated. For instance, from Figure 4.17, observe the top and bottom scales of the graph to notice that with 4 data points per group 22 groups are formed. To form 10 groups of data then each group should contain 16 data points. Therefore, the fewer the data points per group, the more the groups formed, the more the overlaps and the high the accuracy of prediction.

It can now be confidently concluded that the data grouping concept in improving OGM(1,1)'s accuracy is worthy in time series smoothing and forecasting.

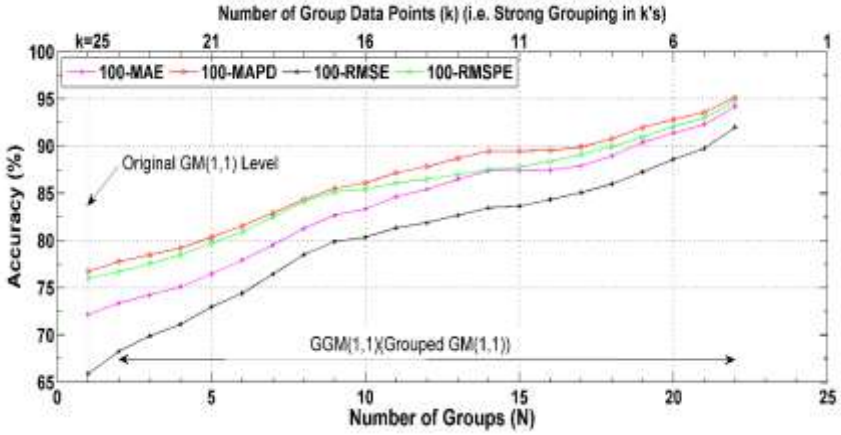

Figure 4.17: Vehicle Volume; GGM (1,1) Strong Grouping Technique's Performance Analysis

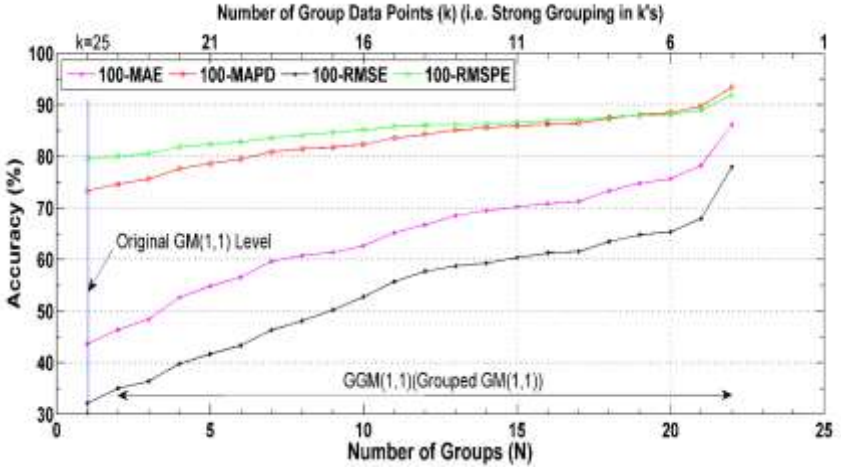

Figure 4.18: CO_2 Emission; GGM (1,1) Strong Grouping Technique's Performance Analysis

iii) Weak Grouping (WG) Technique Performance Analysis

Similarly, the weak grouping technique was analyzed as follows. Data was grouped in 4s, 5s, 7s, 9s and 13s which corresponds to 8, 6, 4, 3, and 2 groups, respectively, and the accuracies were computed by error indicators discussed in section 4.1.3d(i). In Figures 4.19 and 4.20 the accuracy levels of OGM(1,1) are not shown as it forms 1 group consisting of all the 25 data points but GGM (1,1)'s accuracy levels are observed to improve as the number of groups increases from 2 to 8. Unfortunately, the weak

grouping technique results in a fewer number of groups. In this case it was possible to form only 8 groups unlike in strong grouping technique which resulted to 22 groups. Notice again that the more the number of groups the more the accuracy is improved. Thus, the strong grouping technique is more accurate than the weak grouping technique.

From both graphs (i.e., from Figures 4.19 and 4.20) again note that data grouping in 4s is the most accurate. In this case it involves eight groups and four data points per group.

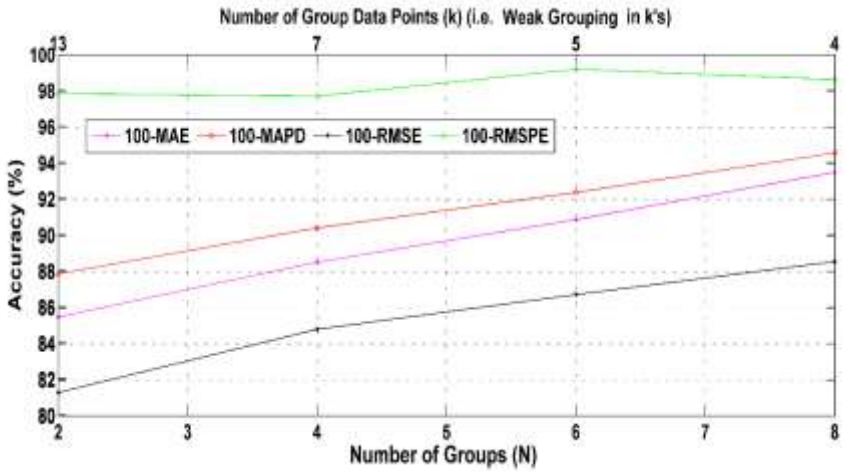

Figure 4.19: Vehicle Volume; GGM(1,1) Weak Grouping Technique's Performance Analysis

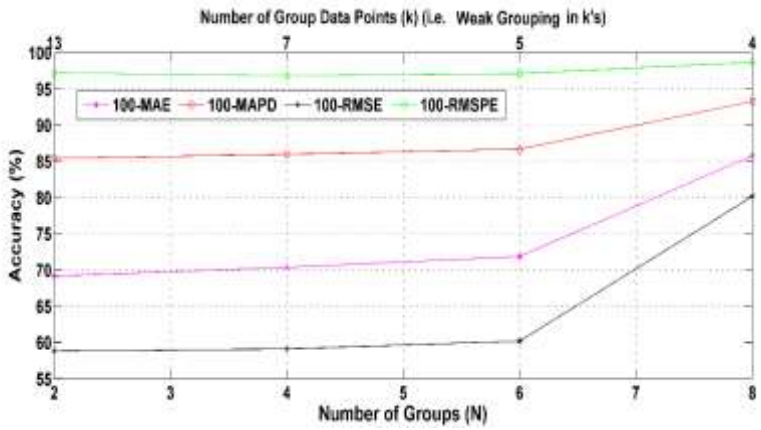

Figure 4.20: CO_2 Emission; GGM(1,1) Weak Grouping Technique's Performance Analysis

Surely, the data grouping concept in OGM(1,1)'s precision improvement is worthy as it has been shown and more especially in SG techniques. The data grouping performance analysis and traffic parameter forecast in this section have shown that GGM(1,1) is applicable and reliable in time series forecasting.

4.2 Assessment of Grouping Technique Based Univariate Grey Model, GM(1,1), on Energy Consumption and Carbon Dioxide Emissions

In order to assess the performance of the DGT in improving the accuracy of the OGM(1,1) two application results are discussed. The DGT is applied in vehicular CO_2 emission and electricity consumption modelling and forecasting. Firstly, the proposed new approach of DGT and the past approach of MIC are applied and compared in improving the accuracy of the original grey model (OGM(1,1)). This is for assessing the DGT's ability in enhancing the OGM(1,1). These techniques have been discussed in sections 3.1.2a and 3.1.2d. Consequently, the GGM(1,1) and MICGM(1,1) are used in modelling and forecasting CO_2 emissions. Vehicular CO_2 emission time-series data is utilized to model and forecast CO_2 emission by these models. Secondly, the GGM(1,1) is applied in modelling and forecasting of electricity consumption.

4.2.1 GGM(1,1) and MICGM in Vehicular CO_2 Emission Modelling and Forecasting

The OGM(1,1)'s precision is improved by, firstly, grouping CO_2 emission data and establishing a GGM(1,1) as before. Secondly, MIC (Madhi & Mohamed, 2017c) is introduced to establish the MICGM(1,1). So, modelling and forecasting CO_2 emission is done by these improved grey models.

a) Data Source

The CO_2 emission data used in this section were as recorded in Table 3.3. The data in Table 3.3, now recorded again in Table 4.10, were used in modelling and forecasting CO_2 emissions and the empirical results are presented in the following paragraphs. The first 22 data points were used in modelling and the remaining three data points were used in forecasting the CO_2 emissions. In other words, the 22 data points were used as historical data and the three as the future data to be forecast.

b) Vehicular CO_2 Emission Modelling Empirical Results

By application of the AGO, IAGO and MGO operations on the original data sequence CO_2 emission is modelled and the OGM(1,1)'s parameters were obtained as a= -0.0692 and b=100.1470. Thus the structure of the OGM(1,1) from (3.9) was found to be:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 1{,}447.2110\, e^{0.0692r} - 1{,}447.2110, \quad r = 0,1,2,\cdots,m-1 \tag{4.17}$$

The fitting values computed from (4.17) were tabulated in Table 4.10. The fitting accuracies based on MAPD calculation was found to be 76.4233%. See Table 4.11. Figure 4.21 shows the OGM(1,1)'s actual, fitted (predicted) and residual curves. The residual is the fitting deviation (model error).

For the MICGM(1,1) the computed parameter values were a= -0.0692 and b=100.1470 and the optimized initial condition value, C= 1.2827e+03. Therefore, the

resulting modified initial condition model structure of the grey model from (3.46) is given as:

$$\hat{x}_{(r)}^{(0)} = 1.2827 * e^{3} * \left(e^{0.0692r} - e^{0.0692(r-1)}\right), \;\; r = 2,3,..m \qquad (4.18)$$

The fitting values computed from (4.18) were tabulated in Table 4.10 and the corresponding fitting accuracies was found to be 78.2432% in terms of MAPD calculation, as in Table 4.11. Figure 4.22 is an indication of how well the actual and predicted values of the MICGM(1,1) fit onto each other and the residual curve.

As stated before for the GGM(1,1) the parameters a and b are different for each formed group and therefore the model structures are also different. Note that in this model and from (3.34) nineteen groups are formed. Hence this will result in nineteen model structures of the GGM(1,1) with nineteen different values of a and b. Hence this model involves a lot of computations but with MATLAB software it is easier to superimpose the simulations and show that this model has excellent accuracy. In this simulation the fitting accuracy obtained was 95.6668% as obtained in terms of MAPD computation, as shown in Table 4.11. In Figure 4.23 notice that GGM(1,1)'s actual and predicted curves fit onto each other with a high accuracy compared with those of OGM(1,1) and MICGM(1,1). Moreover, the residual curve indicates that the fitting deviation is low.

Table 4.10: Original and Improved Grey Models' CO_2 Emission Fitting and Forecasting Values

Raw Data		Traditional	Improved GM(1,1)	
Data Point	**Actual Data**	**OGM(1,1)**	**MICGM(1,1)**	**GGM(1,1)**
Historical Data			**Model Values**	
1	0.0	0.0	0.0	0.0
2	13.62	103.6935	98.4937	19.6619
3	56.06	111.1233	105.5509	51.6320
4	67.88	119.0854	113.1138	69.1628
5	83.08	127.6181	121.2186	83.5470
6	96.13	136.7621	129.9041	88.2273
7	84.74	146.5614	139.2119	98.9877
8	181.12	157.0627	149.1866	162.1046
9	128.72	168.3165	159.8761	132.4549
10	132.58	180.3766	171.3315	140.4709
11	249.10	193.3009	183.6076	240.3059
12	311.93	207.1512	196.7634	307.5334
13	300.76	221.9939	210.8618	298.0975
14	271.66	237.9001	225.9704	285.4907
15	331.52	254.9461	242.1615	312.2848
16	250.40	273.2133	259.5128	264.6029
17	275.48	292.7895	278.1073	272.3879
18	311.09	313.7683	298.0341	311.2325
19	321.01	336.2503	319.3887	309.5854
20	258.03	360.3432	342.2734	281.5715
21	338.53	386.1624	366.7978	318.0301
22	378.92	413.8315	393.0795	384.5836
Future Data			**Forecast Values**	
23	174.47	443.4832	421.2442	335.5412
24	330.33	475.2594	451.4271	449.5600
25	339.23	509.3125	483.7725	663.1573

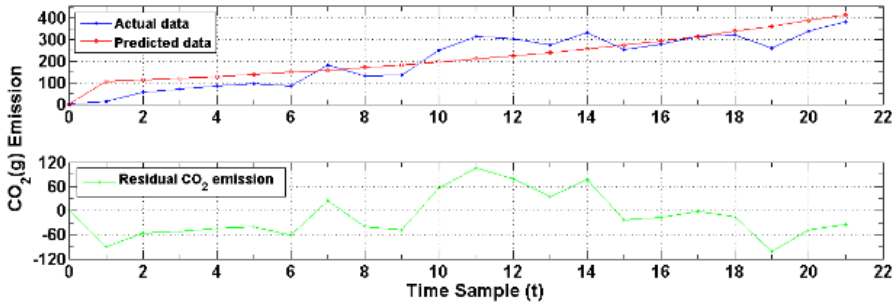

Figure 4.21: OGM(1,1) CO_2 Emission Modelling

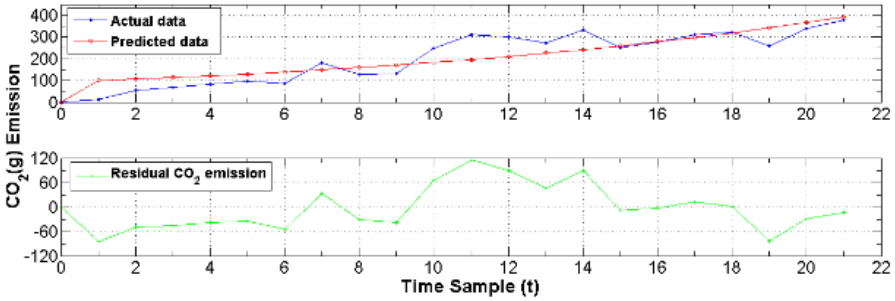

Figure 4.22: MICGM(1,1) CO2 Emission Modelling

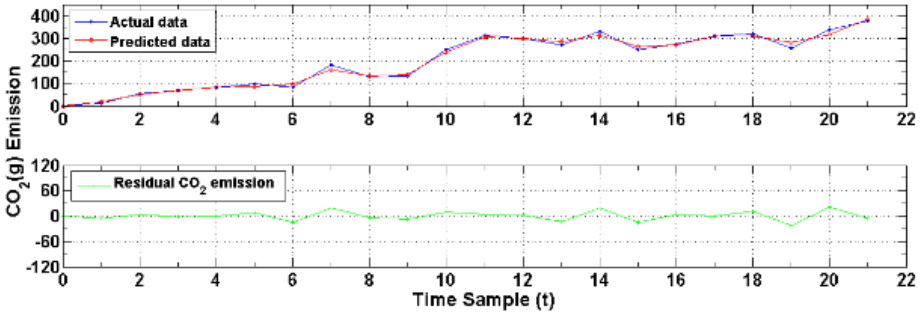

Figure 4.23: GGM(1,1) CO_2 Emission Modelling

c) Vehicular CO_2 Emission Forecasting Empirical Results

The models developed in this section are extrapolated to forecast future trends of CO_2 emissions with an assumption that the models fits a "best curve" to the historical data and that the future will follow that curve. The extrapolation in this context constitutes short-term forecasting of CO_2 emission.

Future forecasts of CO_2 emissions for the original grey model are found by extrapolating the model of (4.17), three points beyond m. Figure 4.24 shows the forecasting at time samples $t = 22, 23, 24$. The future CO_2 emissions were obtained and recorded in Table 4.10. These are the last three values of the OGM(1,1) column and the forecasting accuracy was 69.1399% as per the MAPD error indicator, see Table 4.12.

Similarly, Figure 4.25 shows the forecasting at time samples $t = 22, 23, 24$ as a result of extrapolating the model of (4.18). The three future forecast values as recorded in Table 4.10 for the MICGM(1,1) were obtained at a forecasting accuracy of 72.0238%, based on the MAPD error indicator as seen in Table 4.12.

For the GGM(1,1) all the nineteen models are extrapolated and the simulations superimposed in MATLAB to obtain three forecasts of CO_2 emissions at time samples $t = 22, 23, 24$, as recorded in Table 4.10 and plotted in. Figure 4.26. Here the forecasting accuracy as computed by MAPD error indication was 76.7674%, as shown in Table 4.12.

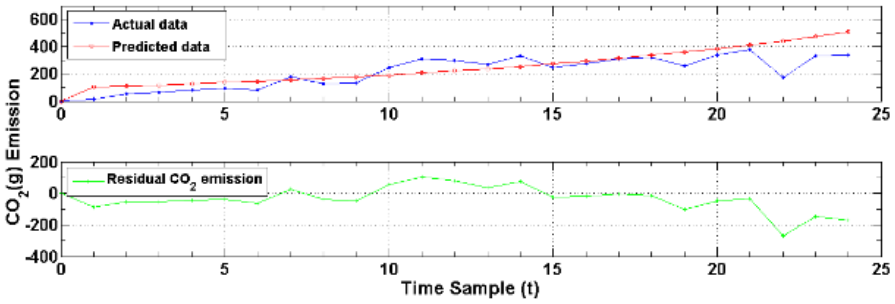

Figure 4.24: OGM(1,1) Future CO_2 Emission Forecasting

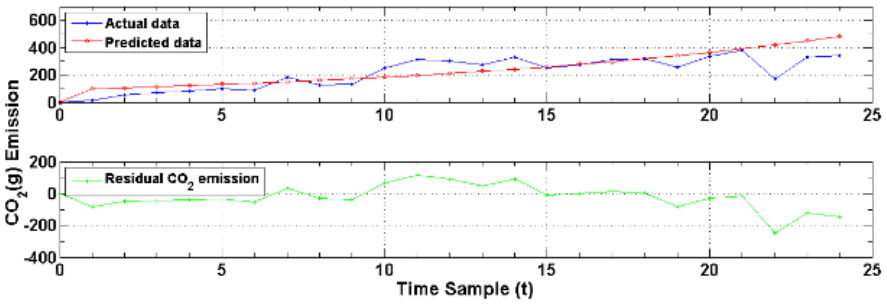

Figure 4.25: MICGM(1,1) Future CO_2 Emission Forecasting

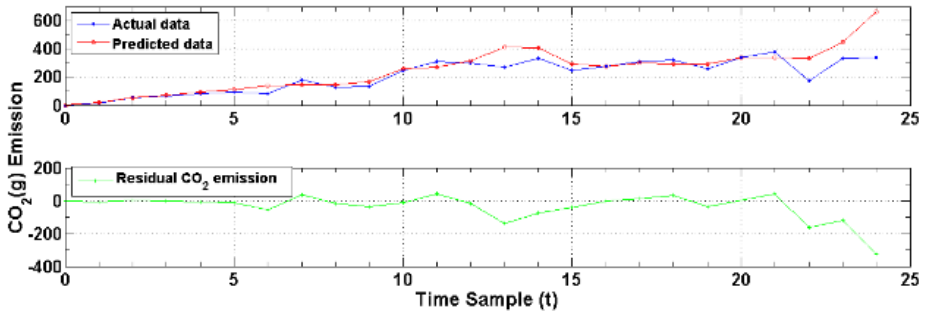

Figure 4.26: GGM(1,1) Future CO_2 Emission Forecasting

d) Error and Accuracy Analysis

The performance of the grey models were evaluated by the MAE and MAPD error indicators and the errors were tabulated in Tables 4.11 and 4.12. Referring to Table 4.6 one can notice that the fitting errors (in terms of MAPD) of both OGM(1,1) and MICGM(1,1) from Table 4.11 are in the range of 20 to 50% and that of the GGM(1,1) is less than 10%. From Table 4.12, GGM(1,1)'s short-term forecast error (in terms of MAPD) ranges between 20 to 50%. Thus the GGM(1,1) has high fitting accuracy with reasonable forecasting accuracy. See the criteria for MAPD in Table 4.6 (Lotfalipour et al., 2013; Zhang et al., 2015).

Table 4.11: CO_2 Emission Modelling Error and Accuracy Evaluation

Error/ Accuracy Indicator	Grey Model		
	Traditional GM(1,1)	Improved GM(1,1)	
	OGM(1,1)	MICGM(1,1)	GGM(1,1)
Error			
MAE	47.6073	43.9326	8.7499
MAPD	23.5767	21.7568	4.3332
Accuracy			
100-MAPD	76.4233	78.2432	95.6668

Table 4.12: CO_2 Emission Forecasting Error and Accuracy Evaluation

Error/ Accuracy Indicator	Grey Model		
	Traditional GM(1,1)	Improved GM(1,1)	
	OGM(1,1)	MICGM(1,1)	GGM(1,1)
Error			
MAE	65.2554	59.1572	49.1267
MAPD	30.8601	27.9762	23.2326
Accuracy			
100-MAPD	69.1399	72.0238	76.7674

Therefore, among the proposed grey models in this section the GGM(1,1) has the highest modelling accuracy with reasonable forecasting accuracy. Thus these results show and validate the claim that the DGT is a reliable technique in improving the precision of the OGM(1,1) as opposed to MIC. The OGM(1,1) has, generally, low accuracy in modelling and forecasting CO_2 emissions.

4.2.2 GGM (1,1), in Modelling Medium-Term Forecasting of Electricity Consumption

From the discussions in sections 4.1.1 and 4.1.2 the GGM(1,1) has emerged as the most accurate technique in time series forecasting. Therefore, in this section the GGM(1,1) was applied in modelling medium-term forecasting of electricity consumption. GGM(1,1) was subjected to electricity consumption data scenario to validate its applicability in time series data forecasting. An analysis of an empirical

example is given. Hence the accuracy of the prediction on electricity consumption in this section of this thesis is improved based on the DGT.

a) The GGM(1,1)

GGM(1,1) is an improved version of the OGM(1,1). The DGT in GGM(1,1) prioritizes new information in a given data series and is useful in reducing the inherent prediction errors in the conventional GM(1,1). So in this section the GGM(1,1) of the strong grouping (SG) technique is used, because of its high accuracy, (see Figure 3.3). In Figure 3.3 seven groups of 4s are formed based on the "new information prior using" principle (Mahdi & Mohamed, 2017b).

Based on this SG technique the electricity consumption data as shown in Table 4.13 can be grouped in 4s as follows. The first group includes data for the years 2000 to 2003 and the last group is composed of data from the year 2016 to 2019. Thus, a total of 17 groups are formed. The conventional GM prediction process is applied on each group separately and the fitted data are superimposed in MATLAB, at points where the groups overlap. The resulting prediction is termed as the GGM(1,1) prediction.

Table 4.13: Kenya's Electricity - Consumption Data (Billion kWh)

Year	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**
Electricity	4.08	4.08	4.43	3.98	3.98	4.34	4.24	5.46	5.12
Year	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
Electricity	4.86	4.86	4.86	5.74	5.52	6.15	6.39	6.63	7.60
Year	**2018**	**2019**							
Electricity	7.86	7.86							

b) Data Source

The data used in this section consists of Kenya's total electricity consumption, expressed in kilowatt-hours (kWh), for the period from the year 2000 to 2019 except for the year 2015 whose value was approximated by computing the mean of the adjacent data. The data were retrieved from the Central Intelligence Agency (CIA) World Factbook(www.indexmundi.com/g/g.aspx?c=ke&v=81) on December 29, 2019 and tabulated as in Table 4.13.

A numerical example is presented by training and testing both the conventional and improved GM(1,1). Thus data of Table 4.13 was subdivided into training and test data sets for estimating and evaluating the GM(1,1)s, respectively (Liu & Cocea, 2017). Consequently, data from the year 2000 to 2016 were used for estimating whereas data from 2017 to 2019 were used for evaluation purposes.

c) Training the Grey Models

In training the grey models the sequence of (3.3) was generated from Table 4.13 data and with (3.20) and (3.23) the matrix A and vector y were formed. Consequently, the parameters a and b were calculated from (3.33). For the OGM(1,1) a= -0.0336, b= 3.6656 and the time function of (3.9) simplified to the model structure:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 113.1752\, e^{0.0336r} - 109.0952, \quad r = 0,1,2,\cdots,m-1 \tag{4.19}$$

To restore the original data IAGO was applied on (4.19) to obtain the fitted data for the years 2000 to 2016 as; {4.0800 3.8674 3.9996 4.1363 4.2777 4.4239 4.5751 4.7315 4.8933 5.0605 5.2335 5.4124 5.5974 5.7887 5.9866 6.1912 6.4028}.

Computation by the GGM(1,1) results to different model structures as each group demonstrates unique values of parameters a and b which may not be necessarily the same. By superimposing the predictions from all the groups, the fitted data for the GGM(1,1) was; {4.0800 4.1459 4.3177 4.0344 4.0095 4.2266 4.4319 5.2753 5.1446 4.9132 4.8449 5.0130 5.5766 5.6517 6.0573 6.4192 6.6306}.

Figure 4.27 shows plots of the real, fitted and fitting deviation data for the OGM(1,1) and the GGM(1,1). The fitting deviation curves demonstrates the model errors in fitting the original data sequence. Note that GGM(1,1)'s errors are lower compared to those of the OGM(1,1). See their corresponding fitting accuracies in Table 4.14.

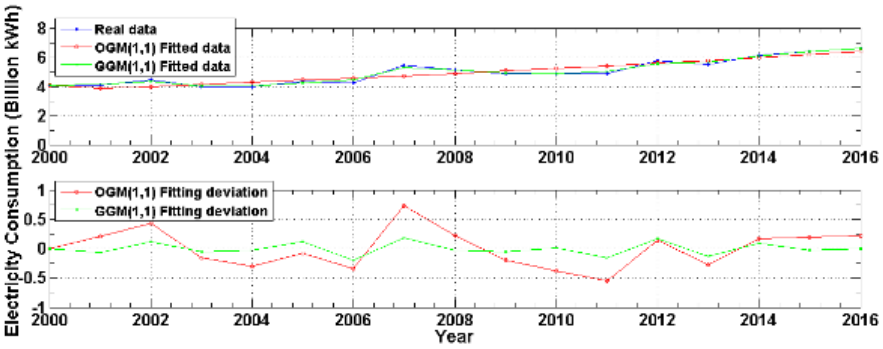

Figure 4.27: Grey Model Training

d) Testing the Grey Models

Testing of the models involved using the constructed models to anticipate the electricity consumption for the years 2017 to 2019. In this respect the OGM(1,1) structure of (4.19) was extrapolated (refer to section 3.1.2(i)) by three points into the future as shown in Figure 4.28. Similarly, extrapolation in GGM(1,1) is illustrated in Figure 4.28. The model simulation values were recorded in Table 4.15 and the performance evaluation accuracies were as given in Table 4.16.

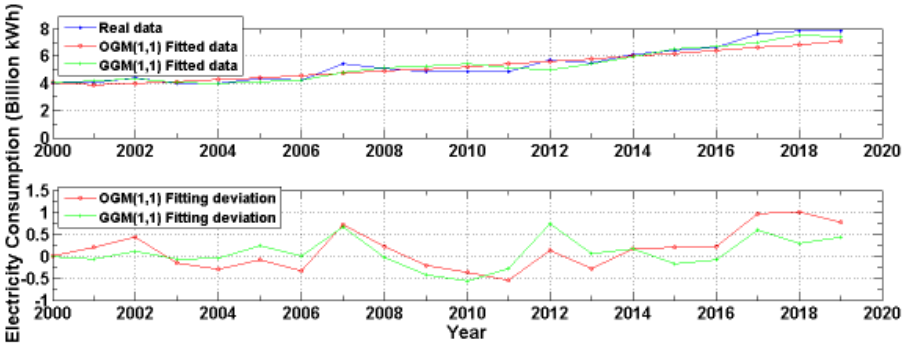

Figure 4.28: Grey Model Extrapolation-Testing

e) Medium-Term Forecasting

Short-term forecasting is daily up to months in the future. It usually involves processes that show results within a year. It can also involve forecasting at predetermined intervals of time say in minutes or hours. On the other hand, Medium-term forecasting tends to be several months up to 2 years into the future. Its period extends from one to

two years (Alvarez et al., 2010) and in this study the electricity consumption for the years 2020 and 2021 was forecasted. From the above discussion it was evident that the GGM(1,1) is more accurate than the OGM(1,1). Consequently, the GGM(1,1) was extrapolated beyond the year 2019 to anticipate the future electricity consumption for the years 2020 and 2021 which were found to be 8.0353 and 8.1023 billion kWh respectively, as read from Figure 4.29.

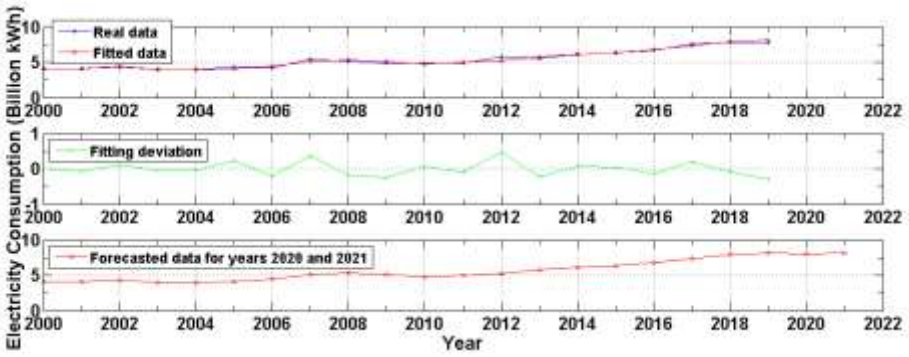

Figure 4.29: GGM (1,1) in Medium-Term Forecasting

It is worthy to note that from CIA the consumption for the year 2020 was 7.86 billion KWh. The prediction by this thesis was 8.0353 Billion KWh. Thus, there was a variance. This variance can be attributed to the fact that during the year 2020, the coronavirus disease 2019 (COVID 19) had struck most parts of the world, Kenya inclusive. So that the rate of electricity consumption in Kenya reduced due to the situation that most of the electricity consuming economic activities were not/not fully in operation. Otherwise, the consumption could reach 8.0353 billion KWh (or near about) as predicted.

f) Accuracy Evaluation

The error measures mentioned in section 3.1.3 were computed and their corresponding accuracies recorded in Tables 4.14 and 4.16. From Table 4.6 notice that both models have high accuracy. However, the GGM(1,1) is slightly more accurate than the OGM(1,1) in both fitting and forecasting as seen in Tables 4.14 and 16.

Table 4.14: Model Training Accuracy Evaluation

Accuracy Indicator	Traditional GM	Grouped GM
100-RMSE	99.6801	99.8956
100-RMSPE	99.5481	99.9730
100-MAE	99.7295	99.9167
100-MAPD	94.5723	98.3270

Table 4.15: Model Evaluation

Test data		Model Data	
YEAR	Raw data	OGM(1,1)	GGM(1,1)
2017	7.60	6.6217	7.0104
2018	7.86	6.8480	7.5712
2019	7.86	7.0821	7.4213

Table 4.16: Model Testing Accuracy Evaluation

Accuracy Indicator	Traditional GM	Grouped GM
100-RMSE	99.5349	99.6572
100-RMSPE	96.1884	97.9596
100-MAE	99.6317	99.7492
100-MAPD	93.1817	95.3579

Therefore, it has been shown that for both OGM(1,1) and GGM(1,1) the fitting and forecasting errors are below 10% and this shows high accuracy (see Tables 4.14 and 4.16). However, the OGM(1,1)'s accuracy has been slightly improved by the DGT and thus the GGM(1,1) is suitable, reliable and valid in electricity consumption modelling and forecasting. This validates GGM(1,1) 'a applicability in time series modelling and forecasting.

4.3 Case Study

In the previous sections of this thesis various case scenarios have been presented, based on a variety of time series data for the purpose of validating the proposed DGT method in improving the precision of the original grey models. In this section a case study is presented to further validate the effectiveness and applicability of the proposed DGT method in modelling and short-term forecasting of vehicle traffic flow. The criteria used in determining the accuracy improvement include the Root Mean Square Error

(RMSE), Root Mean Square Percentage Error (RMSPE), Mean Absolute Error (MAE) and the Mean Absolute Percentage Deviation (MAPD), which are calculated as before. Additionally, in this case study the TSA, discussed in section 3.2.7 of this thesis, was applied in improving the precision of the GM(1,*n*).

4.3.1 Data Source

The case study is based on traffic flow data which were manually collected from Nairobi CBD, Kenya. In order to capture hourly pattern of data that accurately reflects the real-world traffic situation in the CBD area the traffic count was conducted throughout a day. The collected data included counting the number of vehicles (VEHs), pedestrians (PEDs) and motorcycles (MOTs) passing a point of study. The data were collected for three consecutive days, from 16th to 18th February 2021, from seven major counting sites within the CBD. Usually, days counted include Tuesdays, Wednesdays and Thursdays. These are days that are likely to have the same pattern of vehicle, pedestrian and motorcycle volumes within a week and these days are representative enough for one week (Roads and Highways Department, 2001). Mondays and Fridays have different volume patterns (Regehr et al., 2015) and higher volumes are expected on weekends. Fridays and Saturdays are usually different and have higher volumes due to traditional activities like “after works”, recreational and due to the night life (Sampson, 2017). However, these days cause a bias on the volume of traffic (e.g. pedestrians) in comparison to the remaining days of the week; consequently it is better to count in other days rather than these (Miranda et al., 2011). Therefore, data were collected at intervals of 5 minutes for a duration of 15 hours per day for three consecutive days, i.e. from Tuesday to Thursday. For each day data were collected as from 06:00 to 21:00 in all the directions of a road intersection. The traffic data were collected from the seven sites mentioned in section 3.5.4. In this thesis day one traffic data corresponds to data collected on 16th February, day two traffic data is data collected on 17th February, whereas day three traffic data is data collected on the last day of data collection. For economic reasons the manual count method was adopted in this research. In Appendix III shown are some photos of clerks collecting traffic data at the stated sites.

From the photos in Appendix III, it is clear that the weather conditions were favorable for the traffic counting exercise. The Nairobi CBD area was calm and sunny throughout the three days and no incident occurred. Therefore, the traffic counting activity was conducted smoothly and successfully.

4.3.2 The Collected Traffic Data

For empirical analysis and for clear visual interpretation of the traffic data, the data were tabulated in table form as in Appendix IV for each site. In Appendix IV are examples of the traffic flow data sets collected from seven different sites for the three days of the week and at 5-minute interval. Data shown is from 6:00 am to 10:30 am of the day. This is to avoid making this thesis so voluminous. Full day traffic flow data are available upon request from the author. Blank columns in the tabulated traffic data of Appendix IV means that data were not well or fully captured during that time. This was because of failure of the clerk to report in time or the concerned clerk was completely absent due to sickness or any other unavoidable circumstance. However, the collected data is a good representative of the whole CBD traffic system.

Based on these data the prediction and forecasting accuracies of the original and improved grey models were compared. In this section the case study was considered under two headings, namely "Formulating GM(1,1), in modelling and short-term forecasting of vehicle traffic flow" and "Formulating the multivariate grey model, GM(1,n), in vehicle flow modelling". These models' precisions have been improved in the previous sections and in this case study, sections 4.3.3 and 4.3.4 presents more results for the conventional and the proposed improved grey models.

4.3.3 Formulating GM(1,1) in Modelling and Short-Term Forecasting of Vehicle Traffic Flow

Vehicle traffic flow was modelled based on the OGM(1,1) and the proposed MBVGM(1,1), GGM(1,1) and MBVGGM(1,1). The structure of GM(1,1) has been provided in section 3.1.1 part I and the procedure of the proposed grey models are discussed in section 3.1.2 of this thesis. Now from existing results (Ajiboye, et al., 2015) a model trained with the largest size of training sets is the most accurate and it

consistently delivers much better and stable results. Thus, in this thesis a large portion of the collected vehicle volume time series data is used to train these models and the near future data of that time series is used to test the models. The collected vehicle traffic data used were 30 data points and the last 3 points of this time series data was used for testing purposes. That is to say the first 27 data points were used for fitting and the last 3 data points were used for testing purpose. Thus, the grey models were extrapolated three points into the future in order to forecast the last 3 data points. Hence short-term forecasting. As data were collected for three consecutive days sampling was done so that four cases are presented per day in sections 4.3.3.1, 4.3.3.2 and 4.3.3.3. This was for the purpose of increasing the generalizing capability of the models and allow for the models to be most accurate and reliable. Additionally, for each case one direction of traffic flow was randomly selected and its performance evaluated.

4.3.3.1 Day One Vehicle Traffic Flow Modelling and Short-Term Forecasting

Based on vehicle traffic flow data collected on 16th February 2021, four cases are presented in modelling and short-term forecasting of vehicle traffic flow for four sites. These sites include;

- Day 1 Site 1 (D1S1): Haile Selassie Roundabout,
- Day 1 Site 4 (D1S4): Kenyatta Avenue-Moi Avenue-Mondlane Street Junction,
- Day 1 Site 5 (D1S5): Moi Avenue-Slip Road Junction and
- Day 1 Site 7 (D1S7): Haile Selassie Avenue-Moi Avenue Roundabout.

Consequently, data from the said four sites were employed. Note that the data used is from 6:00 am to 8:30 am of the day and these constitutes 30 data points. The data were sub-divided as in Figure 4.30 such that the first 27 data points are for training the models and the last 3 points were used for testing the models (Ajiboye et al., 2015). These four cases are presented in Figures 4.31, 4.33, 4.35 and 4.37 which illustrates the training of the conventional GM(1,1) and the improved GM(1,1), (GGM(1,1)) whereas Figures 4.32, 4.34, 4.36 and 4.38 presents the short-term forecasting (i.e. forecasting of the last 3 data points through grey model extrapolation) by these models.

In both cases of training and forecasting notice that the proposed GGM(1,1) has good performance compared to OGM(1,1) and since it performs well in the four cases of different sites of the Nairobi CBD from which data were collected it is evident that this model is portable and applicable in vehicle flow modelling and forecasting.

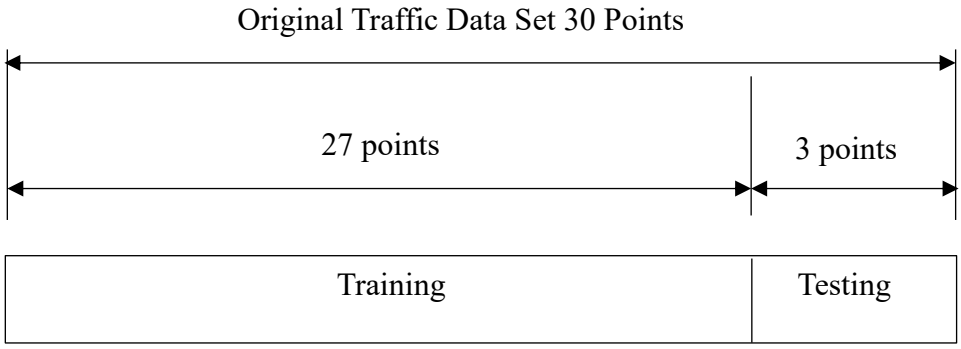

Figure 4.30: Two-Way 30-Point Data Splitting

I. Site 1: Haile Selassie Roundabout

a) Traffic Flow Training

The Haile Selassie roundabout is a four-way type of junction. In this study the northward direction was randomly selected and its results are presented in detail. The AGO, MGO and IAGO operations were applied on the first 27 data points (of Appendix IV Table 1) of each direction and the model parameters a and b were computed and tabulated in Table 4.17.

Table 4.17: Day 1 Site 1 OGM(1,1) Model Parameters

Traffic Flow Direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0153	144.4327
Southward	-0.0207	75.2663
Eastward	-0.0156	57.9234
Westward	-0.0072	85.0393

The time response function of (3.9) for the OGM(1,1) in the northward direction was obtained as:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 9{,}500.0458\, e^{0.0153r} - 9{,}440.0458,\ r = 0,1,2,\cdots,m-1 \tag{4.20}$$

where m=27. From (4.20) the simulated vehicle flow data were recorded in Table 4.19. The real and simulated data for the OGM(1,1) in all directions were as plotted in Figure 4.31a. Similarly, the time response functions for the other three directions can be obtained by substituting the model parameters of Table 4.17 into (3.9).

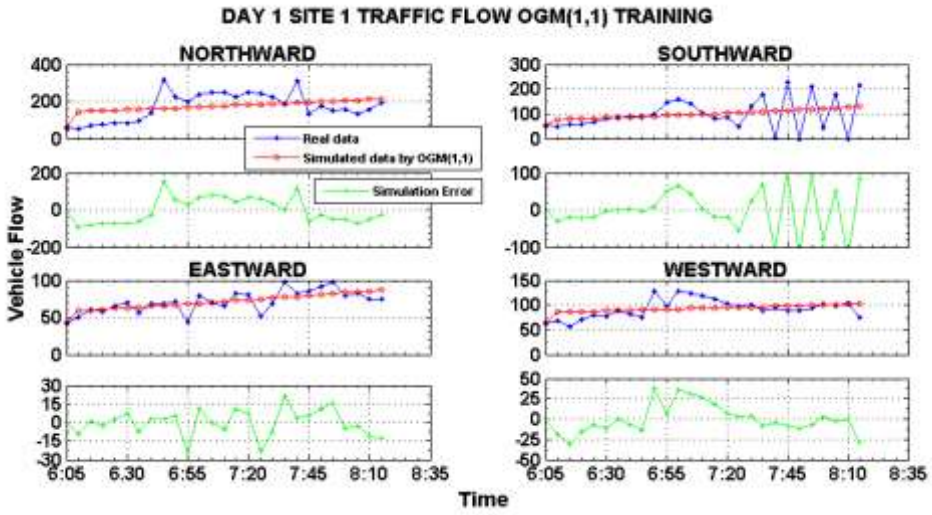

Figure 4.31a: Vehicle Flow OGM(1,1) Training

Modifying the background value resulted to model parameters indicated in Table 4.18. Thus the time response function of (3.9) for the MBVGM(1,1) in the northward direction was obtained as:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 9{,}576.8355\, e^{0.0152r} - 9{,}516.8355,\ r = 0,1,2,\cdots,m-1 \tag{4.21}$$

Table 4.18: Day 1 Site 1 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0152	144.6559
Southward	-0.0192	77.3774
Eastward	-0.0156	57.9595
Westward	-0.0071	85.0742

The final fitting values from (4.21) were recorded in Table 4.19 and Figure 4.31b shows the real and simulated data plots.

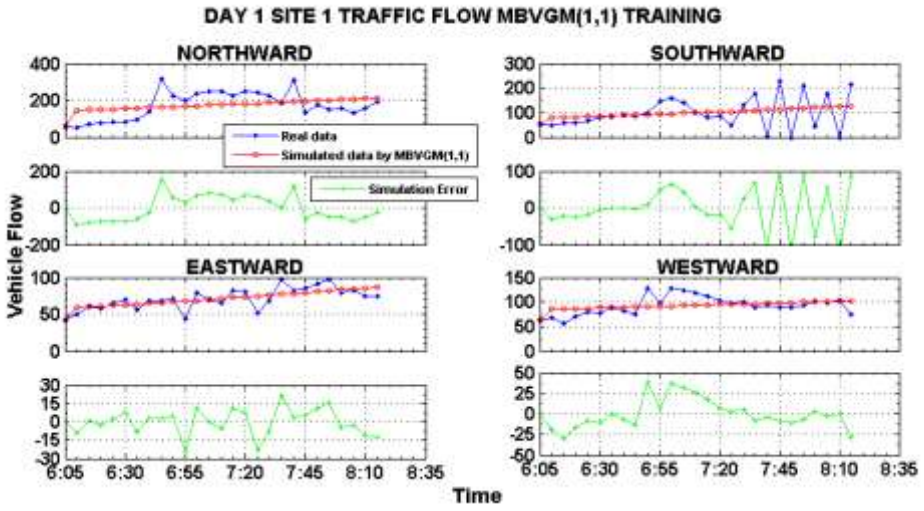

Figure 4.31b: Vehicle Flow MBVGM(1,1) Training

Introduction of the DGT in OGM(1,1) and MBVGM(1,1) results to many groups being formed and hence many time response functions are formulated. In this case with 27 data points 24 groups of data are formed. This means that 24 time response functions are formulated each with unique values of the parameters a and b. Therefore, for GGM(1,1) and MBVGGM(1,1) the model parameters and the corresponding time response functions are not provided here. Figures 4.31c and 4.31d shows the plots of real and simulate data for the GGM(1,1) and MBVGGM(1,1) respectively. The simulated data for these models were also recorded in Table 4.19.

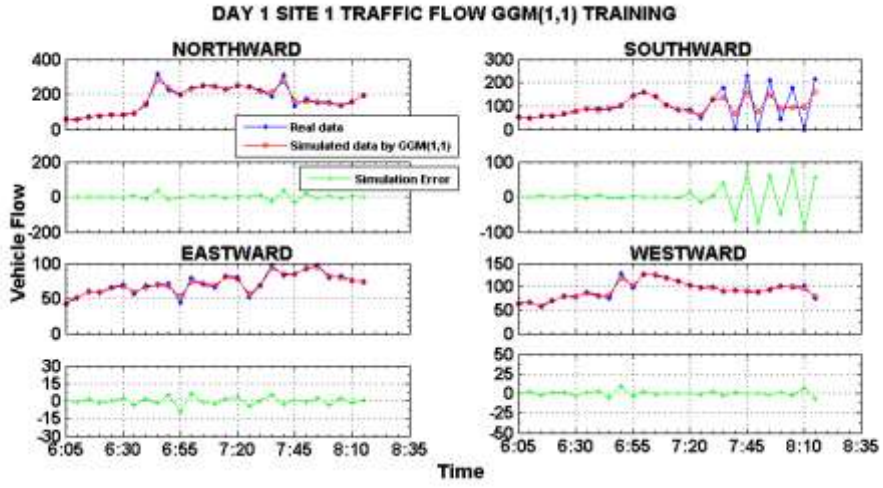

Figure 4.31c: Vehicle Flow GGM(1,1) Training

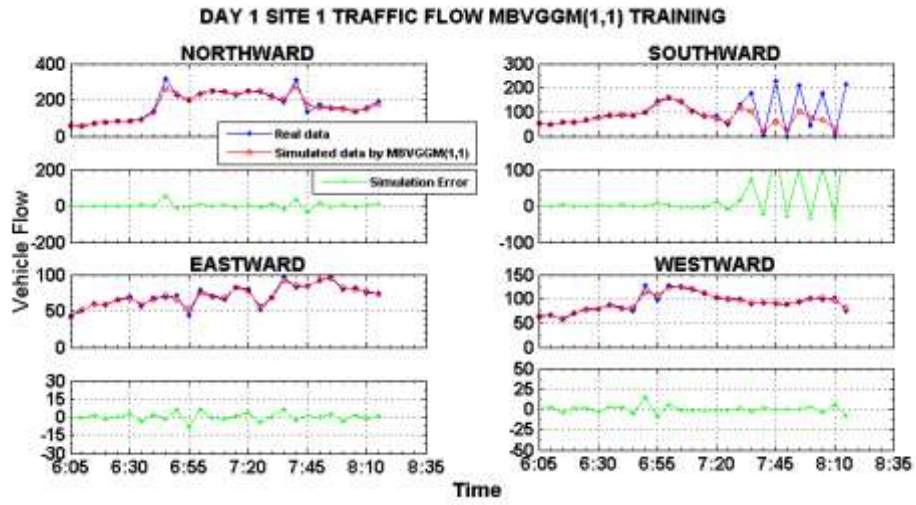

Figure 4.31d: Vehicle Flow MBVGGM(1,1) Training

The simulation errors in Figures 4.31a to 4.31d shows that the OGM(1,1) and MBVGM(1,1) have almost the same fitting accuracy whereas the GGM(1,1) and MBVGGM(1,1) have also almost the same fitting accuracy. However, the GGM(1,1) and MBVGGM(1,1) are more accurate than the OGM(1,1) and MBVGM(1,1). It implies that the DGT has a great impact in improving the fitting accuracy of the OGM(1,1). The accuracy improvement by the MBV is very minimal. Avery poor prediction performance is noticed at the last data point in Figure 4.31d, southward direction. An error of -22.4995 at the final point by MBVGGM(1,1) predicted

occurred. Generally, traffic flow fitting by MBVGGM(1,1) in the said direction was not good.

b) Testing the Grey Models in Short-Term Forecasting

For short-term forecasting the formulated time response functions were extrapolated three points into the future (i.e. the last 3 data points were estimated). For the OGM(1,1) and MBVGM(1,1) (4.20) and (4.21) were extrapolated. In short-term forecasting the parameters a and b remain the same as those obtained in training the models. Figures 4.32 a-4.32d shows the short-term forecasting by the conventional and improved grey models.

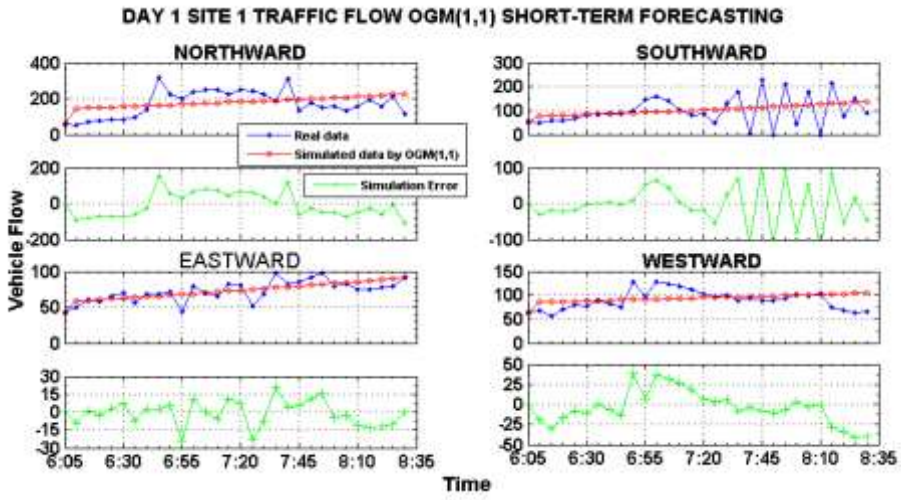

Figure 4.32a: Short-Term Vehicle Flow Forecast by OGM(1,1)

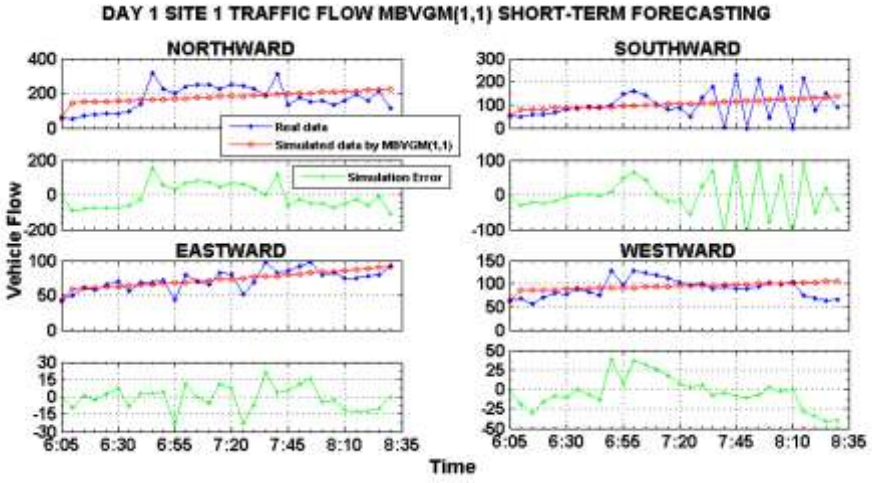

Figure 4.32b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

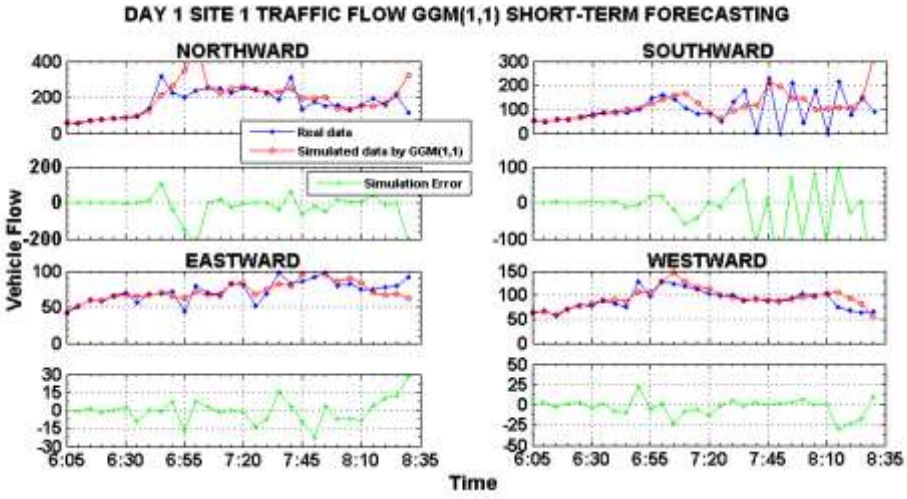

Figure 4.32c: Short-Term Vehicle Flow Forecast by GGM(1,1)

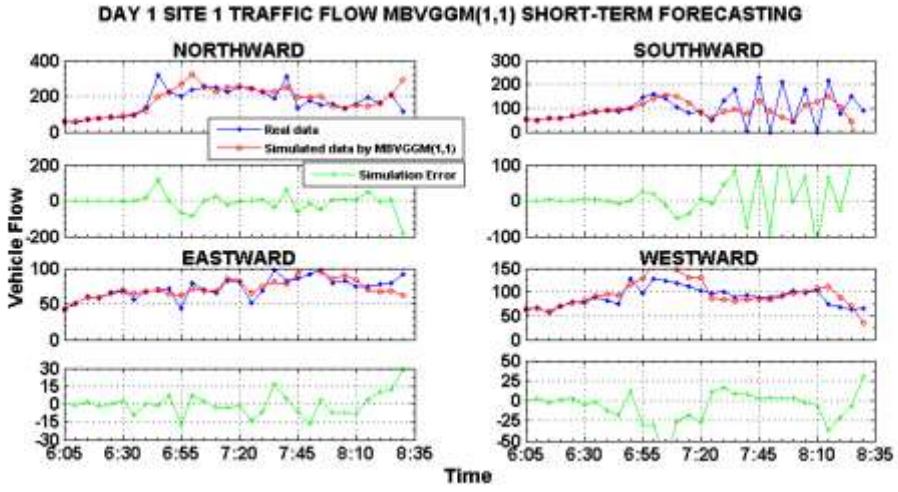

Figure 4.32d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

Comparison of the simulation errors in Figure 4.32a to Figure 4.32d reveals that the forecasting errors are a bit large. However, the simulation error curves in the northward direction indicates that MBVGGM(1,1) is the most accurate. Also the simulation error curves in the westward direction indicates that the GGM(1,1) is more accurate. It again shows that the DGT is good in improving the accuracy of the OGM(1,1).

c) Evaluation of the Grey Models

The final day 1 site 1 traffic flow fitting and forecasting values (for the northward direction only) were obtained and recorded in Table 4.19. Moreover, the overall training and short-term forecasting errors for this site were computed and tabulated in Table 4.20.

Table 4.19: Original and Improved Grey Models' Simulation Values (D1S1 Northward Direction)

Raw Data		Grey Model			
Data Point	Real Value	OGM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training		Fitted Values			
1	60	60.0000	60.0000	60.0000	60.0000
2	55	146.4728	146.6851	55.8896	55.9144
3	70	148.7361	148.9375	69.0092	68.8541
4	77	151.0343	151.2246	77.3329	77.2115
5	82	153.3680	153.5467	80.8203	80.3928
6	81	155.7377	155.9046	81.5486	80.5066
7	97	158.1441	158.2986	90.8216	87.9868
8	137	160.5877	160.7294	150.0240	134.5100
9	319	163.0690	163.1975	280.2892	262.0274
10	223	165.5887	165.7035	234.4838	232.4074
11	198	168.1473	168.2480	201.7108	199.3349
12	238	170.7454	170.8316	233.5593	229.9553
13	251	173.3837	173.4548	251.4464	249.4446
14	247	176.0627	176.1184	243.5463	243.8568
15	222	178.7832	178.8228	229.2007	228.7993
16	252	181.5456	181.5687	246.9259	247.1138
17	244	184.3508	184.3569	245.3689	248.4210
18	226	187.1993	187.1878	216.7087	214.6110
19	189	190.0918	190.0622	213.3698	205.4722
20	312	193.0290	192.9807	274.7988	275.2537
21	134	196.0116	195.9441	166.0507	172.7909
22	176	199.0403	198.9530	158.1449	159.9347
23	151	202.1157	202.0080	156.7823	157.7140
24	154	205.2387	205.1100	149.0264	149.5950
25	135	208.4100	208.2596	140.9255	138.6573
26	160	211.6302	211.4576	156.4693	151.4517
27	192	214.9002	214.7047	191.0012	177.8547
Testing		Short-Term Forecasted Values			
28	159	218.2208	218.0017	168.5547	163.2340
29	215	221.5926	221.3492	217.2907	207.0257
30	112	225.0166	224.7482	324.2162	294.1527

Table 4.19 shows that GGM(1,1)'s fitting values are close to the real data as compared to the other models. And therefore, GGM (1,1) was the most accurate improved grey model. From Table 4.20 it can be clearly seen that GGM(1,1) had less than 10% fitting error in the northbound, eastbound and westbound directions. This is an indication of high fitting accuracies as per the criteria for MAPD and RMSPE shown in Table 4.6.

The fitting error in the southbound direction was in the range of 20 to 50% (see Table 4.20). Based on the criteria in Table 4.6, this is a reasonable accuracy.

Table 4.20: Day 1 Site 1 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional	Improved		
		GM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training					
Northbound	RMSE	67.5900	67.5913	14.2515	16.6227
	RMSPE	34.1376	34.1186	8.1721	9.9934
	MAE	59.3191	59.3094	8.9264	9.9768
	MAPD	34.2079	34.2024	5.1476	5.7534
Southbound	RMSE	60.8495	60.8640	39.2054	65.6346
	RMSPE	40.9329	41.1642	27.3430	67.1580
	MAE	45.7421	45.9943	24.2648	32.4198
	MAPD	46.1523	46.4068	24.4824	32.7105
Eastbound	RMSE	10.3898	10.3898	3.3926	3.4657
	RMSPE	13.4687	13.4672	4.1655	4.4015
	MAE	8.1478	8.1442	2.6427	2.6597
	MAPD	11.4519	11.4469	3.7143	3.7382
Westbound	RMSE	17.2639	17.2639	3.2907	4.4726
	RMSPE	19.1635	19.1599	3.5312	5.0760
	MAE	12.9414	12.9428	2.4323	3.0431
	MAPD	13.9488	13.9503	2.6216	3.2800
Short-Term Forecasting					
Northbound	RMSE	68.2325	68.2118	72.0811	49.8643
	RMSPE	33.9027	33.8805	34.5042	23.4735
	MAE	59.3482	59.3151	38.0688	28.3097
	MAPD	34.4513	34.4322	22.0988	16.4336
Southbound	RMSE	59.2578	59.1766	73.2609	70.4907
	RMSPE	40.2786	40.4874	36.6983	49.0172
	MAE	45.0665	45.2239	45.2428	44.7315
	MAPD	45.1267	45.2843	45.3032	44.7912
Eastbound	RMSE	10.2457	10.2434	9.9066	9.4550
	RMSPE	13.0247	13.0200	13.9423	13.2000
	MAE	8.0665	8.0599	7.0535	6.8747
	MAPD	11.1570	11.1479	9.7559	9.5086
Westbound	RMSE	20.4125	20.4055	10.8469	21.0894
	RMSPE	20.8070	20.8004	11.1116	24.8212
	MAE	15.4888	15.4863	7.1965	14.2263
	MAPD	17.1971	17.1943	7.9902	15.7953

In short-term forecasting, the last three time sample points in Figures 4.32a - 4.32d (i.e. at 8:20, 8:25 and 8:30 AM) are the extrapolated points of focus. For the

northbound direction the forecasted values are recorded in Table 4.19. The errors involved in this short-term forecast are indicated in Table 4.20. The GGM(1,1) is the most accurate in fitting because its fitting errors are lower compared to those of the other three models. In the westbound direction the GGM(1,1) emerged the most accurate in short-term forecasting. It had a MAPD of 7.9902 which translates to an accuracy of 92.0098%.

II. Site 4: Kenyatta Avenue-Moi Avenue-Mondlane Street Junction

a) Traffic Flow Training

Vehicle flow data of site 4 (data of Appendix IV Table 4) in all directions were considered and Figures 4.33a-4.33d and 4.34a-4.34d shows the grey model training and short-term forecasting, respectively.

Table 4.21 shows the OGM(1,1)'s model parameters for this site. Note that site 4 is a three-way junction.

Table 4.21: Day 1 Site 4 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0190	62.0909
Southward	-0.0203	22.3530
Eastward	-0.0320	9.0775

Thus the time response function of (3.9) for the OGM(1,1) in the southward direction was obtained as:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 1{,}120.1330\, e^{0.0203r} - 1{,}101.1330,\ r = 0,1,2,\cdots,m-1 \qquad (4.22)$$

From (4.22) OGM(1,1)'s simulation values were tabulated in Table 4.23 and Figure 4.33a is the plot of the real and the simulation data.

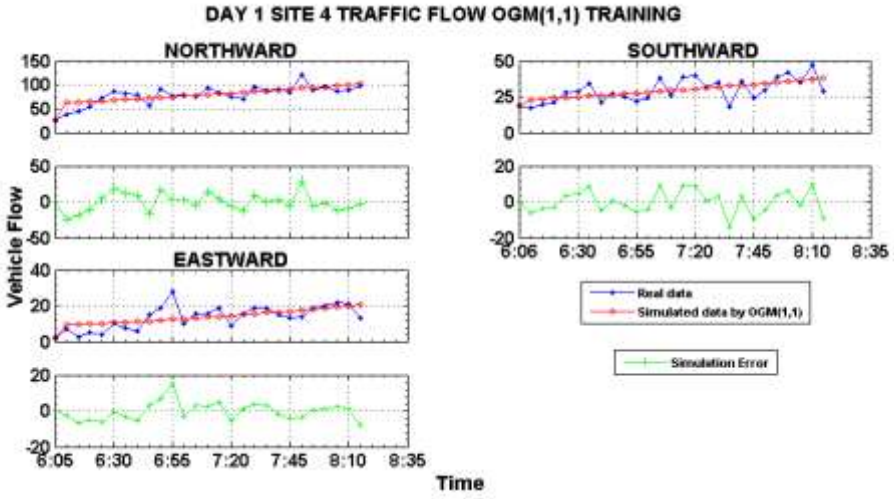

Figure 4.33a: Vehicle Flow OGM(1,1) Training

Vehicle flow MBVGM(1,1) training resulted to the model parameters of Table 4.22 and considering the southward direction the time response function was obtained as:

Table 4.22: Day 1 Site 4 MBVGM(1,1) Model Parameters

Traffic Flow Direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0190	62.1289
Southward	-0.0202	22.3833
Eastward	NaN	NaN

$$\hat{x}^{(1)}_{(r+1)} \triangleq 1{,}127.0842\, e^{0.0202r} - 1{,}108.0842,\ r = 0,1,2,\cdots,m-1 \qquad (4.23)$$

In MATLAB simulation of the eastward direction data the values for the model parameters were not computed because the data matrix A was singular to working precision. So, MATLAB gives a "NaN" result which stands for "Not a Number". The "NaN" results can be attributed to poor data collection. This is indicated in Table 4.22. Therefore, there were no results for the eastward direction as can be seen in Figure 4.33b.

In Table 4.23 recorded are the MBVGM(1,1)'s simulation values as computed from (4.23). The plots of the real, simulated and error values of the MBVGM(1,1) are shown in Figure 4.33b.

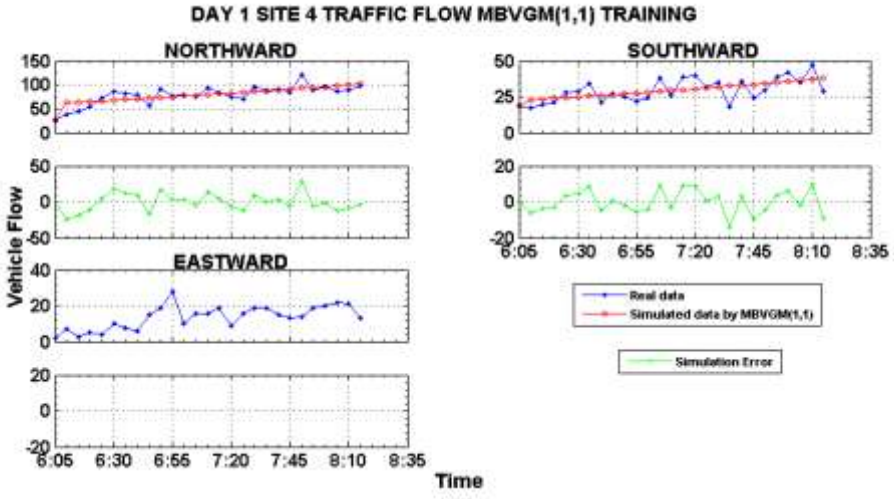

Figure 4.33b: Vehicle Flow MBVGM(1,1) Training

The plots of the real, simulated and error values of the GGM(1,1) and MBVGGM(1,1) are shown in Figure 4.33c and 4.33d, respectively. Their simulation values were as recorded in Table 4.23. As it was mentioned earlier their time response functions and model parameters are many and thus not provided here.

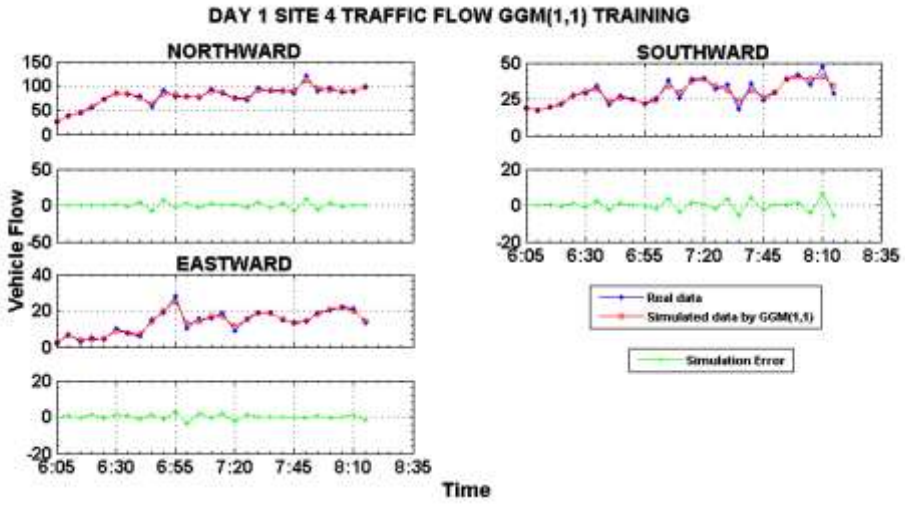

Figure 4.33c: Vehicle Flow GGM(1,1) Training

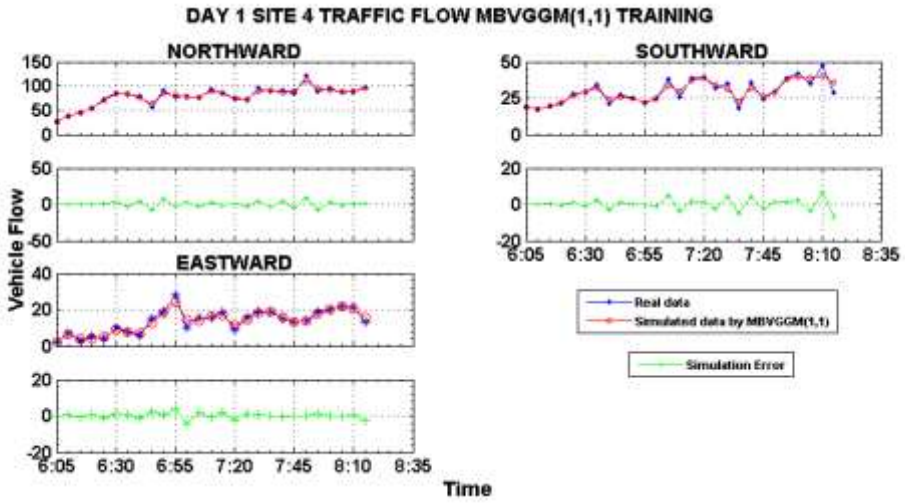

Figure 4.33d: Vehicle Flow MBVGGM(1,1) Training

It can be observed from Figures 4.33a – 4.33d that the error curves for the OGM(1,1) and MBVGM(1,1) are not as smooth as those of the GGM(1,1) and MBVGGM(1,1). The error curves for GGM(1,1) and MBVGGM(1,1) are approaching zero. This is a good indication that their fitting accuracy is higher compared to those of the OGM(1,1) and MBVGM(1,1).

b) Testing the Grey Models in Short-Term Forecasting

For the OGM(1,1) and MBVGM(1,1) (4.22) and (4.23) were extrapolated three points into the future. In short-term forecasting the parameters a and b remained the same as those obtained in training the models (i.e. as in Tables 4.21 and 4.22). Figures 4.34a – 4.34d show the short-term forecasting by the conventional and improved grey models.

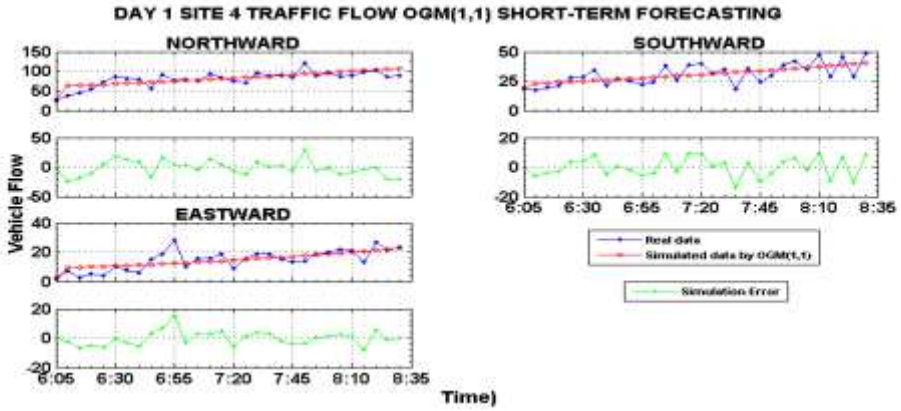

Figure 4.34a: Short-Term Vehicle Flow Forecast by OGM(1,1)

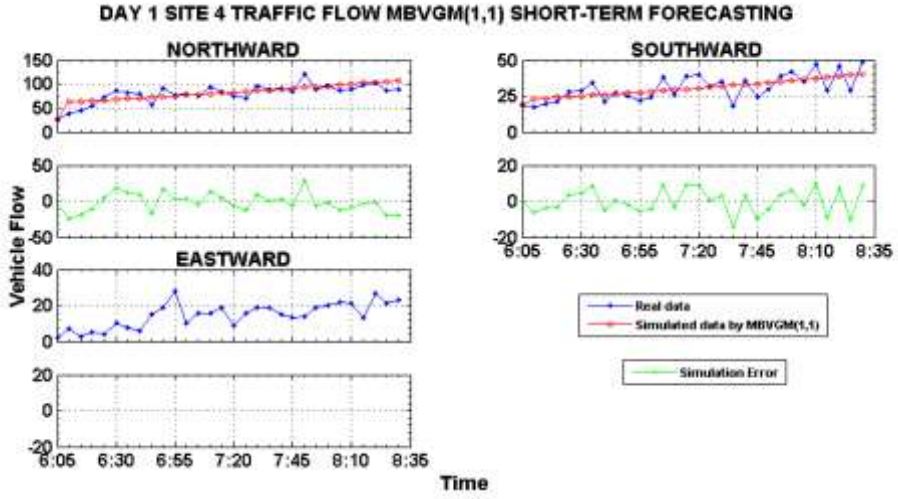

Figure 4.34b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

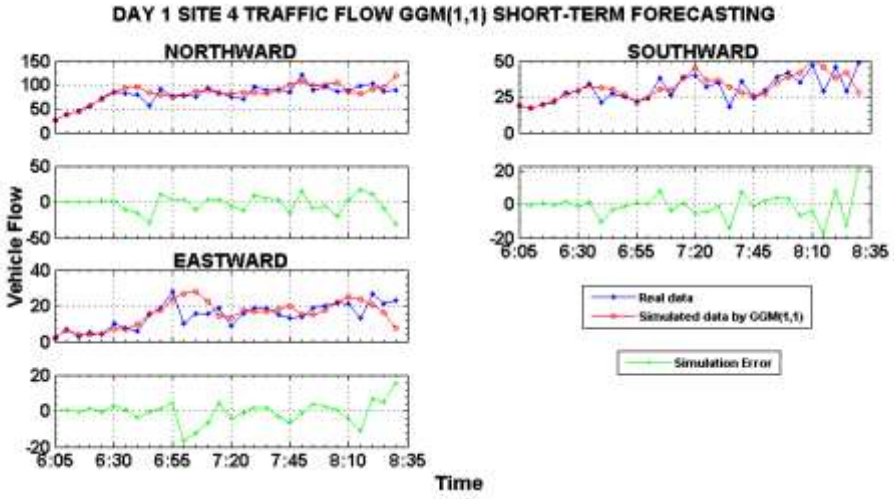

Figure 4.34c: Short-Term Vehicle Flow Forecast by GGM(1,1)

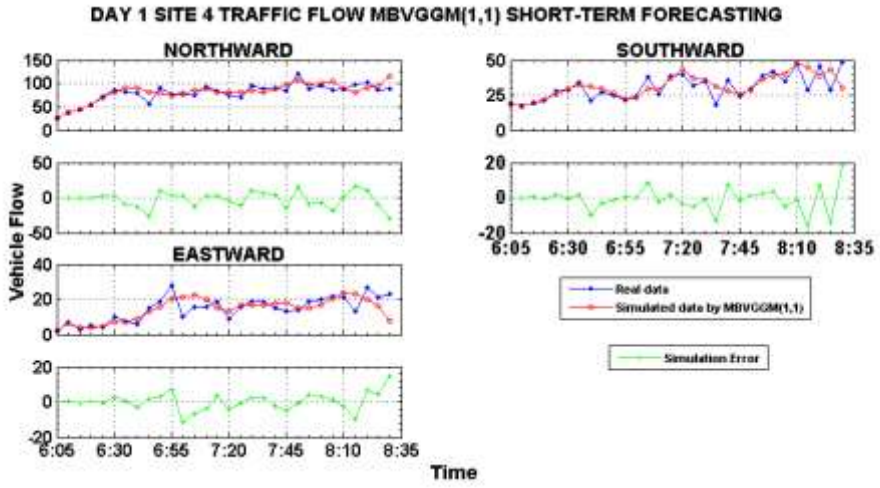

Figure 4.34d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

From Figures 4.34a - 4.34d the following observations are made. Firstly, there were no results for the eastward direction of Figure 4.34b as MATLAB simulation indicated a NaN result, as explained earlier. Secondly, close observation of the error curves can reveal that MBVGGM(1,1) is more accurate in short-term forecasting compared with the other three models. This is very clear from the evaluation point of view discussed next.

c) Evaluation of the Grey Models

The simulation data for the four grey models in the southward direction were as tabulated in Table 4.23. Both training and short-term forecasting simulation data are as shown. The overall training and short-term forecasting errors for site 4 were as well tabulated in Table 4.24. Table 4.24 reveals that GGM(1,1) is more accurate in vehicle flow fitting because its fitting errors are lower as compared to those of the other three grey models. The forecasting errors of the MBVGGM(1,1) in the northward and southward directions are lower compared to the other models' errors. It means it was the most accurate in short-term forecasting with an accuracy of 100-MAPD=89.3487%, in the northward direction. The reasons for the "NaN" results in Table 4.24 are as explained before.

Table 4.23: Original and Improved Grey Models' Simulation Values (D1S4 Southward Direction)

Raw Data		Grey Model			
Data Point	**Real Value**	**OGM(1,1)**	**MBVGM(1,1)**	**GGM(1,1)**	**MBVGGM(1,1)**
Training			**Fitted Values**		
1	19	19.0000	19.0000	19.0000	19.0000
2	17	22.9708	22.9991	17.1955	17.1960
3	20	23.4416	23.4689	19.3804	19.3044
4	21	23.9221	23.9482	21.7964	21.6037
5	28	24.4125	24.4373	27.0530	26.6290
6	29	24.9129	24.9364	30.1842	29.9394
7	34	25.4236	25.4457	31.6298	31.5882
8	21	25.9447	25.9654	23.7549	23.9301
9	27	26.4765	26.4957	25.4714	25.7444
10	25	27.0193	27.0369	24.9985	25.0042
11	22	27.5731	27.5891	21.8850	21.7478
12	24	28.1383	28.1525	25.5738	24.9828
13	38	28.7151	28.7275	34.3203	33.1899
14	26	29.3037	29.3143	29.7824	29.4705
15	39	29.9044	29.9130	37.2040	37.1623
16	40	30.5174	30.5239	39.2192	38.9238
17	32	31.1429	31.1473	33.9989	34.5356
18	35	31.7813	31.7835	31.2479	30.7829
19	18	32.4328	32.4326	23.4046	22.8792
20	36	33.0976	33.0950	31.3851	31.6139
21	24	33.7760	33.7709	26.4991	26.4954
22	30	34.4684	34.4607	29.5393	29.0559
23	39	35.1749	35.1645	38.6860	37.7344
24	42	35.8959	35.8827	40.5355	39.7874
25	35	36.6317	36.6155	39.2244	38.9251
26	47	37.3826	37.3634	40.4875	40.4193
27	29	38.1489	38.1265	34.4040	35.5681
Testing			**Short-Term Forecasted Values**		
28	46	38.9309	38.9052	37.9711	38.7661
29	29	39.7289	39.6998	41.7102	43.3160
30	49	40.5433	40.5106	27.7948	30.5772

III. Site 5: Moi Avenue-Slip Road Junction

a) Traffic Flow Training

Vehicle flow data of site 5 (of Appendix IV Table 5) in all directions were simulated and Figures 4.35a-4.35d and 4.36a-4.36d shows the grey model training and short-term forecasting respectively. The first 27 data point were used for training the grey models, ass before.

Table 4.24: Day 1 Site 4 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional		Improved	
		GM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training					
Northbound	RMSE	12.1582	12.1582	4.0208	4.0717
	RMSPE	14.3974	14.3916	5.0627	5.2877
	MAE	9.6951	9.6904	3.0609	3.2367
	MAPD	12.3069	12.3009	3.8854	4.1086
Southbound	RMSE	6.3247	6.3247	2.8533	3.0006
	RMSPE	19.0390	19.0368	9.2014	9.7922
	MAE	5.2938	5.2950	2.1769	2.3276
	MAPD	17.9340	17.9379	7.3746	7.8854
Eastbound	RMSE	4.9101	NaN	1.3525	1.5890
	RMSPE	33.6235	NaN	8.0072	9.2317
	MAE	3.8160	NaN	1.0173	1.1489
	MAPD	27.9982	NaN	7.4635	8.4297
Short-Term Forecasting					
Northbound	RMSE	12.5901	12.5870	11.8193	11.2687
	RMSPE	14.6467	14.6378	13.5735	13.1212
	MAE	10.0775	10.0704	8.7971	8.5317
	MAPD	12.5811	12.5723	10.9827	10.6513
Southbound	RMSE	6.6248	6.6256	7.2162	6.7671
	RMSPE	18.8943	18.9028	22.5237	20.4914
	MAE	5.6396	5.6416	4.8673	4.5144
	MAPD	18.3701	18.3766	15.8544	14.7048
Eastbound	RMSE	4.7742	NaN	6.0297	5.0596
	RMSPE	29.2685	NaN	31.2832	29.2517
	MAE	3.6647	NaN	4.1953	3.7187
	MAPD	25.0436	NaN	28.6695	25.4124

AGO and MGO operations on the first 27 data points of site 5 resulted to OGM(1,1)'s model parameters shown in Table 4.25. Note that site 5 is a three-way junction.

Thus the time response function of (3.9) for the OGM(1,1) in the eastward direction simplified to:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 2{,}112.0187\, e^{0.0267r} - 2{,}097.0187,\ r = 0,1,2,\cdots,m-1 \qquad (4.24)$$

Table 4.25: Day 1 Site 5 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0118	45.4337
Eastward	-0.0267	55.9904
Westward	-0.0752	15.6012

From (4.24) OGM(1,1)'s simulation values were tabulated in Table 4.27 and Figure 4.35a is the plot of the real data, simulation data and the error curve.

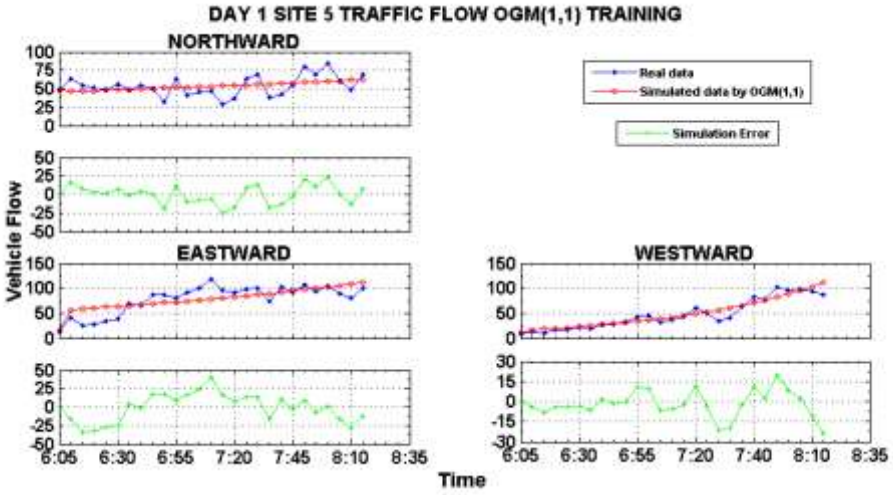

Figure 4.35a: Vehicle Flow OGM(1,1) Training

The MBVGM(1,1) vehicle flow training of the first 27 data points resulted to the model parameters of Table 4.26 and considering the eastward direction the time response function was formulated as:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 2{,}113.5506\, e^{0.0267r} - 2{,}098.5506,\ r = 0,1,2,\cdots,m-1 \qquad (4.25)$$

Table 4.26: Day 1 Site 5 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0117	45.4876
Eastward	-0.0267	56.0313
Westward	-0.0752	15.6482

In Table 4.27 recorded are the MBVGM(1,1)'s simulation values as computed from (4.25). The plots of the real, simulated and error values of the MBVGM(1,1) are shown in Figure 4.35b.

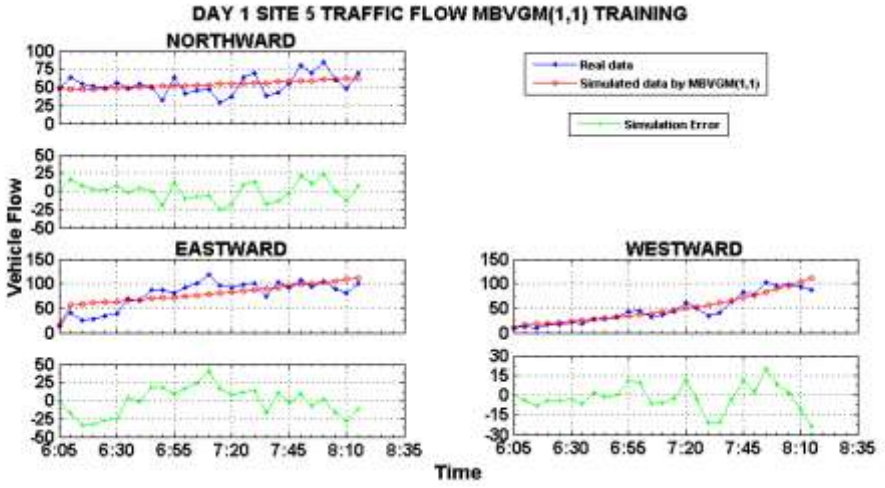

Figure 4.35b: Vehicle Flow MBVGM(1,1) Training

The plots of the real, simulated and error values of the GGM(1,1) and MBVGGM(1,1) are shown in Figures 4.35c and 4.35d, respectively. Their simulation values were as recorded in Table 4.27. Because of the same reasons given earlier their time response functions and model parameters are many and thus not provided here.

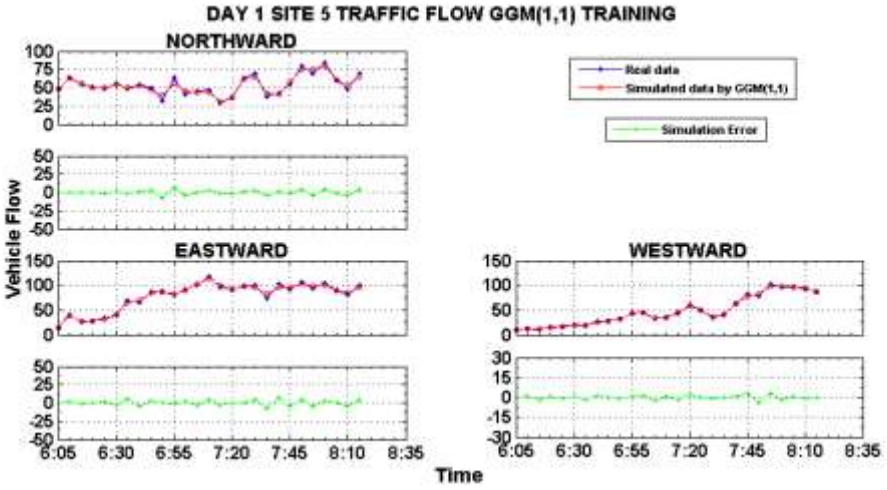

Figure 4.35c: Vehicle Flow GGM(1,1) Training

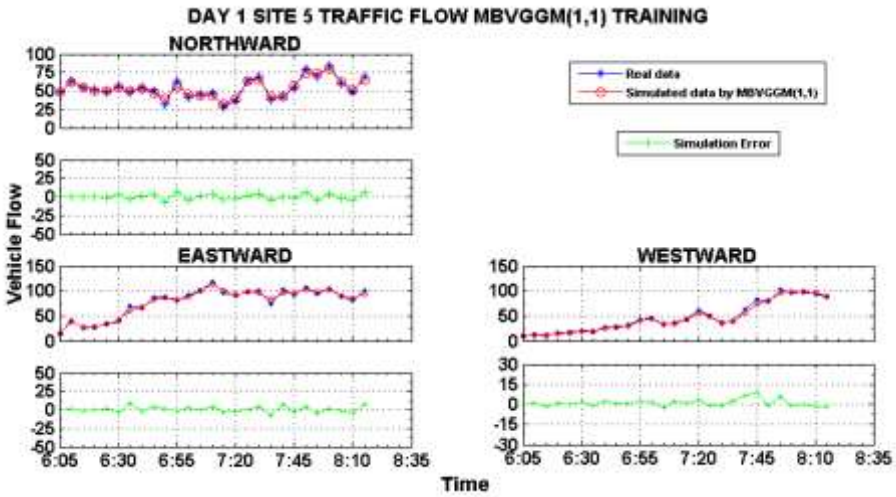

Figure 4.35d: Vehicle Flow MBVGGM(1,1) Training

Similarly, as for site 4, it can be observed from Figures 4.35a - 4.35d show that the error curves for the OGM(1,1) and MBVGM(1,1) are not as smooth as those of the GGM(1,1) and MBVGGM(1,1). The error curves for GGM(1,1) and MBVGGM(1,1) are smooth and approaching zero. This indicates that their fitting accuracy is higher compared to those of the OGM(1,1) and MBVGM(1,1).

b) Testing the Grey Models in Short-Term Forecasting

For the OGM(1,1) and MBVGM(1,1) (4.24) and (4.25) were extrapolated three points into the future. The parameters a and b remained the same as those obtained in training the models (i.e. as in Tables 4.25 and 4.26). Figures 4.36a-4.36d show the short-term forecasting by the conventional and improved grey models.

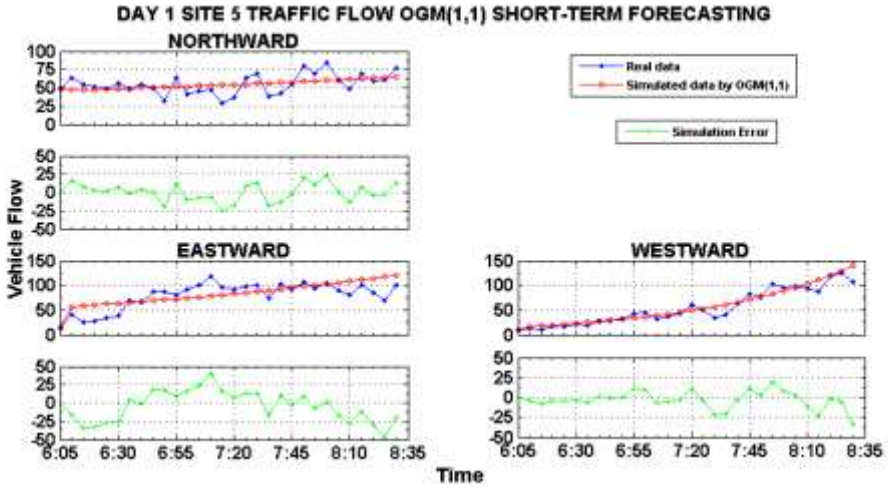

Figure 4.36a: Short-Term Vehicle Flow Forecast by OGM(1,1)

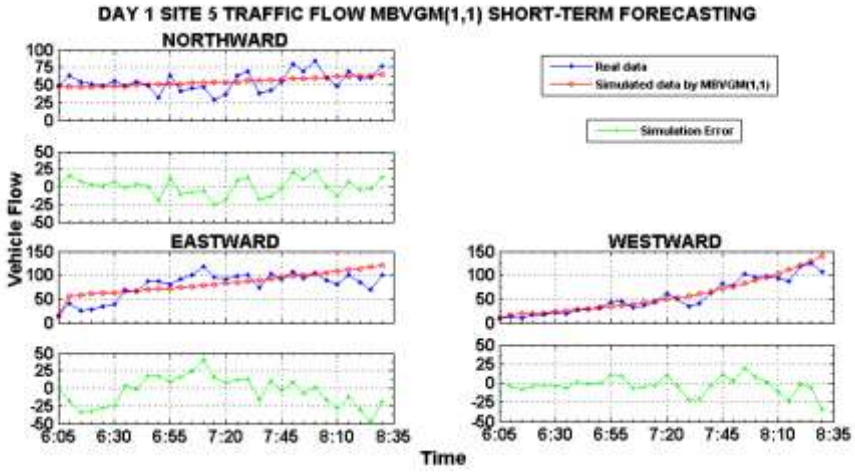

Figure 4.36b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

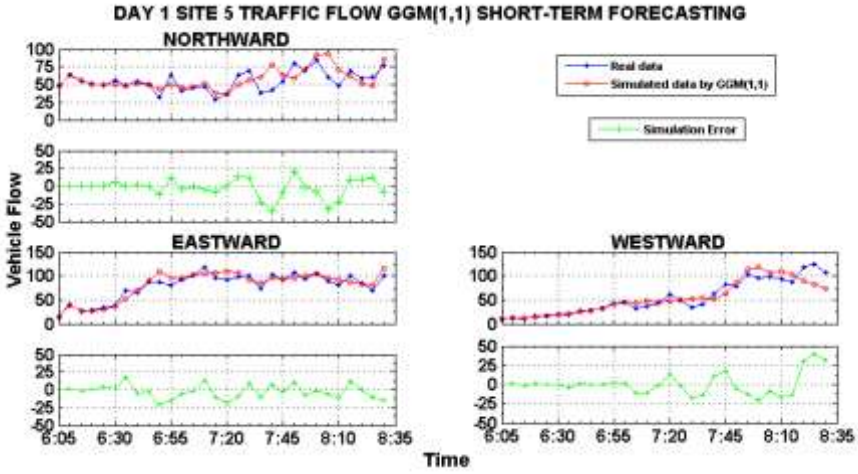

Figure 4.36c: Short-Term Vehicle Flow Forecast by GGM(1,1)

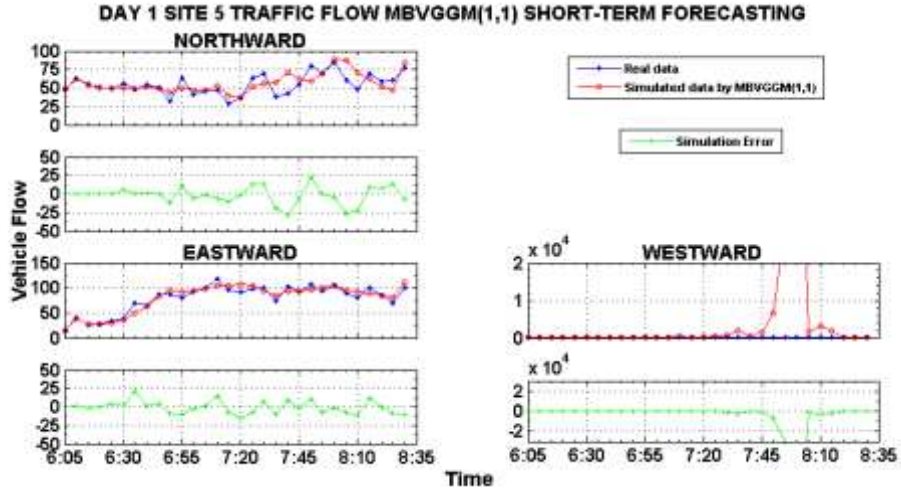

Figure 4.36d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

From Figures 4.36a-4.36d the following observations are made. Firstly, close observation of the error curves can reveal that MBVGGM(1,1) is more accurate in short-term forecasting compared with the other three models. This is particularly in the northward and eastward directions. The MBVGGM(1,1) fitting in those directions is good. Similar result are exhibited by the GGM(1,1). Secondly, the OGM(1,1) had a good fit in the westward direction. The other models, especially the MBVGGM(1,1), had poor fitting in this direction. See Figure 4.36d.

c) Evaluation of the Grey Models

The simulation data for the four grey models in the eastward direction were as tabulated in Table 4.27. Both training and short-term forecasting simulation data are as shown in that table. The overall training and short-term forecasting errors for site 5 were as well tabulated in Table 4.28.

Again Table 4.28 reveals that GGM(1,1) is more accurate in vehicle flow fitting because its fitting errors are lower as compared to those of the other three grey models. Its fitting accuracy as indicated by RMSPE can be shown to be 100-2.6674=97.3326%, in the westward direction. The forecasting errors of the MBVGGM(1,1) in the northward and eastward directions are lower compared to the other models' errors. It means it was the most accurate in short-term forecasting with the highest accuracy of 100-MAPD=91.3983%, in the eastward direction. The OGM(1,1) emerged the most accurate, in the westward direction, because it had the lowest errors whereas the MBVGGM(1,1) was the poorest in short term forecasting of vehicle flow in that direction.

Table 4.27: Original and Improved Grey Models' Simulation Values (D1S5 Eastward Direction)

Raw Data		Grey Model			
Data Point	**Real Value**	**OGM(1,1)**	**MBVGM(1,1)**	**GGM(1,1)**	**MBVGGM(1,1)**
Training			**Fitted Values**		
1	15	15.0000	15.0000	15.0000	15.0000
2	40	57.1503	57.1904	38.8045	38.6735
3	25	58.6965	58.7359	26.5687	27.0070
4	28	60.2846	60.3231	27.5881	27.9089
5	35	61.9156	61.9532	32.9612	33.4278
6	38	63.5908	63.6274	41.5194	40.6947
7	70	65.3113	65.3468	64.1996	61.4665
8	65	67.0783	67.1127	69.4464	67.0840
9	87	68.8931	68.9262	84.4278	82.7373
10	88	70.7571	70.7888	87.3784	86.3037
11	81	72.6715	72.7018	82.8340	82.4817
12	91	74.6376	74.6664	89.6920	88.0121
13	100	76.6570	76.6841	102.3486	99.8184
14	119	78.7310	78.7564	114.6704	114.1660
15	97	80.8611	80.8846	99.3144	100.5934
16	91	83.0488	83.0704	91.5713	91.9812
17	98	85.2958	85.3152	98.2261	98.3724
18	101	87.6035	87.6207	96.1714	96.1120
19	74	89.9736	89.9885	81.9878	81.8919
20	103	92.4079	92.4202	96.2314	95.3456
21	92	94.9081	94.9177	96.7396	95.5833
22	106	97.4759	97.4827	101.7348	101.5812
23	93	100.1131	100.1170	97.1807	97.2632
24	104	102.8217	102.8224	100.6233	101.7907
25	89	105.6036	105.6010	89.2184	90.1996
26	80	108.4608	108.4547	84.2895	84.6780
27	100	111.3953	111.3855	95.4800	93.0706
Testing			**Short-Term Forecasted Values**		
28	84	114.4091	114.3955	83.4791	83.9357
29	70	117.5045	117.4868	80.8643	79.6307
30	101	120.6837	120.6616	115.9550	111.3272

Table 4.28: Day 1 Site 5 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional	Improved		
		GM(1,1)	MBVGM(1,1)	GGM(1,1	MBVGGM(1,1)
Training					
Northbound	RMSE	12.6352	12.6352	3.3856	3.5377
	RMSPE	21.2517	21.2571	5.7929	5.9251
	MAE	10.2713	10.2671	2.7259	2.9039
	MAPD	19.1522	19.1444	5.0828	5.4148
Eastbound	RMSE	18.6457	18.6463	3.6385	4.0083
	RMSPE	18.7725	18.7589	4.1000	4.5681
	MAE	15.5184	15.5156	2.9734	3.2006
	MAPD	19.8577	19.8540	3.8048	4.0956
Westbound	RMSE	10.0481	10.0691	1.6733	2.8186
	RMSPE	15.1881	15.1824	2.6674	4.6669
	MAE	7.5897	7.6085	1.3123	1.9730
	MAPD	16.0346	16.0743	2.7725	4.1682
Short-Term Forecasting					
Northbound	RMSE	12.2374	12.2387	13.1021	11.7994
	RMSPE	20.1397	20.1488	20.5937	18.7787
	MAE	9.8795	9.8743	9.0299	8.4351
	MAPD	18.0174	18.0078	16.4679	15.3832
Eastbound	RMSE	20.7811	20.7790	9.8209	8.5487
	RMSPE	20.8597	20.8456	11.3304	9.8501
	MAE	17.2198	17.2155	7.8180	6.7810
	MAPD	21.8433	21.8378	9.9171	8.6017
Westbound	RMSE	11.3874	11.4281	14.1859	2.5120e+04
	RMSPE	15.2518	15.3058	23.7447	3.9842e+04
	MAE	8.1746	8.2111	9.6439	6.1242e+03
	MAPD	15.0730	15.1404	17.7822	1.1292e+04

IV. Site 7: Haile Selassie Avenue-Moi Avenue Roundabout

a) Traffic Flow Training

Lastly, training and short-term forecasting of vehicle flow of day 1 site 7 (see Appendix IV Table 7) are as illustrated in Figures 4.37a-4.37d and 4.38a-4.38d respectively. Note that this site is a three-way junction.

The first 27 data points were used to train the OGM(1,1) and the parameters of this model were indicated in Table 4.29.

From Table 4.29 the time response function of (3.9) for the eastward direction simplified to:

$$\hat{x}^{(1)}_{(r+1)} \triangleq -3{,}972.9063\, e^{-0.0224r} + 4{,}002.9063,\ r = 0,1,2,\cdots,m-1 \tag{4.26}$$

The computed final model values of (4.26) were recorded in Table 4.31. The OGM(1,1)'s simulation plots are as in Figure 4.37a.

Table 4.29: Day 1 Site 7 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Southward	-0.0081	65.5375
Eastward	0.0224	89.6651
Westward	-0.0021	95.3182

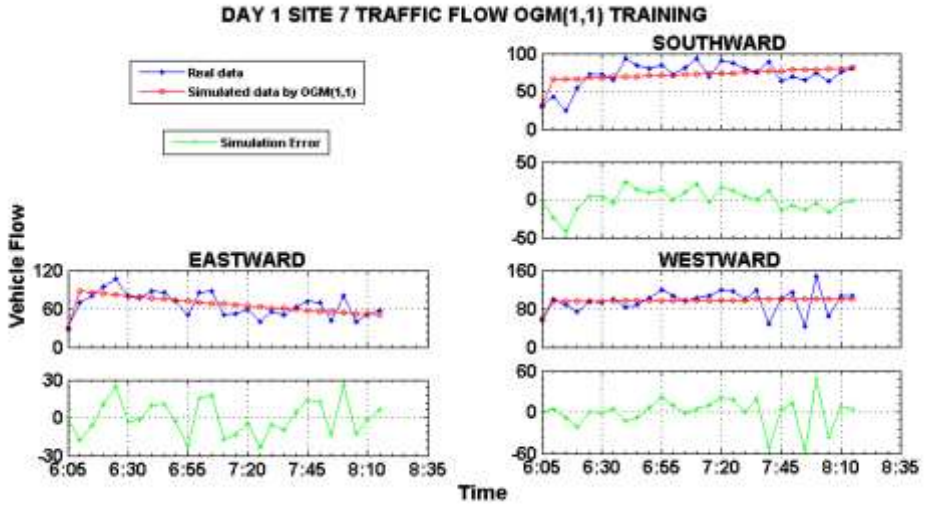

Figure 4.37a: Vehicle Flow OGM(1,1) Training

When the background value was modified the generated parameter values were as indicated in Table 4.30. The MBVGM(1,1)'s eastward direction time response function simplified to:

$$\hat{x}^{(1)}_{(r+1)} \triangleq -3{,}973.8616\, e^{-0.0224r} + 4{,}003.8616,\ r = 0,1,2,\cdots,m-1 \tag{4.27}$$

MBVGM(1,1)'s eastward simulation values from (4.27) were recorded in Table 4.31. Its plots are illustrated in Figure 4.37b.

Vehicle flow training by GGM(1,1) and MBVGGM(1,1) were performed and Figures 4.37c and 4.37d are their plots, respectively.

Table 4.30: Day 1 Site 7 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Southward	0.0081	65.5597
Eastward	0.0224	89.6865
Westward	-0.0020	95.4515

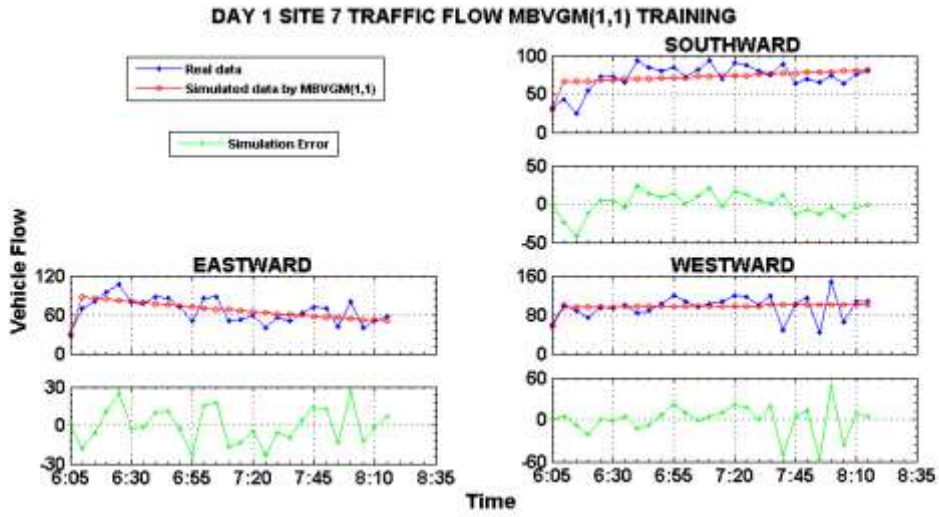

Figure 4.37b: Vehicle Flow MBVGM(1,1) Training

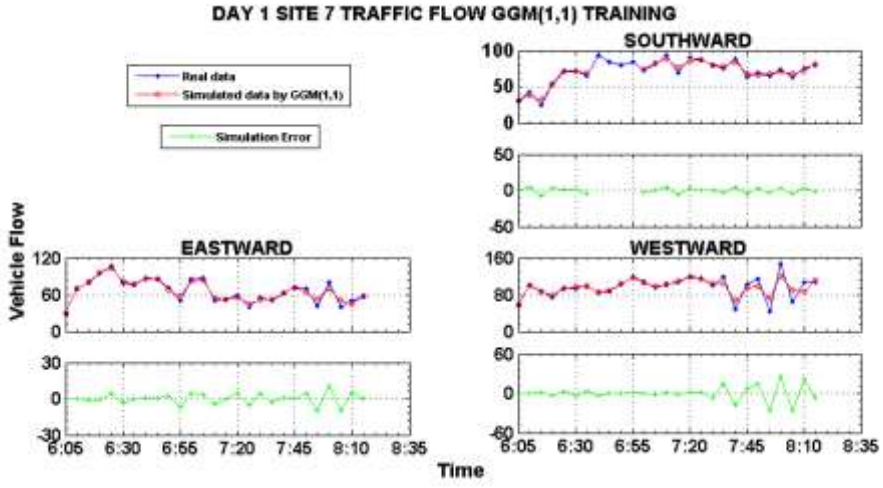

Figure 4.37c: Vehicle Flow GGM(1,1) Training

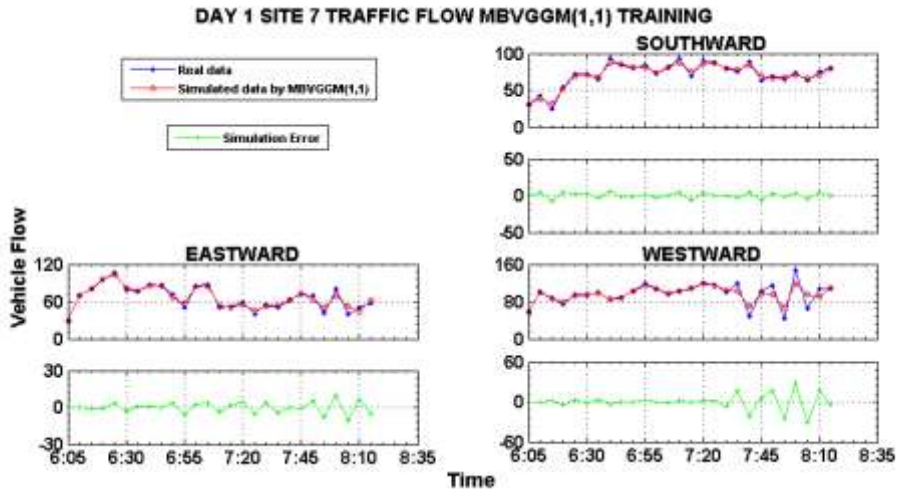

Figure 4.37d: Vehicle Flow MBVGGM(1,1) Training

From the error curves in Figure 4.37a - 4.37d, one can notice that the fit of OGM(1,1) and MBVGM(1,1) are similar and not as accurate as those of GGM(1,1) and MBVGM(1,1). GGM(1,1) and MBVGM(1,1) had a good fit except for GGM(1,1) which resulted to "NaN" in the southward direction. Because of this "NaN" result the GGM(1,1)'s simulated data curve is not complete. See Figure 4.37c.

b) Testing the Grey Models in Short-Term Forecasting

The grey models were extrapolated three points into the future. For OGM(1,1) and MBVGM(1,1) (4.26) and (4.27) were extrapolated and the forecasted values were

recorded in Table 4.31. The plots of real, simulated and error curves of these models are illustrated in Figures 4.38a - 4.38d.

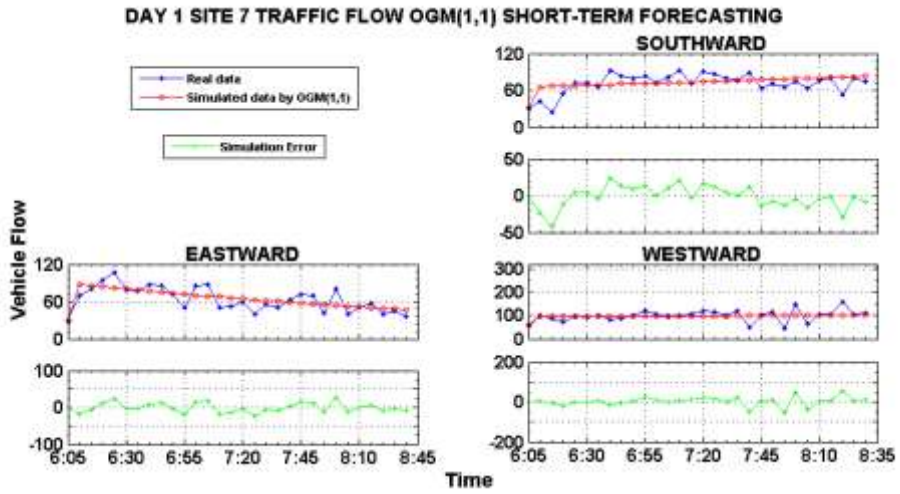

Figure 4.38a: Short-Term Vehicle Flow Forecast by OGM(1,1)

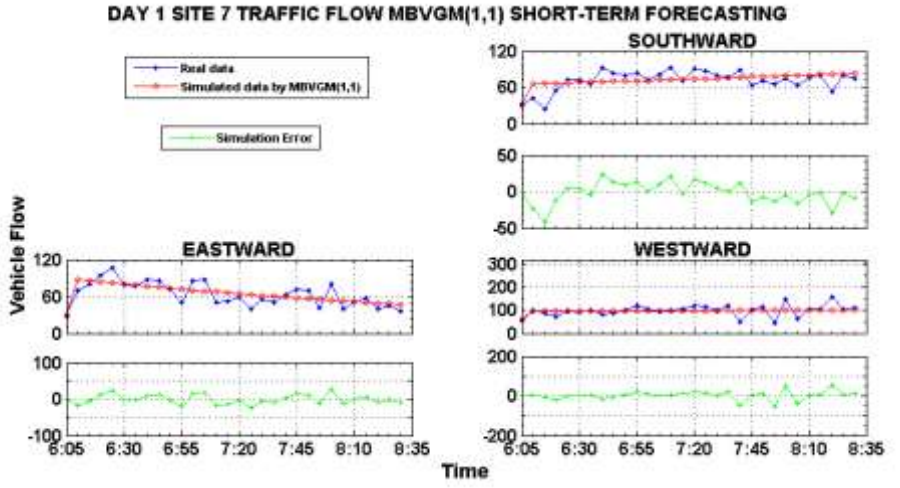

Figure 4.38b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

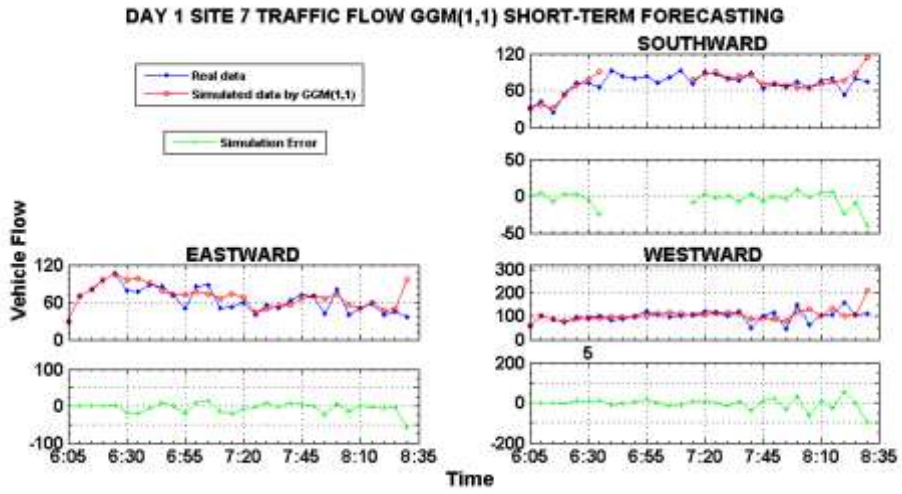

Figure 4.38c: Short-Term Vehicle Flow Forecast by GGM(1,1)

Observation of Figures 4.38a - 4.38d reveal some facts. OGM(1,1) and MBVGM(1,1) had a poor fitting compared to that of GGM(1,1) and MBVGGM(1,1). However, perdition of the 27th data point by OGM(1,1) and MBVGM(1,1) had a good fit. That of the GGM(1,1) and MBVGGM(1,1) was poor. GGM(1,1) simulation had a "NaN" result in the southward direction. This can be attributed to poor data collection.

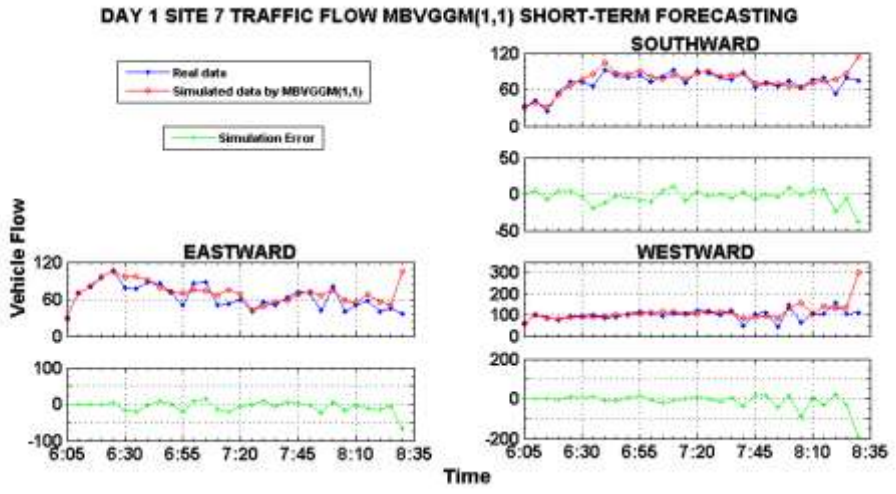

Figure 4.38d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

c) Evaluation of the Grey Models

The fitting and forecasting values for all the four grey models were as recorded in Table 4.31. Note that these were eastward direction values. Moreover, the overall training and short-term forecasting errors for site 7 were evaluated as tabulated in Table 4.32.

Examination of Table 4.32 shows that GGM(1,1)'s fitting accuracy was good compared to the other models. This was true especially in the eastward and westward directions. In the eastward direction GGM(1,1) had the highest fitting accuracy of 100-MAPD=100-5.1272=94.8728%. Notice from Table 4.31 that data point 5 was accurately predicted by both GGM(1,1) and MBVGGM(1,1). The real value of data point 5 is 107. GGM(1,1) and MBVGGM(1,1) fitted values were 102.6946 and 103.0843, respectively, which are close to this real value.

The GGM(1,1) MATLAB simulation resulted to "Not a Number" (NaN) error as seen in Table 4.32. This is because the data matrix A was singular to working precision. Hence it was impossible to compute "good" values of the parameters a and b that minimizes the sum of squared errors (see section 3.1.1 part IV). The "NaN" results can be attributed to poor data collection so that the southward vehicle flow data must have been poorly collected. Nevertheless, the GGM(1,1) can do well in real time, in case of automatic data collection.

Table 4.31: Original and Improved Grey Models' Simulation Values (D1S7 Eastward Direction)

Raw Data		Grey Model			
Data Point	Real Value	OGM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training		Fitted Values			
1	30	30.0000	30.0000	30.0000	30.0000
2	70	88.0055	88.0245	69.6269	69.9691
3	80	86.0582	86.0740	80.5368	80.6188
4	95	84.1539	84.1669	95.9502	95.8271
5	107	82.2917	82.3019	102.6946	103.0843
6	78	80.4708	80.4783	81.4393	81.4947
7	77	78.6901	78.6951	77.0997	76.4949
8	87	76.9489	76.9514	86.1724	86.3429
9	86	75.2462	75.2463	85.4360	85.6654
10	71	73.5811	73.5790	68.0892	67.0451
11	50	71.9529	71.9487	57.1353	56.6118
12	86	70.3608	70.3545	81.9277	83.1959
13	87	68.8038	68.7956	83.5280	83.8157
14	50	67.2813	67.2712	54.5151	53.2910
15	52	65.7925	65.7807	51.6447	49.9878
16	60	64.3367	64.3231	55.9546	55.8467
17	40	62.9130	62.8979	45.3341	45.3337
18	56	61.5209	61.5042	51.7243	52.5498
19	50	60.1596	60.1414	52.6531	54.4520
20	63	58.8284	58.8088	62.2499	63.3723
21	72	57.5266	57.5058	71.9596	72.5508
22	69	56.2537	56.2316	64.2224	63.4375
23	42	55.0089	54.9856	51.9225	50.4312
24	80	53.7917	53.7673	69.5096	70.4331
25	40	52.6014	52.5759	49.4545	51.1907
26	50	51.4374	51.4109	44.2132	43.7003
27	57	50.2992	50.2718	57.4321	62.5299
Testing		Short-Term Forecasted Values			
28	40	49.1862	49.1579	46.9205	55.2923
29	45	48.0978	48.0687	48.4346	50.0987
30	37	47.0335	47.0036	96.0113	105.2051

Table 4.32: Day 1 Site 7 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional	Improved		
		GM(1,1)	MBVGM(1,1	GGM(1,1)	MBVGGM(1,1)
Training					
Southbound	RMSE	14.1245	14.1245	NaN	3.6355
	RMSPE	15.8435	15.8382	NaN	4.4256
	MAE	10.8251	10.8217	NaN	3.1028
	MAPD	15.1126	15.1078	NaN	4.3317
Eastbound	RMSE	13.6036	13.6036	4.5799	4.6689
	RMSPE	18.4850	18.4878	5.2296	5.1629
	MAE	11.4187	11.4181	3.3896	3.5976
	MAPD	17.2720	17.2711	5.1272	5.4418
Westbound	RMSE	21.2909	21.2910	11.4169	11.7899
	RMSPE	18.1481	18.1698	10.3432	10.7383
	MAE	14.6333	14.6535	7.3335	7.3984
	MAPD	15.1611	15.1820	7.5980	7.6652
Short-Term Forecasting					
Southbound	RMSE	14.5403	14.5385	NaN	10.7860
	RMSPE	16.1147	16.1081	NaN	13.3218
	MAE	11.0964	11.0912	NaN	7.4117
	MAPD	15.5556	15.5485	NaN	10.3902
Eastbound	RMSE	13.1545	13.1528	15.1472	16.7721
	RMSPE	18.4852	18.4873	14.9964	15.8796
	MAE	11.0208	11.0173	9.6974	10.7480
	MAPD	17.3373	17.3319	15.2555	16.9082
Westbound	RMSE	22.6508	22.6645	28.1042	42.1754
	RMSPE	21.1885	21.2270	26.0519	37.7661
	MAE	15.4365	15.4688	17.8315	21.7208
	MAPD	15.5557	15.5883	17.9693	21.8886

4.3.3.2 Day Two Vehicle Traffic Flow Modelling and Short-Term Forecasting

Similarly, for day two four cases are presented which include modelling of vehicle traffic flow moving in all directions for;

- Day 2 Site 2 (D2S2): Kenyatta Avenue Uhuru Highway Roundabout,
- Day 2 Site 5 (D2S5): Moi Avenue-Slip Road Junction,
- Day 2 Site 6 (D2S6): City Hall Way-Wabera Street T-Roundabout and
- Day 2 Site 7 (D2S7): Haile Selassie Avenue-Moi Avenue Roundabout.

Consequently, data from the four sites were employed. The four cases are presented in Figures 4.39a-4.46d from which it is observed that the proposed GGM(1,1) model has good performance in modelling and short-term forecasting of traffic flow. However, forecasting of westbound traffic flow of site 2 (see Figure 4.40b) shows a poor performance of the proposed model and this can be attributed to some data collection errors such as not collecting data continuously at specified intervals of time. It can also be caused by in accuracy in data collection. Thus the 19th data point of the westbound traffic flow might have not been zero (0) number of vehicles as this is not realistic.

I. Site 2: Kenyatta Avenue Uhuru Highway Roundabout

a) Traffic Flow Training

The AGO series was generated from the real vehicle flow data of day 2 site 2 (see Appendix IV Table 9). Based on the AGO the MGO was obtained. Then from OGM(1,1) simulation the parameters a and b were obtained and Table 4.33 shows these parameter values.

Table 4.33: Day 2 Site 2 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	0.0036	110.9944
Southward	-0.0001	182.3804
Eastward	-0.0128	84.3814
Westward	-0.0171	54.4240

Based on Table 4.33 (3.9) can be simplified to obtain the time response functions for each direction of vehicle flow. For the eastward direction the time response function simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 6{,}667.2969\, e^{0.0128r} - 6{,}592.2969,\ r = 0,1,2,\cdots,m-1 \qquad (4.28)$$

Generation of the IAGO from (4.28) gives the final fitted data which is recorded in Table 4.35. The real data, simulation data and the fitting errors for the OGM(1,1) were as plotted in Figure 4.39a.

The background value of the OGM(1,1) was modified and Table 4.34 shows MBVGM(1,1)'s model parameters from which (3.9), in the eastward direction, simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 6{,}724.2362\, e^{0.0127r} - 6{,}649.2362,\ r = 0,1,2,\cdots,m-1 \qquad (4.29)$$

The IAGO operation was applied on the series generated from (4.29) and the final MBVGM(1,1)'s fitted data recorded in Table 4.35. The real data, simulation data and the fitting errors for the MBVGM(1,1) were as plotted in Figure 4.39b.

Table 4.34: Day 2 Site 2 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, *a*	Control Variable, *b*
Northward	0.0054	113.6233
Southward	0.0003	183.4814
Eastward	-0.0127	84.4453
Westward	-0.0157	55.6651

Similarly, the real data, simulation data and the fitting errors for GGM(1,1) and MBVGGM(1,1) were as plotted in Figures 4.39c and 4.39d, respectively.

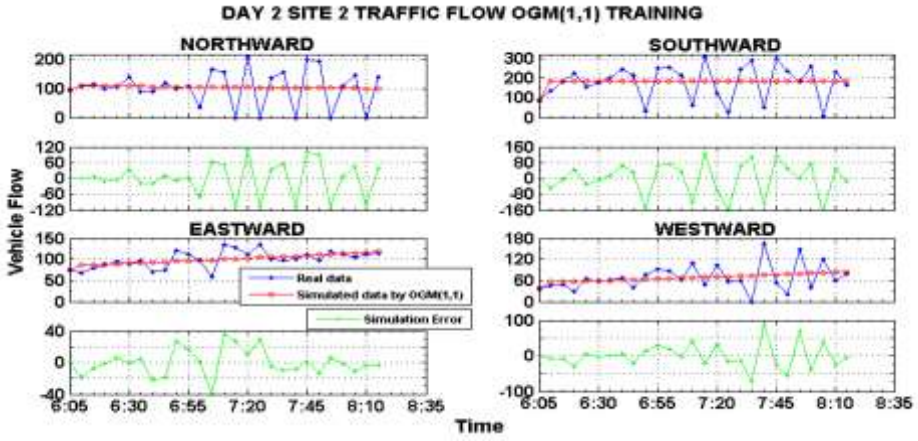

Figure 4.39a: Vehicle Flow OGM(1,1) Training

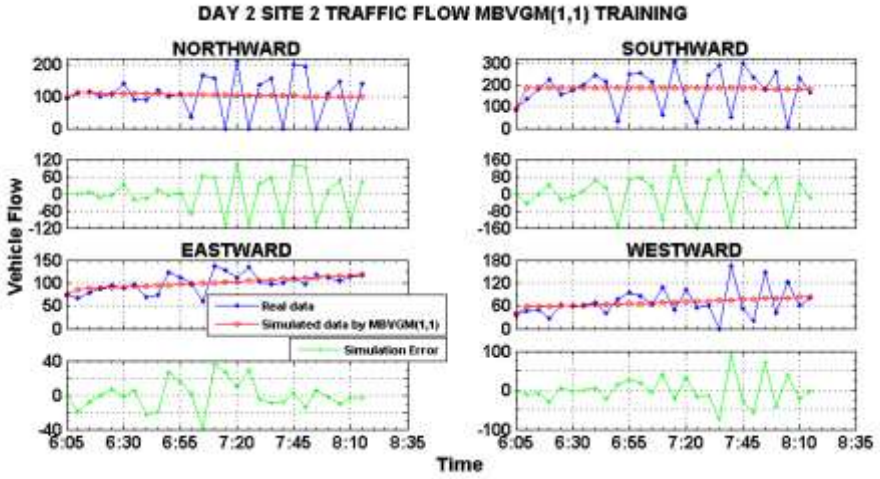

Figure 4.39b: Vehicle Flow MBVGM(1,1) Training

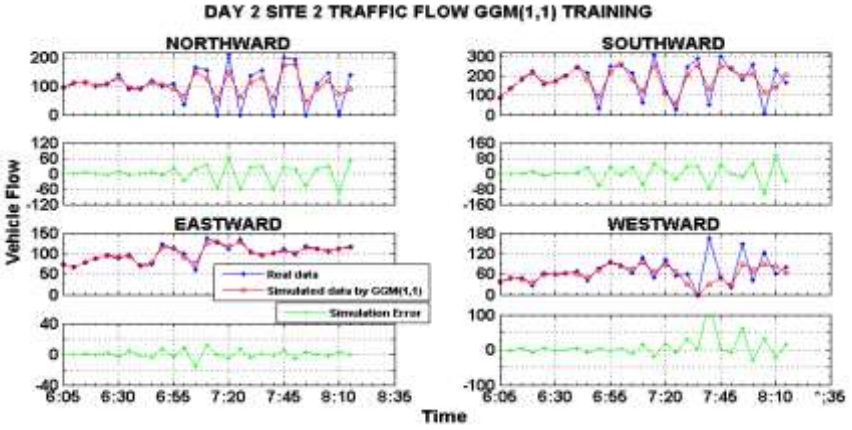

Figure 4.39c: Vehicle Flow GGM(1,1) Training

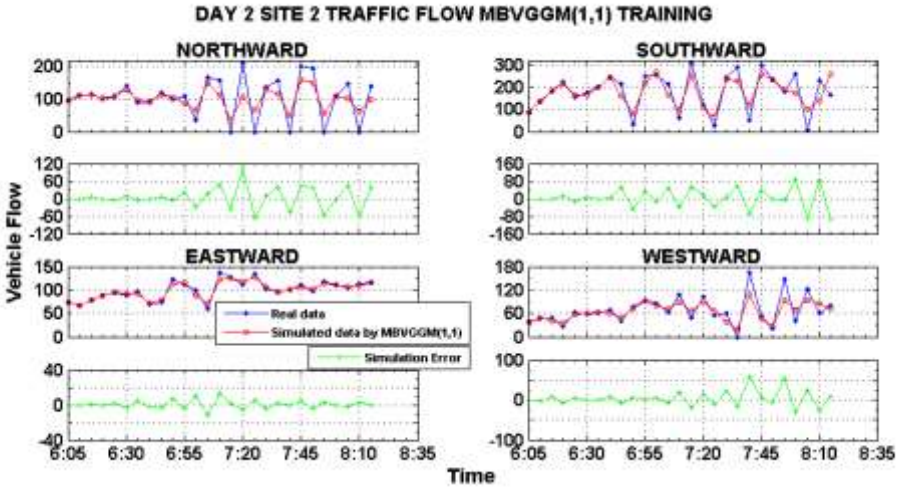

Figure 4.39d: Vehicle Flow MBVGGM(1,1) Training

Notice from Figures 4.39a - 4.39d that both OGM(1,1) and MBVGM(1,1) have poor fitting compared to GGM(1,1) and MBVGGM(1,1). GGM(1,1) and MBVGGM(1,1) have neat fitting especially on the eastward data. The simulation error curves for both GGM(1,1) and MBVGGM(1,1) are approaching zero level, especially on the eastward direction.

b) Testing the Grey Models in Short-Term Forecasting

OGM(1,1) and MBVGM(1,1) (4.28) and (4.29) were extrapolated three points into the future to estimate the last three data points. Their short-term forecasted values for the eastward direction were as recorded in Table 4.35. Plots of the real, simulation and forecasting errors are shown in Figures 4.40a and 4.40b. Similarly, the time response functions in GGM(1,1) and MBVGGM(1,1) were extrapolated and their forecasted values recorded in Table 5.1d2s2. the corresponding plots for GGM(1,1) and MBVGGM(1,1) are as shown in Figures 4.40c and 4.40d, respectively.

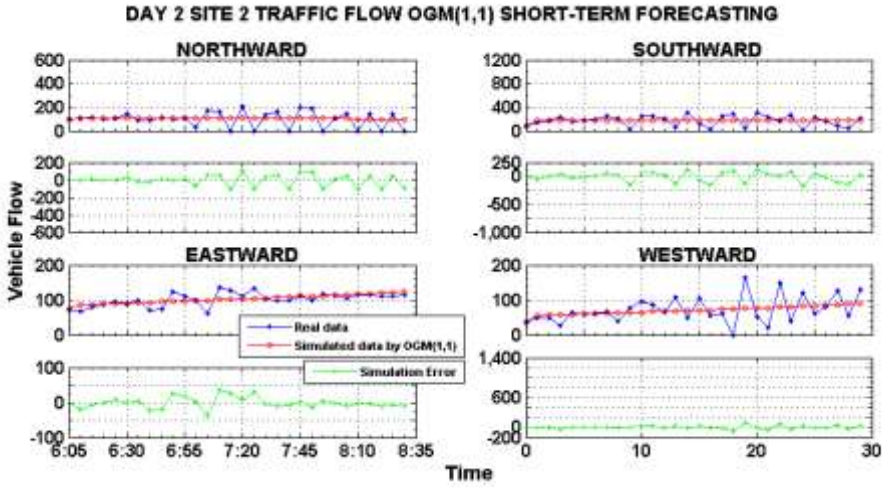

Figure 4.40a: Short-Term Vehicle Flow Forecast by OGM(1,1)

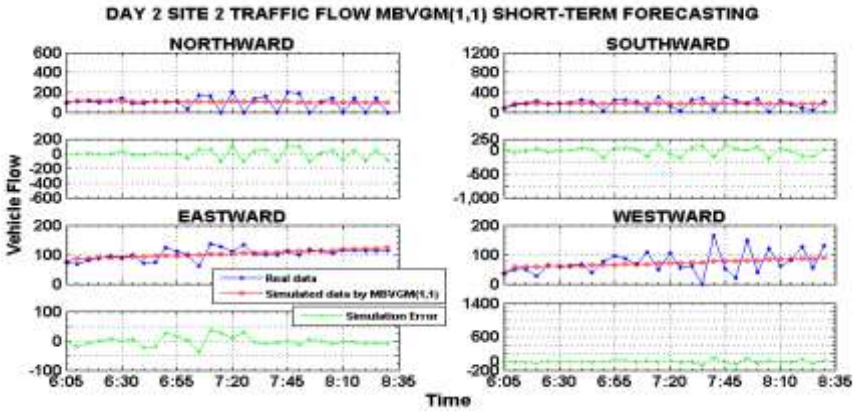

Figure 4.40b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

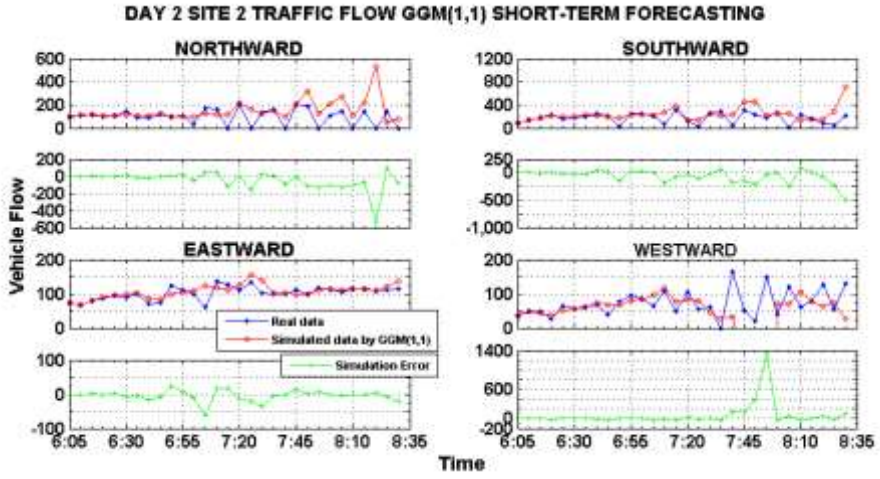

Figure 4.40c: Short-Term Vehicle Flow Forecast by GGM(1,1)

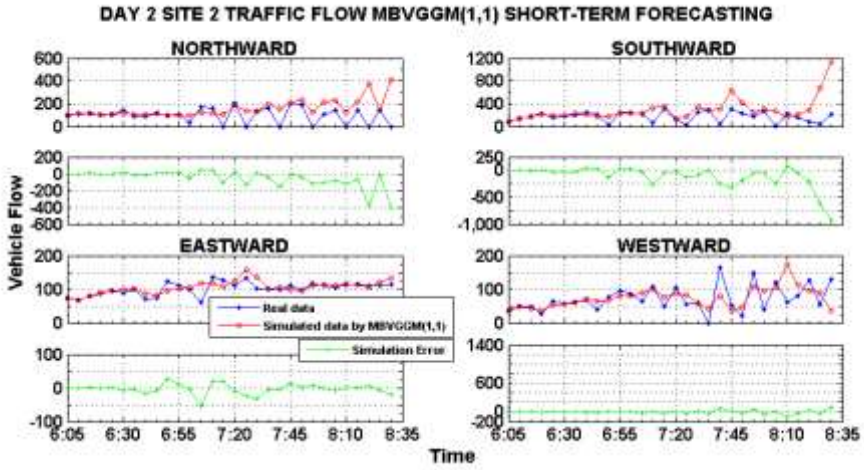

Figure 4.40d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

Short-term forecasting by OGM(1,1) and MBVGM(1,1), as seen in Figures 4.40a and 4.40b, looks similar. This is evident from the simulation error curves. On the other hand, in short-term forecasting all the models seemed to forecast the last three points with high accuracy, in the eastward direction. Observe and compare the forecasts of the last three points in Figures 4.40a - 4.40d.

c) Evaluation of the Grey Models

The simulation data for all the four grey models were as recorded in table 4.35. Note that these data are for the eastward direction only. The overall training and short-term forecasting error evaluation for day 2 site 2 is as tabulated in table 4.36.

Table 4.35: Original and Improved Grey Models' Simulation Values (D2S2 Eastward Direction)

Raw Data		**Grey Model**			
Data Point	**Real Value**	**OGM(1,1)**	**MBVGM(1,1)**	**GGM(1,1)**	**MBVGGM(1,1)**
Training			**Fitted Values**		
1	75	75.0000	75.0000	75.0000	75.0000
2	67	85.8889	85.9479	67.4505	67.4402
3	79	86.9945	87.0504	78.3453	78.1589
4	87	88.1145	88.1671	87.8238	87.6743
5	96	89.2488	89.2981	94.0071	93.8102
6	89	90.3977	90.4436	92.2379	92.4492
7	97	91.5615	91.6038	92.2427	92.6274
8	70	92.7402	92.7788	72.0566	72.0546
9	75	93.9341	93.9690	78.7954	77.5355
10	122	95.1433	95.1744	115.5470	114.5733
11	112	96.3681	96.3953	114.5628	115.7679
12	99	97.6087	97.6318	90.5658	88.3939
13	60	98.8653	98.8842	75.5630	71.8613
14	136	100.1380	100.1527	124.6558	123.2686
15	128	101.4271	101.4374	128.3664	126.4051
16	112	102.7329	102.7386	117.7553	117.7343
17	134	104.0554	104.0565	126.8743	128.3583
18	101	105.3949	105.3914	105.1820	105.9514
19	97	106.7517	106.7433	95.9372	95.4903
20	100	108.1260	108.1126	101.3479	101.1362
21	111	109.5179	109.4994	107.1038	106.5742
22	97	110.9278	110.9041	102.1744	101.7038
23	118	112.3558	112.3267	114.3808	114.5219
24	112	113.8022	113.7676	112.1374	112.0603
25	104	115.2673	115.2270	106.4402	106.1792
26	113	116.7512	116.7051	110.3678	109.9767
27	115	118.2541	118.2022	116.1274	115.3424
Testing			**Short-Term Forecasted Values**		
28	110	119.7765	119.7185	106.5689	106.1209
29	112	121.3184	121.2542	119.9326	120.0411
30	113	122.8802	122.8096	134.5849	134.1905

From Table 4.36 the errors involved in training the OGM(1,1) and MBVGM(1,1) are large compared to those of GGM(1,1) and MBVGGM(1,1). Generally, the GGM(1,1) training errors are the smallest in value and therefore GGM(1,1) was more accurate in fitting the real data in all directions of traffic flow except the westward direction. In the westward direction MBVGGM(1,1) was the most accurate with an accuracy of 100-MAPD=100-20.5771=79.4229%. In short-term forecasting no one specific model emerged as the most accurate. But in the eastward direction the MBVGGM(1,1) performed well. And in the southward direction MBVGM(1,1) was the most accurate among the four models.

Table 4.36: Day 2 Site 2 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional		Improved	
		GM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training					
Northbound	RMSE	62.0528	62.0691	33.7346	37.6636
	RMSPE	39.6257	40.0932	18.3727	28.0829
	MAE	47.6995	48.0677	25.6133	27.5952
	MAPD	45.3162	45.6660	24.3335	26.2164
Southbound	RMSE	83.7442	83.7464	43.6047	45.8546
	RMSPE	30.4697	30.5342	15.7895	18.8121
	MAE	66.8685	66.9053	32.6690	34.1508
	MAPD	37.3335	37.3541	18.2395	19.0668
Eastbound	RMSE	16.5449	16.5450	5.1675	5.0757
	RMSPE	15.8976	15.8916	4.7021	4.9082
	MAE	12.2612	12.2585	3.7405	3.7678
	MAPD	12.2340	12.2313	3.7322	3.7595
Westbound	RMSE	34.7333	34.7413	31.6858	19.9880
	RMSPE	44.0293	44.1954	54.4613	29.5262
	MAE	25.7608	25.7357	17.0022	13.9924
	MAPD	37.8835	37.8466	25.0032	20.5771
Short-Term Forecasting					
Northbound	RMSE	64.7117	64.5011	119.0153	119.6668
	RMSPE	39.2706	39.7792	37.0482	24.8176
	MAE	50.9361	51.1746	65.9250	67.7025
	MAPD	51.2436	51.4835	66.3230	68.1112
Southbound	RMSE	85.3689	85.2755	139.1228	239.1549
	RMSPE	30.3785	30.4423	52.5835	90.2041
	MAE	68.9119	68.9108	86.4059	133.1119
	MAPD	39.9953	39.9947	50.1485	77.2559
Eastbound	RMSE	15.9905	15.9867	16.8365	16.0693
	RMSPE	15.1992	15.1897	13.7939	13.7450
	MAE	12.0009	11.9921	10.7183	10.5722
	MAPD	11.8391	11.8304	10.5738	10.4297
Westbound	RMSE	35.0611	35.1349	260.2773	37.8887
	RMSPE	41.7766	42.0734	398.3535	42.3894
	MAE	26.9509	26.9744	85.4140	26.2252
	MAPD	37.6233	37.6562	119.2378	36.6104

II. Site 5: Moi Avenue-Slip Road Junction

a) Traffic Flow Training

AGO and MGO operations on the first 27 data points of Appendix IV Table 5 led to the computation of OGM(1,1)'s a and b parameters as tabulated in Table 4.37. Note that site 5 is a three-way junction. Considering the westward direction (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 1{,}690.3785\, e^{0.0288r} - 1{,}650.3785,\ r = 0,1,2,\cdots,m-1 \tag{4.30}$$

The OGM(1,1)'s simulation data as computed from (4.30) were recorded in Table 4.39. The corresponding OGM(1,1)'s plot is shown in Figure 4.41a.

Table 4.37: Day 2 Site 5 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0387	32.1944
Eastward	-0.0113	54.1796
Westward	-0.0288	47.5309

In MBVGM(1,1) the background value was modified and its a and b parameters were calculated. See Table 4.38. For this model, in the westward direction, (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 1{,}929..3310\, e^{0.0287r} - 1{,}889.3310,\ \ r = 0,1,2,\cdots,m-1 \tag{4.31}$$

MBVGM(1,1)'s simulation data as obtained from IAGO on (4.31) were recorded in Table 4.39 and Figure 4.41b shows plot of its real, simulated and fitting error data.

Table 4.38: Day 2 Site 5 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0387	32.2526
Eastward	-0.0112	54.2238
Westward	-0.0287	47.6158

Figures 4.41c and 4.41d are the plots for GGM(1,1) and MBVGGM(1,1), respectively. Their simulation data were recorded in Table 4.39.

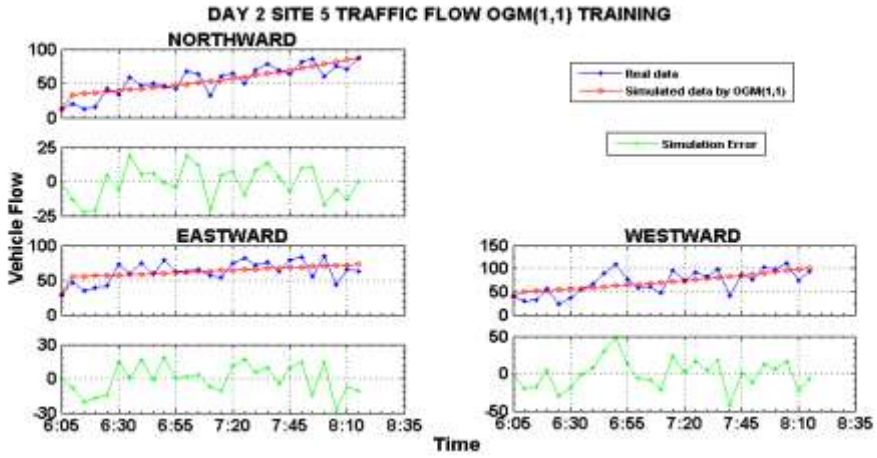

Figure 4.41a: Vehicle Flow OGM(1,1) Training

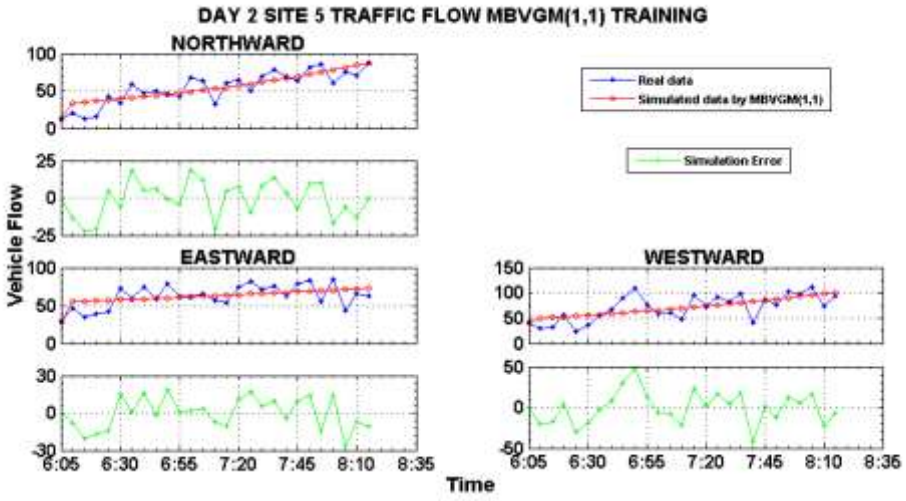

Figure 4.41b: Vehicle Flow MBVGM(1,1) Training

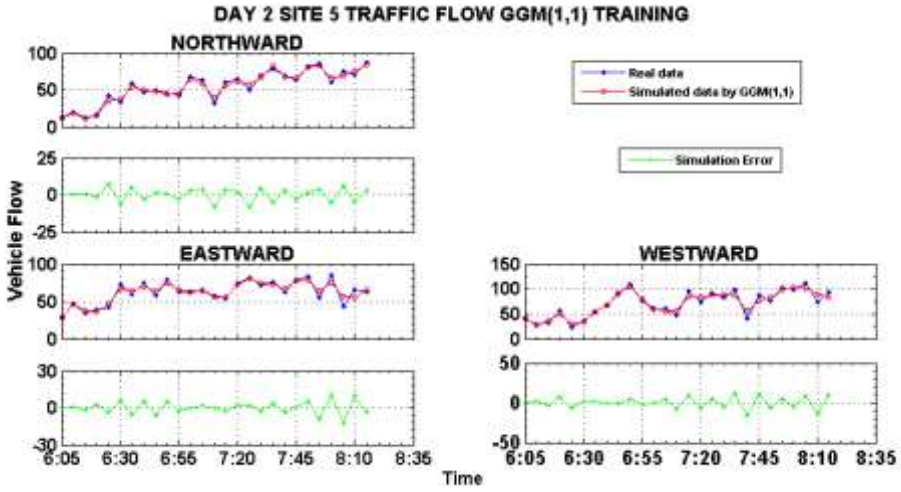

Figure 4.41c: Vehicle Flow GGM(1,1) Training

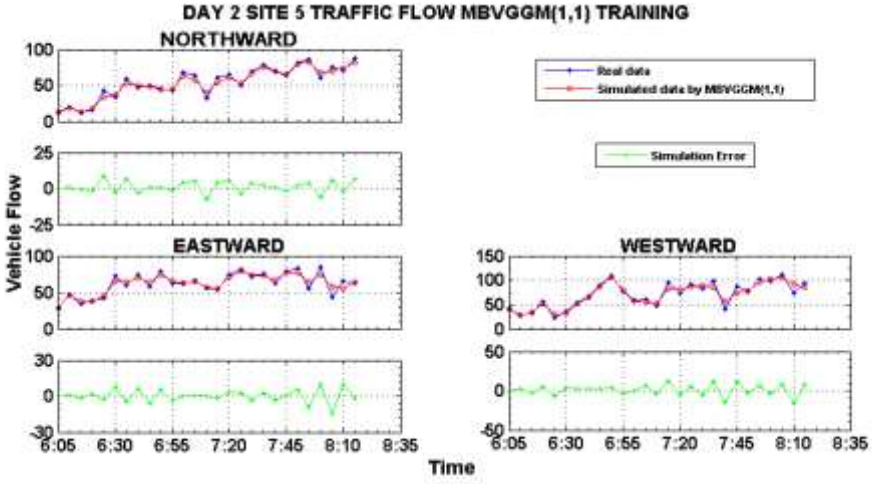

Figure 4.41d: Vehicle Flow MBVGGM(1,1) Training

The simulation error curves for OGM(1,1) and MBVGM(1,1) in Figures 4.41a and 4.41b are similar. Thus, these models have similar performance. Then it means MBV's impact on OGM(1,1)'s fitting accuracy is not significant enough to improve its accuracy. On the other hand the simulation curves for GGM(1,1) and MBVGGM(1,1), in Figures 4.41c and 4.41d, are almost approaching the zero level meaning that these models are accurate compared to OGM(1,1) and MBVGM(1,1). This is because the DGT has a great effect on OGM(1,1)'s fitting accuracy. DGT has actually improved OGM(1,1)'s fitting accuracy.

b) Testing the Grey Models in Short-Term Forecasting

To compute the short-term forecasts for the OGM(1,1) and MBVGM(1,1) (4.30) and (4.31) were extrapolated three points into the future. Also the time response functions of GGM(1,1) and MBVGGM(1,1) were extrapolated during simulation. The simulation data from these four models were tabulated in Table 4.39. Plotted in Figures 4.42a to 4.42d are the real, simulated and forecasted errors.

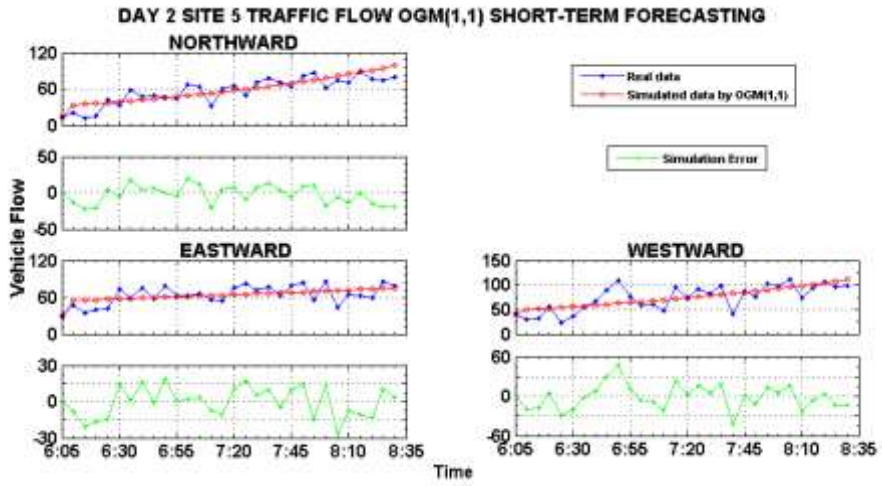

Figure 4.42a: Short-Term Vehicle Flow Forecast by OGGM(1,1)

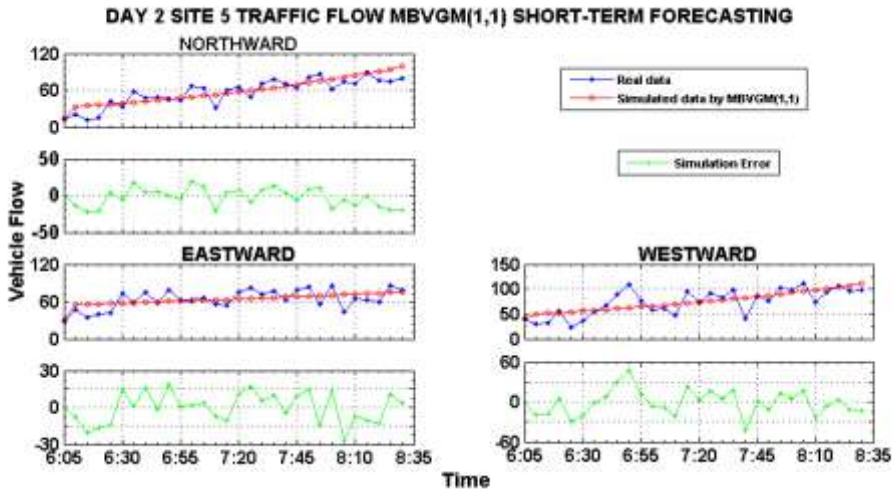

Figure 4.42b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

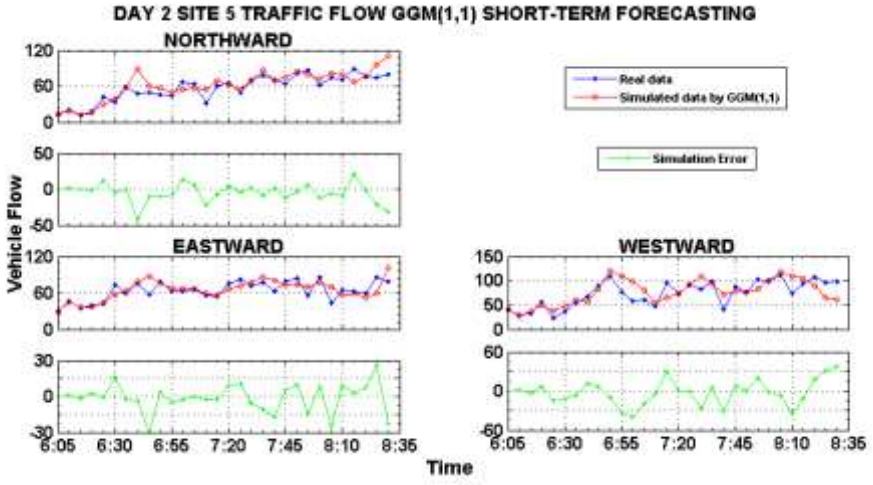

Figure 4.42c: Short-Term Vehicle Flow Forecast by GGM(1,1)

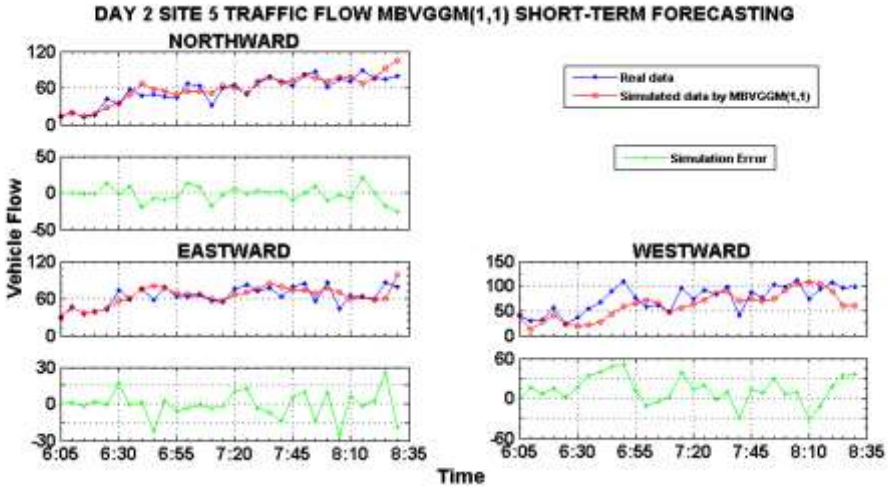

Figure 4.42d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

The short-term forecasting simulation error curves of OGM(1,1) and MBVGM(1,1) are not good enough to deduce that these models are good in short-term forecasting. Even their fitting of the simulated data onto the real data is not promising. See Figures 4.42a and 4.42b. In the northward and eastward directions the MBVGGM(1,1) had a good forecasting compared to its forecasting in the westward direction. This is in consideration of how well the simulated and real curves fit onto each other. See Figure 4.42d.

Table 4.39: Original and Improved Grey Models' Simulation Values (D2S5 Westward Direction)

Raw Data		Grey Model			
Data Point	Real Value	OGM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
	Training		Fitted Values		
1	40	40.0000	40.0000	40.0000	40.0000
2	29	49.3882	49.4691	27.0257	26.7847
3	33	50.8294	50.9086	37.0917	35.0883
4	57	52.3126	52.3899	50.0132	51.2142
5	24	53.8391	53.9143	29.7797	30.4660
6	36	55.4101	55.4830	34.0938	32.1999
7	55	57.0270	57.0974	53.1194	51.7536
8	67	58.6910	58.7588	67.9218	64.2509
9	90	60.4036	60.4685	90.8902	87.1323
10	110	62.1662	62.2280	105.0620	106.4460
11	76	63.9802	64.0387	78.3997	80.2238
12	59	65.8472	65.9020	60.3726	59.5565
13	60	67.7686	67.8196	54.9497	53.3005
14	48	69.7461	69.7930	56.0156	52.3703
15	95	71.7813	71.8238	85.6378	82.1448
16	75	73.8759	73.9136	81.6227	80.4531
17	92	76.0316	76.0643	87.2150	86.6750
18	83	78.2502	78.2776	88.4329	88.3497
19	98	80.5336	80.5553	86.6147	85.3680
20	40	82.8836	82.8992	56.0016	55.4777
21	87	85.3021	85.3113	75.4192	74.7753
22	76	87.7912	87.7937	82.5664	79.3328
23	103	90.3530	90.3482	97.9621	95.8523
24	98	92.9895	92.9771	102.6840	101.9626
25	112	95.7030	95.6825	102.4037	103.8674
26	75	98.4956	98.4666	88.9277	91.5262
27	94	101.3697	101.3317	83.7825	85.4466
Testing			Short-Term Forecasted Values		
28	107	104.3277	104.2802	89.2362	88.5903
29	95	107.3720	107.3145	64.2311	59.7076
30	98	110.5051	110.4370	60.8991	61.0283

c) Evaluation of the Grey Models

Simulation data for all the grey models were as recorded in Table 4.39. These data were for the westward direction. The overall training and short-term forecasting error evaluation for day 2 site 5 were as tabulated in Table 4.40.

In obtaining the fitted values, shown in Table 4.39, of the four models the corresponding fitting errors indicated in Table 4.40 clearly shows that GGM(1,1) is more accurate in vehicle flow fitting. GGM(1,1)'s fitting errors were lower in all directions except for the northward direction for which case MAPD value indicated that MBVGGM(1,1) had a lower error compared to GGM(1,1). The highest fitting accuracy obtained by GGM(1,1) was 100-RMSPE=100-5.9430= 94.0570%, in the northward direction. In short-term forecasting MBVGGM(1,1) emerged the most accurate in both northward and eastward directions. It had the highest accuracy of 100-MAPD=100-12.1235=87.8765% in the eastward direction.

Table 4.40: Day 2 Site 5 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional		Improved	
		GM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training					
Northbound	RMSE	11.8009	11.8027	4.2200	4.2362
	RMSPE	15.2739	15.2491	5.9430	6.3853
	MAE	9.8062	9.8030	3.6005	3.4991
	MAPD	18.1223	18.1164	6.6539	6.4664
Eastbound	RMSE	12.4874	12.4875	5.2222	5.3362
	RMSPE	17.6632	17.6592	7.6029	7.7609
	MAE	10.4683	10.4671	4.0595	4.1095
	MAPD	16.8742	16.8724	6.5437	6.6242
Westbound	RMSE	19.3844	19.3849	7.2472	7.4807
	RMSPE	22.6057	22.5819	8.5235	8.9327
	MAE	15.2601	15.2608	5.9781	6.1332
	MAPD	21.5493	21.5502	8.4418	8.6608
Short-Term Forecasting					
Northbound	RMSE	12.6156	12.6162	13.3648	10.3546
	RMSPE	17.2737	17.2538	19.6135	16.2785
	MAE	10.6523	10.6489	9.3550	7.6323
	MAPD	18.8983	18.8922	16.5966	13.5405
Eastbound	RMSE	12.2757	12.2760	11.8812	10.8036
	RMSPE	16.9024	16.9013	17.5509	16.1098
	MAE	10.3532	10.3539	8.6342	7.6742
	MAPD	16.3557	16.3569	13.6401	12.1235
Westbound	RMSE	18.6743	18.6723	19.0545	23.8856
	RMSPE	20.6849	20.6599	21.3681	29.4039
	MAE	14.6524	14.6504	14.4098	18.9653
	MAPD	19.8721	19.8694	19.5431	25.7214

III. Site 6: City Hall Way-Wabera Street T-Roundabout

a) Traffic Flow Training

The real vehicle flow data for day 2 site 6 (see Appendix IV Table 13), for which the eastward direction data was as recorded in Table 4.43, were used. The first 27 data points were subjected to OGM(1,1) training. The AGO and MGO on the data resulted to the parameters in Table 4.41. From Table 4.41 and considering the eastward direction (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 939.2968\, e^{0.0310r} - 912.2968,\ r = 0,1,2,\cdots,m-1 \tag{4.32}$$

From (4.32) OGM(1,1)'s simulated values were computed and indicated in Table 4.43. Figure 4.43a is a plot of the real, simulated and fitting error data for OGM(1,1).

Table 4.41: Day 2 Site 6 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Southward	-0.0354	14.4567
Eastward	-0.0310	28.2812
Westward	-0.0316	22.7895

Next OGM(1,1)'s background value was modified and the developed MBVGM(,1)'s a and b parameters were as shown in Table 4.42. For southward and westward directions MATLAB simulation output was "NaN". Therefore, considering the eastward direction (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 944.2071\, e^{0.0309r} - 917.2071,\ r = 0,1,2,\cdots,m-1 \tag{4.33}$$

The IAGO operation was applied on the computed values of (4.33). Thus MBVGM(,1)'s simulation values in the eastward direction were calculated and recorded in Table 4.43. The corresponding plots are shown in Figure 4.43b.

Table 4.42: Day 2 Site 6 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Southward	NaN	NaN
Eastward	-0.0309	28.3417
Westward	NaN	NaN

Figures 4.43c and 4.43d are the plots of GGM(1,1) and MBVGGM(1,1) respectively. The simulation values of GGM(1,1) and MBVGGM(1,1) were as tabulated in Table 4.43.

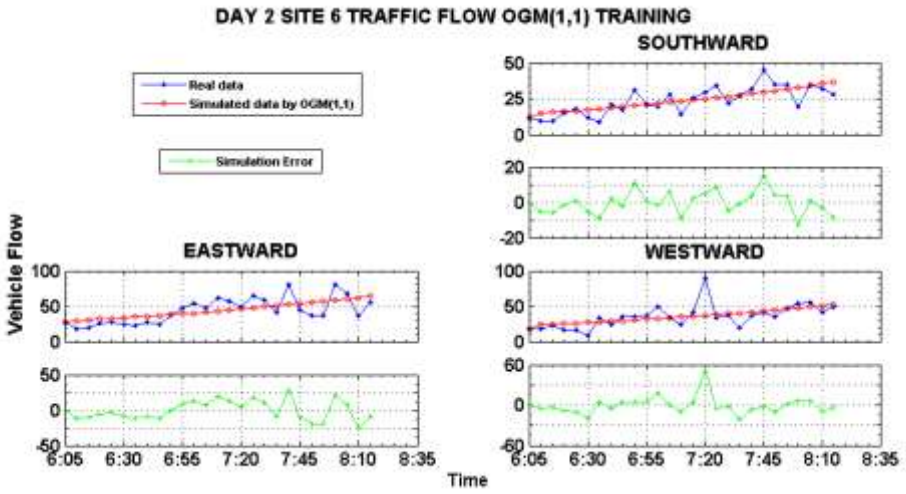

Figure 4.43a: Vehicle Flow OGM(1,1) Training

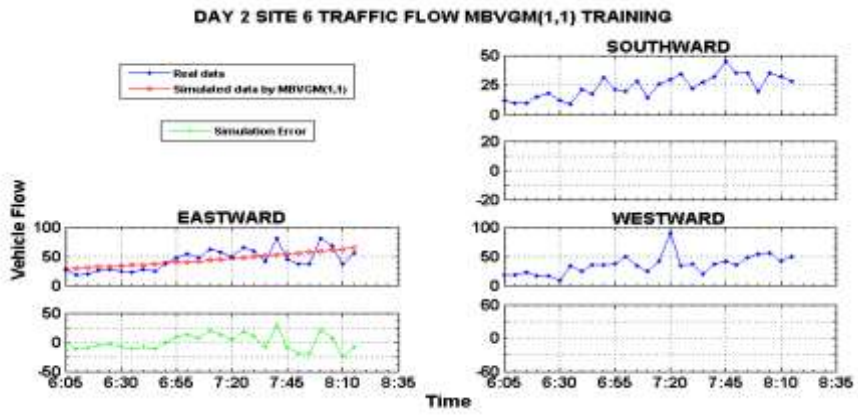

Figure 4.43b: Vehicle Flow MBVGM(1,1) Training

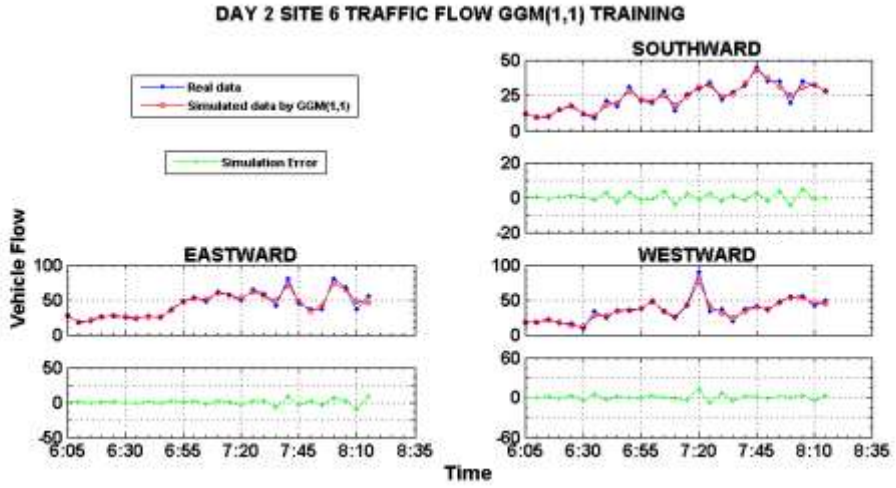

Figure 4.43c: Vehicle Flow GGM(1,1) Training

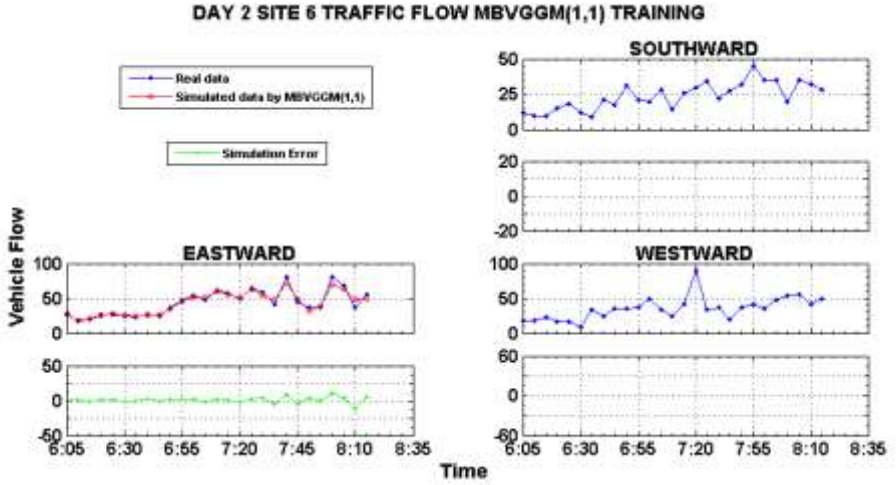

Figure 4.43d: Vehicle Flow MBVGGM(1,1) Training

The following observations can be derived from Figures 4.43a-4.43d. The MBVGM(1,1) and MBVGGM(1,1) did not yield any simulation result in both southward and westward directions (see Figures 4.43b and 4.43d). For this case the MATLAB output was "NaN". This "Not a Number" result from MATLAB means that the data matrix A was singular to working precision. And, therefore, the parameters a and b could not be computed. However, in the eastward direction, MBVGGM(1,1) had a very smooth fitting (see Figure 4.43d). So it was more accurate in fitting compared to the MBVGM(1,1). Thus the DGT real improves the fitting accuracy of the OGM(1,1). Comparison of OGM(1,1) and GGM(1,1) reveals the power of the DGT in

improving the fitting accuracy of the OGM(1,1). From Figures 4.43a and 4.43c one can see that OGM(1,1) has poor fitting result whereas GGM(1,1) has very neat fitting result in all directions of vehicle flow. Moreover, in the eastward direction GGM(1,1) and MBVGGM seems to have similar results in vehicle flow fitting.

b) Testing the Grey Models in Short-Term Forecasting

In short-term forecasting the formulated time response functions were extrapolated three points into the future. This included extrapolation of (4.32) and (4.33) for OGM(1,1) and MBVGM(1,1), respectively. As a result, plotted in Figures 4.44a-4.44d are the real, simulated and error values for the four grey models. Note that in short-term forecasting the parameters a and b remain the same as those obtained in training of the models.

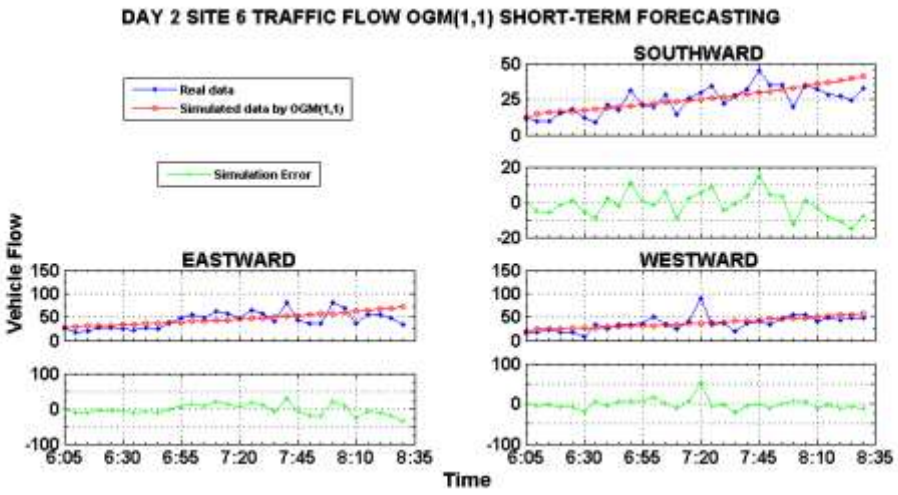

Figure 4.44a: Short-Term Vehicle Flow Forecast by OGM(1,1)

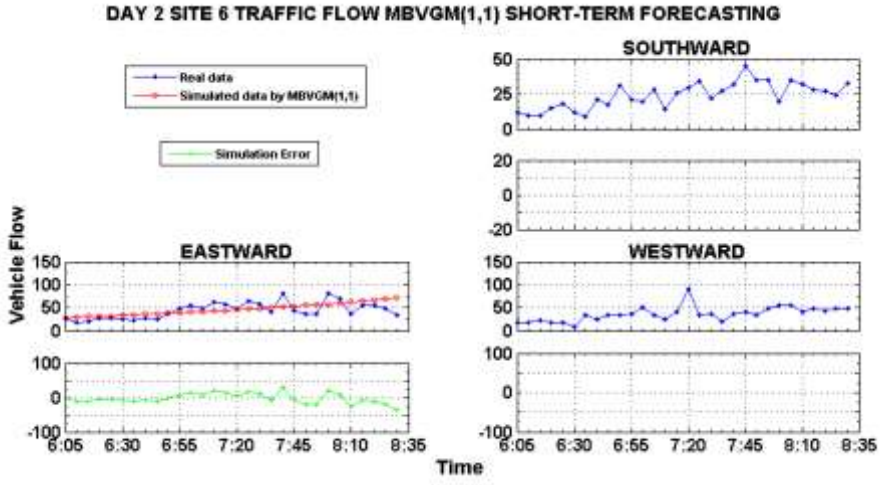

Figure 4.44b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

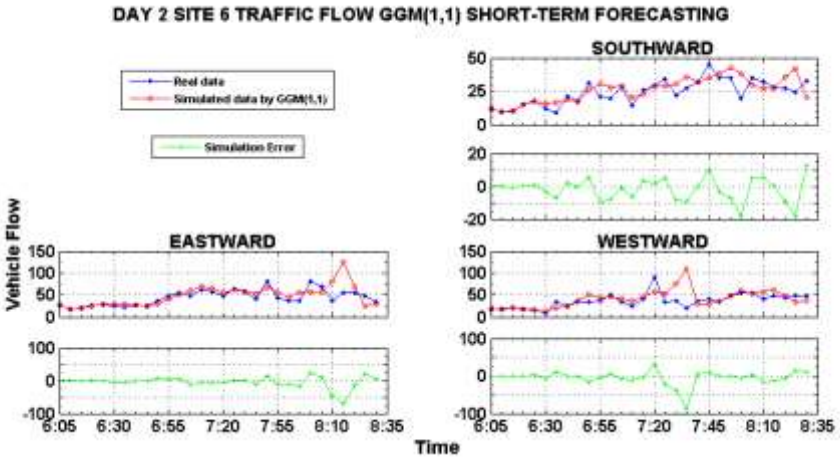

Figure 4.44c: Short-Term Vehicle Flow Forecast by GGM(1,1)

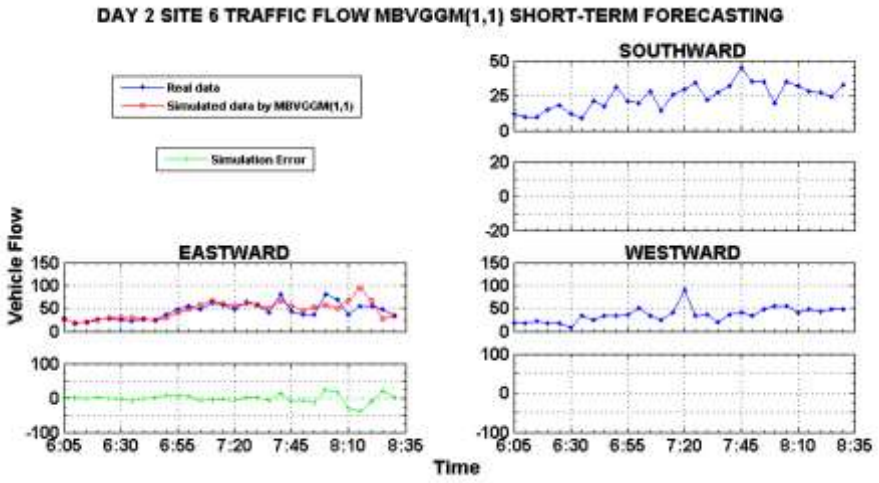

Figure 4.44d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

Again in this short-term forecasting the MBVGM(1,1) and MBVGGM(1,1) did not give any result for the southward and westward directions because of the "NaN" condition (see Figures 4.44b and 4.44d). However, close observation of the results of Figures 4.44c and 4.44d shows that GGM(1,1) and MBVGGM(1,1) had a good short-term forecasting fit, especially at the last point of forecast (i.e. the 30th point).

c) Evaluation of the Grey Models

To evaluate the grey models the simulation data were recorded in Table 4.43. The training data set included the first 27 data points as seen in Table 4.43. Therefore, the testing data set was the last three data points of the real data series. Now, the overall training and short-term forecasting error evaluation for day 2 site 6 was as tabulated in Table 4.44.

Table 4.43: Original and Improved Grey Models' Simulation Values (D2S6 Eastward Direction)

Raw Data		Grey Model			
Data Point	**Real Value**	**OGM(1,1)**	**MBVGM(1,1)**	**GGM(1,1)**	**MBVGGM(1,1)**
Training			**Fitted Values**		
1	27	27.0000	27.0000	27.0000	27.0000
2	18	29.5738	29.6310	17.6442	17.5813
3	20	30.5046	30.5606	20.6083	20.4476
4	26	31.4647	31.5194	25.6660	25.6026
5	28	32.4550	32.5082	27.6943	27.7968
6	25	33.4765	33.5281	25.0108	25.0025
7	23	34.5302	34.5799	23.7179	23.5505
8	27	35.6170	35.6648	25.5859	25.3000
9	25	36.7380	36.7837	26.4837	25.8400
10	37	37.8943	37.9377	35.0886	34.8107
11	48	39.0870	39.1279	47.4085	45.9277
12	54	40.3172	40.3554	51.7846	50.9404
13	48	41.5862	41.6215	50.4922	50.1665
14	62	42.8951	42.9272	59.7389	59.6591
15	57	44.2452	44.2740	56.3449	55.9214
16	49	45.6377	45.6630	52.1678	51.4649
17	65	47.0742	47.0955	62.4608	62.9380
18	58	48.5558	48.5730	55.2356	53.9676
19	42	50.0840	50.0969	49.7973	46.9571
20	80	51.6604	51.6685	71.1773	72.1542
21	44	53.2864	53.2895	48.0478	48.7943
22	36	54.9635	54.9613	33.3354	31.4997
23	37	56.6935	56.6856	41.2661	37.8098
24	80	58.4778	58.4640	73.6278	70.1540
25	68	60.3184	60.2982	65.2852	63.7643
26	37	62.2169	62.1899	47.7931	48.2412
27	55	64.1751	64.1409	45.7920	48.7831
Testing			**Short-Term Forecasted Values**		
28	54	66.1950	66.1532	68.8852	65.8112
29	47	68.2784	68.2286	25.1986	25.8452
30	35	70.4274	70.3691	29.8434	34.7702

Table 4.44: Day 2 Site 6 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional	Improved		
		GM(1,1)	MBVGM(1,1	GGM(1,1)	MBVGGM(1,1)
Training					
Southbound	RMSE	6.3061	NaN	2.2967	NaN
	RMSPE	23.5178	NaN	7.9946	NaN
	MAE	4.9129	NaN	1.8740	NaN
	MAPD	20.7586	NaN	7.9183	NaN
Eastbound	RMSE	13.4899	13.4906	4.2154	4.1756
	RMSPE	27.2923	27.2792	8.5107	8.9343
	MAE	11.5859	11.5929	2.9820	2.9805
	MAPD	26.6002	26.6164	6.8465	6.8430
Westbound	RMSE	13.2014	NaN	4.2114	NaN
	RMSPE	44.6108	NaN	12.1900	NaN
	MAE	8.4321	NaN	3.0319	NaN
	MAPD	23.9145	NaN	8.5989	NaN
Short-Term Forecasting					
Southbound	RMSE	7.0524	NaN	7.2741	NaN
	RMSPE	25.3940	NaN	24.2167	NaN
	MAE	5.5609	NaN	5.3925	NaN
	MAPD	23.0742	NaN	22.3753	NaN
Eastbound	RMSE	15.0222	15.0147	17.8207	12.6729
	RMSPE	28.7732	28.7518	34.2894	25.4816
	MAE	12.7240	12.7253	10.5803	8.5695
	MAPD	29.0944	29.0976	24.1927	19.5948
Westbound	RMSE	12.8779	NaN	20.4949	NaN
	RMSPE	42.7865	NaN	35.7079	NaN
	MAE	8.5294	NaN	11.3961	NaN
	MAPD	23.4538	NaN	31.3368	NaN

There were no results from MBVGM(1,1) and MBVGGM(1,1) simulation in the southward and westward directions because of the "NaN" condition. Table 4.43 shows the simulation results for the eastward direction and from Table 4.44 it is noticeable that MBVGGM(1,1) was the most accurate in the eastward direction. MBVGGM(1,1) had the highest forecasting accuracy of 100-MAPD=100-19.5948=80.4052%, in the eastward direction. In the southward direction GGM(1,1) was the most accurate in short-term forecasting. It had the highest accuracy of 100-22.3753=77.6247% compared with OGM(1,1) which had an accuracy of 100-23.0742=76.9258%. That is in consideration of the MAPD error indicator. In fitting of vehicle flow GGM(1,1) had lower fitting errors compared to the OGM(1,1), in the southward direction. In the

eastward direction, MBVGGM(1,) emerged the best in vehicle flow fitting with an accuracy of 100-MAPD=100-6.8430=93.1570%.

IV. Site 7: Haile Selassie Avenue-Moi Avenue Roundabout

a) Traffic Flow Training

The real vehicle flow data for day 2 site 7 (see Appendix IV Table 14) were used. The westward direction data were as recorded in Table 4.47. The first 27 data points were subjected to OGM(1,1) training. The AGO and MGO on the data resulted to the parameters in Table 4.45. From Table 4.45 and considering the westward direction (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 9{,}164.8152\, e^{0.0092r} - 9{,}112.8152,\ r = 0,1,2,\cdots,m-1 \tag{4.34}$$

From (4.34) OGM(1,1)'s simulated values were computed and indicated in Table 4.47. Figure 4.45a is a plot of the real, simulated and fitting error data for OGM(1,1).

Table 4.45: Day 2 Site 7 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Southward	-0.0035	76.3657
Eastward	0.0134	89.4151
Westward	-0.0092	83.8379

Next OGM(1,1)'s background value was modified and the developed MBVGM(,1)'s a and b parameters were as shown in Table 4.46. For the southward direction MATLAB simulation output was "NaN". Considering the westward direction (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 9{,}278.2418\, e^{0.0091r} - 9{,}226.2418,\ r = 0,1,2,\cdots,m-1 \tag{4.35}$$

The IAGO operation was applied on the computed values of (4.35). Thus MBVGM(,1)'s simulation values in the westward direction were calculated and recorded in Table 4.47. The corresponding plots are shown in Figure 4.45b.

Table 4.46: Day 2 Site 7 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Southward	NaN	NaN
Eastward	0.0135	89.5711
Westward	-0.0091	83.9588

Figures 4.45c and 4.45d are the plots of GGM(1,1) and MBVGGM(1,1) traffic flow training, respectively. The simulation values of GGM(1,1) and MBVGGM(1,1) were as well tabulated in Table 4.47.

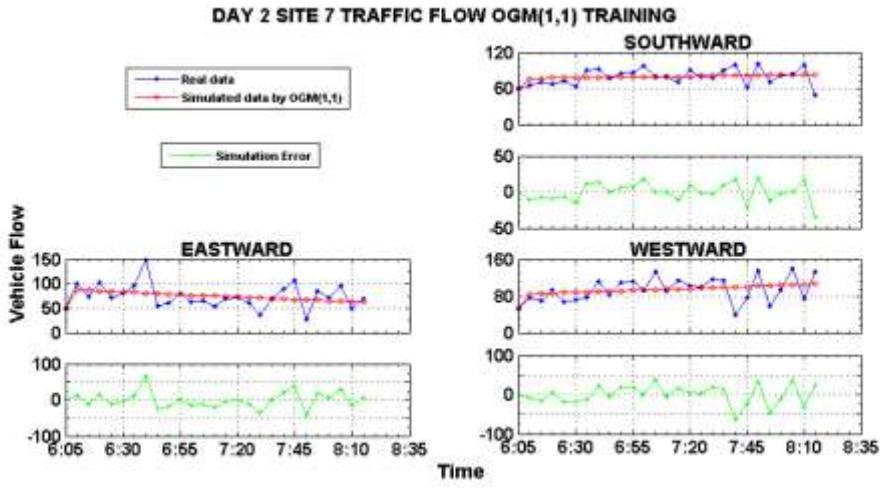

Figure 4.45a: Vehicle Flow OGM(1,1) Training

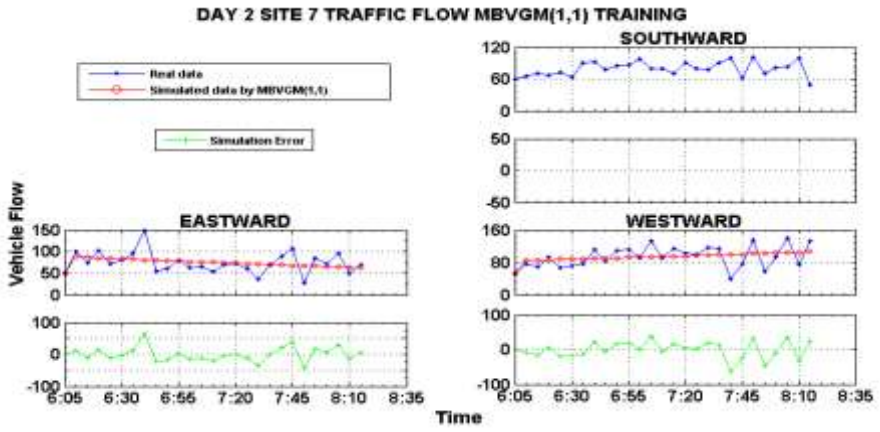

Figure 4.45b: Vehicle Flow MBVGM(1,1) Training

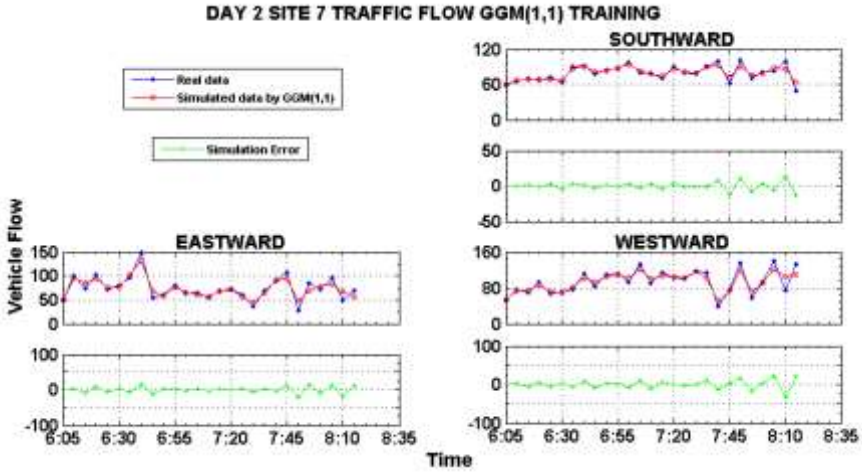

Figure 4.45c: Vehicle Flow GGM(1,1) Training.

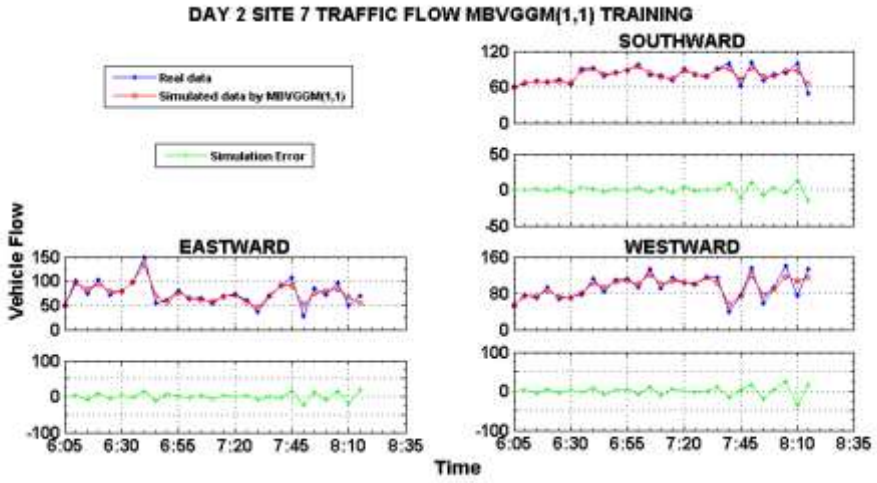

Figure 4.45d: Vehicle Flow MBVGGM(1,1) Training

The following observations can be derived from Figures 4.45a-4.45d. The MBVGM(1,1) did not yield any simulation result in the southward direction. This was because the MATLAB output was a "NaN" condition. As before it means that the data matrix A was singular to working precision. And, therefore, the parameters a and b could not be computed. Once more, comparison of OGM(1,1) and GGM(1,1) reveals the ability of the DGT in improving the fitting accuracy of the OGM(1,1). From Figures 4.45a and 4.45c one can see that OGM(1,1) has poor fitting result whereas GGM(1,1) has very neat fitting result in all directions of vehicle flow. Thus the DGT

real improves the fitting accuracy of the OGM(1,1). Moreover, in all directions GGM(1,1) and MBVGGM seems to have similar results in vehicle flow fitting.

b) Testing the Grey Models in Short-Term Forecasting

In short-term forecasting the formulated time response equations were extrapolated three points into the future. This included extrapolation of (4.34) and (4.35) for OGM(1,1) and MBVGM(1,1), respectively. As a result, plotted in Figures 4.46a-4.46d are the real, simulated and error values for the four grey models. The parameters a and b remains the same as those obtained in training of the models.

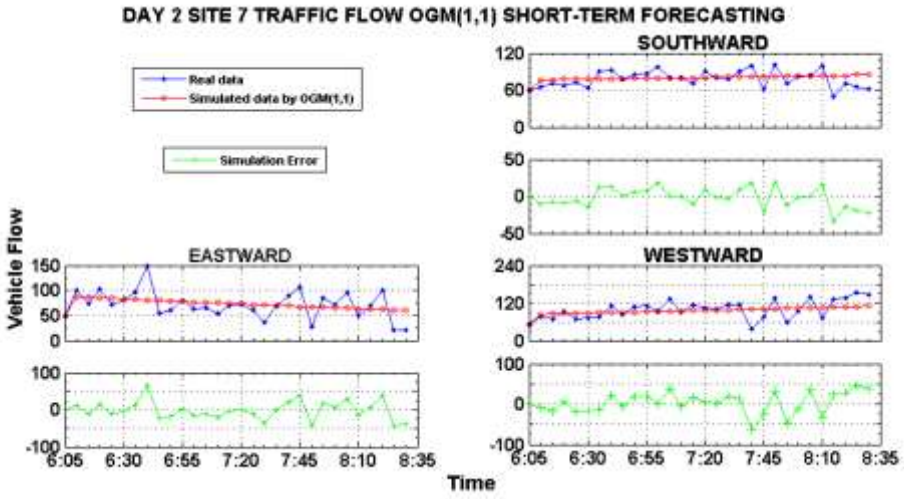

Figure 4.46a: Short-Term Vehicle Flow Forecast by OGM(1,1)

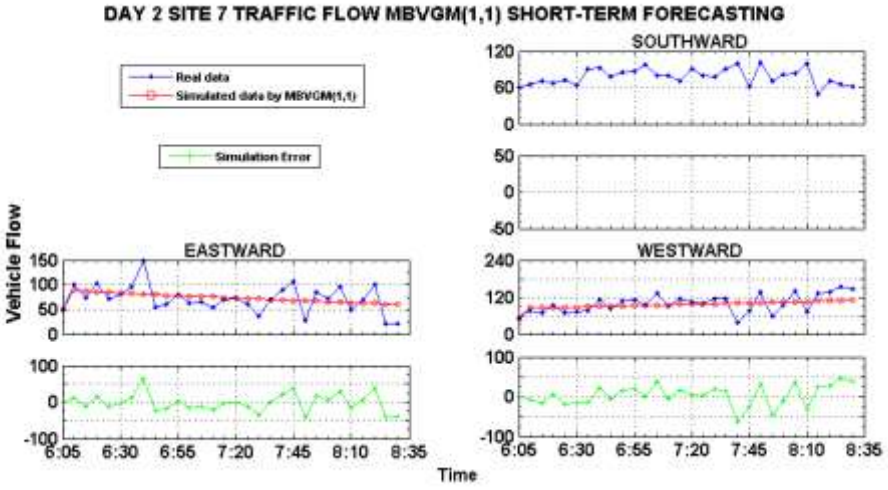

Figure 4.46b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

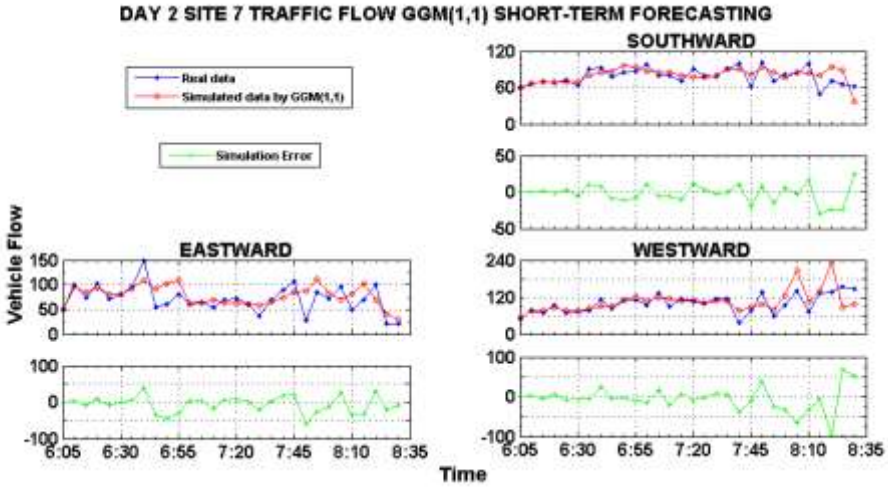

Figure 4.46c: Short-Term Vehicle Flow Forecast by GGM(1,1)

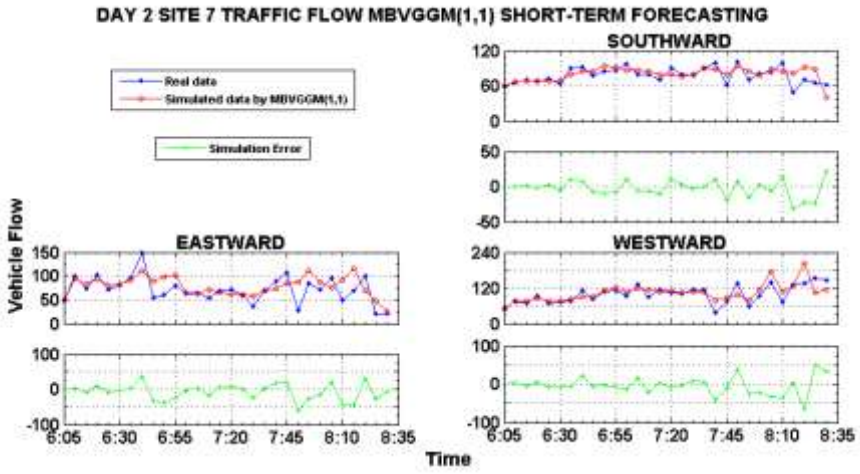

Figure 4.46d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

Notice, from Figure 4.46b, that in this short-term forecasting the MBVGM(1,1) did not give any result in the southward direction because of the "NaN" condition. Moreover, close observation of Figures 4.46c and 4.46d shows that GGM(1,1) and MBVGGM(1,1) had a good short-term forecasting fit, especially at the last point of forecast (i.e. the 30^{th} point).

c) Evaluation of the Grey Models

To evaluate the grey models the simulation data were recorded in Table 4.47. The training data set included the first 27 data points as seen in Table 4.47. Therefore, the testing data set was the last three data points of the real data series. Now, the overall training and short-term forecasting error evaluation for day 2 site 7 was as tabulated in Table 4.48.

Table 4.47: Original and Improved Grey Models' Simulation Values (D2S7 Westward Direction)

Raw Data		Grey Model			
Data Point	**Real Value**	**OGM(1,1)**	**MBVGM(1,1)**	**GGM(1,1)**	**MBVGGM(1,1)**
Training		**Fitted Values**			
1	52	52.0000	52.0000	52.0000	52.0000
2	77	84.7052	84.8180	74.1842	74.3690
3	69	85.4879	85.5940	74.9112	74.2435
4	92	86.2779	86.3771	85.9147	86.2077
5	68	87.0751	87.1673	72.1525	72.6156
6	72	87.8797	87.9648	69.1600	69.0576
7	76	88.6918	88.7696	80.8114	80.1021
8	112	89.5113	89.5817	103.3332	102.6636
9	84	90.3384	90.4013	92.0796	92.3201
10	109	91.1732	91.2283	105.6060	105.4393
11	112	92.0157	92.0629	108.2952	107.6048
12	93	92.8660	92.9052	102.0390	100.8066
13	132	93.7241	93.7552	120.5943	120.4702
14	90	94.5902	94.6129	100.1951	100.9096
15	113	95.4642	95.4785	107.0503	107.3884
16	103	96.3464	96.3520	103.8286	103.8715
17	99	97.2366	97.2335	101.3463	100.8253
18	116	98.1352	98.1231	115.5773	115.2391
19	113	99.0420	99.0208	103.3312	102.7761
20	38	99.9572	99.9267	51.7533	54.7926
21	76	100.8808	100.8409	73.9851	74.2658
22	136	101.8130	101.7635	118.5632	119.3986
23	57	102.7538	102.6945	72.2067	75.9019
24	93	103.7033	103.6340	90.5787	89.2475
25	141	104.6616	104.5821	120.0440	115.7626
26	73	105.6287	105.5389	104.6700	108.2316
27	132	106.6048	106.5045	109.8781	114.8672
Testing		**Short-Term Forecasted Values**			
28	135	107.5898	107.4788	230.3233	201.3029
29	155	108.5840	108.4621	86.8386	105.1925
30	147	109.5874	109.4544	95.3443	113.2102

Table 4.48: Day 2 Site 7 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional	Improved		
		GM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training					
Southbound	RMSE	12.5224	NaN	5.7827	5.9604
	RMSPE	14.3681	NaN	6.8284	6.9715
	MAE	9.6038	NaN	4.2813	4.3584
	MAPD	12.1000	NaN	5.3941	5.4912
Eastbound	RMSE	22.4726	22.4727	9.3273	9.9330
	RMSPE	31.6569	31.6562	10.4956	10.4962
	MAE	16.9399	16.9449	7.4719	7.7622
	MAPD	22.8689	22.8757	10.0871	10.4789
Westbound	RMSE	24.0552	24.0554	11.2396	11.9574
	RMSPE	21.3298	21.3400	11.0219	11.3301
	MAE	19.1415	19.1479	8.3666	8.7357
	MAPD	20.4438	20.4507	8.9359	9.3300
		Short-Term Forecasting			
Southbound	RMSE	13.2950	NaN	12.5009	12.1167
	RMSPE	15.2178	NaN	13.4132	12.9865
	MAE	10.4975	NaN	9.5725	9.3526
	MAPD	13.4584	NaN	12.2725	11.9905
Eastbound	RMSE	24.7296	24.7197	23.1889	23.6075
	RMSPE	32.2171	32.2267	24.3525	23.8018
	MAE	19.2058	19.2046	17.7577	17.8862
	MAPD	26.9114	26.9097	24.8824	25.0624
Westbound	RMSE	25.7741	25.7919	30.9565	23.6103
	RMSPE	23.0066	23.0423	33.2970	23.7609
	MAE	20.9353	20.9533	20.7724	17.1612
	MAPD	21.1824	21.2006	21.0176	17.3638

There were no results from MBVGM(1,1) simulation in the southward direction because of the "NaN" condition. Table 4.47 shows the simulation results for the westward direction and from Table 4.48 it is noticeable that GGM(1,1) was the most accurate in all directions. GGM(1,1) had the highest fitting accuracy of 100-MAPD=100-5.3941=94.6059%, in the southward direction. In the southward direction MBVGGM(1,1) was the most accurate in short-term forecasting. It had the highest accuracy of 100-11.9905=88.0095% compared with OGM(1,1) which had an accuracy of 100-13.4584=86.5416%. That is in consideration of the MAPD error indicator.

4.3.3.3 Day Three Vehicle Traffic Flow Modelling and Short-Term Forecasting

Similarly, for day three vehicle traffic flow moving in all directions for the following four sites were modelled.

- Day 3 Site 1 (D3S1): Haile Selassie Roundabout,
- Day 3 Site 2 (D3S2): Kenyatta Avenue Uhuru Highway Roundabout,
- Day 3 Site 4 (D3S4): Kenyatta Avenue-Moi Avenue-Mondlane Street Junction and
- Day 3 Site 7 (D3S7): Haile Selassie Avenue-Moi Avenue Roundabout.

Data from the four sites were employed and Figures 4.47a to 4.54d presents the four cases in training and short-term forecasting.

I. Site 1: Haile Selassie Roundabout

a) Traffic Flow Training

The AGO and MGO operations were performed on the real vehicle flow data of Appendix IV Table 15, in order to create the data matrix A. Then the parameters a and b for OGM(1,1) were computed and Table 4.49 contains these parameters. Using the parameters in Table 4.49 the time response equation for each direction can be developed. For instance, in the northward direction (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 89{,}461.3636\, e^{0.0022r} - 89{,}380.3636,\ r = 0,1,2,\cdots,m-1 \qquad (4.36)$$

Based on IAGO, OGM(1,1)'s simulation data was obtained from (4.36) and recorded in Table 4.51. Figure 4.47a is the plot of real, simulated and error data of OGM(1,1).

Table 4.49: Day 3 Site 1 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	**Control Variable, b**
Northward	-0.0022	196.6368
Southward	-0.0157	87.4806
Eastward	-0.0360	41.4252
Westward	-0.0109	75.6861

Modification of OGM(1,1)'s background value develops the hybrid model MBVGM(11) whose parameters were obtained and recorded in Table 4.50. From (3.9) and using the parameters in Table 4.50 the time response equation for each direction can be obtained. In the northward direction the time response of MBVGM(1,1) was found to be:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 93{,}866.7143\, e^{0.0021r} - 93{,}785.7143,\ r = 0,1,2,\cdots,m-1 \qquad (4.37)$$

MBVGM(1,1)'s simulation data was obtained from (4.37) through IAGO operation. These data were recorded in Table 5.51. The real, simulated and error data were plotted in Figure 4.47b.

Table 4.50: Day 3 Site 1 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	**Control Variable, b**
Northward	-0.0021	196.9500
Southward	-0.0156	87.7911
Eastward	-0.0358	41.5409
Westward	-0.0110	75.6207

The GGM(1,1) and MBVGGM(1,1) involve a number of time response equations, each with different values of parameters a and b which are not necessarily the same. The time response equations are as many as the number of data groups formed. The parameters a and b are also as many as the number of data groups formed. Therefore, not provided here as before. Shown in Figures 4.47c and 4.47d are the plots of real, simulation and error data for GGM(1,1) and MBVGGM(1,1), respectively.

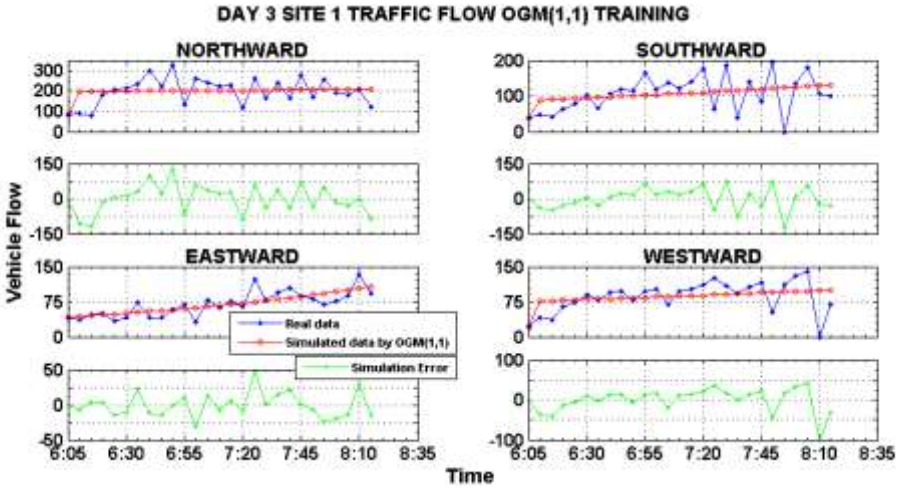

Figure 4.47a: Vehicle Flow OGM(1,1) Training

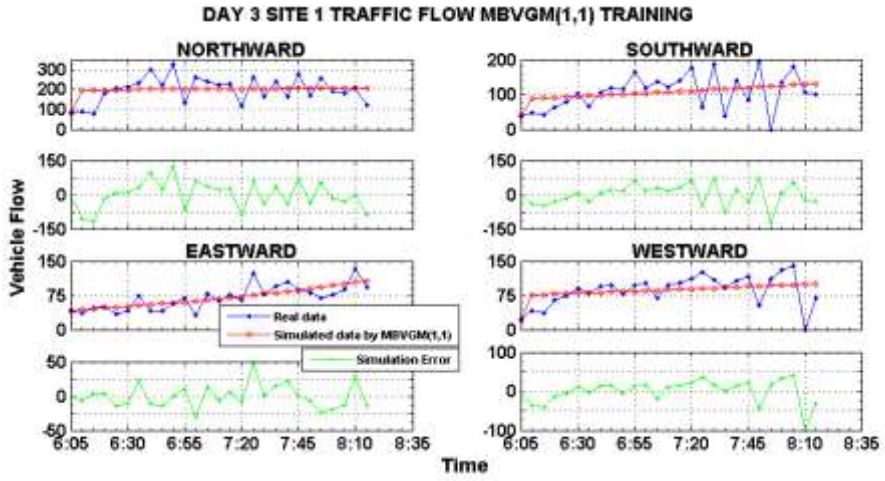

Figure 4.47b: Vehicle Flow MBVGM(1,1) Training

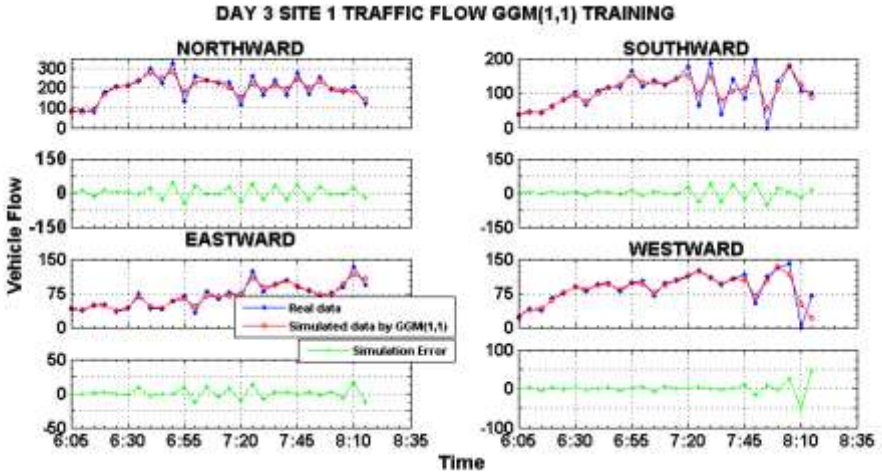

Figure 4.47c: Vehicle Flow GGM(1,1) Training

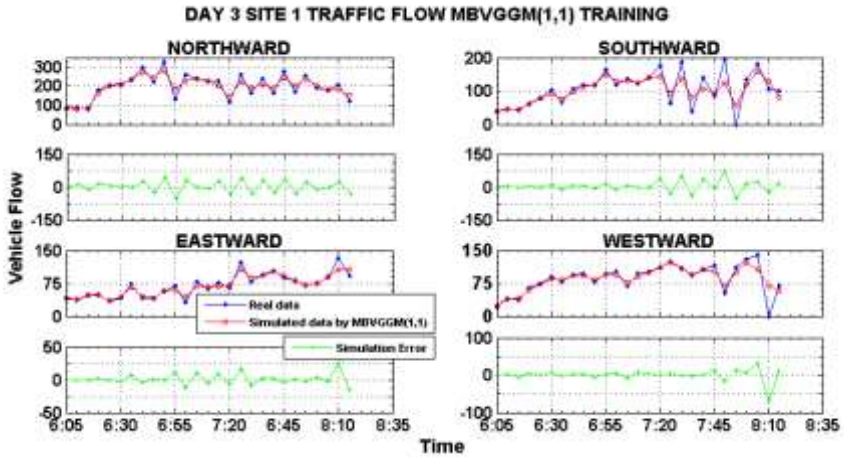

Figure 4.47d: Vehicle Flow MBVGGM(1,1) Training

As before the OGM(1,1) and MBVGM(1,1) have similar performance in vehicle flow training. This is evident from Figures 4.47a and 4.47b. Thus it can be deduced that MBV does not greatly improve the performance of the OGM(1,1). Its improvement on OGM(1,1)'s performance is too minute as compared to that of DGT. On the other hand GGM(1,1) and MBVGGM(1,1) have similar performance in vehicle flow fitting. Notice from Figures 4.47c and 4.47d that the simulation error curves are almost at zero level. This shows that the DGT greatly improves OGM(1,1)'s performance in vehicle flow fitting.

b) Testing the Grey Models in Short-Term Forecasting

The grey models were tested by extrapolating their time response equations. For OGM(1,1) and MBVGM(1,1) (4.36) and (4.37) were extrapolated three points into the future to forecast the last three points. The short-term forecasted values for each model were as tabulated in Table 4.51. Figures 4.48a-4.48d show the short-term forecasting plots for the four grey models.

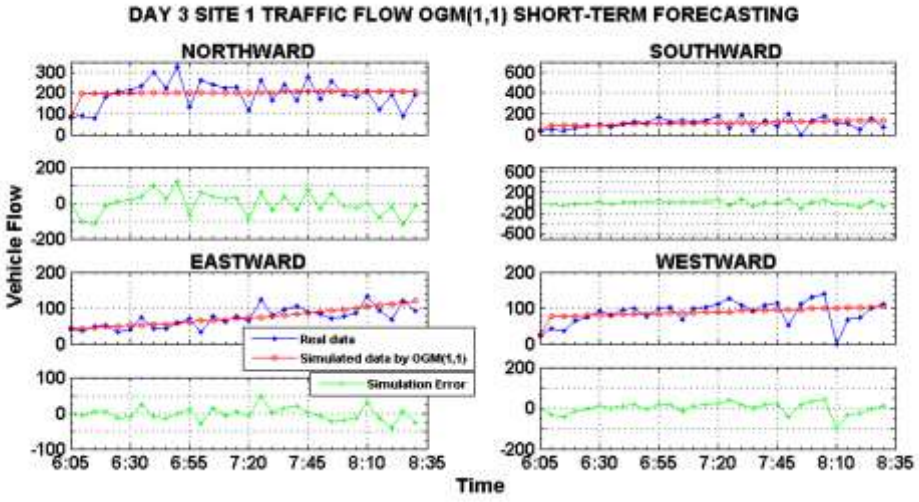

Figure 4.48a: Short-Term Vehicle Flow Forecast by OGM(1,1)

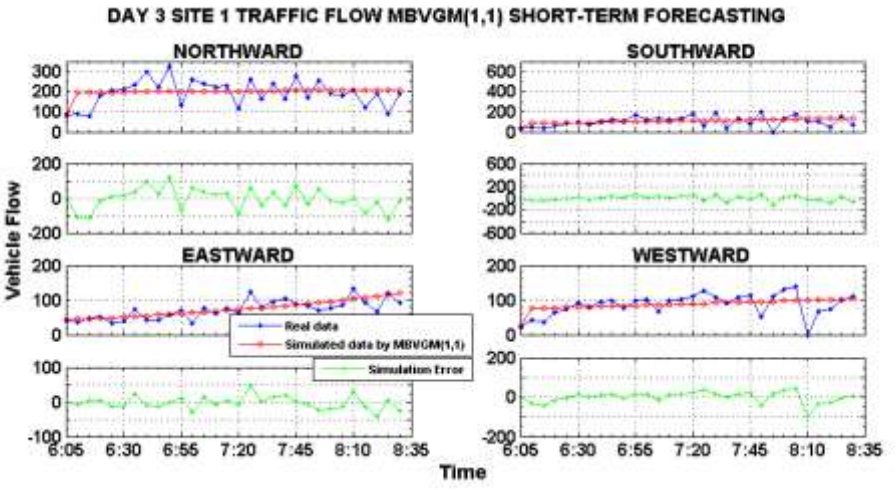

Figure 4.48b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

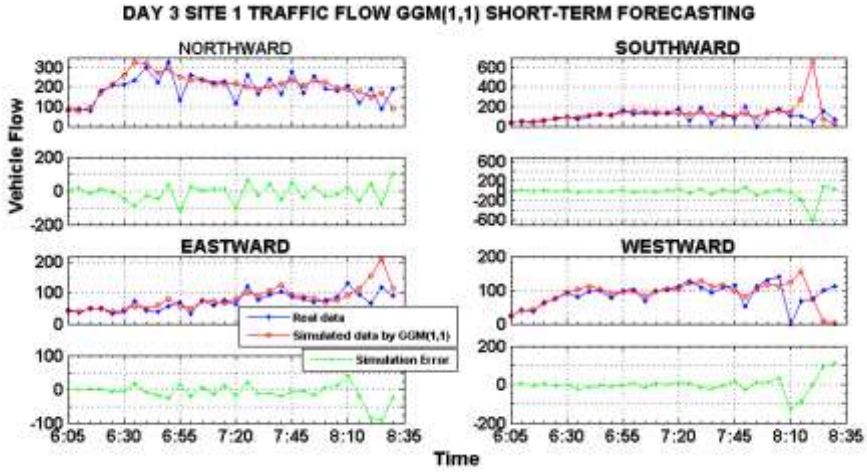

Figure 4.48c: Short-Term Vehicle Flow Forecast by GGM(1,1)

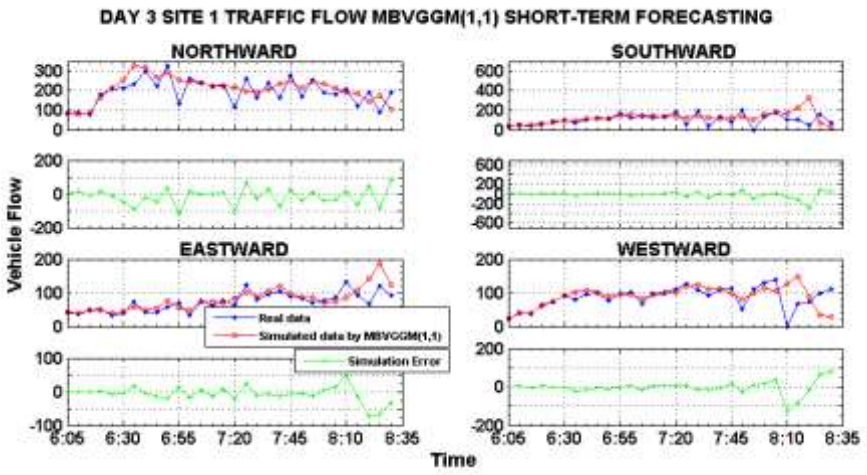

Figure 4.48d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

Examine the last three data points of the error curves in Figures 4.48a-4.48d and notice that MBVGM(1,1) is the most accurate in vehicle flow short-term forecasting, in the southward direction. This is because its error curve is almost at zero level. The performance of the rest of the models is not clearly distinctive. The performance is made clear in part (c) below.

c) Evaluation of the Grey Models

Table 4.51 shows the simulation performance (both training and short-term forecasting) of the four grey models, in the northward direction. In this simulation the

errors involved were calculated and the overall training and short-term forecasting error evaluation for day 3 site 1 was as tabulated in Table 4.52.

Table 4.51: Original and Improved Grey Models' Simulation Values (D3S1 Northward Direction)

Raw Data		Grey Model			
Data Point	Real Value	OGM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training		Fitted Values			
1	81	81.0000	81.0000	81.0000	81.0000
2	90	197.0363	197.3291	77.3545	77.2398
3	79	197.4753	197.7460	94.1042	90.9715
4	180	197.9153	198.1638	167.7417	164.2319
5	206	198.3563	198.5824	205.8256	199.9224
6	212	198.7983	199.0019	211.8578	207.3093
7	232	199.2412	199.4223	240.0815	233.8624
8	297	199.6852	199.8436	275.7347	269.3707
9	221	200.1301	200.2658	252.4715	247.1235
10	325	200.5760	200.6889	281.2705	281.5169
11	131	201.0229	201.1129	180.3573	183.7034
12	260	201.4708	201.5378	230.6736	231.4312
13	237	201.9197	201.9635	241.8403	239.5545
14	222	202.3696	202.3902	227.4029	229.4472
15	228	202.8205	202.8178	203.9778	200.0433
16	114	203.2725	203.2463	154.3551	145.2742
17	263	203.7254	203.6756	225.0999	222.2530
18	162	204.1793	204.1059	192.6028	191.1329
19	240	204.6342	204.5371	210.6845	207.4535
20	165	205.0902	204.9692	196.3511	190.7474
21	280	205.5472	205.4023	246.0748	243.8765
22	169	206.0052	205.8362	201.7839	201.4542
23	258	206.4642	206.2711	234.2487	238.9841
24	190	206.9242	206.7068	198.4787	200.1683
25	179	207.3853	207.1435	185.3850	181.1938
26	207	207.8473	207.5811	184.2423	184.9904
27	121	208.3104	208.0197	143.8172	151.5959
Testing		Short-Term Forecasted Values			
28	189	208.7746	208.4591	145.6179	142.0339
29	88	209.2398	208.8995	168.0954	174.0508
30	192	209.7060	209.3409	90.7493	105.9078

Table 4.52: Day 3 Site 1 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional		Improved	
		GM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training					
Northbound	RMSE	60.0896	60.0899	25.5816	25.7114
	RMSPE	25.7361	25.7224	11.1759	11.4111
	MAE	48.5454	48.5022	21.4165	21.5410
	MAPD	24.5041	24.4823	10.8103	10.8732
Southbound	RMSE	46.3823	46.3834	22.3996	25.8152
	RMSPE	32.0459	32.0232	15.5012	21.5986
	MAE	37.3073	37.2921	16.4326	18.1481
	MAPD	35.2202	35.2059	15.5133	17.1328
Eastbound	RMSE	17.0142	17.0149	7.3407	8.3307
	RMSPE	23.6716	23.6529	10.1071	12.8315
	MAE	13.3250	13.3268	5.5295	5.8316
	MAPD	19.1370	19.1395	7.9414	8.3751
Westbound	RMSE	29.3854	29.3854	15.0617	16.1311
	RMSPE	21.8539	21.8197	10.4906	10.8039
	MAE	21.8389	21.8244	8.3409	8.2001
	MAPD	25.5924	25.5755	9.7745	9.6095
Short-Term Forecasting					
Northbound	RMSE	61.3444	61.3155	51.1668	51.2457
	RMSPE	25.6281	25.6106	19.1836	18.5851
	MAE	48.9815	48.9087	40.3371	39.5804
	MAPD	25.2569	25.2193	20.7995	20.4093
Southbound	RMSE	48.5652	48.5480	123.8618	69.1543
	RMSPE	31.7502	31.7299	44.4531	32.4986
	MAE	39.5376	39.5169	53.2710	40.9068
	MAPD	37.7868	37.7670	50.9120	39.0954
Eastbound	RMSE	18.6887	18.6808	27.6995	23.6732
	RMSPE	23.4848	23.4617	37.7256	31.8691
	MAE	14.5110	14.5100	17.7799	16.2264
	MAPD	20.1542	20.1528	24.6943	22.5366
Westbound	RMSE	28.3786	28.3821	39.9783	35.6490
	RMSPE	21.1844	21.1543	33.3100	27.3552
	MAE	21.0007	20.9916	22.2562	20.9722
	MAPD	24.3439	24.3334	25.7993	24.3109

From Table 4.51 the fitting data obtained by the GGM(1,1) is more close to the original (real) data and this is evident in Table 4.52, where the fitting errors for GGM(1,1) are actually smaller compared to those of the other three models. In short-term forecasting MBVGGM(1,1) outperformed the other models with an accuracy of 100-

RMSPE=100-18.5851=81.4149%, in the northward direction. This is good accuracy as seen from Table 4.6. In both southward and eastward directions MBVGM(1,1)'s RMSPE and MAPD are in the range of 20-50%, this is reasonable accuracy as can be seen from Table 4.6. This model outperformed the rest in these two directions of vehicle flow.

II. Site 2: Kenyatta Avenue Uhuru Highway Roundabout

a) Traffic Flow Training

In training the OGM(1,1), AGO and MGO of the real vehicle flow data (of Appendix IV Table 16) was obtained from which the parameters a and b were calculated and recorded in Table 4.53. Thus from (3.9) OGM(1,1)'s time response equation for the southward direction can be given as:

$$\hat{x}^{(1)}_{(r+1)} \triangleq -54{,}341.6429\, e^{-0.0028r} + 54{,}429.6429,\ r = 0,1,2,\cdots,m-1 \qquad (4.38)$$

From (4.38) the simulation values of OGM(1,1) were calculated and recorded in Table 4.55. Further, Figure 4.49a shows the plots of real, simulation and error data curves.

Table 4.53: Day 3 Site 2 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	0.0109	123.8431
Southward	0.0028	152.4030
Eastward	-0.0090	79.2429
Westward	-0.0214	55.0473

The model parameters for the MBVGM(1,1) were as shown in Table 4.54. This is after modification of OGM(1,1)'s background value. Now, from (3.9) MBVGM(1,1)'s time response equation, in the southward direction, is given by:

$$\hat{x}^{(1)}_{(r+1)} \triangleq -49{,}244.0645\, e^{-0.0031r} + 49{,}332.0645,\ r = 0,1,2,\cdots,m-1 \qquad (4.39)$$

MBVGM(1,1)'s simulation values as obtained from (4.39) were as recorded in Table 4.55. Its plots of real, simulated and error curves are shown in Figure 4.49b.

Table 4.54: Day 3 Site 2 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	0.0129	126.6548
Southward	0.0031	152.9294
Eastward	-0.0089	79.3214
Westward	-0.0213	55.2208

Similarly, GGM(1,1)'s and MBVGGM)(1,1)'s simulation values (in the southward direction) were obtained and recorded in Table 4.55. Their real, simulated and error curves are as shown in Figures 4.49c and 4.49d, respectively.

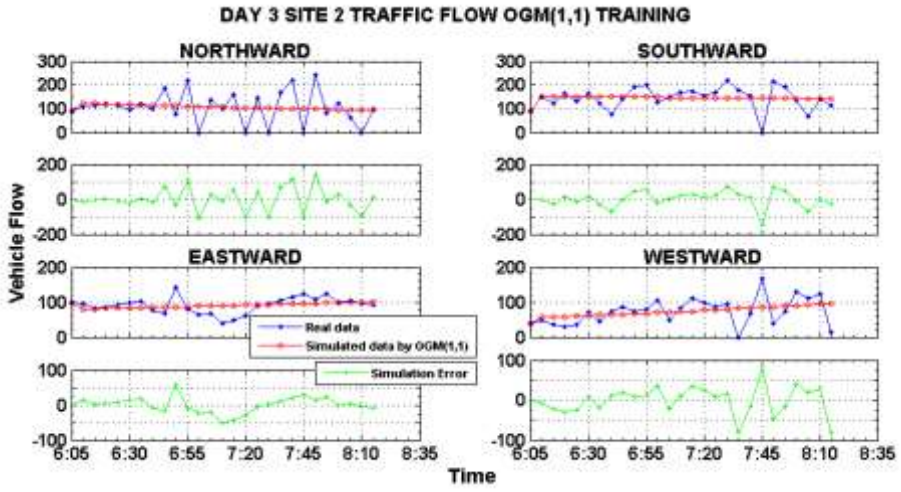

Figure 4.49a: Vehicle Flow OGM(1,1) Training

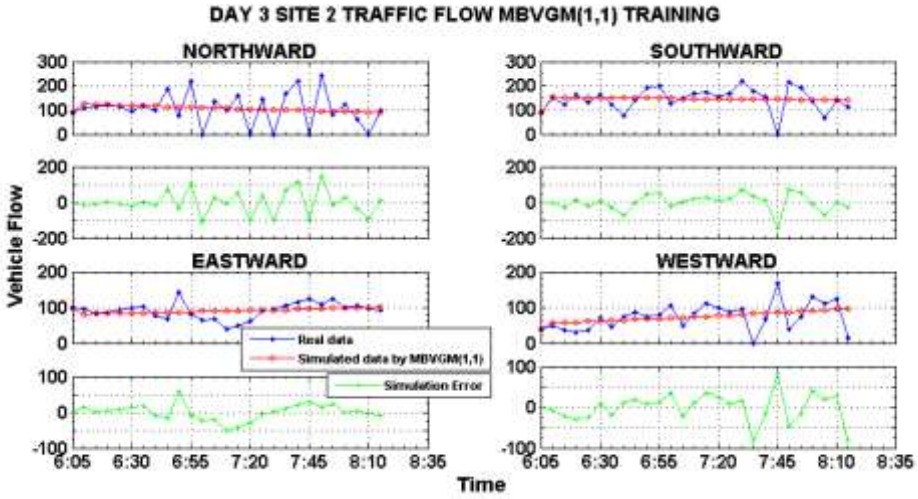

Figure 4.49b: Vehicle Flow MBVGM(1,1) Training

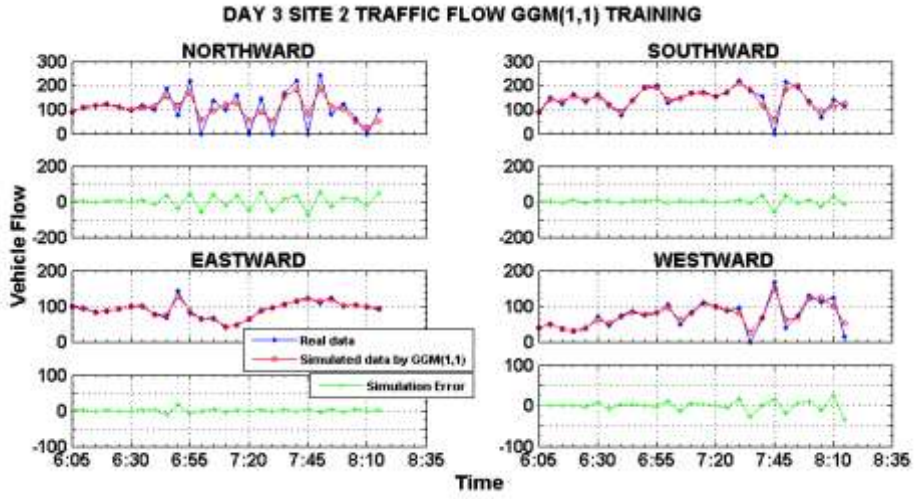

Figure 4.49c: Vehicle Flow GGM(1,1) Training

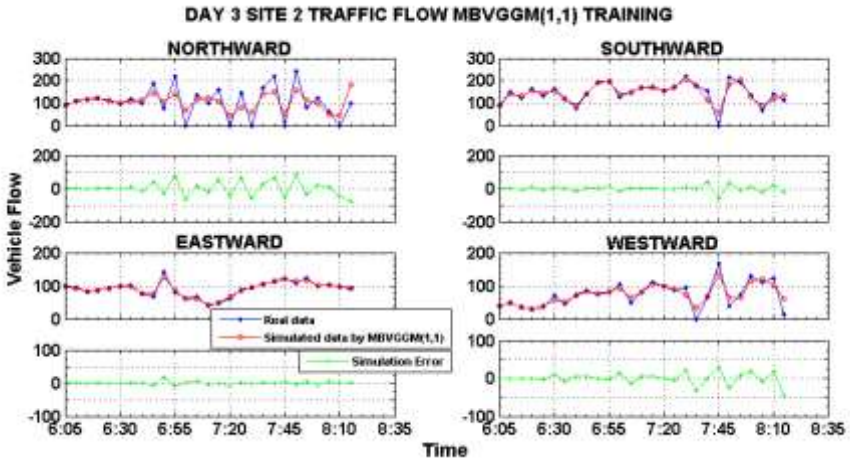

Figure 4.49d: Vehicle Flow MBVGGM(1,1) Training

Observe the error curves in Figures 4.49a-4.49d to notice that for GGM(1,1) and MBVGGM(1,1) the error curve is almost at zero level unlike with the OGM(1,1) and MBVGM(1,1). It means that both GGM(1,1) and MBVGGM(1,1) have good vehicle flow fitting compared to OGM(1,1) and MBVGM(1,1). Probably, this is because of the ability of the DGT in improving the fitting accuracy of the OGM(1,1).

b) Testing the Grey Models in Short-Term Forecasting

To test the grey models the time response functions of the grey models were extrapolated. For instance (4.38) and (4.39) were extrapolated for OGM(1,1) and MBVGM(1,1), respectively. The forecasted values for all the four models were then recorded in Table 4.55. Note that Table 4.55 shows values for the southward direction only. Figures 4.50a-4.50d show the real, simulated and error curves for these grey models.

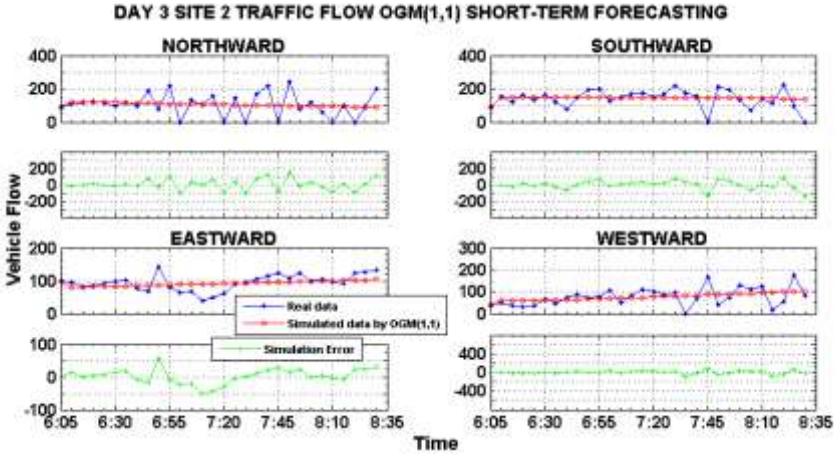

Figure 4.50a: Short-Term Vehicle Flow Forecast by OGM(1,1)

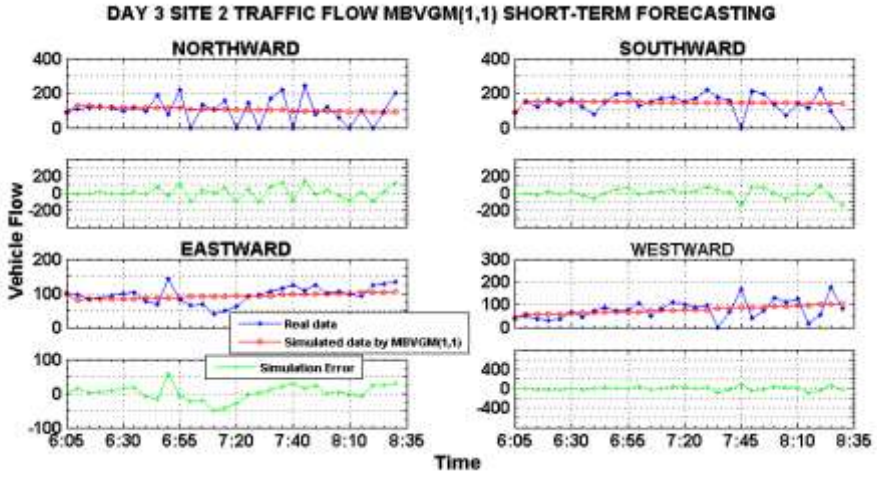

Figure 4.50b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

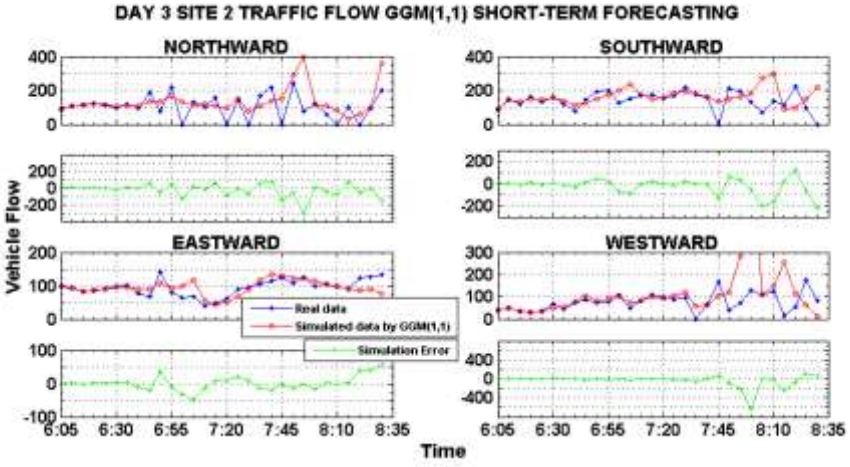

Figure 4.50c: Short-Term Vehicle Flow Forecast by GGM(1,1)

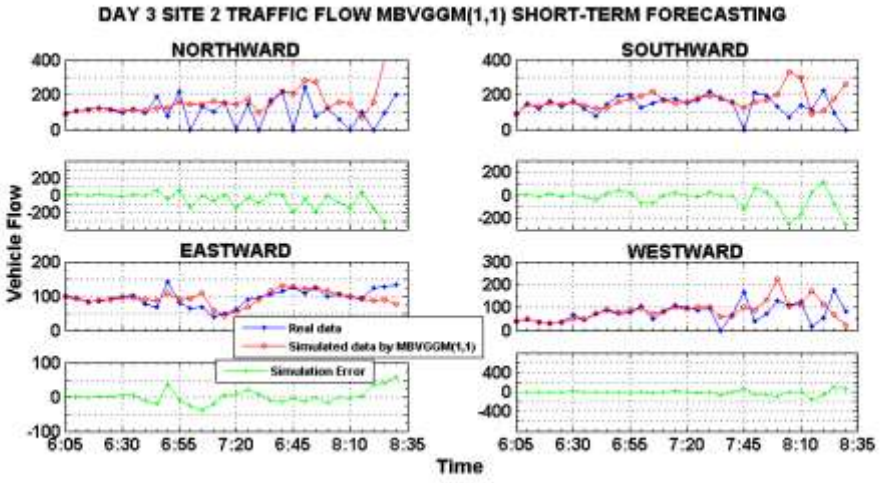

Figure 4.50d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

From Figure 4.50c the following observation can be made. For GGM(1,1), short-term forecast of 757.2532 at the 7th last data point was obtained, in the westward direction. This was too inaccurate. For MBVGGM(1,1), a northward direction forecast of 1.0e+03 *0.4118 and 1.0e+03 *1.7561 at the last two data point were obtained, respectively. This is a too high forecast. The MBVGGM(1,1) was not accurate in this case.

Table 4.55: Original and Improved Grey Models' Simulation Values (D3S2 Southward Direction)

Raw Data		Grey Model			
Data Point	**Real Value**	**OGM(1,1)**	**MBVGM(1,1)**	**GGM(1,1)**	**MBVGGM(1,1)**
Training		**Fitted Values**			
1	88	88.0000	88.0000	88.0000	88.0000
2	148	151.9458	152.427	142.1899	142.6015
3	123	151.5228	151.9627	134.4356	133.5144
4	164	151.1010	151.4991	152.7004	152.7649
5	132	150.6804	151.0370	143.0762	143.0876
6	163	150.2609	150.5763	154.6982	154.7098
7	123	149.8426	150.1170	119.1037	118.7950
8	77	149.4255	149.6591	89.4635	90.7545
9	143	149.0095	149.2025	137.2645	139.9474
10	193	148.5947	148.7474	192.3308	192.2816
11	203	148.1810	148.2937	193.4688	193.9792
12	127	147.7685	147.8413	139.2992	140.3661
13	149	147.3572	147.3903	144.2782	143.4691
14	167	146.9470	146.9408	167.1153	166.7675
15	173	146.5379	146.4925	169.8008	169.9238
16	153	146.1300	146.0457	155.7136	156.0159
17	169	145.7232	145.6002	173.2335	171.5189
18	217	145.3175	145.1560	207.8388	207.6875
19	176	144.9130	144.7133	183.2731	179.0215
20	154	144.5096	144.2718	117.2700	111.2778
21	0	144.1073	143.8317	58.6761	56.7312
22	214	143.7062	143.3930	181.5777	181.8777
23	193	143.3061	142.9556	200.0580	205.0285
24	135	142.9072	142.5195	128.2747	125.4675
25	68	142.5094	142.0848	93.1578	89.7345
26	139	142.1126	141.6514	111.0508	115.5722
27	113	141.7170	141.2193	126.6762	137.5969
Testing		**Short-Term Forecasted Values**			
28	223	141.3225	140.7885	97.6830	104.8835
29	93	140.9291	140.3591	153.2434	174.9376
30	0	140.5368	139.9309	218.4783	262.4073

c) Evaluation of the Grey Models

The fitting and forecasted data for the four grey models were computed and recorded in Table 4.55. This was for the southward direction. The errors which were involved in the simulation process were as well calculated and recorded in Table 4.56.

Table 4.56: Day 3 Site 2 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional		Improved	
		GM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training					
Northbound	RMSE	65.8868	65.9048	36.0221	43.1865
	RMSPE	46.8358	47.2605	20.7942	30.3828
	MAE	49.6257	49.8337	29.0376	34.0396
	MAPD	46.6699	46.8655	27.3081	32.0121
Southbound	RMSE	45.3105	45.3114	17.9132	18.2807
	RMSPE	22.9012	22.9503	8.4069	8.8186
	MAE	32.2579	32.2770	12.3086	12.6018
	MAPD	22.3095	22.3227	8.5126	8.7154
Eastbound	RMSE	21.9446	21.9447	4.8197	4.7215
	RMSPE	22.0605	22.0569	5.6781	5.4827
	MAE	16.2335	16.2264	3.2887	3.2559
	MAPD	17.9339	17.9260	3.6332	3.5969
Westbound	RMSE	34.9147	34.9153	12.4437	15.6906
	RMSPE	36.0106	35.9905	10.4758	14.8323
	MAE	26.9081	26.9021	8.6073	10.7011
	MAPD	36.8044	36.7961	11.7729	14.6368
Short-Term Forecasting					
Northbound	RMSE	67.8272	67.8971	85.9174	301.3335
	RMSPE	47.7119	48.2782	41.9303	260.6571
	MAE	51.5958	51.8889	55.7738	115.7676
	MAPD	48.8597	49.1372	52.8161	109.6284
Southbound	RMSE	52.9626	52.9203	77.3037	86.8158
	RMSPE	25.0932	25.1698	32.5474	32.5602
	MAE	38.0369	38.0327	50.0153	52.9555
	MAPD	27.0404	27.0374	35.5559	37.6461
Eastbound	RMSE	22.2953	22.3048	21.2363	19.9741
	RMSPE	21.5428	21.5545	22.2564	21.5535
	MAE	17.1205	17.1223	14.8458	13.8632
	MAPD	18.1682	18.1701	15.7543	14.7116
Westbound	RMSE	37.0398	37.0414	133.2588	47.0540
	RMSPE	38.3584	38.3588	151.1395	45.9697
	MAE	28.9735	28.9654	56.8899	28.0897
	MAPD	38.0230	38.0123	74.6586	36.8632

In vehicle flow fitting GGM(1,1) was the most accurate (except in the eastward direction) since its simulation resulted to small error values (see Table 4.56). It had the highest accuracy, in the southward direction, of 100-RMSPE=100-8.4069=91.5931%. This is high accuracy (see Table 4.6). In the eastward direction MBVGGM(1,1) had

the highest accuracy in fitting. It had an accuracy of 100-MAPD=100-3.5969=96.4031%. This accuracy is also high.

In short-term forecasting, in the northward direction, GGM(1,1) emerged the accurate grey model with an accuracy of 100-MAPD=100-41.9303=58.0697%. This is reasonable accuracy as seen from Table 4.6. In the eastward direction, MBVGGM(1,1) had an accuracy of 100-MAPD=100-14.7116=85.2884%. According to Table 4.6 this is good forecasting accuracy.

Generally, from this evaluation the DGT seems to do great in improving the fitting accuracy of the OGM(1,1).

III. Site 4: Kenyatta Avenue-Moi Avenue-Mondlane Street Junction

a) Traffic Flow Training

Generation of the AGO and MGO series from the real data of day 3 site 4 (see Appendix IV Table 18), led to the computation of the parameters a and b for the OGM(1,1). These parameters are shown in Table 4.57. Note that site 4 is a three-way junction.

Table 4.57: Day 3 Site 4 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0301	38.0664
Southward	-0.0165	25.3203
Eastward	-0.0165	10.9007

From Table 4.57 the time response equations for each direction can be formulated. In the southward direction the time response equation from (3.9) was found to be:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 1{,}558.5636\, e^{0.0165r} - 1{,}534.5636,\ r = 0,1,2,\cdots,m-1 \qquad (4.40)$$

IAGO was applied on the time series data generated from (4.40) to obtain OGM(1,1)'s fitted data. These fitted data was recorded in Table 4.59. And the real, simulated and error curves were plotted in Figure 4.51a.

The background value of the OGM(1,1) was modified and the developed MBVGM(1,1)'s a and b parameters were as recorded in Table 4.58. Thus its time response equation for the southward direction was obtained from (3.9) as:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 1{,}559.2e^{0.0165r} - 1{,}535.2,\ r = 0,1,2,\cdots,m-1 \tag{4.41}$$

Then through the IAGO process MBVG(1,1)'s fitted data were calculated and recorded in Table 4.59. Figure 4.51b shows MBVG(1,1)'s real, simulated and error curves.

Similarly, for GGM(1,1) and MBVGGM(1,1) Figures 4.51c and 4.51d were plotted. Their simulated values were as well recorded in Table 4.59.

Table 4.58: Day 3 Site 4 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Northward	-0.0301	38.0956
Southward	-0.0165	25.3308
Eastward	NaN	NaN

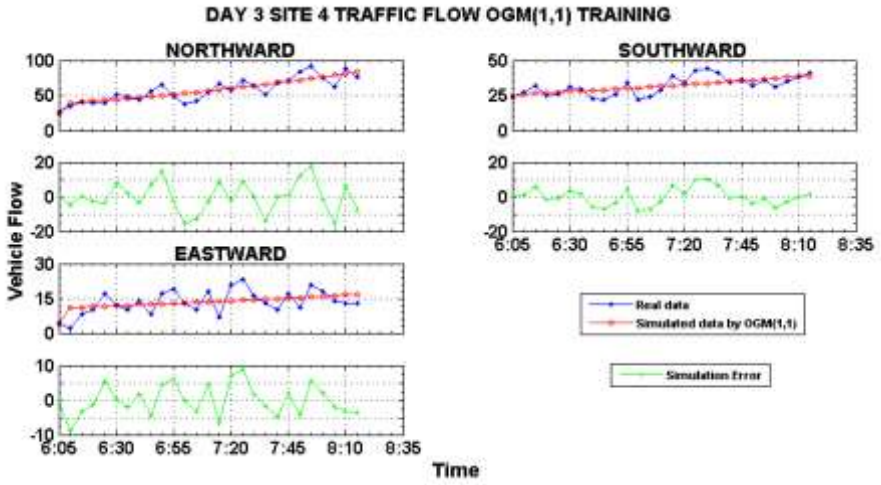

Figure 4.51a: Vehicle Flow OGM(1,1) Training

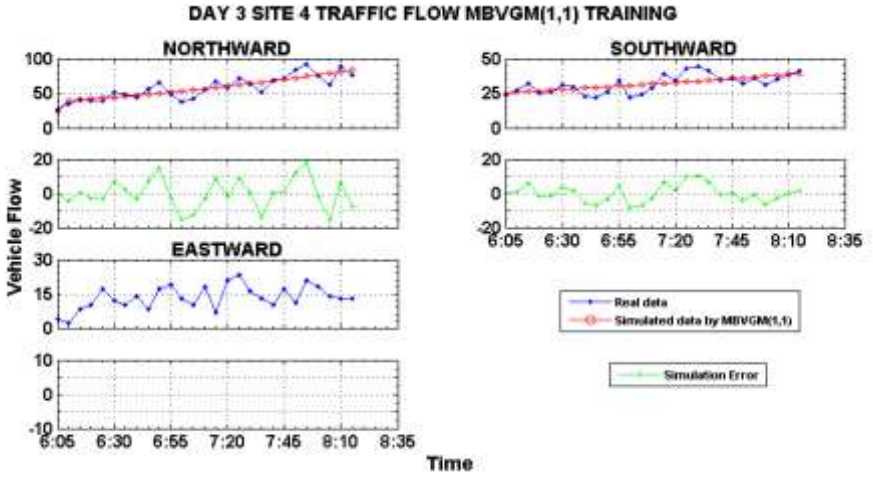

Figure 4.51b: Vehicle Flow MBVGM(1,1) Training

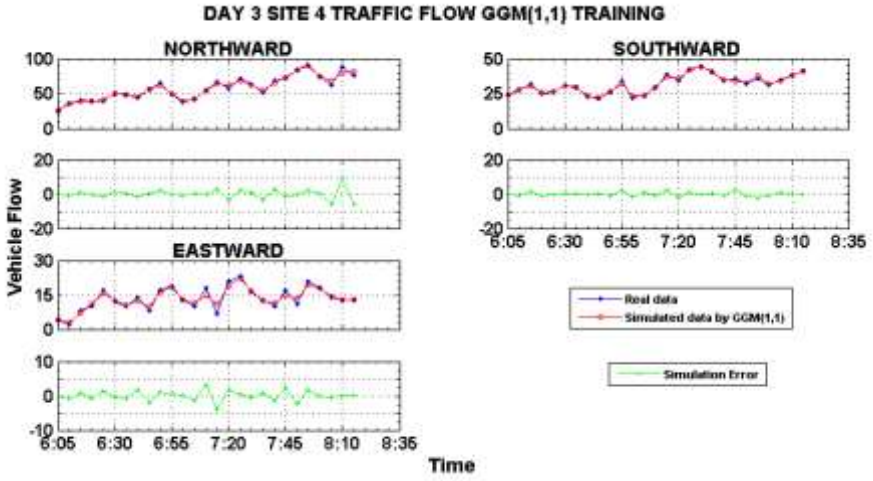

Figure 4.51c: Vehicle Flow GGM(1,1) Training

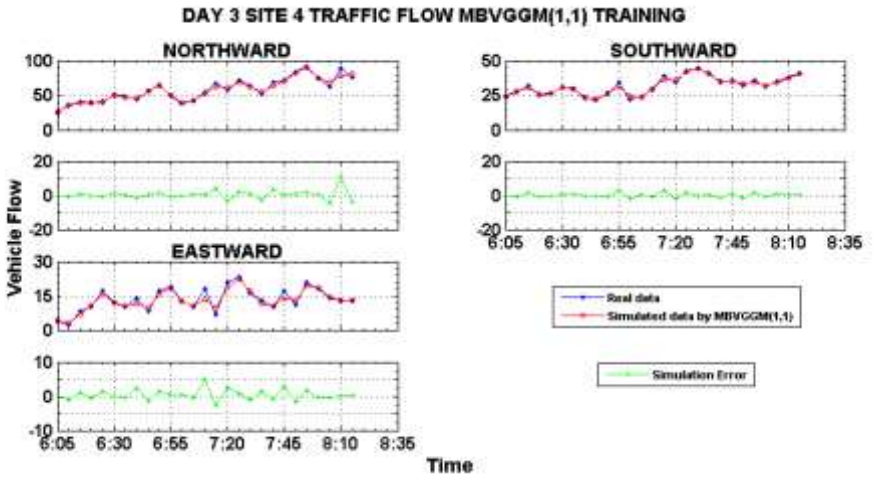

Figure 4.51d: Vehicle Flow MBVGGM(1,1) Training

From Figure 4.51b notice that there were no eastward direction result for MBVGM(1,1). This was because of the "NaN" condition. It means that the data matrix A was singular to working precision. And therefore the parameters a and b could not be computed. Also from Figures 4.51a-4.51d observe that the GGM(1,1) and MBVGGM(1,1) have very neat vehicle flow fitting compared to the other two models.

b) Testing the Grey Models in Short-Term Forecasting

Testing of the grey models involved extrapolating the time response equations to forecast the last three data points of the real data series. So (4.40) and (4.41) were extrapolated for OGM(1,1) and MBVGM(1,1) respectively. Their forecasts were as recorded in Table 4.59. Figures 4.52a-4.52d show the plots of the real, simulated and error data curves for all the four grey models.

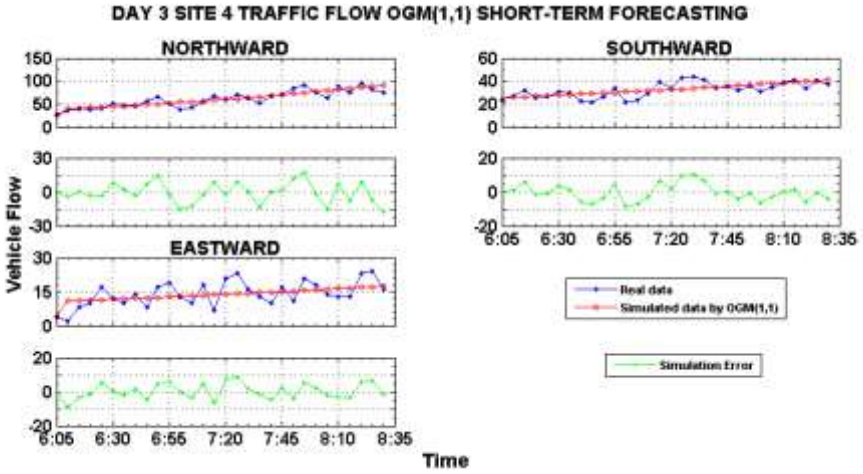

Figure 4.52a: Short-Term Vehicle Flow Forecast by OGGM(1,1)

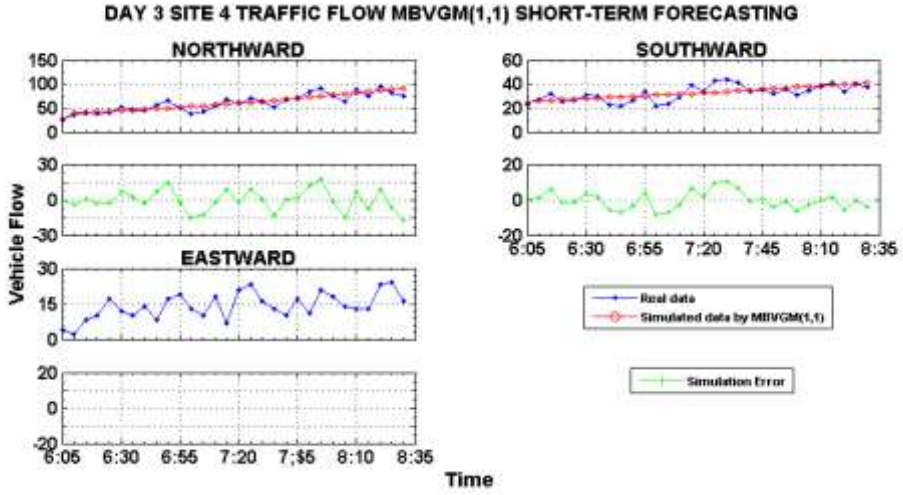

Figure 4.52b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

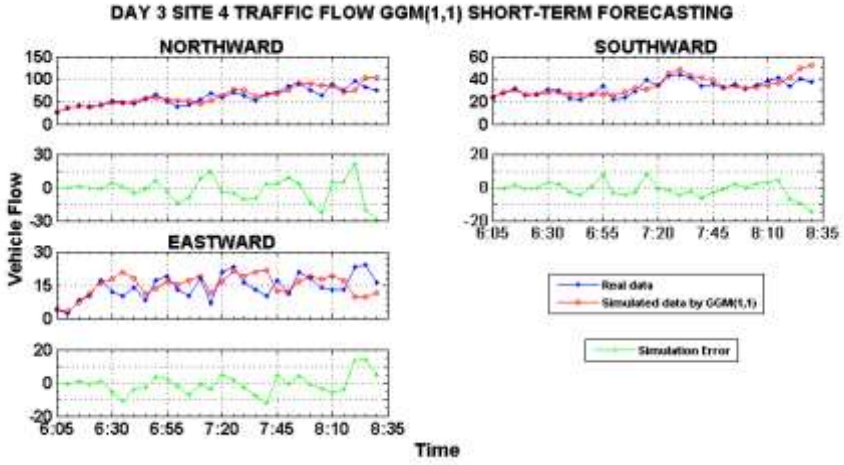

Figure 4.52c: Short-Term Vehicle Flow Forecast by GGM(1,1)

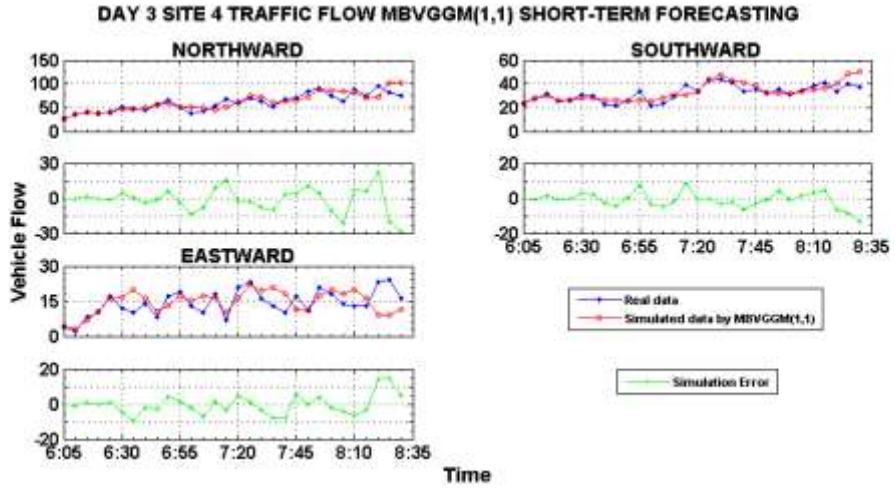

Figure 4.52d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

Figure 4.52b shows that there were no result of the eastward direction because of the 'NaN" condition as obtained during simulation. OGM(1,1) and MBVGM(1,1) have similar performance as observed from their error curves. Same observation can be deduced from the GGM(1,1) and MBVGGM(1,1) curves, that they also have similar forecasting performance.

Table 4.59: Original and Improved Grey Models' Simulation Values (D3S4 Southward Direction)

Raw Data		Grey Model			
Data Point	**Real Value**	**OGM(1,1)**	**MBVGM(1,1)**	**GGM(1,1)**	**MBVGGM(1,1)**
Training			**Fitted Values**		
1	24	24.0000	24.0000	24.0000	24.0000
2	27	25.9301	25.9399	27.9680	27.9795
3	32	26.3618	26.3712	30.2722	30.1875
4	25	26.8007	26.8097	25.9690	26.0279
5	26	27.2470	27.2554	26.2044	26.3220
6	31	27.7006	27.7086	30.5656	30.5853
7	30	28.1618	28.1693	29.4817	29.3481
8	23	28.6307	28.6377	23.5147	23.4367
9	22	29.1073	29.1139	21.9740	22.2174
10	26	29.5919	29.5979	26.7643	26.8687
11	34	30.0846	30.0901	31.7091	31.4636
12	22	30.5855	30.5904	23.8119	23.9859
13	24	31.0947	31.0990	23.3284	23.7033
14	29	31.6124	31.6161	29.7625	29.7552
15	39	32.1388	32.1418	36.9330	36.4275
16	34	32.6738	32.6762	36.1319	35.8833
17	43	33.2178	33.2195	41.8870	41.7286
18	44	33.7709	33.7719	44.1221	44.0856
19	41	34.3331	34.3334	40.4317	40.6202
20	34	34.9048	34.9043	34.6833	35.3563
21	36	35.4859	35.4846	34.1004	34.8506
22	32	36.0767	36.0746	32.6277	33.3187
23	36	36.6773	36.6745	37.8801	34.5411
24	31	37.2880	37.2842	32.1852	32.1820
25	35	37.9088	37.9042	34.3400	34.2761
26	38	38.5400	38.5344	38.0317	37.5214
27	41	39.1816	39.1751	41.0110	40.3920
Testing			**Short-Term Forecasted Values**		
28	34	39.8340	39.8265	40.9989	40.5556
29	40	40.4972	40.4887	49.7992	48.1725
30	37	41.1714	41.1619	51.9644	50.4153

c) Evaluation of the Grey Models

In evaluating the four grey models, their simulation data were as indicated in Table 4.59. These data were for the southward direction. The overall traffic flow training and short-term forecasting error evaluation for day 3 site 4 was as tabulated in Table 4.60.

Table 4.60: Day 3 Site 4 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional	Improved		
		GM(1,1)	MBVGM(1,1)	GGM(1,1)	MBVGGM(1,1)
Training					
Northbound	RMSE	8.5393	8.5394	2.8477	2.9165
	RMSPE	13.6806	13.6758	5.1650	5.4773
	MAE	6.5203	6.5208	1.9071	1.8780
	MAPD	11.2996	11.3005	3.3051	3.2546
Southbound	RMSE	4.9317	4.9317	1.1549	1.2097
	RMSPE	15.0597	15.0573	3.4083	3.5793
	MAE	3.9268	3.9265	0.9128	0.9916
	MAPD	12.3428	12.3416	2.8690	3.1169
Eastbound	RMSE	4.4169	NaN	1.5028	1.6315
	RMSPE	28.5696	NaN	8.4766	10.9102
	MAE	3.6890	NaN	1.1543	1.2221
	MAPD	27.7445	NaN	8.6814	9.1910
Short-Term Forecasting					
Northbound	RMSE	8.9553	8.9543	10.8731	10.5833
	RMSPE	13.5554	13.5502	17.3097	17.1646
	MAE	6.9797	6.9796	7.9574	7.7464
	MAPD	11.5749	11.5747	13.1963	12.8464
Southbound	RMSE	4.8592	4.8587	4.8645	4.5528
	RMSPE	14.4984	14.4943	15.1054	14.0041
	MAE	3.8842	3.8831	3.5818	3.3912
	MAPD	12.0131	12.0095	11.0778	10.4882
Eastbound	RMSE	4.5138	NaN	5.8324	5.6305
	RMSPE	27.8647	NaN	37.1411	37.8498
	MAE	3.7976	NaN	4.4614	4.2514
	MAPD	26.9972	NaN	31.7158	30.2234

Table 4.59 shows the simulation values for the four grey models in the southward direction. From Table 4.60, in the southward direction, GGM(1,1) proved to have a high fitting accuracy of 100-MAPD=100-2.8690=97.1310%. See Table 4.6. It also had the highest fitting accuracy of 100-RMSPE=100-8.4766=91.5234%, in the eastward direction.

In short-term forecasting, MBVGGM(1,1) had the highest forecasting accuracy given by 100-MAPD=100-10.4882=89.5118%, in the southward direction. See Table 4.60. In the northward direction, MBVGM(1,1) had the highest accuracy at 100-MAPD=100-11.5747= 88.4253%. In this simulation a unique result was obtained in

the eastward direction. That is, the OGM(1,1) had the highest short-term forecasting accuracy of 100-MAPD=100-26.9972=73.0028%. This was not expected.

IV. Site 7: Haile Selassie Avenue-Moi Avenue Roundabout

a) Traffic Flow Training

The parameters a and b for OGM(1,1) were computed through the generation of the AGO and MGO series from the real data of day 3 site 7 (see Appendix IV Table 21). These parameters are shown in Table 4.61. Note that site 4 is a three-way junction.

Table 4.61: Day 3 Site 7 OGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, a	Control Variable, b
Southward	-0.0021	76.9635
Eastward	0.0006	84.1522
Westward	-0.0122	88.8354

From Table 4.61 it is possible to formulate the time response equations for each direction based on (3.9). In the westward direction the time response equation was found to be:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 7{,}368.5902\, e^{0.0122r} - 7{,}281.5902,\ r = 0,1,2,\cdots,m-1 \tag{4.42}$$

The IAGO operation was applied on the time series data generated from (4.42) to obtain OGM(1,1)'s fitted data. These fitted data were recorded in Table 4.63. And the real, simulated and error curves were plotted in Figure 4.53a.

Next the background value of OGM(1,1) was modified and the developed MBVGM(1,1)'s a and b parameters were as recorded in Table 4.62. Thus, its time response equation for the westward direction was obtained from (3.9) as:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 7{,}371.8689\, e^{0.0122r} - 7{,}284.8689,\ r = 0,1,2,\cdots,m-1 \tag{4.43}$$

Then through the IAGO process MBVG(1,1)'s fitted data was calculated and recorded in Table 4.63. Figure 4.53b shows MBVG(1,1)'s real, simulated and error curves.

Table 4.62: Day 3 Site 7 MBVGM(1,1) Model Parameters

Traffic flow direction	Grey Model Parameter	
	Development Coefficient, *a*	**Control Variable, *b***
Southward	-0.0020	76.9965
Eastward	0.0006	84.1784
Westward	-0.0122	88.8754

Similarly, for GGM(1,1) and MBVGGM(1,1) Figures 4.53c and 4.53d were plotted. Their simulated values were as well recorded in Table 4.63.

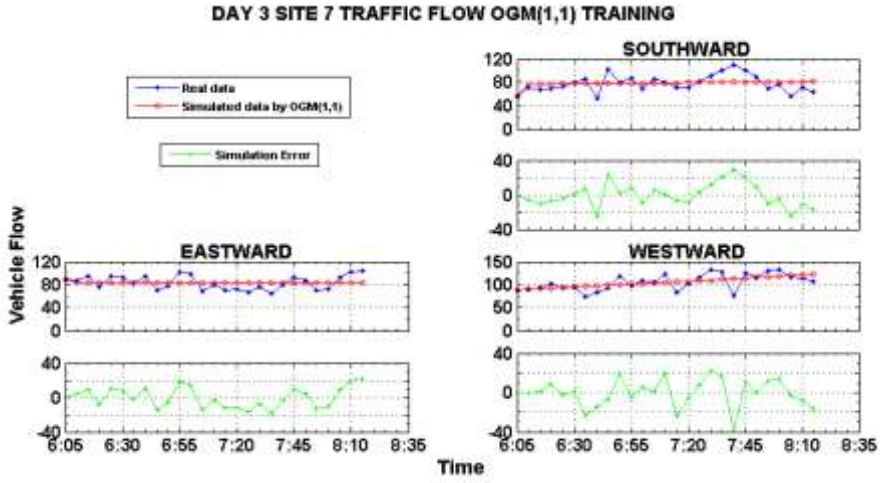

Figure 4.53a: Vehicle Flow OGM(1,1) Training

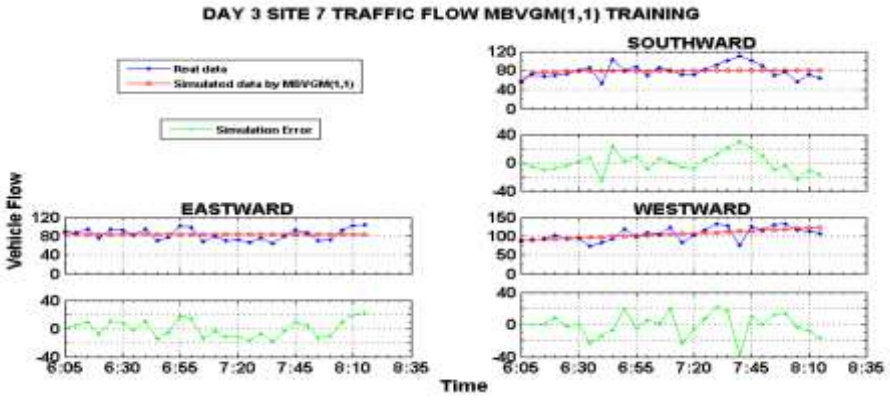

Figure 4.53b: Vehicle Flow MBVGM(1,1) Training

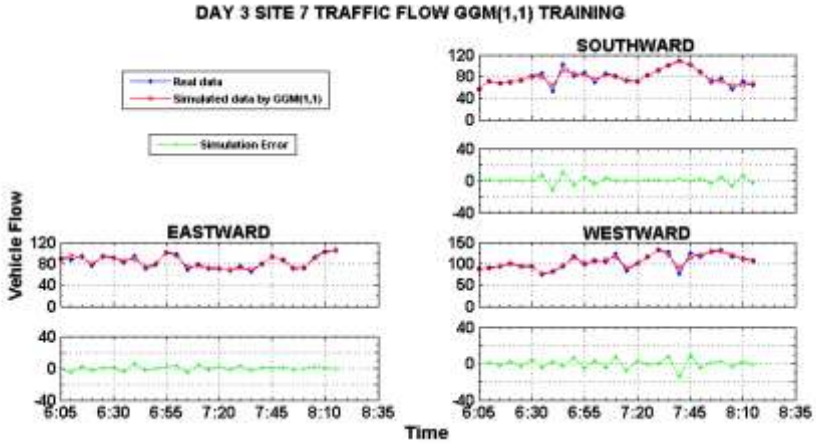

Figure 4.53c: Vehicle Flow GGM(1,1) Training

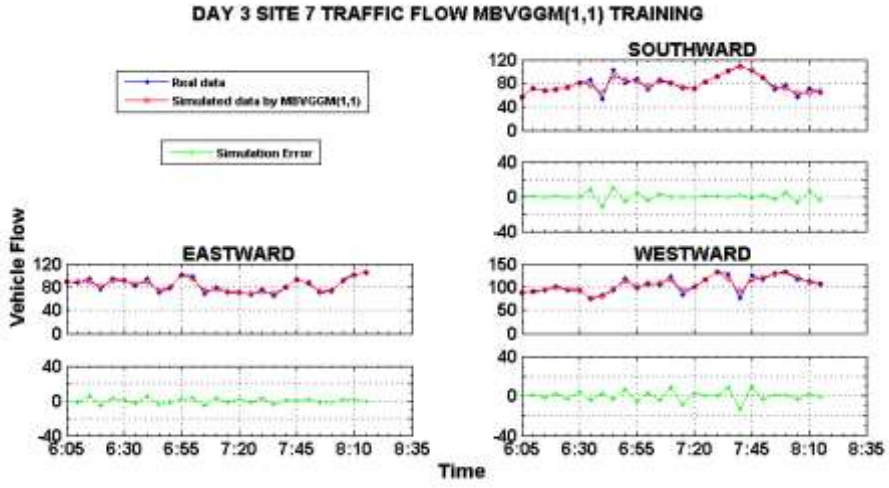

Figure 4.53d: Vehicle Flow MBVGGM(1,1) Training

From Figures 4.53a-4.53d observe that the GGM(1,1) and MBVGGM(1,1) have very neat vehicle flow fitting compared to the other two models. That is to say with GGM(1,1) and MBVGGM(1,1) the simulated data can trace the real data with high accuracy. This is very clear from the corresponding error curves which are approaching zero level. Meaning the deviation between real data and the simulated data is almost zero. Thus the DGT is powerful in improving the fitting accuracy of the OGM(1,1), compared with MBV.

b) Testing the Grey Models in Short-Term Forecasting

Testing of the grey models involved extrapolating the time response equations to forecast the last three data points of the real data series. Therefore, (4.42) and (4.43) were extrapolated for OGM(1,1) and MBVGM(1,1) respectively. Their short-term forecasts were as recorded in Table 4.63. Figures 4.54a-4.54d show the plots of the real, simulated and error data curves for all the four grey models.

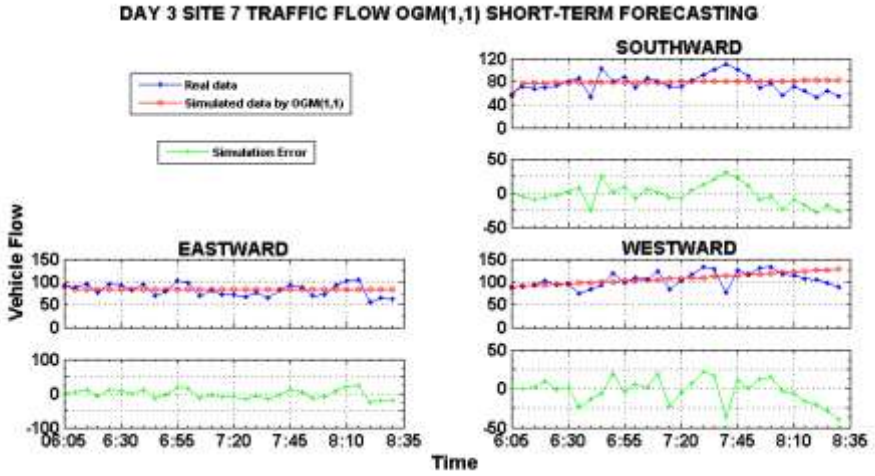

Figure 4.54a: Short-Term Vehicle Flow Forecast by OGGM(1,1)

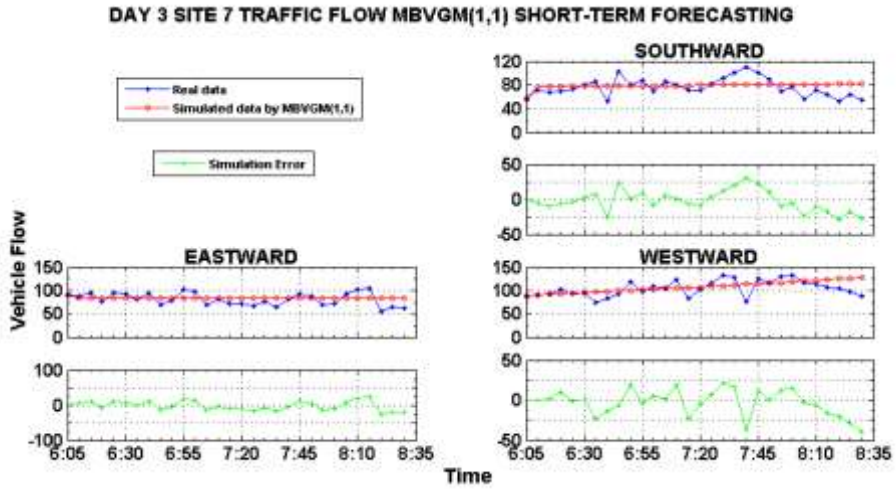

Figure 4.54b: Short-Term Vehicle Flow Forecast by MBVGM(1,1)

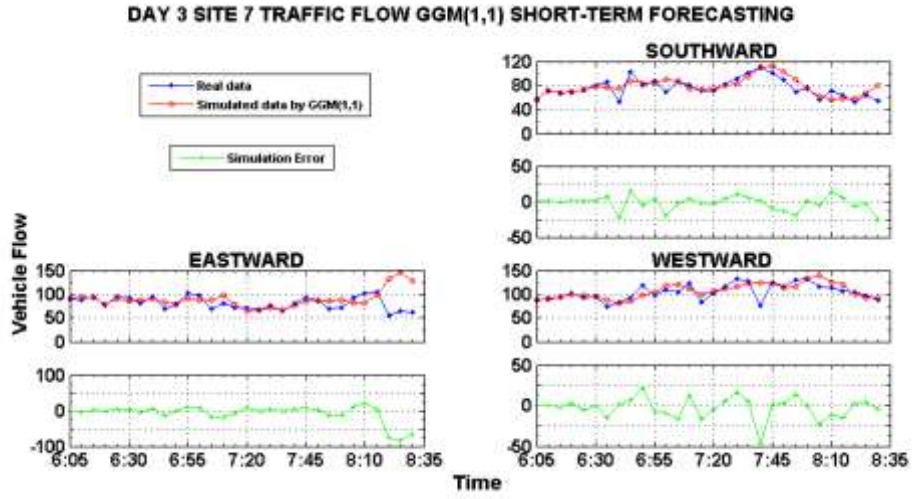

Figure 4.54c: Short-Term Vehicle Flow Forecast by GGM(1,1)

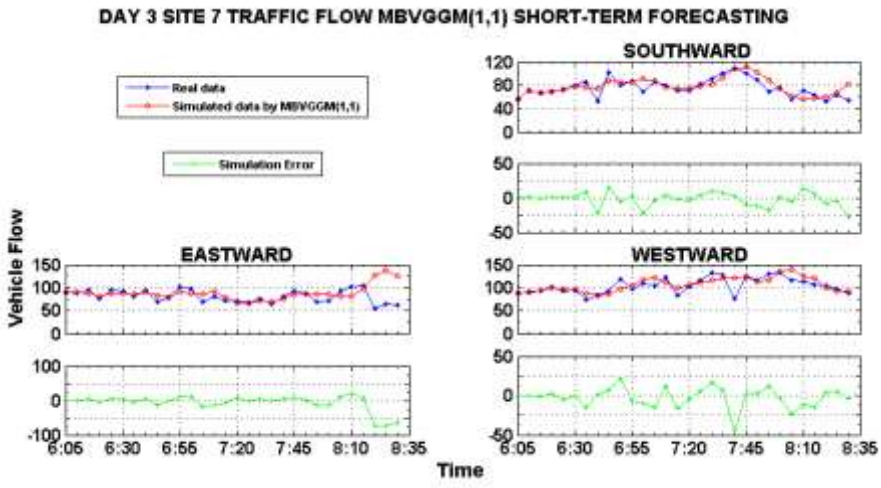

Figure 4.54d: Short-Term Vehicle Flow Forecast by MBVGGM(1,1)

Critical observation of the error curves of OGM(1,1) and MBVGM(1,1) shows that they have similar performance characteristics in short-term forecasting. Same observation can be deduced from GGM(1,1) and MBVGGM(1,1) curves, that they also have similar forecasting performance characteristics. For instance the fitting of the real and simulated curves of GGM(1,1) and MBVGGM(1,1), especially at the last three data points, looks alike. It implies that they have same forecasting accuracy. However, these observations shows that GGM(1,1) and MBVGGM(1,1) are more accurate compared to OGM(1,1) and MBVGM(1,1).

Table 4.63: Original and Improved Grey Models' Simulation Values (D3S7 Westward Direction)

Raw Data		Grey Model			
Data Point	**Real Value**	**OGM(1,1)**	**MBVGM(1,1)**	**GGM(1,1)**	**MBVGGM(1,1)**
Training			**Fitted Values**		
1	87	87.0000	87.0000	87.0000	87.0000
2	91	90.4535	90.4902	90.3969	90.4054
3	93	91.5680	91.6028	94.7812	94.7165
4	102	92.6963	92.7289	99.5821	99.6221
5	92	93.8384	93.8690	94.9811	94.9750
6	96	94.9947	95.0230	91.8038	91.5642
7	73	96.1651	96.1913	77.0426	77.0249
8	83	97.3500	97.3739	80.5262	80.8761
9	92	98.5495	98.5710	95.2334	95.1658
10	119	99.7638	99.7829	112.9703	112.7266
11	97	100.9930	101.0096	102.3218	102.5614
12	108	102.2374	102.2515	104.4047	104.2792
13	105	103.4971	103.5086	109.6131	109.2603
14	124	104.7723	104.7811	115.8587	115.7764
15	83	106.0632	106.0693	90.9539	91.5199
16	102	107.3701	107.3734	98.5037	98.4533
17	116	108.6930	108.6935	117.0218	116.1637
18	132	110.0323	110.0298	131.8930	132.1058
19	128	111.3880	111.3825	120.1338	119.0076
20	76	112.7605	112.7519	90.1990	89.5358
21	125	114.1498	114.1381	114.6857	115.3380
22	116	115.5563	115.5413	120.3096	119.7228
23	129	116.9801	116.9618	128.0916	127.9790
24	133	118.4215	118.3998	131.1774	131.4525
25	117	119.8806	119.8554	119.7824	120.4742
26	114	121.3577	121.3289	111.9856	112.2661
27	106	122.8530	122.8206	106.9030	107.0019
Testing			**Short-Term Forecasted Values**		
28	104	124.3667	124.3306	100.6744	100.4521
29	97	125.8991	125.8591	92.4957	91.6956
30	88	127.4503	127.4065	92.4021	91.8531

c) Evaluation of the Grey Models

To evaluate the four grey models, their simulation data were as indicated in Table 4.63. These data were for the westward direction. The overall vehicle flow training and short-term forecasting error evaluation for day 3 site 7 was as tabulated in Table 4.64.

Table 4.64: Day 3 Site 7 Traffic Flow Training and Forecasting Error Evaluation

Vehicle Traffic Flow Direction	Error Indicator	Grey Model			
		Conventional		Improved	
		GM(1,1)	MBVGM(1,1))	GGM(1,1)	MBVGGM(1,1)
Training					
Southbound	RMSE	13.5506	13.5506	4.5004	4.4866
	RMSPE	17.2462	17.2482	5.0632	5.0993
	MAE	10.8006	10.8002	3.1171	3.1538
	MAPD	13.7815	13.7810	3.9774	4.0242
Eastbound	RMSE	11.8033	11.8033	2.7599	2.9706
	RMSPE	13.8157	13.8172	3.0571	3.3044
	MAE	10.4236	10.4227	2.2749	2.5120
	MAPD	12.4474	12.4464	2.7166	2.9997
Westbound	RMSE	13.8982	13.8982	5.1486	5.1429
	RMSPE	11.9219	11.9215	4.4631	4.4893
	MAE	10.5177	10.5148	3.9678	3.9438
	MAPD	10.0027	10.0000	3.7735	3.7507
Short-Term Forecasting					
Southbound	RMSE	15.1105	15.1049	10.0606	10.3253
	RMSPE	18.3115	18.3098	11.0555	11.3033
	MAE	12.1803	12.1765	7.1725	7.4218
	MAPD	15.9777	15.9726	9.4086	9.7356
Eastbound	RMSE	13.2725	13.2681	25.3377	23.7074
	RMSPE	14.8543	14.8529	22.7293	21.4016
	MAE	11.5977	11.5942	14.2764	13.6548
	MAPD	14.2420	14.2377	17.5314	16.7680
Westbound	RMSE	16.3519	16.3446	13.2749	12.9833
	RMSPE	13.8715	13.8655	11.1758	10.9766
	MAE	12.4231	12.4165	9.1084	9.0948
	MAPD	11.9147	11.9084	8.7357	8.7227

Table 4.63 shows the simulation values for the four grey models in the westward direction. From Table 4.64, in the westward direction, MBVGGM(1,1) proved to have a high fitting accuracy of 100-MAPD=100-3.7507=96.2493%. See Table 4.6. In both northward and eastward directions GGM(1,1) emerged the most accurate vehicle flow fitting model with an accuracy of 100-MAPD=100-2.7166=97.2834%, in the eastward

direction. This shows how DGT greatly improves the fitting accuracy of OGM(1,1). In this vehicle flow fitting OGM(1,1) attained the highest accuracy of 100-MAPD=100-10.0027=89.9973%, in the westward direction.

In short-term forecasting, MBVGGM(1,1) had the highest forecasting accuracy given by 100-MAPD=100-8.7227=91.2773%, in the westward direction. See Table 4.64. In the southward direction, GGM(1,1) had the highest forecasting accuracy at 100-MAPD=100-9.4086=90.5914%. In the eastward direction, GGM(1,1) had the highest short-term forecasting accuracy of 100-MAPD=100-14.2377=85.7623%. This is good accuracy according to Table 4.6. Lastly, in this short-term forecasting the highest accuracy attained by OGM(1,1) was 100-MAPD=100-11.9147=88.0853%, in the westward direction.

4.3.4 Formulating the Multivariate Grey Model, GM(1,*n*), in Vehicle Flow Modelling

In this section focus is on the problem of low prediction accuracy of the conventional multivariate Grey Model, GM(1,*n*). The GM(1,*n*) has inherent low precision commonly because of great variations in the time series of the relative variables used. Therefore, the TSA is proposed and employed to improve the precision of the conventional GM(1,*n*). The TSA consists of smoothing the relative variables by a univariate grey model, GM(1,1), adoption of a DGT in the GM(,*n*) and correction of the prediction error based on Fourier series. The FSECA filters noise components in the prediction error. Thus a Variable Smoothed-Grouped data-Fourier series based multivariate Grey Model denoted as VSGFGM(1,*n*) is developed in this thesis.

In this section of this thesis the prediction accuracy of a conventional first order three variable grey model denoted as GM(1,3) is improved by smoothing its two relative variables in addition to adopting the DGT and Fourier series error correction.

As mentioned above based on RVSA, DGT and FSECA an optimized GM(1,3) is developed and referred to as VSGFGM(1,3). And three empirical examples for improving the prediction accuracy of the conventional GM(1,*n*) are presented in the following sections.

4.3.4.1 Empirical Example 1

An empirical example to validate the TSA method in improving the prediction accuracy of the conventional GM(1,n) is presented. Considered was a vehicle traffic flow system and factors that affect its performance. In the real world vehicle flow is influenced by many factors such as weather conditions, accidents, pedestrians, motorcycles among other factors. Thus, in this thesis prediction of vehicle traffic flow was considered under the influence of pedestrians and motorcycles. Hence considered was a three variable GM, GM(1,3), where $n = 3$.

a) Data Source

The data sets used were collected from Nairobi CBD, Kenya, and in particular the data were collected from the Kenyatta Avenue Uhuru Highway Roundabout in the westward direction on Tuesday, February 16, 2021, from 6:00 am to 7:45 am. This consists of 21 data points. The data includes three variables, namely vehicle (VEH) flow, pedestrian (PED) and motorcycle (MOT). Vehicle traffic flow is the dependent or output variable influenced by pedestrian and motorcycle as relative or input variables. The raw data, as extracted from Appendix IV Table 2, was as tabulated in Table 4.65. Also included in Table 4.65 is the AGO series of each variable and the MGO series of the dependent variable.

Table 4.65: Day 1 Site 2 Vehicle Flow, Related Factors and Grey Generating Operations

Time Sample t	Time	Raw Data			AGO			MGO
		VEH	PED	MOT	VEH	PED	MOT	VEH
		$x_1^{(0)}(t)$	$x_2^{(0)}(t)$	$x_3^{(0)}(t)$	$x_1^{(1)}(t)$	$x_2^{(1)}(t)$	$x_3^{(1)}(t)$	$z_1^{(1)}(t)$
1	06:00-06:05	48	10	6	48	10	6	
2	06:06-06:10	41	4	0	89	14	6	68.5
3	06:11-06:15	39	6	6	128	20	12	108.5
4	06:16-06:20	47	5	2	175	25	14	151.5
5	06:21-06:25	34	6	5	209	31	19	192.0
6	06:26-06:30	64	21	1	273	52	20	241.0
7	06:31-06:35	64	5	3	337	57	23	305.0
8	06:36-06:40	55	12	2	392	69	25	364.5
9	06:41-06:45	79	7	2	471	76	27	431.5
10	06:46-06:50	52	10	2	523	86	29	497.0
11	06:51-06:55	98	3	2	621	89	31	572.0
12	06:56-07:00	72	12	6	693	101	37	657.0
13	07:01-07:05	86	13	4	779	114	41	736.0
14	07:06-07:10	87	8	2	866	122	43	822.5
15	07:11-07:15	94	8	16	960	130	59	913.0
16	07:16-07:20	86	8	10	1046	138	69	1003.0
17	07:21-07:25	75	11	11	1121	149	80	1083.5
18	07:26-07:30	69	7	5	1190	156	85	1155.5
19	07:31-07:35	99	15	8	1289	171	93	1239.5
20	07:36-07:40	57	12	13	1346	183	106	1317.5
21	07:41-07:45	95	10	5	1441	193	111	1393.5

These three variables are defined and presented in GM(1,3) as follows (Shen et al., 2019):

$x_1^{(0)}(t)$ is the dependent time series of vehicle volume,

$x_2^{(0)}(t)$ is the independent relative time series of pedestrian and

$x_3^{(0)}(t)$ is the independent relative time series of motorcycle.

where t is an order of the time series and in this section, it is the time of the day, presented by time sample t as in Table 4.65.

b) Conventional GM(1,3) in Vehicle Flow Prediction/Modelling

In conventional GM(1,3) simulation the three variables' raw data shown in Table 4.65 were used. From (3.24) the AGO for each variable was generated and recorded in Table 4.65. From (3.26) the MGO of the dependent variable was also generated and recorded. The generated AGO and MGO were used in (3.31) to construct the data matrix X. Then

from (3.30) the vector parameter $\hat{B}$ was computed and the model parameters were obtained as a=0.2037, b_2=2.377 and b_3=-0.7368. Thus, the time response function of (3.29) simplified to:

$$\hat{x}_1^{(1)}(k+1) = e^{-0.2037k}\left(x_1^{(1)}(1) - \frac{1}{0.2037}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)\right) + \frac{1}{0.2037}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)$$

$$k = 0,1,2,\cdots,m-1,\ \ m = 21 \tag{4.44}$$

Lastly, from (4.44) and (3.32), the final forecasted (fitted) values of the original sequence were obtained and tabulated in Table 4.68 and the corresponding relative errors are as tabulated in Table 4.69.

This GM(1,3) simulation was done in MATLAB and Figure 4.55 shows the plot of real, simulated and error values.

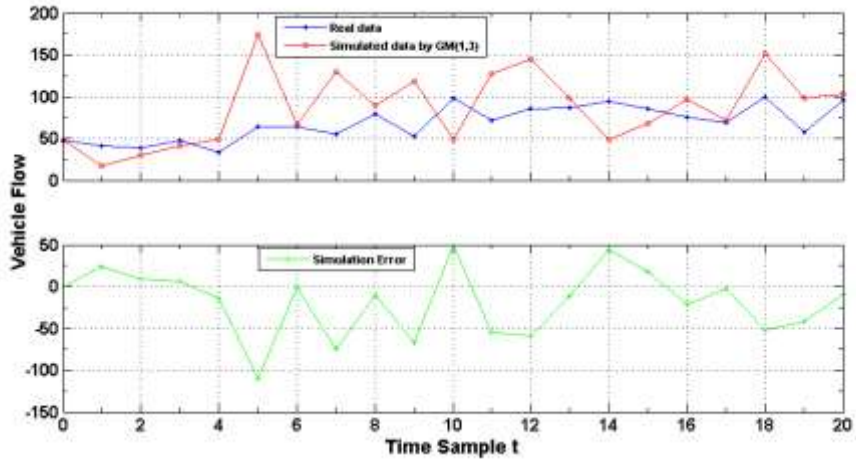

Figure 4.55: Day 1 Site 2 Vehicle Volume Simulation by The Conventional GM(1,3)

Observe the simulation error curve in Figure 4.55. This error curve indicates that the deviation between the real and simulated data is large. That means the fitting accuracy of the conventional GM(1,n) is low. The fitting accuracy of the conventional GM(1,3) was improved in this thesis by the TSA. In (c) below this fitting accuracy improvement is carried out.

c) Improved GM(1,3) in Vehicle Flow Prediction/Modelling

The TSA was employed in improving the conventional GM(1,3)'s fitting accuracy as follows. In the first step, the observed relative variables were smoothed by OGM(1,1). Thus the procedure of OGM(1,1) was adopted in data pre-processing. The AGO of the relative variables are as already provided in Table 4.65. Now from (3.5) or (3.26) the MGO for each variable were calculated and recorded in Table 4.66.

Table 4.66: Day 1 Site 2 Traffic Flow MGO

Time Sample t	Raw Data			MGO		
	VEH	PED	MOT	VEH	PED	MOT
	$x_1^{(0)}(t)$	$x_2^{(0)}(t)$	$x_3^{(0)}(t)$	$z_1^{(1)}(t)$	$z_2^{(1)}(t)$	$z_3^{(1)}(t)$
1	48	10	6			
2	41	4	0	68.5000	12.0000	6.0000
3	39	6	6	108.5000	17.0000	9.0000
4	47	5	2	151.5000	22.5000	13.0000
5	34	6	5	192.0000	28.0000	16.5000
6	64	21	1	241.0000	41.5000	19.5000
7	64	5	3	305.0000	54.5000	21.5000
8	55	12	2	364.5000	63.0000	24.0000
9	79	7	2	431.5000	72.5000	26.0000
10	52	10	2	497.0000	81.0000	28.0000
11	98	3	2	572.0000	87.5000	30.0000
12	72	12	6	657.0000	95.0000	34.0000
13	86	13	4	736.0000	107.5000	39.0000
14	87	8	2	822.5000	118.0000	42.0000
15	94	8	16	913.0000	126.0000	51.0000
16	86	8	10	1003.0000	134.0000	64.0000
17	75	11	11	1083.5000	143.5000	74.5000
18	69	7	5	1155.5000	152.5000	82.5000
19	99	15	8	1239.5000	163.5000	89.0000
20	57	12	13	1317.5000	177.0000	99.5000
21	95	10	5	1393.5000	188.0000	108.5000

Using the MGO in Table 4.66 the data matrix A of (3.16) was constructed for each relative variable and from (3.33) the model parameters a and b were obtained for each relative variable. For PED variable a= -0.0230, b=6.9828 and for MOT variable a= -0.0811 and b=1.6898. Therefore, for the PED variable (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 313.6e^{0.023r} - 303.6, \tag{4.45}$$

And for the MOT variable (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 26.836e^{0.0811r} - 20.836, \quad r = 0,1,2,\cdots,m-1 \tag{4.46}$$

The IAGO was applied on (4.45) and (4.46) and new time series (smoothed relative variables) were generated as tabulated in Table 4.68.

In the second step, the DGT was introduced in GM(1,3) modelling. Note that the DGT involves many data groups and hence many values of the parameters a and b are obtained. For this case, according to (4.16), the 21 data points results to 18 groups of data. Hence, it results to 18 time response equations each with unique values of parameters a and b. Because these response equations and their parameters are many they are not presented in this thesis. However, they are easily computed in MATLAB simulation.

Now, GM(1,3) modelling procedure was followed based on the raw vehicle flow, smoothed relative variables and DGT. This combination of the RVSA and DGT in GM(1,3) develops a Variable Smoothed-Grouped data multivariate Grey Model denoted as VSGGM(1,3). VSGGM(1,3) simulation was carried out in MATLAB and final forecasted (fitted) values of VSGGM(1,3) were obtained and tabulated in Table 4.68. A time series error (i.e. the difference between the raw vehicle flow data and the VSGGM(1,*n*)'s fitted data) was generated and obtained as:

$$\begin{aligned} e=\{&0 \quad 0.1308 \quad 6.2043 \quad 2.4225 \quad 25.4598 \quad 0.0133 \quad 8.7298 \quad 22.1080 \quad 0.6115 \\ &44.9020 \; 2.3101 \quad 31.1330 \quad 11.6173 \quad 15.6295 \quad 12.0428 \quad 16.4169 \; 17.3974 \\ &23.5759 \quad -1.5713 \quad 99.7484 \quad 51.4499\} \end{aligned} \tag{4.47}$$

The real, fitted and error values were plotted in Figure 4.56.

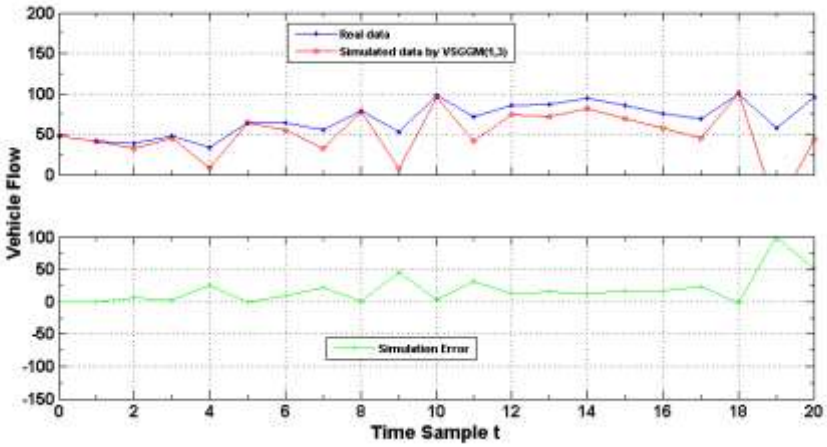

Figure 4.56: Day 1 Site 2 Vehicle Volume Simulation by Improved GM(1,3), VSGGM(1,3).

The error curves in Figures 4.55 and 4.56 are not the same neither similar. It is a clear indication that the fitting accuracy of the conventional GM(1,3) has been improved. The simulation error in Figure 4.56 is approaching zero level. Hence, VSGGM(1,3) is more accurate in vehicle flow fitting compared to GM(1,3). However, at the second last data point the fitted data by VSGGM(1,3) was -42.7484. This error was improved by the FSECA, discussed next.

In the third step, the error set of (4.47) was subjected to FSECA. By Fourier series and truncating the sum over m after some low value $m = N$ the modified random error sequence was approximated by (3.54). From (3.51) T is the period, from (4.44) m=21 and from (3.54) $N = \{(m-1)/2\} - 1$, and $T = m - 1$. Therefore, (3.54) becomes:

$$\hat{e}(k) = \frac{a_0}{2} + \sum_{m=1}^{9} \left[a_m \cos\left(\frac{2\pi m}{20} k\right) + b_m \sin\left(\frac{2\pi m}{20} k\right) \right] \tag{4.48}$$

where a_0, a_m and b_m are Fourier coefficients, $m = 1,2,\cdots,9$ and $k = 2,3,\cdots,21$. To compute the Fourier coefficients (3.55) to (3.58) were used and the values of these coefficients were tabulated in Table 4.67.

Table 4.67: Fourier Coefficients

m	1	2	3	4	5	6	7	8	9
a_0					39.0332				
a_m	7.8512	13.2165	8.5892	11.1317	8.7577	7.4791	3.6882	14.8899	-1.4631
b_m	-2.9362	1.1503	1.0276	4.6905	7.8820	6.6942	3.7229	3.7025	6.5132

These computed values of the Fourier coefficients were substituted back into (4.48) to obtain the modified random error sequence as:

$$\hat{e}(k) = \{0 \quad -5.9607 \quad 12.2957 \quad -3.6690 \quad 31.5512 \quad -6.0782 \quad 14.8213 \quad 16.0166$$
$$6.7030 \quad 38.8106 \quad 8.4015 \quad 25.0416 \quad 17.7088 \quad 9.5381 \quad 18.1343 \quad 10.3255$$
$$23.4888 \quad 17.4845 \quad 4.5201 \quad 93.6570 \quad 57.5413\} \qquad (4.49)$$

This corrected error was added to VSGGM(1,*n*)'s fitted data to give the final vehicle flow prediction. With this Fourier series error correction a hybrid grey model was established and referred to as a Variable Smoothed-Grouped data-Fourier series based multivariate Grey Model denoted as VSGFGM(1,*n*). Therefore, the final restored value of the dependent valuable was obtained based on (3.59) and tabulated in Table 4.68 as VSGFGM(1,3)'s fitted data. These results were also plotted in Figure 4.57.

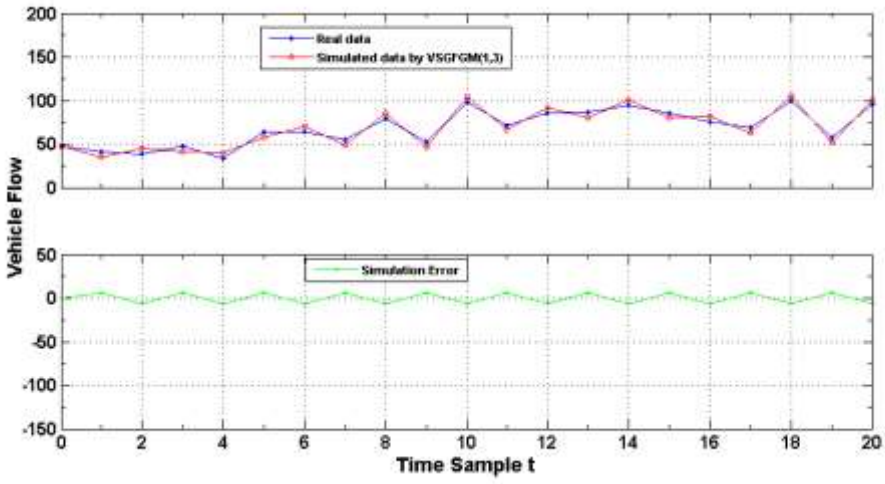

Figure 4.57: Day 1 Site 2 Vehicle Volume Simulation by Improved GM(1,3), VSGFGM(1,3)

Figures 4.55, 4.56 and 4.57 clearly demonstrates that the TSA improves the conventional GM(1,3)'s fitting accuracy. VSGFGM(1,3)) is the most accurate

compared to VSGGM(1,3)) and the conventional GM(1,3). Therefore, the accuracy of the conventional GM(1,3) has been improved as observed from the error curves in Figures 4.55, 4.56 and 4.57. The simulation error curve of Figure 4.57 is approaching zero which is an indication of improved accuracy. Moreover, in Figure 4.57 the simulation data almost fits onto the real data noting that the fitting error that occurred at the second last data point was now corrected.

d) Evaluating the Performance of the Improved GM (1,3) in Traffic Flow Modelling

For the purpose of evaluating the grey models the fitted data by each model were as tabulated in Table 4.68. Moreover, Table 4.69 shows prediction error and accuracy evaluation.

From Table 4.68 the fitted data by the conventional GM(1,3) at time sample 20 was 98.6602 and that of VSGGM(1,3) was -42.7484. After Fourier series error correction VSGFGM(1,3)'s fitted value at this data point was 50.9086. This was a great improvement on the fitting accuracy noting that the real data at this data point is 57. The fitted value of 50.9086 by VSGFGM(1,3) is more closer to 57 compared to GM(1,3)'s and VSGGM(1,3)'s fitted values of 98.6602 and -42.7484, respectively. Thus VSGFGM(1,3) is the most accurate in this context.

Table 4.68: Traffic Flow, Smoothed Variables and Simulated Data

Time Sample t	Raw Data	Existing Model GM(1,3)			Improved Models VSGGM(1,3) and VSGFGM(1,3)			
		Raw Relative Variables		Fitted Value	Smoothed Relative Variables		VSGGM Fitted Value	VSGFGM Fitted Value
	VEH	PED	MOT	VEH	PED	MOT	VEH	VEH
1	48	10	6	48.0000	10.0000	6.0000	48.0000	48.0000
2	41	4	0	17.2645	7.2964	2.2674	40.8692	34.9086
3	39	6	6	30.2513	7.4661	2.4590	32.7957	45.0914
4	47	5	2	40.7850	7.6398	2.6669	44.5775	40.9086
5	34	6	5	48.2576	7.8176	2.8923	8.5402	40.0914
6	64	21	1	174.2489	7.9995	3.1368	63.9867	57.9086
7	64	5	3	65.8822	8.1856	3.4020	55.2702	70.0914
8	55	12	2	129.8792	8.3761	3.6895	32.8920	48.9086
9	79	7	2	89.3849	8.5710	4.0014	78.3885	85.0914
10	52	10	2	118.7329	8.7704	4.3396	7.0980	45.9086
11	98	3	2	49.2143	8.9744	4.7064	95.6899	104.0914
12	72	12	6	126.8544	9.1833	5.1043	40.8670	65.9086
13	86	13	4	144.8652	9.3969	5.5357	74.3827	92.0914
14	87	8	2	98.1569	9.6156	6.0036	71.3705	80.9086
15	94	8	16	49.3499	9.8393	6.5111	81.9572	100.0914
16	86	8	10	67.8505	10.0682	7.0615	69.5831	79.9086
17	75	11	11	96.5661	10.3025	7.6584	57.6026	81.0914
18	69	7	5	71.5265	10.5422	8.3057	45.4241	62.9086
19	99	15	8	150.8273	10.7875	9.0078	100.5713	105.0914
20	57	12	13	98.6602	11.0385	9.7692	-42.7484	50.9086
21	95	10	5	103.4767	11.2953	10.5950	43.5501	101.0914

Table 4.69: Day 1 Site 2 Vehicle Flow Prediction Error and Accuracy Evaluation

Error Indicator	Multivariate Grey Models		
	Existing Model	Improved Models	
	GM(1,3)	VSGGM(1,3)	VSGFGM(1,3)
Error			
RMSE	43.3561	29.6811	5.9446
RMSPE	53.1629	33.4921	7.4041
MAE	32.3621	18.7369	5.8014
MAPD	47.1619	27.3057	8.4544
Accuracy			
100-RMSPE	46.8371	66.5079	92.5959
100-MAPD	52.8381	72.6943	91.5456

Accuracy improvement was explored through error analysis and tabulated in Table 4.69 are the simulation errors and accuracies as obtained by various error indicators. The prediction errors of the conventional and improved grey models were measured and evaluated by Root Mean Square Error (RMSE), Root Mean Square Percentage Error (RMSPE), Mean Absolute Error (MAE) and Mean Absolute Percentage Deviation (MAPD). These error indicators are computed as in section 3.1.3 of this

thesis. From Table 4.6 and Table 4.69 note that, as computed by RMSPE, VSGFGM(1,3) has a higher fitting accuracy at 92.5959% as compared with that of the conventional GM(1,3) at 46.8371%.

4.3.4.2 Empirical example 2

Based on the TSA method employed in section 4.3.4.1 a second empirical example is presented but this time different traffic data is employed for the purpose of validating the portability and applicability of the improved GM(1,3) in vehicle traffic flow modelling and forecasting. Following a similar procedure, as in section 4.3.4.1, simulation was performed by the conventional GM(1,3), VSGGM(1,3) and VSGFGM(1,3) and the results were plotted as shown in Figures 4.58, 4.59 and 4.60, respectively. In Table 4.74 the error and accuracy analysis of this case scenario was tabulated.

a) Data Source

Used is the Nairobi CBD traffic data collected from the Haile Selassie Avenue-Moi Avenue Roundabout in the eastward direction (see Appendix IV Table 14). The data were collected on Wednesday, February 17, 2021, from 6:00 am to 7:45 am. Note that this is day 2 data and as before the data includes three variables, namely vehicle flow, pedestrian and motorcycle. The data were as tabulated in Table 4.70 and defined as in section 4.3.4.1. Also included in Table 4.70 is the AGO series for each variable and the background value (i.e. MGO) of the dependent variable.

b) Conventional GM(1,3) in Vehicle Flow Prediction/Modelling

The three variables' raw data shown in Table 4.70 were used. From (3.24) the AGO for each variable was generated and recorded in Table 4.70. From (3.26) the MGO of the dependent variable was also generated and recorded. The generated AGO and MGO were used in (3.31) to construct the data matrix X. Then the vector parameter $\hat{B}$ was computed from (3.30) and the model parameters were obtained as a= -0.0986, b_2= -0.1595 and b_3=7.6001. Thus the time response function of (3.29) simplified to:

$$\hat{x}_1^{(1)}(k+1) = e^{0.0986k}\left(x_1^{(1)}(1) + \frac{1}{0.0986}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)\right)$$

$$-\frac{1}{0.0986}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)$$

$$k = 0,1,2,\cdots,m-1,\ \ m = 21 \qquad (4.50)$$

Table 4.70: Day 2 Site 7 Vehicle Flow, Related Factors and Grey Generating Operations

Time Sample t	Time	Raw Data			AGO			MGO
		VEH	PED	MOT	VEH	PED	MOT	VEH
		$x_1^{(0)}(t)$	$x_2^{(0)}(t)$	$x_3^{(0)}(t)$	$x_1^{(1)}(t)$	$x_2^{(1)}(t)$	$x_3^{(1)}(t)$	$z_1^{(1)}(t)$
1	06:00-06:05	50	128	5	50	128	5	
2	06:06-06:10	100	133	4	150	261	9	100.0
3	06:11-06:15	75	192	3	225	453	12	187.5
4	06:16-06:20	102	171	4	327	624	16	276.0
5	06:21-06:25	72	181	3	399	805	19	363.0
6	06:26-06:30	80	200	3	479	1005	22	439.0
7	06:31-06:35	95	148	3	574	1153	25	526.5
8	06:36-06:40	148	213	4	722	1366	29	648.0
9	06:41-06:45	55	161	5	777	1527	34	749.5
10	06:46-06:50	60	176	4	837	1703	38	807.0
11	06:51-06:55	80	285	4	917	1988	42	877.0
12	06:56-07:00	62	223	2	979	2211	44	948.0
13	07:01-07:05	65	197	2	1044	2408	46	1011.5
14	07:06-07:10	53	186	3	1097	2594	49	1070.5
15	07:11-07:15	70	342	3	1167	2936	52	1132.0
16	07:16-07:20	72	195	6	1239	3131	58	1203.0
17	07:21-07:25	60	204	3	1299	3335	61	1269.0
18	07:26-07:30	36	206	9	1335	3541	70	1317.0
19	07:31-07:35	70	192	7	1405	3733	77	1370.0
20	07:36-07:40	90	562	5	1495	4295	82	1450.0
21	07:41-07:45	106	362	5	1601	4657	87	1548.0

From (4.50) and (3.32), the final forecasted (fitted) values of the original sequence were obtained and tabulated in Table 4.73 and the corresponding relative errors are as tabulated in Table 4.74. Figure 4.58 shows the plot of real, simulated and error values of GM(1,3).

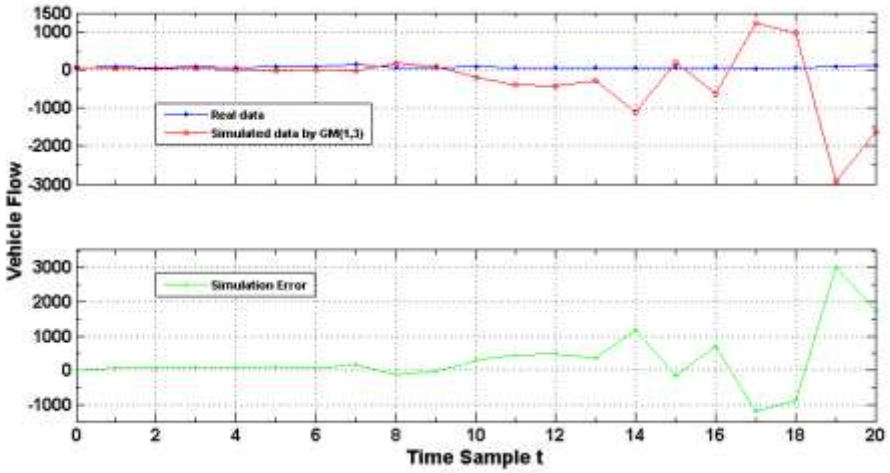

Figure 4.58: Day 2 Site 7 Vehicle Volume Simulation by the Conventional GM(1,3)

Conventional GM(1,3)'s simulation plot of Figure 4.58, shows that this model's fitting accuracy was too low. Its simulation error was in the range of -1500 to 3000 as can be seen in Figure 4.58. Improvement of its fitting accuracy is explained and carried out below.

c) Improved GM(1,3) in Vehicle Flow Prediction/Modelling

The fitting accuracy of the conventional GM(1,3) was improved by smoothing the observed relative variables using OGM(1,1) as a data pre-processing tool. Using the AGO series of Table 4.70 in (3.5) or (3.26) the MGO for each variable were calculated and recorded in Table 4.71.

Table 4.71: Day 1 Site 7 Traffic Flow MGO

Time Sample t	Raw Data			MGO		
	VEH	PED	MOT	VEH	PED	MOT
	$x_1^{(0)}(t)$	$x_2^{(0)}(t)$	$x_3^{(0)}(t)$	$z_1^{(1)}(t)$	$z_2^{(1)}(t)$	$z_3^{(1)}(t)$
1	50	128	5			
2	100	133	4	100.0000	194.5000	7.0000
3	75	192	3	187.5000	357.0000	10.5000
4	102	171	4	276.0000	538.5000	14.0000
5	72	181	3	363.0000	714.5000	17.5000
6	80	200	3	439.0000	905.0000	20.5000
7	95	148	3	526.5000	1079.0000	23.5000
8	148	213	4	648.0000	1259.5000	27.0000
9	55	161	5	749.5000	1446.5000	31.5000
10	60	176	4	807.0000	1615.0000	36.0000
11	80	285	4	877.0000	1845.5000	40.0000
12	62	223	2	948.0000	2099.5000	43.0000
13	65	197	2	1011.5000	2309.5000	45.0000
14	53	186	3	1070.5000	2501.0000	47.5000
15	70	342	3	1132.0000	2765.0000	50.5000
16	72	195	6	1203.0000	3033.5000	55.0000
17	60	204	3	1269.0000	3233.0000	59.5000
18	36	206	9	1317.0000	3438.0000	65.5000
19	70	192	7	1370.0000	3637.0000	73.5000
20	90	562	5	1450.0000	4014.0000	79.5000
21	106	362	5	1548.0000	4476.0000	84.5000

The MGO values of Table 4.71 were used to construct the data matrix A of (3.16) for each relative variable. By (3.33) the model parameters a and b for PED variable were a= -0.0464, b=130.2636 and for the MOT variable a= -0.0361 and b=2.6021. Therefore, for the PED variable (3.9) simplified to:

$$\hat{x}_{(r+1)}^{(1)} \hat{=} 2{,}935.4052\, e^{0.0464r} - 2{,}807.4052, \quad r = 0,1,2,\cdots,m-1 \tag{4.51}$$

And for the MOT variable (3.9) simplified to:

$$\hat{x}_{(r+1)}^{(1)} \hat{=} 77.0803\, e^{0.0361r} - 72.0803, \quad r = 0,1,2,\cdots,m-1 \tag{4.52}$$

Then the smoothed relative variables were generated from (4.51) and (4.52) by the IAGO process. See Table 4.73.

Next the DGT was introduced in GM(1,3) modelling and as it was mentioned earlier the DGT involves many data groups and consequently many model equations are established. The number of equations formed are as many as the number of data groups

formed. Because these equations and their a and b parameters are many they are not presented in this thesis.

Both the smoothed relative variables and the DGT were introduced in the conventional GM(1,3) and the developed grey model i.e. the VSGGM(1,3) resulted to the fitted data tabulated in Table 4.73. The corresponding simulation error was obtained as:

$$e = \{0 \quad 0.4188 \quad 27.0469 \quad -2.0426 \quad 11.2406 \quad 0.5427 \quad 11.8035 \quad -1.2825 \quad 30.6845 \quad 5.1863 \quad -4.3606 \quad 4.9863 \quad -4.7813 \quad 9.9100 \quad -1.4667 \quad -4.4943 \quad -2.7054 \quad 43.9862 \quad 61.3923 \quad -12.4342 \quad -10.0449\} \quad (4.53)$$

VSGGM(1,3)'s real, fitted and error values were as plotted in Figure 4.59.

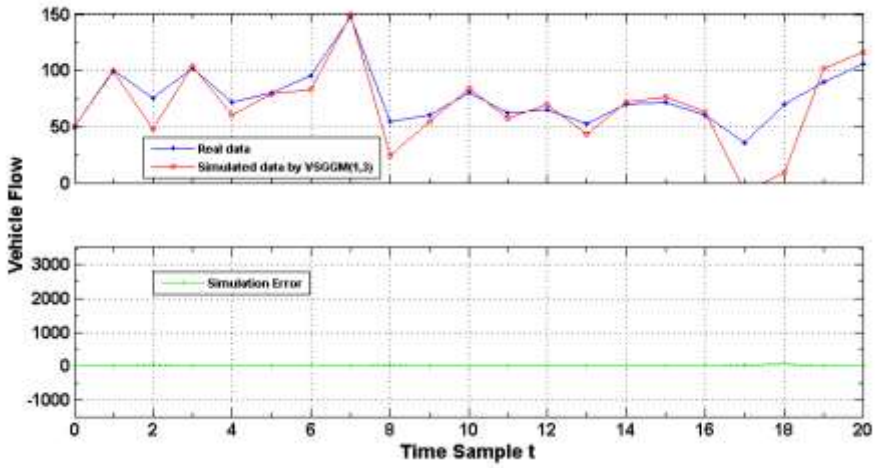

Figure 4.59: Day 2 Site 7 Vehicle Volume Simulation by Improved GM(1,3), VSGGM(1,3)

Comparison of Figures 4.58 and 4.59 clearly demonstrates that VSGGM(1,3) is a great improvement from the conventional GM(1,3). VSGGM(1,3)'s simulation error curve is approaching zero level. Perhaps, much of the inaccuracy was as a result of the 18th data point prediction which was observed at -7.9862 in Table 4.73 and in Figure 4.59. This error was corrected by the FSECA.

The error of (4.53) was corrected by Fourier series. Starting with (4.48) and from (3.55) to (3.58) the Fourier coefficients in Table 4.72 were determined.

Table 4.72: Fourier Coefficients

m	1	2	3	4	5	6	7	8	9
a_0					16.3586				
a_m	4.4036	4.7822	-1.8216	-3.6557	-7.5311	-5.3361	-1.4690	-3.8921	-2.3921
b_m	0.3079	-7.8591	-7.1162	-14.2175	-7.0022	-7.5846	1.5502	-2.1540	7.4819

These Fourier coefficients were substituted back into (4.48) to obtain the modified random error sequence as:

$$\hat{e}(k) = \{0 \quad 4.1205 \quad 23.3453 \quad 1.6590 \quad 7.5390 \quad 4.2443 \quad 8.1019 \quad 2.4191 \quad 26.9829$$
$$8.8879 \quad -8.0622 \quad 8.6879 \quad -8.4829 \quad 13.6117 \quad -5.1684 \quad -0.7927 \quad -6.4070$$
$$47.6878 \quad 57.6907 \quad -8.7326 \quad -13.7465\} \qquad (4.54)$$

This modified error of (4.54) was added to VSGGM(1,*n*)'s fitted data (according to (3.59)) to give the final vehicle flow prediction as tabulated in Table 4.73. This was VSGFGM(1,3)'s fitted data. These results were also plotted in Figure 4.60.

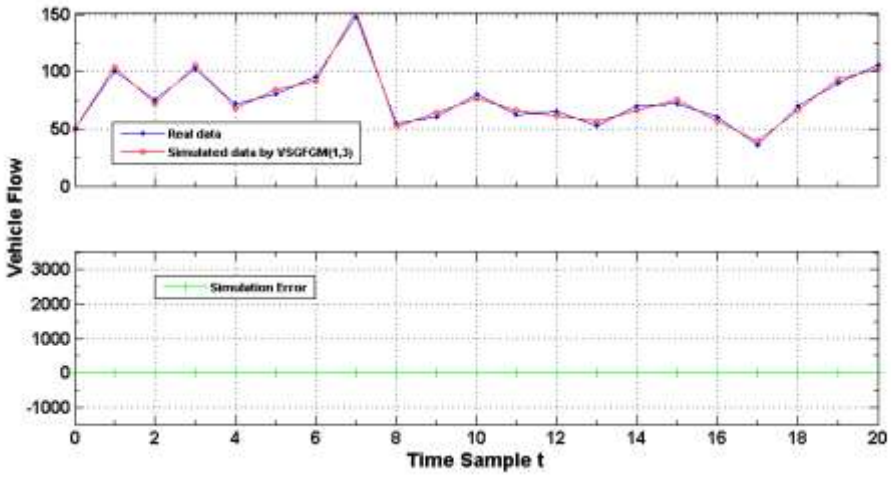

Figure 4.60: Day 2 site 7 Vehicle Volume Simulation by Improved GM(1,3), VSGFGM(1,3).

The fitting of the real and the simulated data in Figures 4.58 to 4.60 can be observed to change from poor fitting in Figure 4.58 to neat fitting in Figure 4.60. In other words the fitting accuracy of the conventional GM(1,3) was improved significantly by introducing the TSA in its modelling procedure. The simulation error in Figure 4.58 is large compared to that of Figure 4.60. Thus VSGFGM(1,3) is more accurate in vehicle

flow fitting compared to the conventional GM(1,3). VSGGM(1,3)'s fitting accuracy lies between that of the conventional GM(1,3) and VSGFGM(1,3).

d) Evaluating the Performance of the Improved GM (1,3) in Traffic Flow Modelling

For easy of comparison and evaluation the fitted data of the three grey models were tabulated in Table 4.73. Moreover, error and accuracy evaluation was tabulated in Table 4.74.

Table 4.73: Traffic Flow, Smoothed Variables and Simulated Data

Time Sample t	Raw Data	Existing Model GM(1,3)			Improved Models VSGGM(1,3) and VSGFGM(1,3)			
		Raw Relative Variables		Fitted Value	Smoothed Relative Variables		VSGGM Fitted Value	VSGFGM Fitted Value
	VEH	PED	MOT	VEH	PED	MOT	VEH	VEH
1	50	128	5	50.0000	128.0000	5.0000	50.0000	50.0000
2	100	133	4	33.3000	139.4117	2.8331	99.5812	103.7016
3	75	192	3	19.4000	146.0326	2.9371	47.9531	71.2984
4	102	171	4	41.4000	152.9678	3.0449	104.0426	105.7016
5	72	181	3	8.3000	160.2324	3.1567	60.7594	68.2984
6	80	200	3	-26.3000	167.8421	3.2726	79.4573	83.7016
7	95	148	3	13.7000	175.8131	3.3927	83.1965	91.2984
8	148	213	4	-15.2000	184.1626	3.5173	149.2825	151.7016
9	55	161	5	165.5000	192.9087	3.6464	24.3155	51.2984
10	60	176	4	79.3000	202.0702	3.7802	54.8137	63.7016
11	80	285	4	-200.6000	211.6668	3.9190	84.3606	76.2984
12	62	223	2	-385.2000	221.7191	4.0628	57.0137	65.7016
13	65	197	2	-414.6000	232.2487	4.2120	69.7813	61.2984
14	53	186	3	-283.3000	243.2785	4.3666	43.0900	56.7016
15	70	342	3	-1097.5000	254.8321	4.5269	71.4667	66.2984
16	72	195	6	212.7000	266.9344	4.6931	76.4943	75.7016
17	60	204	3	-628.2000	279.6114	4.8653	62.7054	56.2984
18	36	206	9	1243.1000	292.8904	5.0439	-7.9862	39.7016
19	70	192	7	964.6000	306.8002	5.2291	8.6077	66.2984
20	90	562	5	-2922.3000	321.3704	5.4210	102.4342	93.7016
21	106	362	5	-1630.6000	336.6327	5.6200	116.0449	102.2984

Table 4.74: Day 2 Site 7 Vehicle Flow Prediction Error and Accuracy Evaluation

Error Indicator	Multivariate Grey Models		
	Existing Model	Improved Models	
	GM(1,3)	VSGGM(1,3)	VSGFGM(1,3)
Error			
RMSE	897.0257	19.6709	3.6124
RMSPE	996.4187	16.2186	3.7836
MAE	529.4186	11.9434	3.5253
MAPD	694.4279	15.6659	4.6241
Accuracy			
100-RMSPE	-896.4187	83.7814	96.2164
100- MAPD	-594.4279	84.3341	95.3759

From Table 4.73 most of the existing GM(1,3)'s fitted values were negative values. Prediction values in grey modelling ought to be positive values. Thus, the negative fitted values of the existing model was a sign of inaccuracy in vehicle flow fitting. This inaccuracy of the existing model was improved by relative data smoothing in combination with the DGT, as demonstrated by Figure 4.59. The accuracy increased from -896.4187% to 83.7814%, as obtained by the RMSPE error indicator, in Table 4.74. The accuracy was further improved by Fourier series error correction as observed in Figure 4.60. Because of Fourier series error correction, the accuracy increased from 83.7814% to 96.2164%. Considering the MAPD error indicator, in Table 4.74, the accuracy of the existing model was improved from -594.4279% to 95.3759%. This is great improvement in the fitting accuracy of the existing model.

4.3.4.3 Empirical Example 3

A third empirical example was carried out to assess the TSA in improving the fitting accuracy of the conventional GM(1,3). The results were as plotted in Figures 4.61, 4.62 and 4.63 for the conventional GM(1,3), VSGGM(1,3) and VSGFGM(1,3), respectively. Moreover, the fitting errors and accuracies were tabulated and evaluated.

a) Data Source

This particular example was based on traffic data collected in the Nairobi CBD, during the third day, from the Kenyatta Avenue-Moi Avenue-Mondlane Street Junction in the

northward direction on Thursday, February 18, 2021, from 6:00 am to 7:45 am (see Appendix IV Table 18). These data were recorded in Table 4.75. It included vehicle flow, pedestrian and motorcycle variables as in the previous two cases. As before included in Table 4.75 are the AGO series for each variable and the background value (i.e. MGO) of the dependent variable.

Table 4.75: Day 3 Site 4 Vehicle Flow, Related Factors and Grey Generating Operations

Time Sample t	Time	Raw Data			AGO			MGO
		VEH	PED	MOT	VEH	PED	MOT	VEH
		$x_1^{(0)}(t)$	$x_2^{(0)}(t)$	$x_3^{(0)}(t)$	$x_1^{(1)}(t)$	$x_2^{(1)}(t)$	$x_3^{(1)}(t)$	$z_1^{(1)}(t)$
1	06:00-06:05	26	96	2	26	96	2	
2	06:06-06:10	35	84	2	61	180	4	43.5000
3	06:11-06:15	41	107	1	102	287	5	81.5000
4	06:16-06:20	39	100	4	141	387	9	121.5000
5	06:21-06:25	40	93	6	181	480	15	161.0000
6	06:26-06:30	52	110	1	233	590	16	207.0000
7	06:31-06:35	48	89	5	281	679	21	257.0000
8	06:36-06:40	44	107	4	325	786	25	303.0000
9	06:41-06:45	56	85	5	381	871	30	353.0000
10	06:46-06:50	65	150	4	446	1021	34	413.5000
11	06:51-06:55	49	147	4	495	1168	38	470.5000
12	06:56-07:00	38	126	6	533	1294	44	514.0000
13	07:01-07:05	42	129	4	575	1423	48	554.0000
14	07:06-07:10	54	146	13	629	1569	61	602.0000
15	07:11-07:15	67	123	8	696	1692	69	662.5000
16	07:16-07:20	58	192	13	754	1884	82	725.0000
17	07:21-07:25	71	156	9	825	2040	91	789.5000
18	07:26-07:30	64	98	12	889	2138	103	857.0000
19	07:31-07:35	52	121	18	941	2259	121	915.0000
20	07:36-07:40	68	203	12	1009	2462	133	975.0000
21	07:41-07:45	71	192	8	1080	2654	141	1044.5000

b) Conventional GM(1,3) in Vehicle Flow Prediction/Modelling

The conventional GM(1,n)'s first order AGO for the three variables were computed using (3.24) and recorded in Table 4.75. The background value of the vehicle variable was also obtained from (3.26) and recorded in Table 4.75. This background value and the AGO series of the relative variables were used in (3.31) to construct the data matrix X. Then (3.31) was used in the least square method of (3.30) and the parameters a=0.4590, b_2=0.2537, and b_3= -0.8715 were determined. Hence the time response function of (3.29) simplified to:

$$\hat{x}_1^{(1)}(k+1) = e^{-0.4590k}\left(x_1^{(1)}(1) - \frac{1}{0.4590}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)\right)$$

$$+\frac{1}{0.4590}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)$$

$$k = 0,1,2,\cdots,m-1, \quad m = 21 \tag{4.55}$$

Now (3.32) generated the IAGO of (4.55) and this IAGO was as recorded in Table 4.78. This was the existing GM(1,3)'s vehicle flow fitted value. The real, simulated and error values of the GM(1,3) were plotted in Figure 4.61.

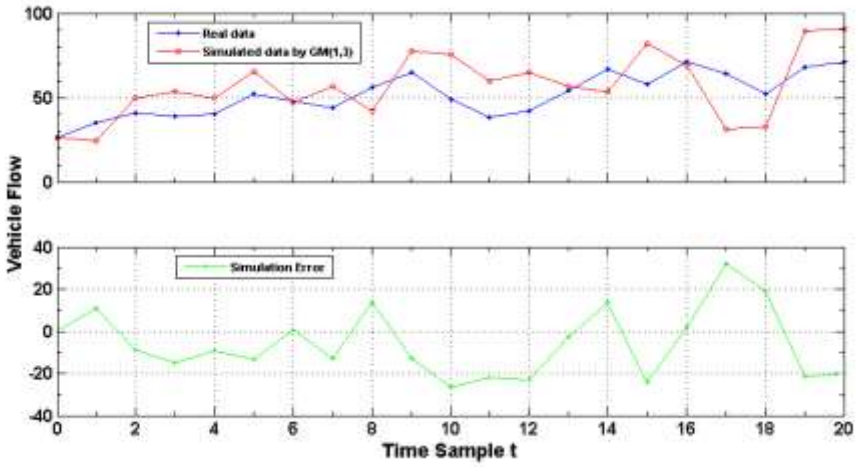

Figure 4.61: Day 3 Site 4 Vehicle Volume Simulation by the Conventional GM(1,3)

It can be noted from Figure 4.61 that there were discrepancies in the fitting of the real and simulated values except at data points $t = 0, 6, 13\ and\ 16$, which are observed to have a good fit. These discrepancies (i.e. errors) were reduced by the TSA in (c) below.

c) Improved GM(1,3) in Vehicle Flow Prediction/Modelling

The discrepancies in the fitting of the conventional GM(1,3) were reduced by first combining the RVSA and the DGT in the GM(1,3)'s procedure. This combination, as

it was mentioned earlier, develops the VSGGM(1,3). Secondly, the FSECA was introduced in the VSGGM(1,3) to further reduce the simulation error.

As before the OGM(1,1), as a data pre-processing tool, was used to smooth the relative variable. The AGO of the relative variables, as already presented in Table 4.75, were used in (3.26) to generate the MGO provided in Table 4.76.

The data matrix A, of (3.16), for each relative variable was constructed from Table 4.76. By (3.33) the model parameters a and b for PED variable were a= -0.0346, b=85.3666 and for the MOT variable a= -0.0853 and b=2.5975. Therefore, the time response equation for the PED variable from (3.9) simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 2{,}563.2428\, e^{0.0346r} - 2{,}467.2428, \quad r = 0,1,2,\cdots,m-1 \tag{4.56}$$

And that of the MOT variable simplified to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 32.4513\, e^{0.0853r} - 30.4513, \quad r = 0,1,2,\cdots,m-1 \tag{4.57}$$

Table 4.76: Day 3 Site 4 Traffic Flow MGO

Time Sample t	Raw Data			MGO		
	VEH	PED	MOT	VEH	PED	MOT
	$x_1^{(0)}(t)$	$x_2^{(0)}(t)$	$x_3^{(0)}(t)$	$z_1^{(1)}(t)$	$z_2^{(1)}(t)$	$z_3^{(1)}(t)$
1	26	96	2			
2	35	84	2	43.5000	138.0000	3.0000
3	41	107	1	81.5000	233.5000	4.5000
4	39	100	4	121.5000	337.0000	7.0000
5	40	93	6	161.0000	433.5000	12.0000
6	52	110	1	207.0000	535.0000	15.5000
7	48	89	5	257.0000	634.5000	18.5000
8	44	107	4	303.0000	732.5000	23.0000
9	56	85	5	353.0000	828.5000	27.5000
10	65	150	4	413.5000	946.0000	32.0000
11	49	147	4	470.5000	1094.5000	36.0000
12	38	126	6	514.0000	1231.0000	41.0000
13	42	129	4	554.0000	1358.5000	46.0000
14	54	146	13	602.0000	1496.0000	54.5000
15	67	123	8	662.5000	1630.5000	65.0000
16	58	192	13	725.0000	1788.0000	75.5000
17	71	156	9	789.5000	1962.0000	86.5000
18	64	98	12	857.0000	2089.0000	97.0000
19	52	121	18	915.0000	2198.5000	112.0000
20	68	203	12	975.0000	2360.5000	127.0000
21	71	192	8	1044.5000	2558.0000	137.0000

Then the smoothed relative variables were generated from (4.56) and (4.57) by the IAGO process given by (3.12). These smoothed variables were as presented in Table 4.78.

As a second step, in the TSA, the DGT was introduced and for the same reasons given earlier the parameters a and b and the resulting time response equations for the DGT are not provided in this section. The developed grey model i.e. the VSGGM(1,3) resulted to the fitted data tabulated in Table 4.78 and the corresponding simulation error was obtained as:

$$e = 0 \quad 0.0230 \quad -2.9822 \quad -0.8475 \quad 1.7957 \quad -3.7941 \quad -1.8244 \quad 3.5114 \quad -1.5771 \quad -6.1531 \quad -0.9981 \quad 2.3289 \quad 0.8272 \quad -1.4166 \quad -6.4437 \quad 5.4865 \quad -5.0493 \quad -3.4346 \quad 11.3829 \quad -4.2635 \quad -8.0608 \tag{4.58}$$

The real, fitted and error values of the VSGGM(1,3) were as plotted in Figure 4.62.

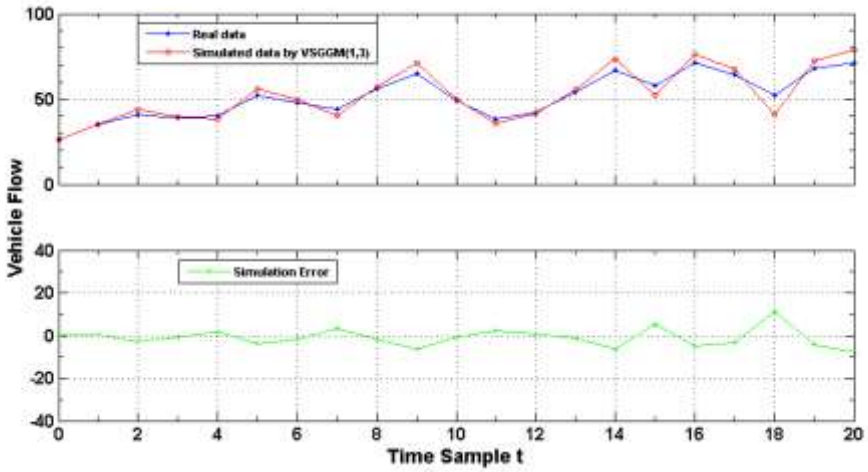

Figure 4.62: Day 3 Site 4 Vehicle Volume Simulation by Improved GM(1,3), VSGGM(1,3)

Comparison of the simulation error curves in Figures 4.61 and 4.62 shows that the fitting error had reduced due to the fact that the relative variables were smoothed and the DGT was introduced in GM(1,3) modelling. It means that the smoothness of the relative variables has an influence on the fitting accuracy of the conventional GM(1,3). Moreover, the DGT was influential in improving the fitting accuracy of the GM(1,3).

From the figures it can be inferred that VSGGM(1,3) is more accurate compared to the conventional GM(1,3).

To further improve the fitting accuracy the third step of the TSA was employed. This step involved application of the Fourier series error correction to modify the error of (4.58). Using (4.48) and (3.55) to (3.58) the Fourier coefficients in Table 4.77 were determined.

Table 4.77: Fourier Coefficients

$\boldsymbol{m}$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
$\boldsymbol{a_0}$					-2.1489				
a_m	-0.1078	-0.1161	0.6936	-0.9671	2.0991	-1.5445	-1.3376	-1.7246	-0.4026
b_m	-0.3979	-0.6763	-1.5085	-1.5615	-1.1199	-2.5815	-3.4152	-0.6439	0.7137

By substituting these Fourier coefficients back into (4.48) the modified random error sequence was obtain as:

$$\hat{e}(k) = \{0 \quad -0.1955 \quad -2.7637 \quad -1.0661 \quad 2.0142 \quad -4.0126 \quad -1.6059 \quad 3.2929 \quad -1.3586 \quad -6.3716 \quad -0.7796 \quad 2.1104 \quad 1.0457 \quad -1.6351 \quad -6.2252 \quad 5.2680 \quad -4.8308 \quad -3.6531 \quad 11.6015 \quad -4.4821 \quad -7.8423\} \qquad (4.59)$$

The modified error of (4.59) was added to VSGGM(1,*n*)'s fitted data (according to (3.59)) to give the final vehicle flow prediction as tabulated in Table 4.78. This was VSGFGM(1,3)'s fitted data. These results were also plotted in Figure 4.63.

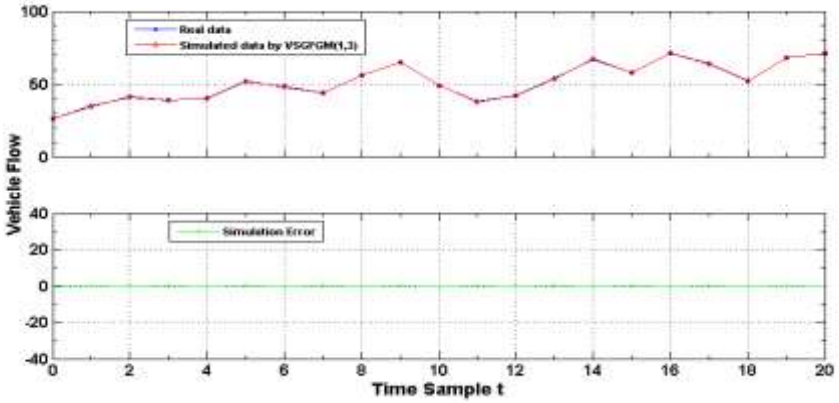

Figure 4.63: Day 3 site 4 Vehicle Volume Simulation by Improved GM(1,3), VSGFGM(1,3)

In Figure 4.63 it can be observed that the fitting of the real and simulated values was very neat and smooth. In addition, the simulation curve is approaching zero level. This is an indication that the FSECA has greatly improved the fitting accuracy of the conventional GM(1,3). So VSGFGM(1,3) was more accurate compared to VSGGM(1,3).

d) Evaluating the Performance of the Improved GM (1,3) in Traffic Flow Modelling

In Table 4.78 the fitted values of the existing and improved grey models were recorded. And in Table 4.79 the fitting errors and accuracies were also tabulated for evaluation purposes.

From the fitted values in Table 4.78 it is evident that VSGFGM(1,3)'s fitted values were so close to the real values. This is also very clear from Table 4.79 which indicates that VSGFGM(1,3)'s fitting accuracy is far much high compared to that of the conventional model. For instance, considering the RMSPE error indicator, the conventional model had an accuracy of 69.7243% which was improved to 99.6281% by VSGFGM(1,3). See Table 4.79.

Table 4.78: Traffic Flow, Smoothed Variable and Simulated Data

Time Sample t	Raw Data	Existing Model GM(1,3)			Improved Models VSGGM(1,3) and VSGFGM(1,3)			
		Raw Relative Variables		Fitted Value	Smoothed Relative Variables		VSGGM Fitted Value	VSGFGM Fitted Value
	VEH	PED	MOT	VEH	PED	MOT	VEH	VEH
1	26	96	2	26.0000	96.0000	2.0000	26.0000	26.0000
2	35	84	2	24.2587	90.2405	2.8896	34.9770	34.7815
3	41	107	1	49.7172	93.4176	3.1469	43.9822	41.2185
4	39	100	4	53.7511	96.7065	3.4272	39.8475	38.7815
5	40	93	6	49.5011	100.1112	3.7323	38.2043	40.2185
6	52	110	1	65.3439	103.6358	4.0647	55.7941	51.7815
7	48	89	5	47.1783	107.2845	4.4266	49.8244	48.2185
8	44	107	4	56.7282	111.0616	4.8208	40.4886	43.7815
9	56	85	5	41.8838	114.9717	5.2500	57.5771	56.2185
10	65	150	4	77.8375	119.0195	5.7175	71.1531	64.7815
11	49	147	4	75.7142	123.2098	6.2266	49.9981	49.2185
12	38	126	6	59.9274	127.5476	6.7811	35.6711	37.7815
13	42	129	4	64.8827	132.0381	7.3849	41.1728	42.2185
14	54	146	13	56.8762	136.6867	8.0425	55.4166	53.7815
15	67	123	8	53.3987	141.4990	8.7586	73.4437	67.2185
16	58	192	13	81.8279	146.4807	9.5385	52.5135	57.7815
17	71	156	9	69.4215	151.6378	10.3879	76.0493	71.2185
18	64	98	12	31.5955	156.9765	11.3128	67.4346	63.7815
19	52	121	18	32.8455	162.5031	12.3202	40.6171	52.2185
20	68	203	12	89.5061	168.2243	13.4172	72.2635	67.7815
21	71	192	8	90.9958	174.1470	14.6119	79.0608	71.2185

Table 4.79: Day 3 Site 4 Vehicle Flow Prediction Error and Accuracy Evaluation

Error Indicator	Multivariate Grey Models		
	Existing Model	Improved Models	
	GM(1,3)	VSGGM(1,3)	VSGFGM(1,3)
Error			
RMSE	16.8536	4.4400	0.2133
RMSPE	30.2757	8.6492	0.3719
MAE	14.4775	3.4381	0.2081
MAPD	28.1507	6.6852	0.4047
Accuracy			
100-RMSPE	69.7243	91.3508	99.6281
100-MAPD	71.8493	93.3148	99.5953

From the results of the three examples above an optimized GM(1,*n*) referred to as Variable Smoothed-Grouped data-Fourier series based multivariate Grey Model denoted as VSGFGM(1,*n*) was established. VSGFGM(1,*n*) was constructed based on the TSA for improving the accuracy of the conventional GM(1,*n*). This TSA consists of smoothing the relative variables by GM(1,1), adoption of the DGT in the GM(,*n*) and correction of the arising error based on Fourier series. With *n*=3 the results

obtained have shown that the VSGFGM(1,3)'s performance was superior to that of the conventional GM(1,3). Consequently, the results have proved that smoothness of the relative variables in GM(1,*n*) is important in predicting and forecasting of systems involving more than one relative variables.

4.4 Assessment of Grouping Technique Based Multivariate Grey Model, GM(1,*n*), on Energy Consumption and Carbon Dioxide Emissions

In order to assess the DGT in improving the fitting accuracy of the OGM(1,3) two application results were presented using two types of data i.e. CO_2 emission and clean energy. OGM(1.3) was improved by the DGT to develop a hybrid grey model called GGM(1,3). The two models were subjected to the two types of data mentioned above. Then the performance of the two grey models were compared for validating the DGT in accuracy improvement.

4.4.1 Data Source

The two numerical examples presented in this section were based on data obtained from (Cheng et al., 2020; Özceylan, 2016). The data from (Özceylan, 2016) consists of CO_2 emission *(mt)* as the dependent or output variable and energy consumption *(mtoe)* and number of motor vehicles *(10^6)* as relative or input variables. On the other hand data from (Cheng et al., 2020) consists of clean energy (10,000 tons of standard coal) as the output variable and economic scale, GDP (CNY 0.1Billion) and population size (10,000 people) as the input variables. These data are as tabulated in Tables 4.80 and 4.82 together with the simulation results.

In reference to section 3.1.2h of this thesis the three variables of the GM(1,3) are defined and presented as follows;

$x_1^{(0)}(t)$ is the dependent time series of CO_2 emission (mt) or clean energy (10,000 tons),

$x_2^{(0)}(t)$ is the independent relative time series of energy consumption (mtoe) or economic scale (GDP) and

$x_3^{(0)}(t)$ is the independent relative time series of the number of motor vehicles (10^6) or population size (10,000 people).

where t is an order of the time series and in this case it is the year, presented by time sample t.

Now, the two numerical examples are presented in the following sections.

4.4.2 GM(1,3) in CO_2 Emission Fitting

From Table 4.80 (Özceylan, 2016) CO_2 emission is the dependent variable and the factors affecting this variable are the energy consumption and the number of motor vehicles. These three factors were used to fit the time series of CO_2 emission.

a) Conventional GM(1,3)'s CO_2 Emission Fitting Results

A similar procedure as in section 4.3.4 was followed in this section. From (3.24) the AGO for each variable was generated. From (3.26) the MGO of the dependent variable was also generated. The generated AGO and MGO were used in (3.31) to construct the data matrix X. Then from (3.30) the vector parameter $\hat{B}$ was computed and the model parameters were obtained as a=0.9002, b_2=2.7189 and b_3=-2.7557. Thus, the time response function of (3.29) simplified to:

$$\hat{x}_1^{(1)}(k+1) = e^{-0.9002k}\left(x_1^{(1)}(1) - \frac{1}{0.9002}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)\right) + \frac{1}{0.9002}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)$$

$$k = 0,1,2,\cdots,m-1,\ \ m = 29 \tag{4.60}$$

Next from (4.60) and (3.32), the final forecasted (fitted) values and the corresponding relative errors were as tabulated in Table 4.80. This OGM(1,3) simulation was done in MATLAB and Figure 4.64 shows the plot of real, simulated and error values.

b) GGM(1,3)'s CO_2 Emission Fitting Results

The DGT was introduced in OGM(1,3) and for the same reasons given earlier the parameters a and b and the resulting time response equations for the DGT are not provided in this section. The developed grey model i.e. the GGM(1,3) resulted to the fitted data and the corresponding relative errors tabulated in Table 4.80. The real, fitted/simulated and error values of the GGM(1,3) were as plotted in Figure 4.65.

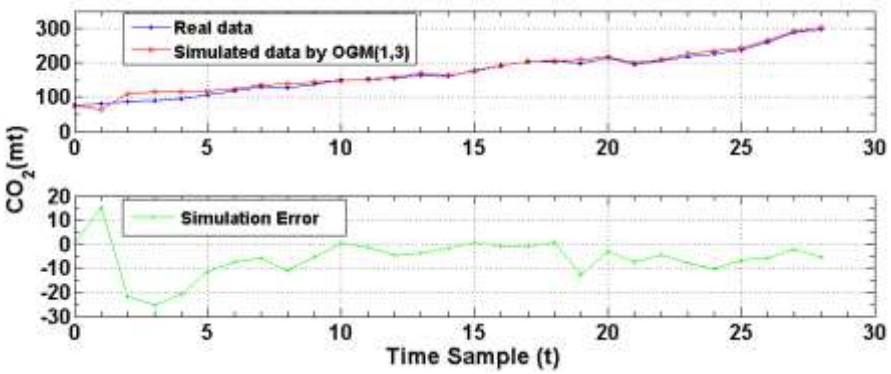

Figure 4.64: CO_2 Emission Simulation by OGM(1,3)

Table 4.80: CO_2 Emission and Related Indicators of Turkey

Time Sample t	Year	CO_2 Emission (mt)	Energy Consumption (mtoe)	Number of Motor Vehicles (10^6)	Existing Model OGM (1,3)		Improved Model GGM(1,3)	
		Output Variable	Input Variable	Input Variable	Fitted Value	Relative Error-%	Fitted Value	Relative Error-%
1	1980	75.702	31.970	1.431	75.7020	0	75.7020	0
2	1981	79.809	32.050	1.513	64.4806	19.2064	75.4781	5.4266
3	1982	86.917	34.390	1.626	108.7593	25.1301	86.6585	0.2974
4	1983	90.468	35.700	1.761	115.9032	28.1151	89.0152	1.6059
5	1984	95.718	37.430	1.927	116.5785	21.7937	89.1734	6.8374
6	1985	106.630	39.400	2.098	118.0837	10.7415	102.814	3.5787
7	1986	116.786	42.470	2.275	124.2462	6.3879	116.3299	0.3905
8	1987	129.801	46.880	2.484	135.4814	4.3762	127.4357	1.8223
9	1988	126.206	47.910	2.705	137.1711	8.6882	123.7014	1.9845
10	1989	139.203	50.710	2.895	144.6574	3.9183	138.3195	0.6347
11	1990	150.667	52.980	3.228	150.3033	0.2414	153.0257	1.5655
12	1991	151.675	54.270	3.548	153.1279	0.9579	153.229	1.0246
13	1992	154.368	56.680	3.974	159.0588	3.0387	153.13	0.8020
14	1993	164.125	60.260	4.568	168.0326	2.3809	154.5696	5.8220
15	1994	161.527	59.120	5.047	163.114	0.9825	157.3057	2.6134
16	1995	176.561	63.680	5.355	175.9391	0.3522	173.3421	1.8231
17	1996	192.342	69.860	5.753	193.3838	0.5416	191.145	0.6223
18	1997	202.722	73.780	6.282	203.6029	0.4345	205.7741	1.5056
19	1998	205.254	74.710	6.923	204.4492	0.3921	205.5053	0.1224
20	1999	196.607	76.770	7.377	209.281	6.4464	199.9555	1.7031
21	2000	215.971	80.500	7.966	218.7433	1.2836	208.5686	3.4275
22	2001	194.379	75.400	8.453	201.8494	3.8432	191.7555	1.3497
23	2002	205.510	78.330	8.612	210.2119	2.2879	201.6107	1.8974
24	2003	218.330	83.840	8.882	226.0269	3.5254	215.4068	1.3389
25	2004	225.222	87.820	9.686	235.5862	4.6018	222.7684	1.0894
26	2005	237.174	91.580	10.666	243.9424	2.8538	237.1868	0.0054
27	2006	261.357	100.580	11.945	267.2093	2.2392	263.8114	0.9391
28	2007	288.445	109.060	12.708	290.4852	0.7073	311.6923	8.0595
29	2008	297.120	113.850	13.512	302.4909	1.8077	306.9714	3.3156

Note: Adapted from "Forecasting CO_2 emission of Turkey: swarm intelligence approaches", by Özceylan, E.**,** 2016, *Int. J. Global Warming, 9*(3), 337-361. DOI: 10.1504/IJGW.2016.075450.

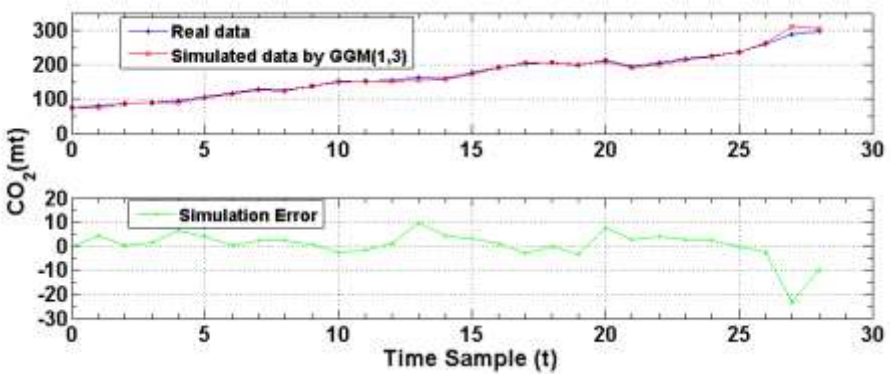

Figure 4.65: CO_2 Emission Simulation by GGM(1,3)

Figures 4.64 and 4.65 clearly demonstrates that the GGM(1,3) is more accurate in CO_2 emission fitting compared to OGM(1,3). This is evident from their corresponding simulation error curves. Therefore, the accuracy of the OGM(1,3) has been improved slightly by the DGT.

c) Evaluating the Performance of OGM (1,3) and GGM (1,3) on CO_2 Emission Fitting

The magnitude of the percentage relative error at each data point was determined from (3.65) and recorded in Table 4.80. Relative error is defined as the ratio of the absolute error of the measurement to the actual measurement. Percent error is the difference between the measured value and the true value, as a percentage of true value. In this case the measured value was the fitted value and the true value was the output variable. From the percentage relative errors it can be observed that the GGM(1,3) had smaller values compared to OGM(1,3)'s error values. Thus GGM(1,3)'s fitted values were close to the real data (i.e. output variable), indicating that GGM(1,3) is more accurate compared to OGM(1,3).

To further explore the accuracy improvement, simulation error analysis was done and tabulated in Table 4.81 are the simulation errors as obtained by various error indicators.

Table 4.81: CO_2 Emission Simulation Error Evaluation

Error Indicator	Grey Model	
	OGM(1,3)	GGM(1,3)
RMSE	9.7098	5.8048
RMSPE	3.2296	3.7689
MAE	7.0877	3.7060
MAPD	4.1552	2.1727

From Tables 4.6 and 4.81 the MAPD error indicator shows that both OGM(1,3) and GGM(1,3) had high fitting accuracy. However, from Table 4.81, it can be noticed that OGM(1,3)'s fitting accuracy had been improved from 100-4.1552=95.8448% to 100-2.1727=97.8273% by the DGT. The error indicators RMSE and MAE also shows that OGM(1,3)'s fitting accuracy was improved. Except with the RMSPE error indicator which indicates that OGM(1,3) had less accuracy compared to GGM(1,3).

4.4.3 GM(1,3) in Clean Energy Consumption Fitting

From Table 4.82 (Cheng et al., 2020), clean energy consumption is the dependent variable and economic scale and population size are the relative variables.

a) Conventional GM(1,3)'s Clean Energy Consumption Fitting Results

As in section 4.4.2 the AGO for each variable was generated from (3.24) whereas from (3.26) the MGO of the dependent variable was generated. These AGO and MGO were used in (3.31) to construct the data matrix X which was used in (3.30) to compute the vector parameter $\hat{B}$. Thus, the model parameters were obtained as a=-0.0918, b_2=-0.0140 and b_3=0.0654. And the time response function of (3.29) simplified to:

$$\hat{x}_1^{(1)}(k+1) = e^{0.0918k}\left(x_1^{(1)}(1) + \frac{1}{0.0918}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)\right) - \frac{1}{0.0918}\sum_{i=2}^{3} b_i x_i^{(1)}(k+1)$$

$$k = 0,1,2,\cdots,m-1,\ \ m = 13 \tag{4.61}$$

Application of the IAGO on (4.61), i.e. use of (3.32), computed the fitted values as tabulated in Table 4.82. The corresponding relative errors were also calculated and tabulated in Table 4.82. Plotted in Figure 4.66 is the real, simulated and error values of the OGM(1,3).

Table 4.82: Clean Energy Consumption and Related Indicators of China

Time Sample t	Year	Clean Energy Consumption	Economic Scale	Population Size	Existing Model OGM(1,3)		Improved Model GGM(1,3)	
		Output Variable	Input Variable	Input Variable	Fitted Value	Relative Error-%	Fitted Value	Relative Error%
1	2006	21,198.56	219,438.5	131,448.0	21199	0.0021	21199	0.0021
2	2007	23,358.15	270,092.3	132,129.0	12893	44.8030	22277	4.6286
3	2008	26,931.32	319,244.6	132,802.0	23360	13.2608	26809	0.4542
4	2009	28,570.71	348,517.7	133,450.0	34024	19.0870	29516	3.3086
5	2010	33,900.91	412,119.3	134,091.0	42490	25.3359	33103	2.3537
6	2011	32,511.61	487,940.2	134,735.0	47727	46.7999	33167	2.0159
7	2012	39,007.39	538,580.0	135,404.0	52284	34.0361	37439	4.0208
8	2013	42,525.13	592,963.2	136,072.0	53953	26.8732	41948	1.3572
9	2014	48,116.08	641,280.6	136,782.0	53426	11.0356	47532	1.2139
10	2015	52,018.51	685,992.9	137,462.0	50220	3.4574	52439	0.8083
11	2016	57,988.00	740,060.8	138,271.0	41278	28.8163	59038	1.8107
12	2017	61,897.02	820,754.3	139,008.0	19692	68.1859	67985	9.8357
13	2018	66,352.00	900,309.5	139,538.0	-11629	117.5262	68653	3.4679

Note: Adapted from "Forecasting Clean Energy Consumption in China by 2025: Using Improved Grey Model GM (1, N)", by Cheng, M., Li, J., Liu, Y., & Liu, B., 2020, *Sustainability, MDPI*, *12*(2), 1-20.

b) GGM(1,3)'s Clean Energy Consumption Fitting Results

Introduction of the DGT in OGM(1,3) resulted to the developed grey model, i.e. the GGM(1,3), whose fitted data and corresponding relative errors were tabulated in Table 4.82. The real, simulated and error values of the GGM(1,3) were as plotted in Figure 4.67.

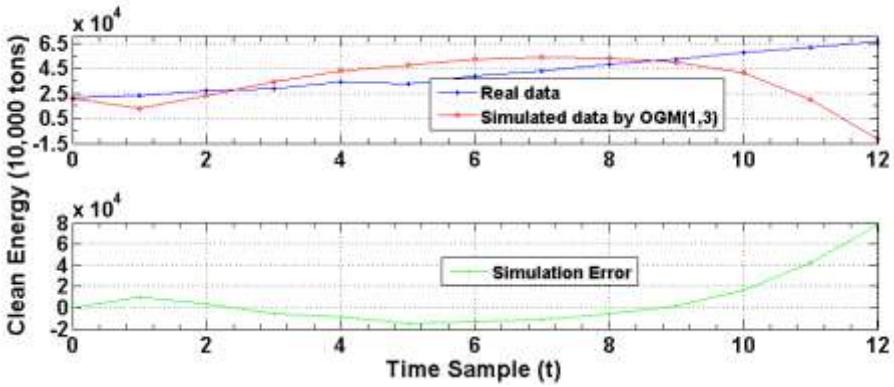

Figure 4.66: Clean Energy Consumption Simulation by OGM(1,3)

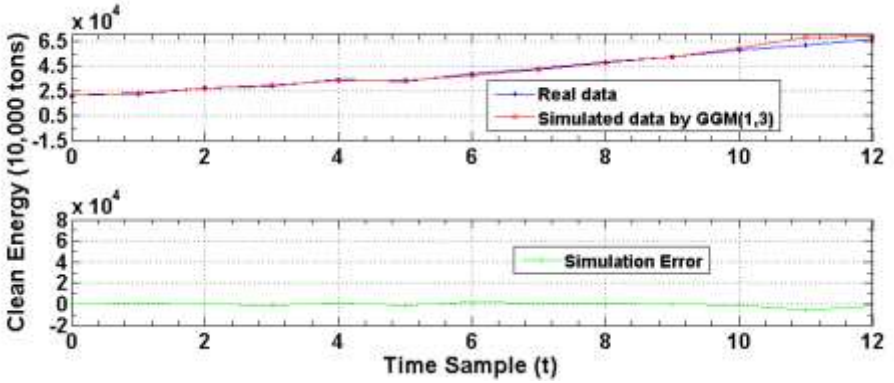

Figure 4.67: Clean Energy Consumption Simulation by GGM(1,3)

From the plots in Figures 4.66 and 4.67, it is clear again that the accuracy of the OGM(1,3) has been improved by the DGT. The GGM(1,3) has higher accuracy because its error curve approaches zero level more closely.

c) Evaluating the Performance of OGM (1,3) and GGM (1,3) on Clean Energy Consumption Fitting

The percentage relative errors at each data point were determined from (3.65) and recorded in Table 4.82. From these percentage relative errors it can be observed that GGM(1,3) had much smaller values compared to OGM(1,3)'s relative error values. Thus GGM(1,3) was more accurate in clean energy consumption fitting compared to OGM(1,3).

The simulation errors by various error indicators were as well computed and recorded in Table 4.83. Considering the MAPD error indicator accuracy was improved from 100-39.6730=60.3270% to 100-3.0294=96.9706%. This is great improvement in OGM(1,3)'s fitting accuracy.

Table 4.83: Clean Energy Consumption Simulation Error Evaluation

Error Indicator	Grey Model	
	OGM(1,3)	**GGM(1,3)**
RMSE	2.6220e+04	1.9594e+03
RMSPE	72.4159	5.1135
MAE	1.6308e+04	1.2452e+03
MAPD	39.6730	3.0294

4.5 Investigating the Effect of Univariate and Multivariate Formulation on Accuracy of Grey Models on Vehicle Traffic Flow Fitting

In this section the conventional univariate and multivariate grey models were compared to improved univariate and multivariate grey models. This comparison established the effect of the relative factors in vehicle flow modelling. The univariate model was used to fit the main variable without considering the relative variables. On the other hand, the multivariate model was used to fit the main variable with consideration of the relative variables. The main variable in this thesis is the vehicle flow time series data whereas the relative variables are pedestrian and motorcycle time series data.

In section 4.3 it was established that GGM(1,1) and VSGFGM(1,3) had the best performance in vehicle flow fitting. These improved models are compared to establish the influence of the relative variables on vehicle flow. Also considered in this comparison were their corresponding conventional grey models, i.e. the OGM(1,1) and the GM(1,3).

Three cases were considered using the same data as in section 4.3.4. Therefore, the fitting results of GM(1,3) and VSGFGM(1,3) are as presented in section 4.3.4 and these results are used in this section. OGM(1,1) and GGM(1,1)'s vehicle flow fitting were carried out as follows.

4.5.1 Day 1 Site 2 Vehicle Flow Fitting by Univariate and Multivariate Grey Models

a) Vehicle Flow OGM(1,1) Fitting

For the conventional GM(1,1)'s vehicle flow fitting the simulation was done using raw vehicle flow data of Table 4.65 (i.e. data of day 1 site 2 westward direction). From the structure of this model its model parameters were obtained as a= -0.0321 and b=48.3724. Thus the time response function was obtained from (3.9) as:

$$\hat{x}^{(1)}_{(r+1)} \triangleq 1{,}554.9283\, e^{0.0321r} - 1{,}506.9283,\ r = 0,1,2,\cdots,m-1 \tag{4.62}$$

where m=21. From (4.62) and with the IAGO process the fitted vehicle flow data were recorded in Table 4.85. Figure 4.68 shows the plot of the real, simulated and simulation error curves for day 1 (D1), site 2 (S2) in the westward (W) direction. Table 4.86 shows the fitting errors and accuracies as a result of this OGM(1,1) simulation.

b) Vehicle Flow GGM(1,1) Fitting

GGM(1,1)'s fitting simulation involved 18 groups of 4 data points. This was as a result of (3.34) with m=21. Hence many values of the parameters a and b were obtained. It means 18 time response equations were obtained. Table 4.84 shows these parameters and the corresponding time response equations. The time response equations were computed from (3.9).

Table 4.84: D1-S2-W GGM(1,1)'s Parameters and Corresponding Time Response Functions

Groups 1 ~ 18	Model Parameters		Time Response Function
	a	b	$r = 0, 1, 2, \cdots, m-1, \quad m = 21$
G1	-0.0737	34.2622	$\hat{x}^{(1)}_{(r+1)} \triangleq 512.8874\, e^{0.0737r} - 464.8874$
G2	0.0574	45.8883	$\hat{x}^{(1)}_{(r+1)} \triangleq -751.4477\, e^{-0.0574r} + 799.4477$
G3	-0.2045	26.6863	$\hat{x}^{(1)}_{(r+1)} \triangleq 178.4954\, e^{0.2045r} - 130.4954$
G4	-0.2523	24.2336	$\hat{x}^{(1)}_{(r+1)} \triangleq 144.0507\, e^{0.2523r} - 96.0507$
G5	0.0720	70.2465	$\hat{x}^{(1)}_{(r+1)} \triangleq -927.6458\, e^{-0.0720r} + 975.6458$
G6	-0.1236	46.4731	$\hat{x}^{(1)}_{(r+1)} \triangleq 423.9960\, e^{0.1236r} - 375.9960$
G7	0.0212	65.3477	$\hat{x}^{(1)}_{(r+1)} \triangleq -3{,}0344387\, e^{-0.0212r} + 3{,}082.4387$
G8	-0.1467	52.3938	$\hat{x}^{(1)}_{(r+1)} \triangleq 405.1403\, e^{0.1467r} - 357.1493$
G9	-0.1155	52.8296	$\hat{x}^{(1)}_{(r+1)} \triangleq 505.3991\, e^{0.1155r} - 457.3991$
G10	0.0761	99.3378	$\hat{x}^{(1)}_{(r+1)} \triangleq -1{,}257.3587\, e^{-0.0761r} + 1{,}305.3587$
G11	-0.0894	62.4038	$\hat{x}^{(1)}_{(r+1)} \triangleq 746.0291\, e^{0.0894r} - 698.0291$
G12	-0.0454	79.7822	$\hat{x}^{(1)}_{(r+1)} \triangleq 1{,}805.3172\, e^{0.0454r} - 1{,}757..3172$
G13	0.0055	90.2010	$\hat{x}^{(1)}_{(r+1)} \triangleq -16{,}352.1818\, e^{-0.0055r} + 16{,}400.1818$
G14	0.1110	109.5115	$\hat{x}^{(1)}_{(r+1)} \triangleq -938.5901\, e^{-0.1110r} + 986.5901$
G15	0.1120	100.7026	$\hat{x}^{(1)}_{(r+1)} \triangleq -851.1304\, e^{-0.1120r} + 899.1304$
G16	-0.1594	49.1900	$\hat{x}^{(1)}_{(r+1)} \triangleq 356.5947\, e^{0.1594r} - 308.5947$
G17	0.0686	88.1284	$\hat{x}^{(1)}_{(r+1)} \triangleq -1{,}236.6706\, e^{-0.0686r} + 1{,}284.6706$
G18	0.0282	89.1933	$\hat{x}^{(1)}_{(r+1)} \triangleq 3{,}114.8830\, e^{-0.0282r} + 3{,}162.8830$

For each group equation the generated data series was subjected to IAGO process to obtain the fitted data. The group fitted data series were superimposed and the final fitted data were as recorded in Table 4.85. Figure 4.69 shows GGM(1,1)'s plot of real, fitted and error curves. Included in Table 4.86 are the fitting errors and accuracies of GGM(1,1).

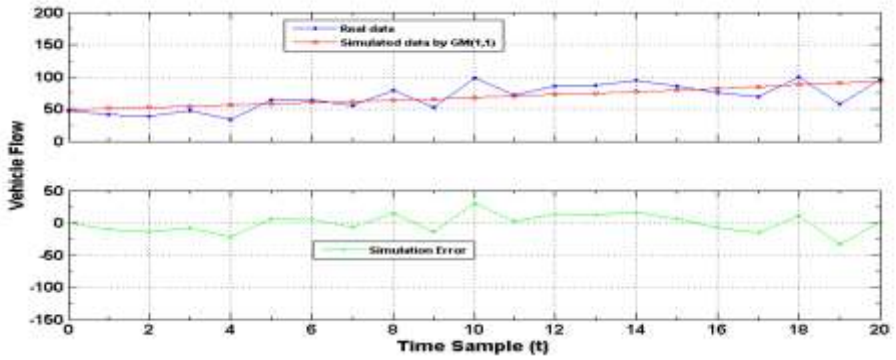

Figure 4.68: D1-S2-W Vehicle Flow OGM(1,1) Fitting

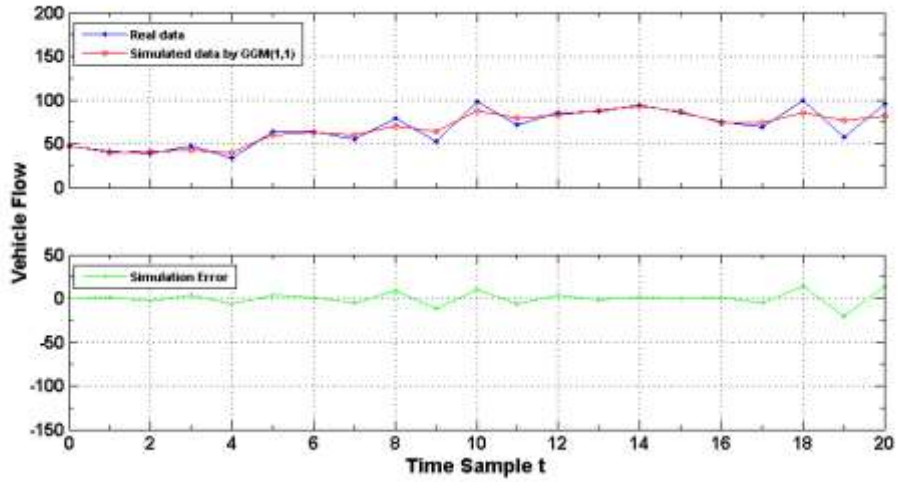

Figure 4.69: D1-S2-W Vehicle Flow GGM(1,1) Fitting

Comparing the vehicle flow fitting in Figures 4.68 and 4.69 shows that GGM(1,1) is more accurate than the OGM(1,1). This is seen from observation of the simulation curves. GGM(1,1)'s simulation error curve is smoother compared to that of OGM(1,1).

c) Evaluation of the Univariate and Multivariate Grey Models in Vehicle Flow Fitting

The fitted values for both the univariate and multivariate grey models were as in Table 4.85. The univariate grey models only involved the main variable which is the VEH variable. Whereas the multivariate grey models involved the relative variables (PED and MOT) in vehicle flow fitting. For the existing GM(1,3) the relative variables were

not smoothed. However, to improve its fitting accuracy the relative variables were smoothed and this resulted to the improved model called VSGFGM(1,3).

Table 4.85: D1-S2-W Vehicle Flow, Relative Variables and Simulated Data

Time Sample *t*	Raw Data	Univariate Grey Models		Multivariate Grey Models					
		Existing OGM (1,1)	Improved GGM (1,1)	Existing Model GM(1,3)			Improved Model VSGFGM(1,3)		
		Fitted Value	Fitted Value	Raw Relative Variables		Fitted Value	Smoothed Relative Variables		Fitted Value
	VEH	VEH	VEH	PED	MOT	VEH	PED	MOT	VEH
1	48	48.0000	48.0000	10	6	48.0000	10.0000	6.0000	48.0000
2	41	50.7237	40.1141	4	0	17.2645	7.2964	2.2674	34.9086
3	39	52.3790	41.1802	6	6	30.2513	7.4661	2.4590	45.0914
4	47	54.0882	42.7189	5	2	40.7850	7.6398	2.6669	40.9086
5	34	55.8532	39.9922	6	5	48.2576	7.8176	2.8923	40.0914
6	64	57.6758	60.0373	21	1	174.2489	7.9995	3.1368	57.9086
7	64	59.5579	62.6871	5	3	65.8822	8.1856	3.4020	70.0914
8	55	61.5014	60.1173	12	2	129.8792	8.3761	3.6895	48.9086
9	79	63.5084	70.0581	7	2	89.3849	8.5710	4.0014	85.0914
10	52	65.5808	63.4444	10	2	118.7329	8.7704	4.3396	45.9086
11	98	67.7209	87.7208	3	2	49.2143	8.9744	4.7064	104.0914
12	72	69.9307	78.5707	12	6	126.8544	9.1833	5.1043	65.9086
13	86	72.2127	82.8111	13	4	144.8652	9.3969	5.5357	92.0914
14	87	74.5692	88.6051	8	2	98.1569	9.6156	6.0036	80.9086
15	94	77.0026	92.6417	8	16	49.3499	9.8393	6.5111	100.0914
16	86	79.5153	86.1024	8	10	67.8505	10.0682	7.0615	79.9086
17	75	82.1101	73.7919	11	11	96.5661	10.3025	7.6584	81.0914
18	69	84.7895	74.3463	7	5	71.5265	10.5422	8.3057	62.9086
19	99	87.5564	84.9093	15	8	150.8273	10.7875	9.0078	105.0914
20	57	90.4136	76.7811	12	13	98.6602	11.0385	9.7692	50.9086
21	95	93.3640	81.3054	10	5	103.4767	11.2953	10.5950	101.0914

Vehicle flow fitting error and accuracy evaluation was as in Table 4.86. From Table 4.86 the following observations were made. GGM(1,1) and VSGFGM(1,3) are more accurate compared to OGM(1,1) and GM(1,3) respectively. GM(1,3) is less accurate compared to OGM(1,1), this indicates that the relative variables have no significance in vehicle flow fitting. However, the relative variables play a great role in improving the fitting accuracy of the conventional multivariate grey model. When the TSA, in which relative variables were smoothed as seen in Table 4.85, was used the accuracy of the conventional GM(1,3) was improved from 46.8371% to 92.5959%. Therefore, the improved multivariate grey model (VSGFGM(1,3)) performed better than the improved univariate grey model (GGM(1,1)). It means that relative variables are

important in traffic flow fitting and, therefore, need to be considered in any time series modelling.

Table 4.86: Day 1 Site 2 Vehicle Flow Prediction Error and Accuracy Evaluation

Error Indicator	Univariate Grey Models		Multivariate Grey Models	
	Existing Model	Improved Model	Existing Model	Improved Model
	OGM(1,1)	GGM(1,1)	GM(1,3)	VSGFGM(1,1)
Error				
RMSE	14.5968	7.8230	43.3561	5.9446
RMSPE	18.4133	10.2630	53.1629	7.4041
MAE	11.8965	5.7783	32.3621	5.8014
MAPD	17.3370	8.4208	47.1619	8.4544
Accuracy				
100-RMSPE	81.5867	89.7370	46.8371	92.5959
100-MAPD	82.6630	91.5792	52.8381	91.5456

4.5.2 Day 2 Site 7 Vehicle Flow Fitting by Univariate and Multivariate Grey Models

a) Vehicle Flow OGM(1,1) Fitting

Similarly, using the raw data of Table 4.70 (i.e. data from day 2 (D2) site 7 (S7) eastward (E) direction) the OGM(1,1)'s simulation gave the model parameters as a=0.0164 and b=91.7669. These parameter values simplified (3.9) and this time response equation became:

$$\hat{x}^{(1)}_{(r+1)} \triangleq -5{,}545.5427\, e^{-0.0164r} + 5{,}595.5427,\ r = 0,1,2,\cdots,m-1 \qquad (4.63)$$

For m=21 in (4.63) and by the IAGO process the fitted data were as in Table 4.88. OGM(1,1)'s results were plotted in Figure 4.70 and the fitting errors and accuracies were as indicated in Table 4.89.

b) Vehicle Flow GGM(1,1) Fitting

As in section 4.5.1 above GGM(1,1)'s fitting simulation resulted to the model parameters and time response equations recorded in Table 4.87.

Table 4.87: D2-S7-E GM(1,1)'s Parameters and Corresponding Time Response Functions

Groups 1 ~ 18	Model Parameters		Time Response Function
	a	b	$r = 0, 1, 2, \cdots, m-1, \quad m = 21$
G1	-0.0119	90.0938	$\hat{X}^{(1)}_{(r+1)} \triangleq 7.620.9076\, e^{0.0119r} - 7{,}570.9076$
G2	0.0162	86.6460	$\hat{X}^{(1)}_{(r+1)} \triangleq -5{,}298..5185\, e^{-0.0162r} + 5{,}348.5185$
G3	0.1400	113.9735	$\hat{X}^{(1)}_{(r+1)} \triangleq -764.0964\, e^{-0.1400r} + 814.0964$
G4	-0.1414	51.5223	$\hat{X}^{(1)}_{(r+1)} \triangleq 414.3727\, e^{0.1414r} - 364.3727$
G5	-0.3323	37.6095	$\hat{X}^{(1)}_{(r+1)} \triangleq 163.1794\, e^{0.3323r} - 113.1794$
G6	0.1594	137.9545	$\hat{X}^{(1)}_{(r+1)} \triangleq -815.4611\, e^{-0.1594r} + 865.4611$
G7	0.5951	239.9187	$\hat{X}^{(1)}_{(r+1)} \triangleq -353.1569\, e^{-0.5951r} + 403.1569$
G8	-0.1993	17.7363	$\hat{X}^{(1)}_{(r+1)} \triangleq 138.9930\, e^{0.1993r} - 88.9930$
G9	-0.0135	65.2290	$\hat{X}^{(1)}_{(r+1)} \triangleq 4{,}881.7778\, e^{0.0135r} - 4{,}831.7778$
G10	0.1143	88.2609	$\hat{X}^{(1)}_{(r+1)} \triangleq -762.1864\, e^{-0.1143r} + 772.1864$
G11	0.0719	72.4452	$\hat{X}^{(1)}_{(r+1)} \triangleq -957.5828\, e^{-0.0719r} + 1{,}007.5828$
G12	-0.0432	56.0069	$\hat{X}^{(1)}_{(r+1)} \triangleq 1{,}346.4560\, e^{0.0432r} - 1{,}296.4560$
G13	-0.1404	43.0664	$\hat{X}^{(1)}_{(r+1)} \triangleq 356.7407\, e^{0.1404r} - 306.7407$
G14	0.0717	78.6169	$\hat{X}^{(1)}_{(r+1)} \triangleq -1{,}046.4700\, e^{-0.0717r} + 1{,}096.4700$
G15	0.3077	107.0769	$\hat{X}^{(1)}_{(r+1)} \triangleq -297.9912\, e^{-0.3077r} + 347.9912$
G16	-0.1084	38.8930	$\hat{X}^{(1)}_{(r+1)} \triangleq 408.7915\, e^{0.1084r} - 358.7915$
G17	-0.3935	10.2452	$\hat{X}^{(1)}_{(r+1)} \triangleq 76.0361\, e^{0.3935r} - 26.0361$
G18	-0.2008	57.1402	$\hat{X}^{(1)}_{(r+1)} \triangleq 334.5627\, e^{0.2008r} - 284.5627$

The forecasted vehicle flow for each group equation were subjected to the IAGO process to obtain the fitted data for each group. The group fitted data were superimposed to obtain the final fitted data which were recorded in Table 4.88. Figure 4.71 shows GGM(1,1)'s plot of real, fitted and error curves. The fitting errors and accuracies of GGM(1,1) are shown in Table 4.89.

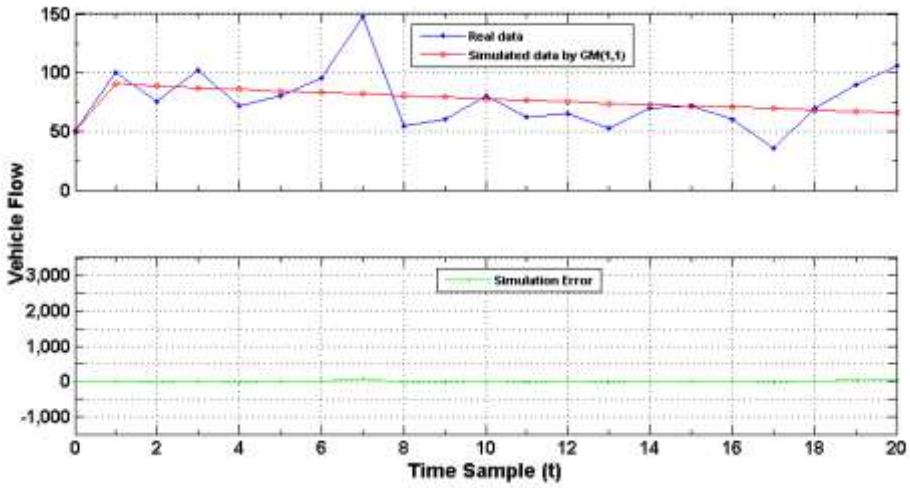

Figure 4.70: D2-S7-E Vehicle Flow OGM(1,1) Fitting

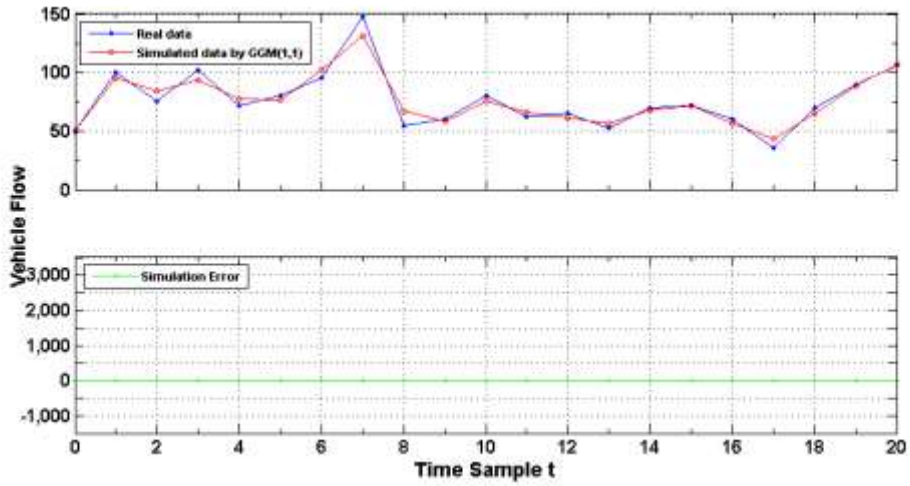

Figure 4.71: D2-S7-E Vehicle Flow GGM(1,1) Fitting

Observation of the error curves in Figures 4.70 and 4.71 indicates that these univariate models were perfect in vehicle flow fitting. This is because both of the error curves are smooth, approaching zero level. However, observing the real and simulated curves shows that GGM(1,1) had good fitting compared to OGM(1.1). These models' performance are evaluated in section (c) below.

c) Evaluation of the Univariate and Multivariate Grey Models in Vehicle Flow Fitting

Tables 4.88 and 4.89 shows the record of the fitted, error and accuracy values for the four grey models. As in the previous section the performance of these models were evaluated and VSGFGM(1,3) emerged as the most accurate among the four models. It had an accuracy of 96.2164%, followed by the GGM(1,1) at an accuracy of 91.6620%. See Table 4.89. These are high accuracies as deduced from Table 4.6.

From Table 4.89 a highly inaccurate vehicle flow estimate by the conventional GM(1,3) was observed at -896.4187%. Of importance is how 'smoothing of the relative variables', DGT and FSECA have played an important role in improving the fitting accuracy of the conventional GM(1,3) from -896.4187% to 96.2164%.

Table 4.88: D2-S7-E Vehicle Flow, Relative Variables and Simulated Data

Time Sample *t*	Raw Data	Univariate Grey Models		Multivariate Grey Models					
		Existing OGM (1,1)	Improved GGM (1,1)	Existing Model GM(1,3)			Improved Model VSGFGM(1,3)		
		Fitted Value	Fitted Value	Raw Relative Variables		Fitted Value	Smoothed Relative Variables		Fitted Value
	VEH	VEH	VEH	PED	MOT	VEH	PED	MOT	VEH
1	50	50.0000	50.0000	128	5	50.0000	128.0000	5.0000	50.0000
2	100	90.2012	95.6164	133	4	33.3000	139.4117	2.8331	103.7016
3	75	88.7301	83.8908	192	3	19.4000	146.0326	2.9371	71.2984
4	102	87.2831	93.7460	171	4	41.4000	152.9678	3.0449	105.7016
5	72	85.8596	77.1113	181	3	8.3000	160.2324	3.1567	68.2984
6	80	84.4594	76.8928	200	3	-26.3000	167.8421	3.2726	83.7016
7	95	83.0820	101.6252	148	3	13.7000	175.8131	3.3927	91.2984
8	148	81.7271	131.6898	213	4	-15.2000	184.1626	3.5173	151.7016
9	55	80.3942	66.9080	161	5	165.5000	192.9087	3.6464	51.2984
10	60	79.0831	58.0622	176	4	79.3000	202.0702	3.7802	63.7016
11	80	77.7934	75.5277	285	4	-200.6000	211.6668	3.9190	76.2984
12	62	76.5247	65.8021	223	2	-385.2000	221.7191	4.0628	65.7016
13	65	75.2767	61.5126	197	2	-414.6000	232.2487	4.2120	61.2984
14	53	74.0491	56.8438	186	3	-283.3000	243.2785	4.3666	56.7016
15	70	72.8414	68.0143	342	3	-1097.5000	254.8321	4.5269	66.2984
16	72	71.6535	71.7609	195	6	212.7000	266.9344	4.6931	75.7016
17	60	70.4849	56.5001	204	3	-628.2000	279.6114	4.8653	56.2984
18	36	69.3354	43.0610	206	9	1243.1000	292.8904	5.0439	39.7016
19	70	68.2047	64.6774	192	7	964.6000	306.8002	5.2291	66.2984
20	90	67.0924	89.1326	562	5	-2922.3000	321.3704	5.4210	93.7016
21	106	65.9982	106.5195	362	5	-1630.6000	336.6327	5.6200	102.2984

Table 4.89: Day 2 Site 7 Vehicle Flow Prediction Error and Accuracy Evaluation

Error Indicator	Univariate Grey Models		Multivariate Grey Models	
	Existing Model	Improved Model	Existing Model	Improved Model
	OGM(1,1)	GGM(1,1)	GM(1,3)	VSGFGM(1,3)
Error				
RMSE	22.2307	6.2278	897.0257	3.6124
RMSPE	32.4625	8.3380	996.4187	3.7836
MAE	16.1430	4.8395	529.4186	3.5253
MAPD	21.1745	6.3478	694.4279	4.6241
Accuracy				
100-RMSPE	67.5375	91.6620	-896.4187	96.2164
100-MAPD	78.8255	93.6522	-594.4279	95.3759

4.5.3 Day 3 Site 4 Vehicle Flow Fitting by Univariate and Multivariate Grey Models

a) Vehicle Flow OGM(1,1) Fitting

Using raw vehicle flow data of Table 4.75 (i.e. data of day 3 (D3) site 4 (S4) northward (N) direction) OGM(1,1)'s simulation gave the model parameters as a= -0.0287 and b=38.2533. Then the time response function of (3.9) reduced to:

$$\hat{x}^{(1)}_{(r+1)} \hat{=} 1{,}358.8676\, e^{0.0287r} - 1{,}332.8676,\ r = 0,1,2,\cdots,m-1 \tag{4.64}$$

where m=21. From (4.64) and with the IAGO process the fitted vehicle flow data were recorded in Table 4.91. Figure 4.72 shows the plot of this OGM(1,1) simulation. The fitting errors and accuracies were as in Table 4.92.

b) Vehicle Flow GGM(1,1) Fitting

GGM(1,1)'s vehicle flow fitting involved 18 groups of 4 data points. The model parameter values a and b were obtained together with the corresponding time response equations. These were as recorded in Table 4.90. These time response equations were computed from (3.9).

Each group's time response equation generated a data series which was subjected to the IAGO process to obtain group fitted data. The group fitted data series were superimposed and the final fitted data were as recorded in Table 4.91. Figure 4.73 shows the plot of GGM(1,1)'s results and in Table 4.92 its fitting errors and accuracies were recorded.

Table 4.90: D3-S4-N GGM(1,1)'s Parameters and Corresponding Time Response Functions

Groups 1 ~ 18	Model Parameters		Time Response Function
	a	b	$r = 0, 1, 2, \cdots, m-1, \quad m = 21$
G1	-0.0504	34.1926	$\hat{X}^{(1)}_{(r+1)} \triangleq 704.4246\, e^{0.0504r} - 678.4246$
G2	0.0127	41.2067	$\hat{X}^{(1)}_{(r+1)} \triangleq -3{,}218.6220\, e^{-0.0127r} + 3{,}244.6220$
G3	-0.1550	27.8300	$\hat{X}^{(1)}_{(r+1)} \triangleq 205.5484\, e^{0.1550r} - 179.5484$
G4	-0.0810	38.0567	$\hat{X}^{(1)}_{(r+1)} \triangleq 495.8358\, e^{0.0810r} - 469.8358$
G5	0.0833	57.5500	$\hat{X}^{(1)}_{(r+1)} \triangleq -664.8764\, e^{-0.0833r} + 690.8764$
G6	-0.0856	38.7762	$\hat{X}^{(1)}_{(r+1)} \triangleq 478.9930\, e^{0.0856r} - 452.9930$
G7	-0.1886	31.7057	$\hat{X}^{(1)}_{(r+1)} \triangleq 194.1108\, e^{0.1886r} - 168.1108$
G8	0.0574	64.2111	$\hat{X}^{(1)}_{(r+1)} \triangleq 1{,}092.6603\, e^{-0.0574r} + 1{,}118.6603$
G9	0.2693	88.6330	$\hat{X}^{(1)}_{(r+1)} \triangleq -303.1237\, e^{-0.2693r} + 329.1237$
G10	0.0863	54.3761	$\hat{X}^{(1)}_{(r+1)} \triangleq -604.0823\, e^{-0.0863r} + 630.0823$
G11	-0.1841	24.2967	$\hat{X}^{(1)}_{(r+1)} \triangleq 157.9756\, e^{0.1841r} - 131.9756$
G12	-0.2298	28.7925	$\hat{X}^{(1)}_{(r+1)} \triangleq 151.2937\, e^{0.2298r} - 125.2937$
G13	-0.0315	55.5602	$\hat{X}^{(1)}_{(r+1)} \triangleq 1{,}789.8159\, e^{0.0315r} - 1{,}763.8159$
G14	-0.0324	60.4513	$\hat{X}^{(1)}_{(r+1)} \triangleq 1{,}891.7809\, e^{0.0324r} - 1{,}865.7809$
G15	-0.0443	57.1790	$\hat{X}^{(1)}_{(r+1)} \triangleq 1{,}316.7223\, e^{0.0443r} - 1{,}290.7223$
G16	0.1501	86.0245	$\hat{X}^{(1)}_{(r+1)} \triangleq -547.1146\, e^{-0.1501r} + 573.1146$
G17	-0.0352	55.6369	$\hat{X}^{(1)}_{(r+1)} \triangleq 1{,}606.5938\, e^{0.0352r} - 1{,}580.5938$
G18	-0.1440	41.6099	$\hat{X}^{(1)}_{(r+1)} \triangleq 314.9576\, e^{0.1440r} - 288.9576$

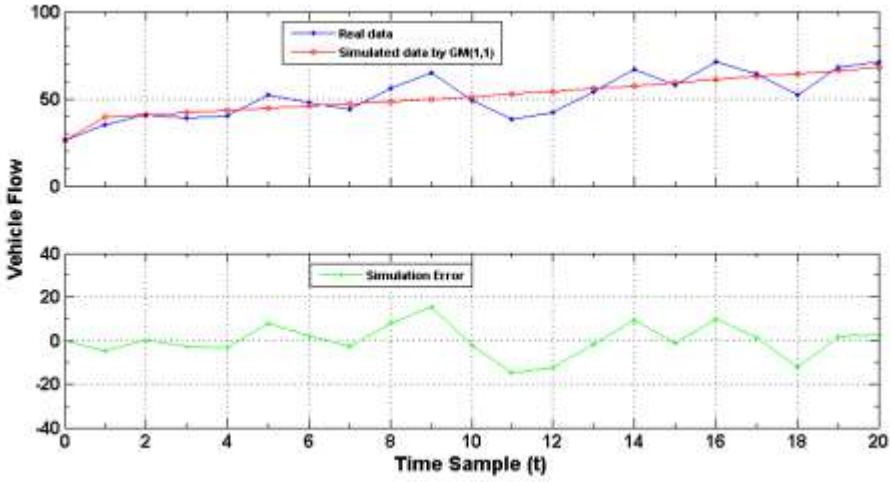

Figure 4.72: D3-S4-N Vehicle Flow OGM(1,1) Fitting

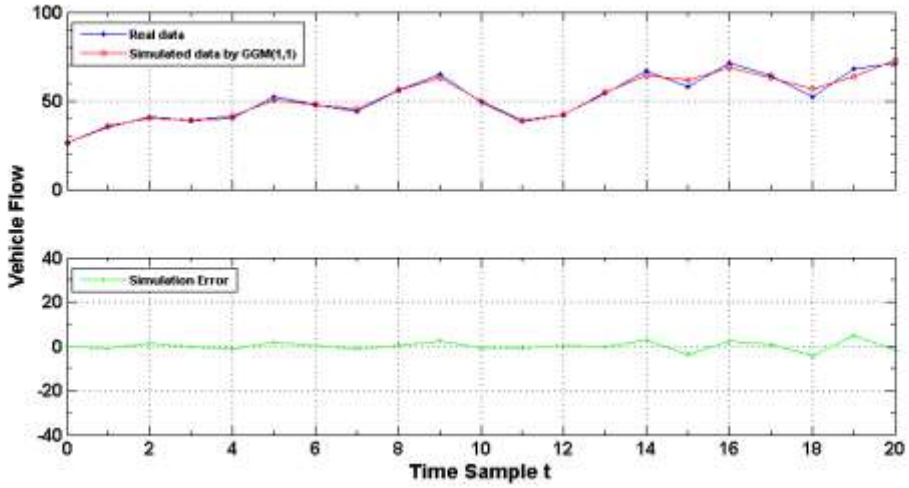

Figure 4.73: D3-S4-N Vehicle Flow GGM(1,1) Fitting

Figures 4.72 and 4.73 indicates that OGM(1,1) has poor vehicle flow fitting whereas GGM(1,1) has good fitting results. This is a clear indication of the role played by the DGT in improving the accuracy of the conventional GM(1,1).

c) Evaluation of the Univariate and Multivariate Grey Models in Vehicle Flow Fitting

Table 4.91 indicates the fitted values of the four grey models. In Table 4.92 are the errors and accuracies of the grey models in this vehicle flow fitting. Tables 4.91 and 4.92 clearly indicates the following observations.

1. The improved models yield good fitting results for the raw vehicle flow data. Considering the RMSPE error indicator; VSGFGM(1,3) had a fitting accuracy of 99.6281% whereas GM(1,3) had a fitting accuracy of 69.7243%. On the other hand GGM(1,1) had an accuracy of 95.9894% whereas GM(1,1)'s accuracy was 86.9847%.
2. The existing multivariate grey model poorly fitted the vehicle flow data as compared to the existing univariate grey model. For instance GM(1,3) had a fitting accuracy of 71.8493% whereas GM(1,1) had an accuracy of 89.1955%, as indicated by the MAPD error indicator. This means that relative variables have no value in vehicle flow fitting. However, as in (1) above when the relative variables are smoothed in VSGFGM(1,3) the accuracy is highly improved.
3. The relative variables as smoothed in the TSA are significant as far as improving the fitting accuracy of the multivariate grey model is concerned. Therefore, it is important to predict vehicle traffic flow under the consideration of relative variables.

Table 4.91: D3-S4-N Vehicle Flow, Relative Variables and Simulated Data

Time Sample t	Raw Data	Univariate Grey Models		Multivariate Grey Models					
		Existing OGM (1,1)	Improved GGM (1,1)	Existing Model GM(1,3)			Improved Model VSGFGM(1,3)		
		Fitted Value	Fitted Value	Raw Relative Variables		Fitted Value	Smoothed Relative Variables		Fitted Value
	VEH	VEH	VEH	PED	MOT	VEH	PED	MOT	VEH
1	26	26.0000	26.0000	96	2	26.0000	96.0000	2.0000	26.0000
2	35	39.5668	35.7063	84	2	24.2587	90.2405	2.8896	34.7815
3	41	40.7209	39.9338	107	1	49.7172	93.4176	3.1469	41.2185
4	39	41.9086	39.0621	100	4	53.7511	96.7065	3.4272	38.7815
5	40	43.1309	41.4001	93	6	49.5011	100.1112	3.7323	40.2185
6	52	44.3889	50.2471	110	1	65.3439	103.6358	4.0647	51.7815
7	48	45.6836	47.8681	89	5	47.1783	107.2845	4.4266	48.2185
8	44	47.0161	45.5154	107	4	56.7282	111.0616	4.8208	43.7815
9	56	48.3874	55.9180	85	5	41.8838	114.9717	5.2500	56.2185
10	65	49.7987	62.8737	150	4	77.8375	119.0195	5.7175	64.7815
11	49	51.2512	49.6057	147	4	75.7142	123.2098	6.2266	49.2185
12	38	52.7460	38.7679	126	6	59.9274	127.5476	6.7811	37.7815
13	42	54.2844	41.8679	129	4	64.8827	132.0381	7.3849	42.2185
14	54	55.8677	54.4323	146	13	56.8762	136.6867	8.0425	53.7815
15	67	57.4972	64.1632	123	8	53.3987	141.4990	8.7586	67.2185
16	58	59.1742	61.5892	192	13	81.8279	146.4807	9.5385	57.7815
17	71	60.9002	68.6332	156	9	69.4215	151.6378	10.3879	71.2185
18	64	62.6764	63.0401	98	12	31.5955	156.9765	11.3128	63.7815
19	52	64.5045	56.3806	121	18	32.8455	162.5031	12.3202	52.2185
20	68	66.3859	63.3156	203	12	89.5061	168.2243	13.4172	67.7815
21	71	68.3222	72.9146	192	8	90.9958	174.1470	14.6119	71.2185

Table 4.92: Day 3 site 4 Vehicle Flow Prediction Error and Accuracy Evaluation

Error Indicator	Univariate Grey Models		Multivariate Grey Models	
	Existing Model	Improved Model	Existing Model	Improved Model
	OGM(1,1)	GGM(1,1)	GM(1,3)	VSGFGM(1,3)
Error				
RMSE	7.3650	2.0346	16.8536	0.2133
RMSPE	13.0153	4.0106	30.2757	0.3719
MAE	5.5566	1.5007	14.4775	0.2081
MAPD	10.8045	2.9179	28.1507	0.4047
Accuracy				
100-RMSPE	86.9847	95.9894	69.7243	99.6281
100-MAPD	89.1955	97.0821	71.8493	99.5953

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This thesis has demonstrated on how to improve the prediction accuracy of the conventional grey models by various methods. The thesis has formulated various methods for improving the prediction accuracy of the conventional grey models. These methods included the newly introduced approaches such as the data grouping technique (DGT), the relative variable smoothing approach (RVSA) and the three-step approach (TSA). Further, these methods were combined with existing methods such as MBV, MIC and FSECA. Thus hybrid grey models were established as a result of introducing these methods into the conventional GM(m,*n*). For instance the DGT in GM(1,1) modelling resulted to a grouped GM(1,1) which was referred to as GGM(1,1) in this thesis. DGT in combination with MBV established the MBVGGM(1,1). Further, the RVSA, DGT and FSECA in the TSA developed an improved grey model which was named as VSGFGM(1,3). The conclusions of this study in line with the objectives are as outlined below.

i. The empirical results demonstrated that the proposed forecasting models outperforms the existing models in terms of accuracy. For instance based on traffic flow data collected from the Nairobi CBD, on day 1 site 1, the GGM(1,1) was the most accurate in traffic flow fitting because its fitting errors were lower compared to three other models. For instance in the westbound direction the GGM(1,1) emerged the most accurate in short-term forecasting with a MAPD value of 7.9902 which translates to an accuracy of 92.0098%, whereas the OGM(1,1) had a MAPD value of 17.1971 which is an 82.8029% accuracy. Based on traffic flow data collected on day 2 site 7 it was noticeable that GGM(1,1) was the most accurate in all directions. It had the highest fitting accuracy of 94.6059%, based on MAPD, in the southward direction. MBVGGM(1,1) was the most accurate in short-term forecasting. It had the highest accuracy of 88.0095% compared with OGM(1,1) which had an accuracy of 86.5416%. That was in consideration of the MAPD error indicator. Finally, based on traffic flow data collected on day 3 site 7,

MBVGGM(1,1) had the highest short-term forecasting accuracy at 91.2773%, in the westward direction. In the southward direction, GGM(1,1) had the highest short-term forecasting accuracy at 90.5914%. This were high accuracies according to Table 4.6. Lastly, in this short-term forecasting the highest accuracy attained by OGM(1,1) was 88.0853%, in the westward direction. Therefore, the DGT has greatly improved the precision of the conventional grey model in traffic flow short-term forecasting. The improvement in accuracy is due to the fact that the DGT adheres to the "difference information" and "new information prior using" principles of the GST. Moreover, the SG technique adheres to the consistency principle that a shorter interval data length leads to more accurate results. Indeed these were promising results and, therefore, the developed grey models are significant in accurate traffic flow modelling and forecasting. Hence the hybrid grey models can enhance the forecasting ability of the ITS.

ii. Firstly, the improved models attained good fitting results on the raw vehicle flow data forecasting. Considering the RMSPE error indicator the hybrid multivariate VSGFGM(1,3) had the highest fitting accuracy of 99.6281% whereas the conventional multivariate GM(1,3) had a fitting accuracy of 69.7243%. On the other hand the hybrid univariate GGM(1,1) attained an accuracy of 95.9894% whereas the conventional univariate GM(1,1)'s accuracy was 86.9847%. Secondly, the existing multivariate grey model poorly fitted the vehicle flow data as compared to the existing univariate grey model. For instance the conventional GM(1,3) had a fitting accuracy of 71.8493% whereas the conventional GM(1,1) had an accuracy of 89.1955%, as was indicated by the MAPD error indicator. This implies that relative variables are not significant in improving vehicle flow fitting accuracy. However, as first stated above when the relative variables in a multivariable GM were smoothed the accuracy of the conventional GM(1,3) was improved and thus VSGFGM(1,3) had the highest accuracy of 99.6281%. Thirdly, and as a conclusion, the relative variables as smoothed in the TSA are significant as far as improving the fitting accuracy of the multivariate grey model is concerned. Therefore, it is important to predict vehicle traffic flow under the consideration of relative variables. An improved multivariate grey model can

attain high traffic flow forecasting accuracies compared with an improved univariate grey model.

iii. The grouping technique based multivariate GM had good performance on both energy consumption and CO_2 emission forecasting. Both OGM(1,3) and GGM(1,3) had high fitting accuracy. However, in one of the three considered cases, OGM(1,3)'s fitting accuracy had been improved from 95.8448% to 97.8273% (based on MAPD) by the DGT on CO_2 emission prediction. On the other hand on energy consumption forecasting, based on the MAPD, the accuracy was improved from 60.3270% to 96.9706%. That is to say OGM(1,3)'s accuracy was at 60.3270% whereas that of GGM(1,3) was at 96.9706%. This was great improvement on OGM(1,3)'s fitting accuracy. Thus the grouping technique based multivariate GM outperformed the conventional multivariate GM on both CO_2 emission and energy consumption forecasting. Similar results were obtained based on the univariate GM as follows. The grouping technique based univariate GM had good performance on both electricity consumption and CO_2 emission forecasting. In modelling medium-term forecasting of electricity consumption the OGM(1,1) at an accuracy of 93.1817% was outperformed by GGM(1,1) at an accuracy of 95.3579% (this was in consideration of the MAPD). On vehicular CO_2 emission forecasting the accuracy of OGM(1,1) was improved from 69.1399% to 76.7674%, by the DGT, based on MAPD. Consequently, the hybrid grey models developed in this thesis are multidisciplinary and can be applied in modelling and forecasting time series data.

5.2 Recommendations and Future Research

5.2.1 Recommendations

In this research it was revealed that the smoothness of the relative variables is a causal to the low accuracy of the multivariate grey model, GM(1,*n*). Therefore, it is recommended that for good results the relative variables of a GM(1,*n*) need to be smoothed before any prediction is done. Use of the conventional GM(1,1) as a data pre-processing tool, as proven in this thesis, can be a great step towards smoothing the relative variables. Additionally, recommended is the incorporation of the newly

improved and developed grey models in PLC and ITS based traffic control systems for the purpose of proactively controlling vehicle traffic flow.

5.2.2 Areas of Future Research

In future the real-time performance of the developed grey models would be evaluated. A systematic study to evaluate and characterize the applicability and boundary of the developed grey model is crucial. In the implementation of the hybrid grey models in ITS, decision support technologies such as Complex Event Processing (CEP), Fuzzy Logic (FL) and Colored Petri Nets (CPN) can be integrated and evaluated.

Further future focus can be based on more state-of-the-art improved forecasting models such as the grey prediction evolution algorithm based on the even grey model (GPEAe), fractional discrete grey model (FDGM(1,1)), Verhulst NGM(1,1), GM(1,1,2), long short-term memory network (LSTM), grey model with cosine term (GM (1,1 | Cos(wt)) and grey model with time power OGM(1,1, tα) can be investigated and compared with the hybrid grey models developed and discussed in this thesis.

In this research unforeseen instances such as when it is raining or lane closures due to issues like a vehicle breakdown/road works etc. were not considered. Therefore, it will be crucial to consider such situations in the future in order to develop more adaptive predictive models.

Traffic flow data during Mondays, Fridays and holidays have different volume patterns and higher volumes are expected on weekends. Traffic flow during those days contains more complex spatial-temporal characteristics, with a large range of changes. For instance, it experiences sudden and irregular characteristics. And, therefore, a future study would be estimation and validation of the proposed models using real-world traffic data collected during such unique days of the week.

Last but not least, in this thesis NaN results were obtained in MATLAB simulations. In future it is important to investigate, solve and get rid of the NaN results in MATLAB. This will enable the proposed grey models to effectively model and forecast

vehicle traffic flow. Additionally, the grey multivariate model is a more reasonable and scientific model compared with the univariate grey model. However, the grey multivariate model still has a shortcoming. When using the grey multivariate model in short-term forecasting of the output variable, the input variable data must be provided; that is to say, to forecast the output variable data for the next 3 instants, the input variable data for the next 3 instants must be available. Now, how to obtain the input variable data for the next 3 instants is the shortcoming. This problem exists not only in the grey multivariate prediction models but also in other prediction models that need to consider influencing factors. As a matter of future investigation, is it possible to use a univariate grey prediction model to predict the data of the input variables before making predictions and then use the predicted data of the input variables to build a multivariate grey prediction model for short-term forecasting?

REFERENCES

Abdulhai, B., Porwal, H., & Recker, W. (2002). Short-Term Traffic Flow Prediction Using Neuro-Genetic Algorithms. *ITS Journal*, *7*(1), 3–41. doi: 10.1080/10248070190048664.

Ajiboye, A. R., Abdullah-Arshah, R., Qin, H., & Isah-Kebbe, H. (2015). Evaluating the Effect of Dataset Size on Predictive Model Using Supervised Learning Technique. *International Journal of Software Engineering & Computer Sciences (IJSECS)*, *1,* 75-84. doi: 10.15282/ijsecs.1.2015.6.0006.

Aldrin, M. (1998). Traffic volume estimation from short-period traffic counts. *Traffic Engineering+ Control, 39*(12), 656-660.

Ali, A., Ayub, N., Shiraz, M., Ullah, N., Gani, A., & Qureshi, MA. (2021). Traffic efficiency models for urban traffic management using mobile crowd sensing: A survey. *Sustainability*. *13*(23), 13068. https://doi.org/10.3390/ su132313068.

Alrawi, F. (2017). The importance of intelligent transport systems in the preservation of the environment and reduction of harmful gases. *Transportation Research Procedia, 24*, 197-203. https://doi.org/10.1016/j.trpro.2017.05.108.

Alvarez, F. M., Troncoso, A., Riquelme, J. C., & Ruiz, J. S. A. (2010). Energy time series forecasting based on pattern sequence similarity. *IEEE Transactions on Knowledge and Data Engineering*, *23*(8), 1230-1243.

Amir, S., Kamal, M. S., Khan, S. S., & Salam, K. A. (2017). PLC based traffic control system with emergency vehicle detection and management," Kannur, India, IEEE, July 6-7, 1467-1472. [Online]. doi**:** 10.1109/ICICICT1.2017.8342786.

Barrera, A., Altava-Ortiz, V., Llasat, M. C., & Barnolas, M. (2007). Heavy rain prediction using deterministic and probabilistic models- the flash flood cases of 11-13 October 2005 in Catalonia (NE Spain). *Advances in Geosciences, 12*, 121-126.

Bas, E., Tekalp, A. M., & Salam, F. S. (2007). Automatic vehicle counting from video for traffic flow analysis. In *2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, June 13-15, 392-397. [Online]. doi:10.1109/IVS.2007.4290146.

Bharti, Redhu, P., & Kumar, K. (2023). Short-term traffic flow prediction based on optimized deep learning neural network: PSO-Bi-LSTM. *Physica A: Statistical Mechanics and its Applications, 625,* 129001. https://doi.org/10.1016/j.physa.2023.129001.

Cai, M. (2010). Non-homogeneous Grey Model NGM(1,1) with initial value modification and its application," *2010 2nd International Conference on Industrial and Information Systems*, Dalian, China, 102-104. doi: 10.1109/INDUSIS.2010.5565904.

Caleb, A. N., Ekwe, E., Muhammad, I., & Fasiu, O. (2022). Grey model analysis of vehicle population, road transport energy consumption and vehicular emissions. *Communications, 25*(1), A1-A13. doi: 10.26552/com.C.2023.006.

Cantarella, G. E., & Fiori, C. (2021). Day-to-Day dynamic multivehicle assignment: deterministic process models. *Discrete Dynamics in Nature and Society, 2021*, 6653905. https://doi.org/10.1155/2021/6653905.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development. 7*(3), 1247-1250. doi: 10.5194/gmd-7-1247-2014.

Chama, C. C. N. (2015). Investigating for contributory factors to traffic congestion in Nairobi City, Kenya. *International Journal of Science and Research (IJSR), 4*(5), 3253-3256.

Chang, C. J., Li, D. C., Dai W. L., & Chen, C. C. (2013). Utilizing an adaptive grey model for short-term time series forecasting: a case study of wafer-level packaging. *Mathematical Problems in Engineering*, *2013*, 1-6.

Changjun, H., Zhengqi, T., Qingshan, Z., & Yuanzhi, C. (2011). Application of improved error GM (1, 1) model on predicting of cultivated land in Yiyang. *Energy Procedia, 5*, 1172-1176.

Chen, G., Shi, Y., She, H., Qin, Y., Zhou, X., Qi, Z., & Guo, W. (2020). Application of improved GM (1, m) model for transformer faults prediction. *IOP Conference Series: Earth and Environmental Science, 558*(5), 052033. doi: 10.1088/1755-1315/558/5/052033.

Chen, L., Ren, Q., Zeng, J., Zou, F., Luo, S., Tian, J., et al. (2023). CSFPre: Expressway key sections based on CEEMDAN-STSGCN-FCM during the holidays for traffic flow prediction. *PLoS ONE, 18*(4), e0283898. https://doi.org/10.1371/journal.pone.0283898.

Chen, X., & Chen, R. (2019). A review on traffic prediction methods for intelligent transportation system in smart cities, *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Suzhou, China, 2019, 1-5. doi: 10.1109/CISP-BMEI48845.2019.8965742.

Cheng, M., Li, J., Liu, Y., & Liu, B. (2020). Forecasting Clean energy consumption in China by 2025: Using improved Grey Model GM (1, N). *Sustainability, MDPI*, *12*(2), 1-20.

Cheng, M., Li, J., Liu, Y., & Xiang, M. (2020). Parameter estimation of modified gray model GM (1, N) and model application. *Communications in Statistics - Simulation and Computation*, *51*(5), 1-18. doi: 10.1080/03610918. 2020. 1771592.

Cheng, Z., Pang, M., & Pavlou, P. (2019). Mitigating traffic congestion: The role of intelligent transportation systems. *Information Systems Research, 31*(3), 1-16. doi: 10.1287/isre.2019.0894.

Chitere, P. O., & Kibua, T. N. (2004). Efforts to improve road safety in Kenya-Achievements and limitations of reforms in the matatu industry. Working

paper, IPAR, Nairobi, Kenya. [Online]. Available: https://www.scribd.com/ document/ 75708374/Road-Safety-Kenya-IPAR.

Claffy, K. C., Polyzos, G. C., & Braun, H. W. (1993). Application of sampling methodologies to network traffic characterization. In *ACM SIGCOMM Computer Communication Review*, New York, NY, USA, *23*(4), September 13-17, 194-203. [Online]. doi:10.1145/166237.166256.

Daniel, O. C. (2016). *Exploring the major causes of road traffic accidents in Nairobi County* [M. S. thesis, CBPS, UoN, Nairobi Kenya]. [Online]. http://hdl.handle.net/11295/98903.

Del, D. G., Lowe, R., Madsen, H., Mikkelsen, P. S., & Rieckermann, J. (2015). Comparison of two stochastic techniques for reliable urban runoff prediction by modeling systematic errors. *Water Resources Research, 51*(7), 5004-5022.

Deng, J. L. (1989). Introduction to Grey System Theory. *The Journal of grey system, 1*(1), 1-24.

Duan, H., Xiao, X., & Pei, L. (2017). Forecasting the short-term traffic flow in the intelligent transportation system based on an inertia nonhomogenous discrete gray model. *2017*, 1-16.

Du, S., Li, T., Gong, X., & Horng, S. J. (2020). A hybrid method for traffic flow forecasting using multimodal deep learning. *Int. J. Comput. Intell. Syst. 13*(1). 85-97.

Farsi, M. (2020). Application of ensemble RNN deep neural network to the fall detection through IoT environment. *Alexandria Engineering Journal, 60*(1), doi: 10.1016/j.aej.2020.06.056.

Gachanja, J. (2015). Mitigating road traffic congestion in Nairobi Metropolitan Region (NMR) (KIPPRA Policy Brief No. 2/2015). Retrieved from https://www.academia.edu/14505158/Mitigating_Road_Traffic_Congestion_in_Nairobi.

Google (n.d.). [Google Maps showing Nairobi city central business district]. Retrieved June 20, 2019.

Guo, H., Yin, J., Zhao, J., Yao, L., Xia, X., & Luo, H. (2015). An ensemble learning for predicting breakdown field strength of polyimide nanocomposite films. *Journal of Nanomaterials, 2015*(7), 11 pages.

Guo-Dong, L., Yamaguchi, D., & Nagai, M. (2006). Non-equidistance grey model based on grey interval weighting accumulated generating operation. in *Proceedings of the 2006 International Conference on Machine Learning; models, Technology and Applications, MLMTA 2006*, Las Vegas, Neveda, USA, June 26 -29, 2006, 180-188.

Hakkert, S., & Gitelman, V. (2005). Measuring the efficiency and effectiveness of traffic police operations: developing guidelines for a systematic monitoring of police enforcement. *In Australasian Road Safety Research Policing Education Conference,* Wellington, New Zealand. https://www.researchgate. net/publication/270645745.

Hamed, T. (2016). Sampling Methods in Research Methodology; How to choose a sampling technique for research. *International Journal of Academic Research in Management, 5*(2), 18-27.

He, F., & Tao, T. (2014). An improved coupling model of grey-system and multivariate linear regression for water consumption forecasting. *Polish Journal of Environmental Studies*, *23*(4), 1165-1174.

Hsu, YT., Liu, MC., Yeh, J., Hung, HF. (2009). Forecasting the turning time of stock market based on Markov–Fourier grey model. *Expert Systems with Applications, 36*(4), 8597–8603.

Hu, YC. (2021). Developing grey prediction with Fourier series using genetic algorithms for tourism demand forecasting. *Qual Quant, 55,* 315–331. https://doi.org/10.1007/s11135-020-01006-5.

Iqelan, B. M. (2017). Forecasts of female breast cancer referrals using grey prediction model GM (1, 1). *Applied Mathematical Sciences*, *11*(54), 2647 -2662.

Ishikawa, R., Honda, N., & Kazama, H. (2005). The wide-area road traffic analysis by microscopic road Traffic Simulator, MITRAM. *IPSJ Transactions on Mathematical Modelling and Its Applications, 46*(SIG17), 46-55.

Jabari, S. E., & Liu, H. (2012). A stochastic model of traffic flow: Theoretical foundations. *Transportation Research Part B Methodological, 46*(1), 156-174. doi: 10.1016/j.trb.2011.09.006.

Jain, S., Jain, S. S., & Jain, G. (2017). Traffic congestion modelling based on origin and destination. *Procedia Engineering, 187*, 442-450.

Jain, V., Sharma, A., & Subramanian, L. (2012). Road traffic congestion in the developing world. *In Proceedings of the 2nd ACM Symposium on Computing for Development.* Atlanta, GA. doi: 10.1145/2160601.2160616.

Javanmardi E., & Liu S. (2019). Exploring grey systems theory-based methods and applications in analyzing socio-economic systems. *Sustainability*, *11*(15), 4192. https://doi.org/10.3390/su11154192.

Jiang, J., Liu, C., Yao, Y., Lu, Y., Xie, W., & Liu, C. (2022). An improved grey model with time power and its application. *Complexity*, *2022*, 1-10. https://doi.org/10.1155/2022/6910865.

Jiang, M., & Liu, Z. (2023). Traffic flow prediction based on dynamic graph spatial-temporal neural network. *Mathematics*, *11*(11), 2528. https://doi.org/10.3390/math11112528.

Jiang, P., Wu, G., Hu, Y-C., Zhang, X., Ren, Y. (2022). Novel fractional grey prediction model with the change-point detection for overseas talent mobility prediction. *Axioms*, *11*(9), 432. https://doi.org/10.3390/axioms11090432.

Jiang, P., Zhou, Q., Jiang, H., & Dong, Y. (2014). An optimized forecasting approach based on grey theory and cuckoo search algorithm: A case study for electricity consumption in new south wales. *Abstract and Applied Analysis*, *2014*, 1-13. https://doi.org/10.1155/2014/183095.

Jiang, S. -Q., Liu, S., & Zhou, X. -C. (2014). Optimization of background value in GM(1,1) based on compound trapezoid formula. *Kongzhi yu Juece/Control and Decision, 29*(12), 2221-2225.

Jong, Y., & Liu, S. (2014). Grey power models based on optimization of initial condition and model parameters. *Grey Syst. Theory Appl., 4*(2), 370-382.

Joubari, O. EL., Othman, J. B., & Veque, V. (2022). A stochastic mobility model for traffic forecasting in urban environments. *Journal of Parallel and Distributed Computing, 165*, 142-155. https://doi.org/10.1016/j.jpdc.2022.03.005.

Kashyap, A. A., Raviraj, S., Devarakonda, A., Nayak K, S. R., Santhosh, K. V., & Bhat, S. J. (2022). Traffic flow prediction models – A review of deep learning techniques, *Cogent Engineering, 9*(1), doi: 10.1080/23311916.2021.2010510.

Khalil, A. Y., JN, A. K., & Ali, S. (2010). Intelligent traffic light flow control system using wireless sensors networks, *J. Inf. Sci. Eng.*, *26*(3), 753-768.

Khuman, A. S., Yang, Y., & John, R. (2013). A New approach to improve the overall accuracy and the filter value accuracy of the GM (1,1) new-information and GM (1,1) metabolic models. *2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Manchester, UK, IEEE, 1282-1287. [Online]. doi: 10.1109/SMC.2013.222.

Kiiru, G. A. (2015). *Technical efficiency analysis of public road transport providers in Nairobi-Kenya* [M. S. thesis, MST-Department of Econometrics and Statistics, KU, Nairobi, Kenya]. [Online]. http://ir-library.ku.ac.ke/handle/123456789/14697.

Kołodziej, J., Hopmann, C., Coppa, G., Grzonka, D., & Widłak, A. (2022). Intelligent transportation systems – Models, challenges, security aspects. In: Kołodziej, J., Repetto, M., Duzha, A. (eds) Cybersecurity of Digital Service Chains. Lecture Notes in Computer Science, *13300*. Springer, Cham. https://doi.org/10.1007/978-3-031-04036-8_3.

Kumar, S. V., (2017). Traffic flow prediction using kalman filtering technique. *Procedia Engineering, 187*, 582-587.

Kumar, S. V., & Vanajakshi, L. (2015). Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review, 7*(3), 21.

Lao, T., Chen, X., & Zhu, J. (2021). The optimized multivariate grey prediction model based on dynamic background value and its application. *Complexity*, *2021*, 1-13. https://doi.org/10.1155/2021/6663773.

Leverger, C., Guyet, T., Malinowski, S., Lemaire, V., Bondu, A., Roze, L., Termier, A., & Marguerie, R. (2021). Probabilistic forecasting of seasonal time series-combining clustering and classification for forecasting. *International conference on Time Series and Forecasting (ITISE).* 1-13. Gran Canaria, Spain.

Leduc, G. (2008). Road traffic data: Collection methods and applications. *Working Papers on Energy, Transport and climate change*, IPTS, Luxembourg, Spain. [Online]. Available: https://www.researchgate.net/publication/254424803.

Li C, Li Y, Xing J. (2023). Multivariate grey prediction model application in civil aviation carbon emission based on fractional order accumulation and background value optimization. *Sustainability*, *15*(11), 9127. https://doi.org/10.3390/su15119127.

Li, H., Zeng B., Wang, J., & Wu, H. (2021). Forecasting the number of new coronavirus infections using an improved grey prediction model. *Iran J Public Health, 50*(9), 1842-1853. doi: 10.18502/ijph.v50i9.7057.

Li, M., Wang, W., De, G., Ji, X. & Tan, Z. (2018). Forecasting carbon emissions related to energy consumption in Beijing-Tianjin-Hebei region based on grey prediction theory and extreme learning machine optimized by support vector machine algorithm, *Energies*, *11*(9), 2475.

Li, Q., & Lin, Y. (2014). Review paper: A briefing to grey systems theory. *Journal of Systems Science and Information*, *2(*2), 178-192.

Li, S., & Wu, N. (2021). A new grey prediction model and its application in landslide displacement prediction. *Chaos, Solitons & Fractals, 147*, 110969. https://doi.org/10.1016/j.chaos.2021.110969.

Ling, B., Gibson, D. R. P., & Middleton D. (2013). Motorcycle detection and counting using stereo camera, IR camera, and microphone array, *Proc. SPIE 8663, Video Surveillance and Transportation Imaging Applications, Burlingame,* California, United States, 86630P (19 March 2013); https://doi.org/10.1117/12.2003220.

Liu, H., & Cocea, M. (2017). Semi-random partitioning of data into training and test sets in granular computing context. *Granular Computing, 2*(4), 357-386.

Liu, H., Zhang, Xy., Yang, Yx. *et al.* (2022). Hourly traffic flow forecasting using a new hybrid modelling method. *J. Cent. South Univ.* 29, 1389–1402. https://doi.org/10.1007/s11771-022-5000-2.

Liu, S., Forrest J., & Yang, Y. (2012). A brief introduction to grey systems theory. *Grey Systems: Theory and Application, 2*(2), 89-104.

Liu, S., Zeng, B., Liu, J., Xie, N., & Yang, Y. (2015). Four basic models of GM(1,1) and their suitable sequences. *Grey Systems: Theory and Application, 5*(2), 141–156.

Liu, X., Peng, H., Bai, Y., Zhu, Y., & Liao, L. (2014). Tourism flows prediction based on an improved grey GM (1, 1) model. *Procedia-Social and Behavioral Sciences, 138*, 767-775.

Liu, Y., & Chen, X. (2009). Design of traffic lights controlling system based on PLC and configuration technology. In *2009 International Conference on Multimedia Information Networking and Security, IEEE*, Hubei, China.

Lotfalipour, M. R., Falahi, M. A., & Bastam, M. (2013). Prediction of CO_2 Emissions in Iran Using Grey and ARIMA Models. *International Journal of Energy Economics and Policy, 3*(3), 229-237.

Lu, M. (2015). Grey system: theory, methods, applications and challenges. *Leverhulme Trust Workshop on Grey Systems and Applications,* Bucharest, Romania.

Lu, Z., Xia, J., Wang, M., Nie, Q., & Ou, J. (2020). Short-term traffic flow forecasting via multi-regime modeling and ensemble learning. *Applied Sciences*, *10*(1), 356. https://doi.org/10.3390/app10010356.

Luo, Y., Zeng, B., & Liao, D. (2013). Non-equidistant GRM(1,1) generated by accumulated generating operation of reciprocal number and its application. *International Journal of Computer Science Issues, 10*(2), 119-123.

Ma, M., Liang, S., Guo, H., & Yang, J. (2017). Short-term traffic flow prediction using a self-adaptive two-dimensional forecasting method. *Advances in Mechanical Engineering, 9*(8), 1-12.

Madhi, M. & Mohamed, N. (2022). Improving GM(1,1) model performance accuracy based on the combination of optimized initial and background values in time series forecasting. *Scientific Research Publishing, 9*(4), e8416.

Madhi, M. & Mohamed, N. (2017c). An Initial condition optimization approach for improving the prediction precision of a GM (1, 1) model. *Mathematical and Computational Applications*, *22*(1), 21.

Mahdi, M. H., & Mohamed, N. (2017a). An improved gm(1,1) model based on modified background value. *Information Technology Journal, 16*(1), 11-16.

Mahdi, M. H., & Mohamed, N. (2017b). Optimized GM(1,1) model based on the modified initial condition. *Journal of Applied Sciences, 17*(2), 90-96.

Manikandan, B. V., Nathan, T. R., Naresh R., & Prasad P. (2023). Traffic flow prediction for intelligent transportation system using machine learning. *International Conference on Smart Engineering for Renewable Energy Technologies (ICSERET-2023), 387*, 1-8. doi: https://doi.org/10.1051/e3sconf/202338705002.

Miranda, S., Carrasco, Y., & Jorge, G. (2011). *Pedestrian volume studies: A case study in the city of Gothenburg* [M. S. thesis, Department of Civil and

Environmental Engineering, Chalmers University of Technology, Gothenburg, Sweden]. [Online]. Available: http://publications.lib.chalmers. se/ records/fulltext/141801.pdf.

Moonchai S., & Rakpuang, W. (2015). A New approach to improve accuracy of grey model gmc(1,n) in time series prediction. *Modelling and Simulation in Engineering*, *2015*, Article ID 126738, 10 pages. https://doi.org/10.1155/ 2015/126738.

Mori, S., Tada, K., Idu, R., & Oya, H. (2013). Traffic jam analysis and CO_2 emission models in Tokushima City via MITRAM. In *(in Japanese) Proc. of 2013 Annual Conf. of Electronics, Information and Systems Society*, *I.E.E. of Japan*, 1046-1051, Kitami.

Mosoti, R., & Moronge, M. (2015). Effect of motor vehicle congestion on the economic performance of Kenya: A case of Nairobi City County. *Strategic Journal of Business & Change Management, 2*(1), 485-502.

Muhammad, A. K. (2011). PLC based intelligent traffic control system. *International Journal of Electrical & Computer Sciences IJECS-IJENS, 11*(6), 69-73.

Nasim, R. (2015). *Architectural evolution of intelligent transport systems (ITS) using cloud computing* [Licentiate dissertation, Karlstad University]. Retrieved from https://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-35719.

National Highway Authority. (2017). *Traffic study report for (Peshawar-Torkham Section) of preliminary design of Peshawar Kabul motorway project.* Islamabad, Pakistan: Associated Consultancy Centre (Pvt.) Ltd. (ACC).

Ngichabe J. K. (2016). *Development of a mobile application for road incident reporting in Kenya* [M. S. thesis, FIT, SU, Nairobi, Kenya]. [Online]. Available: http://su-plus.strathmore.edu/handle/11071/4908.

Nguyen, T. K. L., Le, H. N., Ngo, V. H., & Hoang, B. A. (2020). CRITIC method and grey system theory in the study of global electric cars. *World Electric Vehicle Journal*, *11*(4), 79. https://doi.org/10.3390/wevj11040079.

Niu, D., Wang, P., Wang, Q., Liu, B., Zhang, W., & Meng, X. J. (2014). Research on mid-long term load forecasting based on gray fourier series residual correction model. *Applied Mechanics and Materials, 631-632*, 345-349.

Nizar, H., Ali, K., & Zeinab. F. (2022). Intelligent transportation systems to mitigate road traffic congestion. *Intelligenza Artificiale, 15*(2), 91-104.

Obiri, A. A., Gbeckor, M. S., & Oliver, Y. (2021). Overview and recommendations for road traffic data collection methods and applications in Ghana. *International Journal of Engineering Research and Applications, 11*(2), 1-9.

Ohene-Akoto, J., Twumasi, E., & Frimpong, E. (2023). Enhancement of the prediction accuracy of grey system model using a particle swarm optimized initial condition. *Carpathian Journal of Electrical Engineering, 16*, 106.

Oladimeji, D., Gupta, K., Kose, N. A., Gundogan, K., Ge, L., & Liang, F. (2023). Smart transportation: An overview of technologies and applications. *Sensors, 23*(8). 3880. https://doi.org/10.3390/s23083880.

Özceylan, E. (2016). Forecasting CO_2 emission of Turkey: swarm intelligence approaches. *Int. J. Global Warming, 9*(3), 337-361. doi: 10.1504/IJGW.2016.075450.

Patil, A. G., Kamble, M. S, Sutar M. B, & Bedage P. S. N. (2019). Smart traffic monitoring system using IR sensors based on microcontroller & GSM, GPS technology. *International Research Journal of Engineering and Technology (IRJET), 6*(4), 3181-3187.

Patodiya, D., & Singh, A. P. (2018). A review on PLC and its application on smart traffic control. *Iconic Research and Engineering Journals (IRE Journals), 1*(9), 91-95.

Qin, H., & Zhang, W. (2022). Short-term traffic flow prediction and signal timing optimization based on deep learning. *Wireless Communications and Mobile Computing, 2022*, 1-11. https://doi.org/10.1155/2022/8926445.

Regehr, J., Montufar, J., & Hernandez-Vega, H. (2015). Traffic pattern groups based on hourly traffic variations in urban areas. *Journal of Transportation of the Institute of Transportation Engineers, 7*(1), 1-16.

Ren, C., Chai, C., Yin, C., Ji, H., Cheng, X., Gao, G., & Zhang, H. (2021). Short-term traffic flow prediction: a method of combined deep learnings. *Journal of Advanced Transportation, 2021*, 1-15. https://doi.org/10.1155/2021/9928073.

Roads and Highways Department. (2001). *Manual classified traffic counts instruction guide.* Bangladesh.

Roads Department. (February 2004). *Traffic data collection and analysis* (ISBN 99912-0-417-2). Gaborone, Botswana: Ministry of Works and Transport Roads Department.

Rong, Y., Zhang, X., Feng, X., Ho, T. K., Wei, W., & Xu, D. (2015). Comparative analysis for traffic flow forecasting models with real-life data in Beijing. *Advances in mechanical engineering, 7*(12), 1-9.

Sampson, J. (2017). Traffic flow variations in urban areas. in *36th Southern African Transport Conference (SATC 2017)*, Pretoria, South Africa, July 10-13, 712-727.

Sarraj, Y. R. (2018). Hourly and daily traffic expansion factors on selected roads in Gaza, Palestine. *The Open Civil Engineering Journal, 12*, 355-367. doi: 10.2174/1874149501812010355.

Sayed, S.A., Abdel-Hamid, Y., & Hefny, H. A. (2023). Artificial intelligence-based traffic flow prediction: a comprehensive review. *Journal of Electrical Systems and Information Technology, 10*(13). https://doi.org/10.1186/s43067-023-00081-6.

Schneider, R., Arnold, L., & Ragland, D. (2009). Methodology for counting pedestrians at intersections: use of automated counters to extrapolate weekly

volumes from short manual counts. *Transportation Research Record: Journal of the Transportation Research Board, 2140*(1), 1-12.

Shah, I., Muhammad, I., Ali, S., Ahmed, S., Almazah, M.M.A., & Al-Rezami, A.Y. (2022). Forecasting Day-ahead traffic flow using functional time series approach. *Mathematics, 10*. 4279. https://doi.org/10.3390/math10224279.

Shen, C., Du, F., Zhang, L., & Song, H. (2019). Application of GM(1, N) model in groundwater mineralization in Cheng' an County. *IOP Conference Series: Earth and Environmental Science, 233*(4), 042010. doi: 10.1088/1755-1315/233/4/042010.

Shen, Y., He B., & Qin, P. (2016). Fractional-order grey prediction method for non-equidistant sequences. *Entropy*, *18*(6), 227.

Shen Z. (2022). Short-time traffic flow prediction based on seasonal gray Fourier model. *Research Square*. doi: 10.21203/rs.3.rs-1334110/v1.

Slavek, N., Krmpotić D., & Blažević, D. (2015). Grey system theory approach to quality of intranet. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, *4*(10), 8223-8230.

Smith, B. L., & Demetsky, M. J. (1997). Traffic flow forecasting: comparison of modeling Approaches. *Journal of transportation engineering, 123*(4), 261-266.

Steg, L. (2007). SUSTAINABLE TRANSPORTATION: A Psychological perspective. *IATSS Research*, *31*(2), 58–66. https://doi.org/10.1016/S0386-1112(14)60223-5.

Sun, Z. Y., Hu, Y. J., Li, W., Feng, S. W., & Pei, L. L. (2022). Prediction model for short-term traffic flow based on a K-means-gated recurrent unit combination. *IET Intell. Transp. Syst., 16*(5), 675-690.

Tien, T. L. (2009). A new grey prediction model FGM (1, 1). *Mathematical and Computer Modelling, 49*(7-8), 1416-1426.

Toan, T. D. (2018). Managing traffic congestion in a city: A study of Singapore's experiences. *International Conference on Sustainability in Civil Engineering*, Hanoi, Vietnam.

Toroman, A., & Mujcic, E. (2018). Application of industrial PLC for controlling intelligent traffic lights. In *Telecommunication Forum (TELFOR)*, Belgrade, Serbia, IEEE, January 8, 1-4. [Online]. doi: 10.1109/TELFOR.2017.8249411.

Tu, L., Chen, Y., & Wu, L. (2023). Unequal-order grey model with the difference information and its application. *International Journal of Modeling, Simulation, and Scientific Computing, 0,* 2350001. https://doi.org/10.1142/S1793962323500010.

Twumasi, E., Frimpong, E.A., Kwegyir, D., & Folitse D. (2021). Improvement of grey system model using particle swarm optimization. *Journal of Electrical Systems and Information Technology, 8*(12). https://doi.org/10.1186/s43067-021-00036-9.

Udoakah, Y. N., & Okure, I. G. (2017). Design and implementation of a density-based traffic light control with surveillance system. *Nigerian Journal of Technology, 36*(4), 1239-1248.

Wang, H., Ouyang, M., Meng, Q., & Kong, Q. (2020). A traffic data collection and analysis method based on wireless sensor network. *Journal on Wireless Communications and Networking, 2020*(1), 1-8. https://doi.org/10.1186/s13638-019-1628-5.

Wang Y. (2017). Traffic flow prediction based on combined model of ARIMA and RBF neural network. *Advances in Engineering Research, 138*, 82-86.

Wang, Y. W., Shen, Z. Z., & Jiang, Y. (2018). Comparison of ARIMA and GM (1, 1) models for prediction of hepatitis B in China. *PloS One*, *13*(9), e0201987.

Wang, Y., Wei, F., Sun, C., & Li, Q. (2016). The research of improved grey GM (1, 1) model to predict the postprandial glucose in type 2 diabetes. *BioMed research international, 2016*(2016), 6 pages.

Wasike, S. K. (2001). Road Infrastructure Policies in Kenya: Historical trends and current challenges. Working Paper, KIPPRA Working Paper No. 1, Nairobi, Kenya. [Online]. Available: https://www.researchgate.net /publication /265066196.

Wei, L., & Dang, YG. (2016). The optimized grey model GM(1,N) based on the development trends of driving variables and its application. *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 000235-000240, doi: 10.1109/SMC.2016.7844247.

Willmott, C.J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research, 30*(1), 79-82.

Wu, Z., Liu, M., & Wang, X. (2013). Prediction of electric power consumption based on the improved GM (1, 1). *Telkomnika Indones J Electr Eng, 11*(8), 4669-4675.

Xiao, L., Chen, X., & Wang, H. (2021). Calculation and realization of new method grey residual error correction model. PLoS ONE *16*(7), e0254154. https://doi.org/10.1371/journal.pone.0254154.

Xie, M. N., Liu, S. F., Yang, Y. J., & Yuan, C. Q. (2013). On novel grey forecasting model based on non-homogeneous index sequence. *Applied Mathematical Modelling, 37*(7), 5059-5068.

Xu, N., & Dang, Y. G. (2015). An optimized grey GM (2, 1) model and forecasting of highway subgrade settlement. *Mathematical Problems in Engineering, 2015*(1), 1-6.

Yao, B., Ma, A., Feng, R., Shen X., Zhang, M., & Yao, Y. (2022). A deep learning framework about traffic flow forecasting for urban traffic emission monitoring system. *Front. Public Health, 9*, 804298. doi: 10.3389/fpubh.2021.804298.

Yaping, R., Xingchen, Z., Xuesong, F., Tin-kin, H., Wei, W., & Dejie, X. (2015). Comparative analysis for traffic flow forecasting models with real-life data in Beijing. *Advances in Mechanical Engineering*, *7*(12). doi: 10.1177/1687814015620324.

Yong, L., & Xiuchun, G. U. (2013). Research of pedestrian traffic characteristics in University Campus. in *2013 Fourth International Conference on Digital Manufacturing and Automation (ICDMA)*, Qingdao, China, IEEE, 1022-1024. [Online]. doi: 10.1109/ICDMA.2013.240.

Young, C. C., & Liu, W. C. (2015). Prediction and modelling of rainfall–runoff during typhoon events using a physically-based and artificial neural network hybrid model. *Hydrological Sciences Journal, 60*(12), 2102-2116.

Zaki, J. F., Ali-Eldin, A., Hussein, S. E., Saraya, S. F., & Areed, F. F. (2020). Traffic congestion prediction based on Hidden Markov Models and contrast measure, *Ain Shams Engineering Journal, 11*(3), 535-551. https://doi.org/10.1016/j.asej.2019.10.006

Zaki, J. F., Ali-Eldin, A. M., Hussein, S. E., Saraya, S. F., & Areed, F. F. (2016). Framework for traffic congestion prediction. *International Journal of Scientific & Engineering Research, 7*(6), 1205-1210.

Zapata, J. A., Serna, M. D., & Gomez, R. A. (2013). Information systems applied to transport improvement. *Dyna, 80*(180), 77-86.

Zeng, B., Li, S., Meng, W., & Zhang, D. (2019). An improved gray prediction model for China's beef consumption forecasting. *PloS One*, *14*(9), e0221333.

Zeng, B., Luo, C., Liu, S., Bai, Y., & Li, C. (2016). Development of an optimization method for the GM(1,N) model. *Engineering Applications of Artificial Intelligence*, *55*, 353-362.

Zeng, B., & Li, C. (2018). Improved multi-variable grey forecasting model with a dynamic background-value coefficient and its application. *Computers & Industrial Engineering, 118*, 278-290. https://doi.org/10.1016/j.cie.2018.02.042.

Zeng, B., Zhou, M., Liu, X., & Zhang, Z. (2020). Application of a new grey prediction model and grey average weakening buffer operator to forecast China's shale gas output, *Energy Reports, 6*, 1608-1618. https://doi.org/10.1016/j.egyr.2020.05.021.

Zeng, Y., Luo, M., & Liu, Y. (2018). Future energy consumption prediction based on grey forecast model, *arXiv:1804.01777*.

Zhang, T., Wang, K., & Zhang, X. (2015). Modelling and analyzing the transmission dynamics of HBV epidemic in Xinjiang, China. *PloS one*, *10*(9), e0138765.

Zhang, Y. (2020). Short-term traffic flow prediction methods: A survey. *J. Phys.: Conf. Ser.* *2019 4th International Seminar on Computer Technology, Mechanical and Electrical Engineering (ISCME 2019),* 13-15 December 2019, *1486*, 052018, Chengdu, China. DOI**:** 10.1088/1742-6596/1486/5/052018.

Zhang, Z., Xu, X., & Wang, Z. (2017). Application of grey prediction model to short-time passenger flow forecast. In *AIP Conference Proceedings*, *1839*(1), AIP, May 8, 2017, 020136. Hangzhou, China.

Zheng, Y., Elefteriadou, L., Chase, T., Schroeder, B., & Sisiopiku, V. (2016). Pedestrian traffic operations in urban networks. *Transportation Research Procedia, 15*, 137-149.

Zhou, Q., Chen, N., & Lin, S. (2022). FASTNN: A deep learning approach for traffic flow prediction considering spatiotemporal features. *Sensors*, *22*(18), 6921. https://doi.org/10.3390/s22186921.

Zhuang W., & Cao Y. (2022). Short-term traffic flow prediction based on CNN-BILSTM with multicomponent information. *applied sciences*, *12*(17), 8714. https://doi.org/10.3390/app12178714.

Zou, R., & Wu, H. (2012). Grey GRM(1,1) Model based on reciprocal accumulated generating and its application. *Applied Mathematics, 3*(6), 554-556.

APPENDICES

Appendix I: Traffic Count Tally Sheet

Date		**Counting Site**		
Volunteer (Who will do the count)			**Contacts**	
Type of count				
Road				
Direction	**Northward**	**Southward**	**Eastward**	**Westward**
TIME (5 minute increment)				
06:00-06:05				
06:06-06:10				
06:11-06:15				
06:16-06:20				
06:21-06:25				
06:26-06:30				

Appendix II: Traffic Count Report Sheet

Counter	Counting site	Road	Direction	Date
Shift From:	**To:**	**Count type**		
Weather				
Incident				
Other comments				
Supervisor Name		**Signature:**		
		Date:		

Appendix III: Photos Showing Clerks Collecting Traffic Data from Nairobi CBD

Figure 1: At University Way Uhuru Highway Roundabout

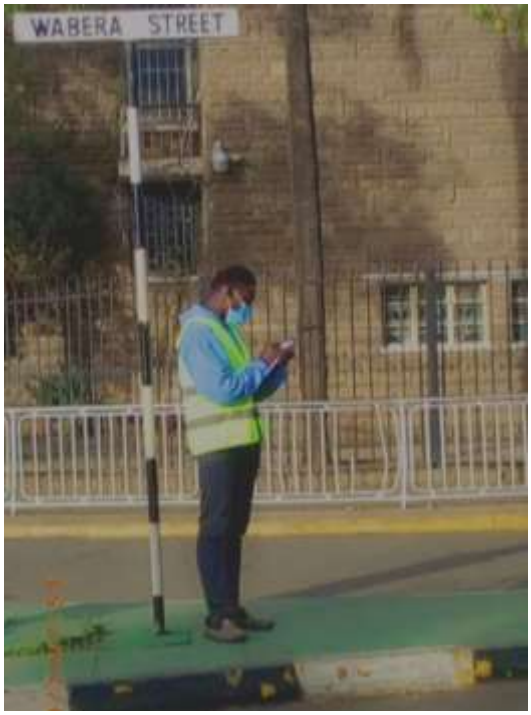

Figure 2: At City Hall Way-Wabera Street T-Roundabout

Figure 3: At Kenyatta Avenue-Moi Avenue Mondlane Street Junction

Figure 4: At City Hall Way-Wabera Street T-Roundabout

Figure 5: At Haile Selassie Roundabout

Figure 6: At Kenyatta Avenue Uhuru Highway Roundabout

Figure 7: At Haile Selassie Avenue-Moi Avenue Roundabout

Figure 8: At Moi Avenue-Slip Road Junction

Appendix IV: Nairobi CBD Traffic Data

Table 1: Day 1 site 1 (Haile Selassie Roundabout) Traffic Data

DAY 1 SITE 1 (HAILE SELASSIE ROUNDABOUT) TRAFFIC DATA												
DIRECTION	**NORTHWARD**			**SOUTHWARD**			**EASTWARD**			**WESTWARD**		
TIME	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**
06:00-06:05	60	5	6	53			43	15	8	63	5	7
06:06-06:10	55	4	9	47			50	9	7	67	7	8
06:11-06:15	70	8	11	60			61	12	4	56	14	8
06:16-06:20	77	7	14	58			58	7	9	71	16	9
06:21-06:25	82	5	12	65			65	8	4	80	20	10
06:26-06:30	81	14	9	80			70	5	7	77	20	7
06:31-06:35	97	13	10	86			56	3	8	89	23	24
06:36-06:40	137	10	14	89			68	11	7	83	26	14
06:41-06:45	319	20	9	86			69	6	3	76	43	14
06:46-06:50	223	26	4	100			72	8	1	129	21	15
06:51-06:55	198	18	11	144			44	12	4	98	14	13
06:56-07:00	238	15	10	161			80	4	3	129	38	14
07:01-07:05	251	16	10	141			70	6	2	124	14	11
07:06-07:10	247	21	8	103			66	3	5	120	26	13
07:11-07:15	222	21	10	83			83	5	11	112	14	12
07:16-07:20	252	17	10	85			81	9	5	102	10	15
07:21-07:25	244	16	27	49			52	4	2	98	17	12
07:26-07:30	226	18	16	132			68	4	11	101	19	13
07:31-07:35	189	15	22	178			98	10	6	89	24	12
07:36-07:40	312	21	14	2			82	6	3	93	10	10
07:41-07:45	134	18	27	229			85	4	5	90	14	12
07:46-07:50	176	24	22	0			92	3	4	88	26	14
07:51-07:55	151	18	14	212			98	7	3	93	32	10
07:56-08:00	154	26	22	42			79	8	5	103	14	12
08:01-08:05	135	33	22	178			82	10	7	98	20	13
08:06-08:10	160	26	19	0			75	4	13	102	12	11
08:11-08:15	192	16	29	213			74	3	7	74	14	12
08:16-08:20	159	16	33	78			77	3	7	69	22	11
08:21-08:25	215	26	33	151			80	2	16	63	34	11
08:26-08:30	112	8	44	91			91	2	10	65	34	12
08:31-08:35	195	21	27	32			93	4	12	0	36	0
08:36-08:40	253	17	28	212			48	4	15	57	14	10
08:41-08:45	201	16	26	0			47	3	6	70	42	10
08:46-08:50	94	18	27	95			115	5	10	56	14	9
08:51-08:55	291	13	32	0			55	3	10	69	26	18
08:56-09:00	154	12	18	111			99	2	7	19	36	16
09:01-09:05	283	5	22	0			69	4	14	63	12	14
09:06-09:10	112	8	35	287			98	13	15	65	16	10
09:11-09:15	354	10	26	8			30	4	16	74	18	0
09:16-09:20	132	18	28	96			105	7	11	63	9	0
09:21-09:25	114	11	39	103			58	8	13	19	0	8
09:26-09:30	322	16	31	5			42	9	19	34	14	6
09:31-09:35	81	1	24	250			62	5	11	12	8	7
09:36-09:40	297	7	33	20			84	5	16	89	14	3
09:41-09:45	103	13	28	86			86	1	13	54	0	0
09:46-09:50	184	5	28	141			67	3	12	0	4	8
09:51-09:55	206	11	25	30			76	1	9	57	14	3
09:56-10:00	241	12	26	169			65	2	15	12	12	0
10:01-10:05	115	10	23	169			94	8	14	23	18	4

DAY 1 SITE 1 (HAILE SELASSIE ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
10:06-10:10	260	10	32	0			75	5	13	46	34	12
10:11-10:15	79	14	27	194			66	7	22	51	24	4
10:16-10:20	242	8	41	75			69	2	18	12	14	2
10:21-10:25	237	10	45	7			61	6	10	76	12	3
10:26-10:30	192	2	53	183			79	8	9	39	26	8

Table 2: Day 1 site 2 (Kenyatta Avenue Uhuru Highway Roundabout) Traffic Data

DAY 1 SITE 2 (KENYATTA AVENUE UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	100	23	1	148	23	3	90	2	6	48	10	6
06:06-06:10	120	10	5	163	31	0	84	2	6	41	4	0
06:11-06:15	150	6	11	187	48	4	71	4	7	39	6	6
06:16-06:20	110	11	6	177	35	4	76	2	3	47	5	2
06:21-06:25	112	15	7	156	57	4	84	1	1	34	6	5
06:26-06:30	115	17	5	193	52	3	97	2	4	64	21	1
06:31-06:35	80	13	4	182	53	6	115	7	0	64	5	3
06:36-06:40	121	13	9	191	71	4	120	7	0	55	12	2
06:41-06:45	110	12	5	209	76	5	120	8	3	79	7	2
06:46-06:50	115	14	6	223	56	7	127	6	1	52	10	2
06:51-06:55	120	15	7	248	42	3	119	5	2	98	3	2
06:56-07:00	150	15	5	251	81	6	136	2	0	72	12	6
07:01-07:05	140	13	6	123	96	8	112	3	0	86	13	4
07:06-07:10	141	14	6	193	83	4	118	6	4	87	8	2
07:11-07:15	120	16	12	88	76	7	124	10	3	94	8	16
07:16-07:20	123	15	7	75	57	5	119	5	9	86	8	10
07:21-07:25	120	18	9	149	53	4	127	6	6	75	11	11
07:26-07:30	130	13	13	0	79	8	0	10	11	69	7	5
07:31-07:35	133	12	13	188	82	6	129	14	7	99	15	8
07:36-07:40	200	23	7	122	54	7	0	6	9	57	12	13
07:41-07:45	180	35	12	97	73	17	137	8	5	95	10	5
07:46-07:50	185	48	13	105	78	19	53	13	15	96	8	11
07:51-07:55	173	47	23	187	83	14	58	10	12	80	4	7
07:56-08:00	148	21	10	222	40	17	71	0	10	72	6	9
08:01-08:05	140	74	13	157	73	31	64	2	5	132	2	7
08:06-08:10	143	61	28	239	39	17	78	8	6	38	10	10
08:11-08:15	169	55	17	35	52	29	112	2	8	110	6	6
08:16-08:20	160	53	19	234	62	21	94	5	5	53	4	17
08:21-08:25	166	64	23	85	51	18	24	2	16	123	5	10
08:26-08:30	87	47	18	214	52	16	88	5	18	86	4	19
08:31-08:35	118	79	16	202	58	25	118	4	6	44	11	21
08:36-08:40	123	69	22	193	43	28	67	4	9	91	7	25
08:41-08:45	0	34	23	222	49	27	84	5	8	97	4	15
08:46-08:50	101	28	14	188	22	23	18	4	9	37	1	4
08:51-08:55	58	31	21	133	37	18	58	2	10	70	9	16
08:56-09:00	132	53	41	158	33	12	64	6	6	36	5	11
09:01-09:05	115	54	27	193	57	41	78	4	8	87	10	22
09:06-09:10	0	41	24	173	26	17	124	5	9	56	6	8
09:11-09:15	111	33	16	223	36	11	113	9	8	36	4	15
09:16-09:20	146	31	27	175	48	14	72	6	7	136	8	11
09:21-09:25	0	39	25	165	21	17	64	7	8	136	5	10
09:26-09:30	139	35	18	216	19	23	127	2	2	41	5	7

DAY 1 SITE 2 (KENYATTA AVENUE UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
09:31-09:35	126	18	7	123	14	32	30	2	10	87	2	20
09:36-09:40	69	32	17	237	21	10	64	2	7	40	7	12
09:41-09:45	115	23	26	191	26	16	73	4	6	144	3	14
09:46-09:50	100	13	30	63	17	6	44	4	10	85	6	11
09:51-09:55	41	28	8	258	46	28	87	2	7	62	5	18
09:56-10:00	73	34	14	209	24	16	88	3	10	116	7	9
10:01-10:05	186	41	23	137	29	41	94	7	12	66	3	25
10:06-10:10	1	13	18	273	11	40	67	3	7	91	4	22
10:11-10:15	103	22	12	293	14	53	82	5	3	65	4	9
10:16-10:20	196	36	23	213	30	21	93	4	5	44	10	15
10:21-10:25	100	33	19	176	22	34	84	4	7	74	3	13
10:26-10:30	21	46	24	143	17	27	78	7	11	111	8	25

Table 3: Day 1 site 3 (University Way Uhuru Highway Roundabout) Traffic Data

DAY 1 SITE 3 (UNIVERSITY WAY UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	102	5	4	0	0	0	47	21	16	0	0	0
06:06-06:10	107	6	2	0	0	0	54	18	14	0	0	0
06:11-06:15	108	10	8	0	0	0	41	34	18	0	0	0
06:16-06:20	98	2	4	0	0	0	53	46	15	0	17	13
06:21-06:25	68	1	10	141	9	3	46	54	8	64	18	14
06:26-06:30	38	9	6	148	5	3	41	31	7	80	11	5
06:31-06:35	63	8	7	140	4	5	70	62	5	56	18	13
06:36-06:40	62	4	4	130	3	6	65	74	16	92	15	11
06:41-06:45	110	6	15	220	4	9	76	39	17	90	14	12
06:46-06:50	61	1	21	156	3	8	84	51	12	105	16	10
06:51-06:55	96	17	18	230	5	7	78	18	8	107	14	7
06:56-07:00	49	9	30	205	4	7	89	29	9	96	13	9
07:01-07:05	47	10	8	183	3	6	51	19	11	93	10	8
07:06-07:10	81	15	9	123	5	3	74	43	14	99	12	9
07:11-07:15	71	2	15	163	5	3	85	14	12	97	11	14
07:16-07:20	32	14	4	174	3	2	70	21	7	87	7	15
07:21-07:25	46	1	6	195	8	7	76	17	8	91	9	11
07:26-07:30	55	1	4	179	29	18	90	39	7	100	8	11
07:31-07:35	23	17	26	17	38	12	85	41	9	110	11	9
07:36-07:40	48	10	22	213	11	21	116	53	4	120	8	12
07:41-07:45	52	20	16	194	20	6	107	46	7	106	7	10
07:46-07:50	63	15	16	176	32	20	130	61	8	158	12	17
07:51-07:55	12	3	26	235	43	24	107	74	2	112	9	12
07:56-08:00	31	5	19	185	33	21	78	42	11	97	6	11
08:01-08:05	82	11	34	198	29	15	113	37	5	71	4	8
08:06-08:10	0	12	29	202	17	25	101	89	14	54	7	13
08:11-08:15	75	19	14	185	21	27	93	54	2	33	5	9
08:16-08:20	55	7	23	168	19	10	109	46	4	66	8	10
08:21-08:25	58	6	21	176	9	23	124	25	7	41	9	12
08:26-08:30	110	8	14	163	18	21	109	18	3	38	6	9
08:31-08:35	42	21	22	158	18	23	135	19	9	36	11	15
08:36-08:40	70	13	11	193	13	24	116	14	7	55	9	13
08:41-08:45	65	12	4	147	28	25	123	27	6	47	8	10
08:46-08:50	84	11	28	177	12	25	146	9	3	78	7	11
08:51-08:55	67	21	23	164	15	30	76	11	4	81	5	14

DAY 1 SITE 3 (UNIVERSITY WAY UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
08:56-09:00	60	7	9	173	12	30	85	7	5	36	8	16
09:01-09:05	46	9	21	161	11	34	91	19	7	79	10	12
09:06-09:10	81	11	20	127	28	32	103	26	14	67	6	9
09:11-09:15	93	11	11	183	17	18	92	17	9	60	9	13
09:16-09:20	48	4	9	173	21	20	96	13	5	76	7	10
09:21-09:25	64	6	2	93	17	20	107	9	3	55	12	23
09:26-09:30	54	21	21	188	18	32	118	19	7	60	8	25
09:31-09:35	66	5	11	179	15	22	88	34	6	57	14	19
09:36-09:40	86	7	6	87	21	14	93	28	7	61	11	21
09:41-09:45	93	10	14	105	17	20	81	39	12	58	6	15
09:46-09:50	57	15	24	97	39	21	76	47	5	40	9	17
09:51-09:55	97	7	22	178	15	27	101	41	11	63	12	20
09:56-10:00	71	14	32	107	14	23	74	22	4	53	8	14
10:01-10:05	67	7	31	174	21	28	90	24	3	49	11	19
10:06-10:10	86	20	6	107	21	37	71	19	5	67	12	27
10:11-10:15	27	13	15	115	34	35	50	14	6	70	9	31
10:16-10:20	60	7	11	125	28	39	63	13	5	41	15	33
10:21-10:25	91	7	20	142	29	23	72	9	14	47	12	38
10:26-10:30	75	7	21	166	21	17	56	22	3	59	9	25

Table 4: Day 1 site 4 (Kenyatta Avenue-Moi Avenue-Mondlane Street Junction) Traffic Data

DAY 1 SITE 4 (KENYATTA AVENUE-MOI AVENUE-MONDLANE STREET JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	26	52	1	19	12	0	2	19	11
06:06-06:10	38	49	1	17	20	2	7	12	6
06:11-06:15	46	76	3	20	29	1	3	44	3
06:16-06:20	55	88	3	21	32	3	5	52	2
06:21-06:25	72	77	5	28	20	0	4	56	7
06:26-06:30	86	73	3	29	22	0	10	63	9
06:31-06:35	82	72	3	34	41	2	8	74	5
06:36-06:40	80	79	1	21	45	1	6	69	6
06:41-06:45	56	88	1	27	32	1	15	77	3
06:46-06:50	91	112	2	25	50	3	19	84	7
06:51-06:55	78	123	5	22	63	4	28	82	4
06:56-07:00	80	97	4	24	50	2	10	70	7
07:01-07:05	74	74	2	38	46	5	16	102	5
07:06-07:10	93	72	3	26	78	8	16	81	13
07:11-07:15	85	82	6	39	60	16	19	59	9
07:16-07:20	76	99	7	40	61	5	9	75	9
07:21-07:25	71	87	8	32	65	5	16	89	4
07:26-07:30	95	123	6	35	45	10	19	88	2
07:31-07:35	88	100	9	18	48	5	19	55	6
07:36-07:40	92	113	7	36	61	7	15	80	6
07:41-07:45	84	101	7	24	64	14	13	59	3
07:46-07:50	122	90	8	30	75	5	14	118	2
07:51-07:55	88	92	10	39	77	13	19	98	5
07:56-08:00	95	91	12	42	60	18	20	109	7
08:01-08:05	86	82	11	35	76	15	22	93	12
08:06-08:10	90	117	13	47	73	23	21	99	8

DAY 1 SITE 4 (KENYATTA AVENUE-MOI AVENUE-MONDLANE STREET JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
08:11-08:15	98	121	12	29	87	15	13	97	7
08:16-08:20	102	87	9	46	58	23	27	49	5
08:21-08:25	86	100	12	29	73	16	21	87	3
08:26-08:30	88	97	13	49	67	13	23	92	5
08:31-08:35	93	93	16	28	92	21	20	96	8
08:36-08:40	72	86	21	36	63	17	20	108	8
08:41-08:45	76	68	22	31	57	29	18	130	15
08:46-08:50	82	92	13	40	73	10	19	127	12
08:51-08:55	79	78	14	21	42	25	30	157	16
08:56-09:00	84	69	9	31	92	22	22	117	8
09:01-09:05	85	93	16	26	73	24	30	104	13
09:06-09:10	78	79	11	24	69	16	40	97	13
09:11-09:15	73	90	15	26	53	10	20	83	12
09:16-09:20	65	74	14	22	83	21	45	91	12
09:21-09:25	71	59	12	26	77	11	21	97	12
09:26-09:30	68	61	20	38	69	9	27	100	14
09:31-09:35	74	42	11	22	94	21	20	92	22
09:36-09:40	67	66	25	25	72	18	35	104	16
09:41-09:45	72	46	11	11	89	23	20	89	20
09:46-09:50	69	73	9	26	96	22	34	124	24
09:51-09:55	72	82	14	25	68	10	25	124	16
09:56-10:00	70	76	13	27	96	16	25	129	24
10:01-10:05	84	67	11	28	59	18	31	126	6
10:06-10:10	76	72	12	31	79	26	27	97	5
10:11-10:15	78	73	26	28	92	20	37	113	12
10:16-10:20	81	71	5	17	94	30	42	98	20
10:21-10:25	61	68	14	11	93	27	40	107	16
10:26-10:30	72	65	19	20	68	26	38	140	16

Table 5: Day 1 site 5 (Moi Avenue-Slip Road Junction) Traffic Data

DAY 1 SITE 5 (MOI AVENUE-SLIP ROAD JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	48	3	1	15	21	5	10	3	0
06:06-06:10	63	5	3	40	27	2	13	5	0
06:11-06:15	55	6	4	25	33	6	10	7	3
06:16-06:20	51	10	2	28	40	3	16	4	0
06:21-06:25	49	12	3	35	38	7	17	11	4
06:26-06:30	56	4	7	38	27	4	20	3	0
06:31-06:35	48	17	3	70	48	3	18	5	0
06:36-06:40	54	21	2	65	39	2	28	10	0
06:41-06:45	50	30	4	87	41	4	27	4	1
06:46-06:50	32	24	3	88	37	6	31	20	3
06:51-06:55	63	25	4	81	38	12	44	27	1
06:56-07:00	41	12	4	91	51	8	46	30	2
07:01-07:05	45	17	2	100	30	4	32	27	0
07:06-07:10	47	21	4	119	47	13	36	10	0
07:11-07:15	29	6	1	97	43	6	43	47	7
07:16-07:20	36	15	3	91	39	5	60	12	1
07:21-07:25	64	7	1	98	39	11	49	25	5
07:26-07:30	69	11	2	101	36	3	35	12	0
07:31-07:35	38	16	2	74	34	19	41	21	1

DAY 1 SITE 5 (MOI AVENUE-SLIP ROAD JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
07:36-07:40	43	9	3	103	51	19	63	15	1
07:41-07:45	55	14	8	92	39	20	82	15	0
07:46-07:50	80	15	4	106	50	21	79	25	2
07:51-07:55	70	20	12	93	42	25	102	10	1
07:56-08:00	84	21	15	104	33	20	97	18	3
08:01-08:05	60	19	16	89	24	21	98	24	3
08:06-08:10	48	26	8	80	26	13	93	17	4
08:11-08:15	70	9	3	100	31	22	88	20	2
08:16-08:20	59	21	16	84	33	20	119	12	5
08:21-08:25	61	13	15	70	26	24	124	8	2
08:26-08:30	77	9	12	101	27	15	106	11	10
08:31-08:35	53	10	8	98	33	18	116	21	0
08:36-08:40	46	6	5	115	18	22	103	16	14
08:41-08:45	75	12	22	98	10	13	97	25	0
08:46-08:50	49	16	18	98	22	19	85	4	11
08:51-08:55	53	9	14	112	23	11	123	5	5
08:56-09:00	43	11	19	98	20	16	103	19	8
09:01-09:05	76	12	20	84	18	17	81	10	6
09:06-09:10	45	7	18	136	20	15	123	7	10
09:11-09:15	71	8	11	111	25	16	90	2	3
09:16-09:20	58	9	17	62	19	14	107	16	4
09:21-09:25	73	22	16	88	30	12	83	11	0
09:26-09:30	52	7	11	78	15	16	101	20	4
09:31-09:35	44	9	13	72	13	15	105	16	0
09:36-09:40	71	14	17	84	18	22	84	23	0
09:41-09:45	59	7	12	40	31	20	107	4	2
09:46-09:50	72	1	19	103	28	29	109	7	0
09:51-09:55	64	6	13	98	27	30	80	12	4
09:56-10:00	56	7	17	84	29	23	127	17	0
10:01-10:05	65	10	16	129	18	13	78	12	1
10:06-10:10	59	9	13	87	23	20	96	21	13
10:11-10:15	92	6	10	80	24	17	110	18	2
10:16-10:20	69	14	19	89	15	13	98	11	0
10:21-10:25	66	25	17	85	8	32	123	4	3
10:26-10:30	64	16	23	89	32	33	87	15	0

Table 6: Day 1 site 6 (City Hall Way-Wabera Street T-Roundabout) Traffic Data

DAY 1 SITE 6 (CITY HALL WAY-WABERA STREET T-ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	6	17	2	7	2	0	12	9	1
06:06-06:10	10	16	0	17	3	1	15	13	1
06:11-06:15	8	12	1	25	1	0	24	17	11
06:16-06:20	19	7	2	26	2	0	18	13	2
06:21-06:25	13	4	1	20	2	0	21	17	0
06:26-06:30	9	10	1	32	3	0	19	13	0
06:31-06:35	14	5	3	26	4	0	23	31	11
06:36-06:40	12	8	0	31	2	0	39	28	1
06:41-06:45	19	4	2	31	3	1	42	31	0
06:46-06:50	25	6	0	24	5	2	35	22	3
06:51-06:55	19	9	0	35	6	0	39	26	3
06:56-07:00	15	15	1	33	8	1	46	27	3

DAY 1 SITE 6 (CITY HALL WAY-WABERA STREET T-ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
07:01-07:05	22	27	3	39	4	0	40	30	0
07:06-07:10	17	27	3	54	3	2	34	23	4
07:11-07:15	27	22	3	43	5	1	31	35	4
07:16-07:20	30	21	0	34	9	3	33	37	1
07:21-07:25	22	17	1	29	6	1	45	46	3
07:26-07:30	16	17	3	44	10	4	36	36	5
07:31-07:35	16	18	3	47	9	2	32	30	10
07:36-07:40	25	21	0	41	7	3	35	30	0
07:41-07:45	26	24	5	40	3	1	46	27	4
07:46-07:50	30	35	3	45	7	1	36	59	9
07:51-07:55	41	76	3	52	12	2	40	21	3
07:56-08:00	28	78	2	61	10	1	45	31	1
08:01-08:05	35	73	7	35	12	2	41	40	4
08:06-08:10	34	39	5	34	15	2	39	31	8
08:11-08:15	38	49	7	50	11	3	51	20	3
08:16-08:20	35	36	2	53	12	4	30	40	5
08:21-08:25	37	37	4	59	13	4	60	27	14
08:26-08:30	48	44	6	35	8	1	36	30	11
08:31-08:35	40	37	11	42	12	4	52	20	15
08:36-08:40	35	52	6	31	12	3	50	26	2
08:41-08:45	45	28	7	51	18	6	43	30	3
08:46-08:50	38	53	4	53	15	4	36	32	12
08:51-08:55	30	32	6	34	14	3	21	24	7
08:56-09:00	43	19	6	66	17	1	29	22	8
09:01-09:05	35	20	11	31	19	4	28	27	9
09:06-09:10	33	34	6	43	14	2	36	21	6
09:11-09:15	43	29	5	32	8	2	31	26	10
09:16-09:20	44	24	4	45	10	3	32	21	10
09:21-09:25	27	27	10	48	16	1	43	27	11
09:26-09:30	38	34	8	38	20	6	38	40	11
09:31-09:35	44	27	8	45	1	3	38	42	5
09:36-09:40	50	37	9	34	22	6	37	62	10
09:41-09:45	44	43	9	49	12	2	35	64	9
09:46-09:50	43	38	10	38	19	2	36	49	7
09:51-09:55	42	44	12	55	18	5	31	56	15
09:56-10:00	55	21	11	41	13	2	42	36	7
10:01-10:05	44	38	12	51	17	4	43	62	8
10:06-10:10	37	45	9	54	15	3	30	51	13
10:11-10:15	34	31	10	48	22	8	44	42	13
10:16-10:20	37	26	10	35	5	2	26	59	9
10:21-10:25	29	45	8	38	8	7	27	75	18
10:26-10:30	40	38	11	45	17	4	40	55	16

Table 7: Day 1 site 7 (Haile Selassie Avenue-Moi Avenue Roundabout) Traffic Data

DAY 1 SITE 7 (HAILE SELASSIE AVENUE-MOI AVENUE ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	31	45	0	30	102	2	57	266	6
06:06-06:10	42	60	4	70	108	3	101	220	3
06:11-06:15	25	31	0	80	103	3	88	298	2

DAY 1 SITE 7 (HAILE SELASSIE AVENUE-MOI AVENUE ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:16-06:20	55	83	0	95	124	4	75	234	1
06:21-06:25	72	49	7	107	166	5	96	276	7
06:26-06:30	73	60	0	78	158	6	94	235	8
06:31-06:35	65	81	0	77	195	4	101	215	12
06:36-06:40	93	110	6	87	152	6	83	229	6
06:41-06:45	84	94	0	86	205	9	89	329	3
06:46-06:50	80	120	4	71	217	3	103	297	3
06:51-06:55	84	105	13	50	265	3	118	316	2
06:56-07:00	72	135	3	86	239	2	107	262	4
07:01-07:05	82	92	1	87	272	3	96	314	6
07:06-07:10	93	84	2	50	182	3	103	364	7
07:11-07:15	70	43	6	52	245	4	108	246	6
07:16-07:20	90	61	0	60	400	8	119	370	3
07:21-07:25	87	74	7	40	214	4	116	326	2
07:26-07:30	80	68	4	56	182	5	99	285	5
07:31-07:35	76	81	10	50	174	7	119	314	4
07:36-07:40	89	74	5	63	406	6	48	980	3
07:41-07:45	63	104	3	72	493	3	103	1028	7
07:46-07:50	70	94	3	69	272	7	113	580	1
07:51-07:55	65	74	6	42	524	5	44	1530	2
07:56-08:00	74	86	5	80	402	7	148	620	4
08:01-08:05	63	107	11	40	334	4	64	520	5
08:06-08:10	76	70	8	50	208	5	108	410	6
08:11-08:15	80	59	9	57	192	7	106	325	7
08:16-08:20	52	80	14	40	184	8	156	296	8
08:21-08:25	80	95	11	45	172	6	104	267	11
08:26-08:30	74	75	2	37	153	6	111	282	3
08:31-08:35	60	115	13	50	182	11	124	253	5
08:36-08:40	54	40	1	44	184	7	112	378	3
08:41-08:45	70	84	0	40	137	8	129	247	7
08:46-08:50	72	63	13	50	125	5	117	262	8
08:51-08:55	69	61	15	45	128	9	106	233	9
08:56-09:00	72	100	17	36	121	8	104	278	9
09:01-09:05	65	47	8	25	118	13	91	235	6
09:06-09:10	64	51	3	30	105	14	99	187	3
09:11-09:15	82	63	10	32	144	10	108	238	2
09:16-09:20	60	49	3	30	137	12	77	286	1
09:21-09:25	82	35	0	60	134	8	82	223	0
09:26-09:30	80	59	15	0	138	9	98	173	5
09:31-09:35	76	54	9	90	126	5	126	226	4
09:36-09:40	85	62	14	43	98	7	68	172	3
09:41-09:45	40	48	13	0	112	5	94	207	2
09:46-09:50	30	32	6	0	88	2	121	272	1
09:51-09:55	116	24	9	0	78	4	105	155	3
09:56-10:00	0	44	16	175	114	7	87	162	10
10:01-10:05	0	27	3	50	102	6	72	174	8
10:06-10:10	155	42	13	0	106	7	70	179	6
10:11-10:15	100	55	14	62	97	6	75	142	5
10:16-10:20	95	35	12	0	89	4	98	197	4
10:21-10:25	30	20	18	162	78	3	117	178	3
10:26-10:30	80	49	17	70	79	10	93	223	2

Table 8: Day 2 Site 1 (Haile Selassie Roundabout) Traffic Data

DAY 2 SITE 1 (HAILE SELASSIE ROUNDABOUT) TRAFFIC DATA												
DIRECTION	**NORTHWARD**			**SOUTHWARD**			**EASTWARD**			**WESTWARD**		
TIME	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**
06:00-06:05	70	6	1	43			29	8	0	8	7	1
06:06-06:10	83	4	1	38			35	10	1	10	12	0
06:11-06:15	95	7	3	69			38	6	3	18	8	3
06:16-06:20	101	5	2	118			31	7	2	49	13	1
06:21-06:25	116	7	3	98			38	6	2	69	16	2
06:26-06:30	178	8	5	120			56	12	4	108	18	3
06:31-06:35	192	10	7	104			47	11	3	114	26	4
06:36-06:40	228	18	8	113			56	8	2	118	23	3
06:41-06:45	213	16	12	131			63	4	5	132	16	2
06:46-06:50	262	17	4	116			62	9	7	148	15	6
06:51-06:55	249	18	14	123			52	8	1	100	24	8
06:56-07:00	265	16	11	53			42	6	3	89	12	4
07:01-07:05	233	23	12	176			74	11	2	69	18	8
07:06-07:10	215	22	17	215			70	8	3	74	21	5
07:11-07:15	291	21	18	80			79	9	12	29	30	7
07:16-07:20	234	21	7	151			75	18	7	24	12	2
07:21-07:25	205	31	17	80			87	9	7	46	21	4
07:26-07:30	193	25	28	108			68	3	6	58	38	7
07:31-07:35	279	25	16	49			80	19	13	116	23	4
07:36-07:40	195	25	26	195			67	13	5	104	26	8
07:41-07:45	150	17	34	220			70	2	13	108	30	3
07:46-07:50	167	37	25	69			84	9	7	113	7	2
07:51-07:55	214	24	26	197			71	5	11	86	19	1
07:56-08:00	258	21	21	2			99	9	14	74	27	0
08:01-08:05	87	42	19	144			90	2	18	44	20	7
08:06-08:10	328	32	33	0			89	4	17	84	12	8
08:11-08:15	118	9	37	205			103	5	8	72	28	2
08:16-08:20	255	12	25	78			106	4	10	109	13	8
08:21-08:25	179	26	31	138			51	2	13	141	7	0
08:26-08:30	182	11	31	73			123	13	12	100	6	4
08:31-08:35	253	10	24	87			89	4	12	161	12	5
08:36-08:40	297	15	25	135			76	5	19	101	15	2
08:41-08:45	190	8	32	50			104	11	14	94	17	0
08:46-08:50	144	6	35	175			85	5	20	84	21	1
08:51-08:55	200	17	46	14			70	5	13	46	2	3
08:56-09:00	202	25	35	176			68	3	9	84	4	0
09:01-09:05	229	7	30	77			78	1	13	96	12	1
09:06-09:10	208	12	37	87			70	4	19	44	3	2
09:11-09:15	278	6	32	144			0	6	13	48	8	4
09:16-09:20	223	18	14	0			151	5	7	40	10	1
09:21-09:25	66	9	37	265			48	1	9	68	7	0
09:26-09:30	234	16	32	0			67	8	24	74	0	2
09:31-09:35	285	9	28	195			60	6	18	66	5	3
09:36-09:40	210	11	36	67			81	10	14	107	10	0
09:41-09:45	164	6	34	95			115	4	15	48	15	2
09:46-09:50	179	9	31	129			68	9	13	148	15	3
09:51-09:55	194	12	32	160			78	3	21	109	8	0
09:56-10:00	201	18	32	90			71	7	15	100	17	4
10:01-10:05	165	10	34	179			88	5	13	108	16	1
10:06-10:10	297	11	47	65			65	7	17	142	17	5
10:11-10:15	168	6	31	148			70	6	9	181	13	4
10:16-10:20	201	14	38	106			86	9	21	109	23	5

DAY 2 SITE 1 (HAILE SELASSIE ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
10:21-10:25	149	14	43	128			79	5	18	112	12	2
10:26-10:30	223	13	25	67			80	4	8	180	15	0

Table 9: Day 2 Site 2 (Kenyatta Avenue Uhuru Highway Roundabout) Traffic Data

DAY 2 SITE 2 (KENYATTA AVENUE UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	96	11	4	87	12	5	75	3	3	37	6	3
06:06-06:10	110	17	3	135	38	2	67	1	1	47	15	3
06:11-06:15	118	12	2	179	23	4	79	2	1	49	19	4
06:16-06:20	100	18	8	226	43	3	87	6	0	28	7	1
06:21-06:25	106	23	7	155	56	6	96	5	0	64	1	4
06:26-06:30	141	30	6	173	33	2	89	8	2	58	14	0
06:31-06:35	89	48	4	197	47	3	97	10	4	61	4	4
06:36-06:40	91	37	5	246	53	7	70	12	0	68	7	2
06:41-06:45	120	22	6	212	42	1	75	2	0	40	9	3
06:46-06:50	101	46	3	30	63	3	122	7	1	77	17	0
06:51-06:55	109	35	7	248	37	5	112	3	1	94	5	6
06:56-07:00	36	37	8	256	58	8	99	7	2	86	4	4
07:01-07:05	168	42	10	215	49	3	60	4	2	63	16	6
07:06-07:10	159	77	10	61	64	8	136	3	3	109	2	4
07:11-07:15	0	71	5	309	58	4	128	5	4	48	7	3
07:16-07:20	211	68	8	120	37	12	112	7	6	104	10	2
07:21-07:25	0	63	10	27	53	9	134	1	1	56	6	9
07:26-07:30	138	101	18	243	68	15	101	1	2	60	5	5
07:31-07:35	159	66	9	290	73	9	97	2	3	0	1	5
07:36-07:40	0	79	14	53	54	14	100	2	1	164	6	4
07:41-07:45	201	53	11	297	65	8	111	1	7	53	2	9
07:46-07:50	193	74	17	232	49	16	97	6	4	20	3	9
07:51-07:55	0	50	10	182	88	13	118	3	7	149	7	16
07:56-08:00	109	83	15	259	38	23	112	9	1	40	6	18
08:01-08:05	146	73	16	8	57	18	104	6	3	121	7	8
08:06-08:10	0	88	18	230	39	28	113	8	5	60	6	12
08:11-08:15	141	46	8	166	43	17	115	4	8	80	7	13
08:16-08:20	0	62	15	74	32	19	110	5	7	128	9	20
08:21-08:25	140	71	22	53	23	12	112	2	5	56	2	21
08:26-08:30	0	61	13	206	16	22	113	5	7	129	6	15
08:31-08:35	151	48	13	254	29	22	132	5	8	55	13	24
08:36-08:40	113	49	24	43	36	14	131	5	8	99	2	16
08:41-08:45	23	61	20	205	22	19	114	16	2	82	7	22
08:46-08:50	280	66	11	142	45	11	116	5	12	8	8	24
08:51-08:55	0	46	26	217	26	19	122	7	4	131	9	17
08:56-09:00	172	44	27	224	16	8	124	2	6	93	4	14
09:01-09:05	61	56	23	214	26	18	136	9	4	115	4	22
09:06-09:10	0	41	18	173	36	9	127	3	1	78	3	20
09:11-09:15	212	69	30	194	13	16	110	9	6	24	10	17
09:16-09:20	210	73	23	53	36	14	102	5	3	77	3	27
09:21-09:25	17	59	26	198	18	12	136	4	0	111	4	27
09:26-09:30	85	41	13	136	27	6	132	8	4	34	6	23
09:31-09:35	96	48	11	142	39	22	126	9	13	40	6	17
09:36-09:40	105	43	8	135	58	36	119	10	12	96	7	8
09:41-09:45	127	63	19	162	23	9	99	2	9	66	9	17

DAY 2 SITE 2 (KENYATTA AVENUE UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	**NORTHWARD**			**SOUTHWARD**			**EASTWARD**			**WESTWARD**		
TIME	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**
09:46-09:50	221	49	24	44	16	21	109	2	16	74	6	25
09:51-09:55	107	53	18	84	24	37	92	2	18	120	9	21
09:56-10:00	247	41	37	186	28	32	110	1	20	98	2	30
10:01-10:05	0	46	27	183	32	14	119	8	9	96	8	29
10:06-10:10	95	37	21	78	16	19	82	5	16	74	8	22
10:11-10:15	186	56	18	115	23	10	93	3	8	57	8	23
10:16-10:20	0	63	14	91	13	23	110	9	6	54	16	29
10:21-10:25	254	30	30	174	16	14	105	6	9	16	0	14
10:26-10:30	0	75	26	121	12	17	94	5	9	154	11	19

Table 10: Day 2 Site 3 (University Way Uhuru Highway Roundabout) Traffic Data

DAY 2 SITE 3 (UNIVERSITY WAY UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	**NORTHWARD**			**SOUTHWARD**			**EASTWARD**			**WESTWARD**		
TIME	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**
06:00-06:05	98	8	4	0	0	0	42	18	7	63	2	1
06:06-06:10	104	4	1	0	0	0	56	12	9	98	4	3
06:11-06:15	119	0	2	0	0	0	48	21	11	81	3	7
06:16-06:20	102	11	3	0	2	1	64	34	8	92	9	11
06:21-06:25	146	15	10	0	1	4	37	27	6	75	12	3
06:26-06:30	131	4	8	187	5	2	29	30	10	93	7	16
06:31-06:35	122	2	11	153	3	3	74	41	11	102	11	8
06:36-06:40	78	1	4	167	8	4	81	31	5	81	8	4
06:41-06:45	113	7	2	156	9	5	35	42	9	72	14	6
06:46-06:50	87	8	1	145	20	2	64	22	4	95	11	13
06:51-06:55	100	17	5	143	27	10	93	36	12	131	9	4
06:56-07:00	114	4	5	157	30	7	84	47	7	0	11	14
07:01-07:05	76	4	7	122	29	9	121	28	9	53	11	13
07:06-07:10	109	10	5	133	23	12	0	39	8	98	8	10
07:11-07:15	89	9	8	143	13	7	97	23	16	68	10	14
07:16-07:20	106	8	4	173	17	20	0	37	12	83	7	12
07:21-07:25	115	14	1	129	12	15	67	24	9	72	9	15
07:26-07:30	119	7	14	133	6	6	69	19	13	65	8	11
07:31-07:35	96	15	8	141	18	11	74	36	8	104	6	9
07:36-07:40	110	8	9	176	16	28	78	18	3	78	8	13
07:41-07:45	107	3	7	143	27	16	86	25	7	81	9	12
07:46-07:50	134	13	8	159	23	19	12	13	2	87	6	15
07:51-07:55	44	15	7	163	15	13	107	16	5	111	10	17
07:56-08:00	124	8	14	97	7	29	114	21	10	103	7	13
08:01-08:05	75	14	5	124	10	18	96	34	13	87	5	16
08:06-08:10	120	28	12	167	13	11	0	27	6	76	8	14
08:11-08:15	79	15	18	138	21	30	79	29	4	55	7	18
08:16-08:20	114	15	13	143	23	34	81	31	14	49	7	15
08:21-08:25	128	9	10	98	10	17	74	25	8	85	9	17
08:26-08:30	80	9	14	133	15	29	78	17	7	60	8	21
08:31-08:35	164	7	5	125	20	23	126	33	12	73	10	19
08:36-08:40	62	6	10	153	11	15	129	39	15	38	10	18
08:41-08:45	102	10	19	165	13	18	87	20	9	59	7	16
08:46-08:50	120	9	18	73	14	24	106	23	11	63	12	19
08:51-08:55	134	15	11	196	21	15	124	18	8	0	6	11
08:56-09:00	104	9	17	177	19	22	109	32	16	60	8	13
09:01-09:05	119	12	12	125	17	15	97	26	7	37	10	18
09:06-09:10	83	13	14	122	20	14	94	12	26	89	9	22

DAY 2 SITE 3 (UNIVERSITY WAY UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
09:11-09:15	110	12	8	151	17	20	86	28	17	57	12	25
09:16-09:20	133	17	15	164	10	37	113	21	13	73	14	20
09:21-09:25	198	15	13	155	17	21	107	11	10	67	10	22
09:26-09:30	132	10	21	171	21	15	74	15	23	181	8	27
09:31-09:35	45	16	21	183	11	9	92	10	19	125	7	31
09:36-09:40	95	14	16	164	20	17	126	12	21	32	10	23
09:41-09:45	156	10	14	195	19	24	67	32	12	0	9	25
09:46-09:50	84	22	10	172	10	22	0	27	19	30	11	33
09:51-09:55	119	21	8	163	17	24	0	28	26	137	15	22
09:56-10:00	201	9	6	188	11	21	56	30	21	83	10	35
10:01-10:05	115	8	12	151	17	32	78	16	11	43	9	29
10:06-10:10	118	12	17	168	16	34	94	28	8	0	14	37
10:11-10:15	107	10	9	172	18	27	69	21	12	39	14	32
10:16-10:20	112	13	11	195	13	29	86	12	14	114	10	33
10:21-10:25	125	6	16	105	21	30	49	31	18	72	12	33
10:26-10:30	58	12	12	115	15	24	62	31	26	132	11	28

Table 11: Day 2 Site 4 (Kenyatta Avenue-Moi Avenue-Mondlane Street Junction) Traffic Data

DAY 2 SITE 4 (KENYATTA AVENUE-MOI AVENUE-MONDLANE STREET JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	29	101	1	15	48		8	15	7
06:06-06:10	35	132	0	19	45		13	8	5
06:11-06:15	40	99	2	10	50		11	20	10
06:16-06:20	48	121	0	26	47		8	18	14
06:21-06:25	50	113	3	28	45		13	14	3
06:26-06:30	45	113	3	20	71		16	10	6
06:31-06:35	52	125	6	25	50		7	22	6
06:36-06:40	54	96	1	22	42		14	18	8
06:41-06:45	60	91	6	34	51		25	20	4
06:46-06:50	58	97	4	30	57		15	28	5
06:51-06:55	55	62	3	33	44		11	30	2
06:56-07:00	46	114	5	25	53		15	28	4
07:01-07:05	58	124	5	34	44		11	21	3
07:06-07:10	62	100	2	37	37		8	29	6
07:11-07:15	68	97	7	32	58		13	58	3
07:16-07:20	65	129	3	41	51		14	67	7
07:21-07:25	72	122	8	34	48		13	54	8
07:26-07:30	88	119	9	46	52		28	51	6
07:31-07:35	86	131	9	40	61		10	111	3
07:36-07:40	85	74	12	24	49		21	110	3
07:41-07:45	78	151	17	34	69		10	72	5
07:46-07:50	134	146	15	33	71		23	93	6
07:51-07:55	125	147	16	37	67		15	101	5
07:56-08:00	98	149	15	32	78		13	126	4
08:01-08:05	84	165	12	51	81		28	134	11
08:06-08:10	91	131	15	41	75		21	148	22
08:11-08:15	88	148	17	39	96		16	151	20
08:16-08:20	96	115	21	32	88		22	130	15
08:21-08:25	80	107	21	28	85		24	129	10

DAY 2 SITE 4 (KENYATTA AVENUE-MOI AVENUE-MONDLANE STREET JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
08:26-08:30	106	111	32	44	97		31	97	6
08:31-08:35	95	142	22	38	83		22	112	12
08:36-08:40	86	115	19	20	72		35	94	12
08:41-08:45	98	139	21	35	84		18	89	13
08:46-08:50	93	136	15	32	69		35	75	15
08:51-08:55	75	157	17	35	75		13	64	14
08:56-09:00	69	110	12	22	73		18	78	13
09:01-09:05	84	89	20	22	80		26	89	10
09:06-09:10	72	77	19	28	71		21	95	9
09:11-09:15	87	79	22	33	68		20	108	9
09:16-09:20	96	119	11	32	82		35	102	10
09:21-09:25	82	78	17	36	73		24	98	15
09:26-09:30	78	118	25	27	76		30	109	18
09:31-09:35	73	102	20	42	82		27	101	11
09:36-09:40	67	95	18	36	88		24	107	6
09:41-09:45	88	107	31	22	54		25	106	10
09:46-09:50	64	92	14	27	101		22	98	8
09:51-09:55	92	65	17	32	68		25	98	8
09:56-10:00	74	71	18	32	73		32	100	7
10:01-10:05	83	104	24	35	69		34	124	12
10:06-10:10	78	98	12	23	81		24	112	18
10:11-10:15	72	101	19	26	83		27	98	10
10:16-10:20	85	62	19	27	77		28	84	10
10:21-10:25	68	110	25	20	61		34	101	16
10:26-10:30	57	71	20	34	72		20	102	22

Table 12: Day 2 Site 5 (Moi Avenue-Slip Road Junction) Traffic Data

DAY 2 SITE 5 (MOI AVENUE-SLIP ROAD JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	13	8	2	29	12	4	40		
06:06-06:10	20	16	3	47	14	6	29		
06:11-06:15	12	12	4	35	10	2	33		
06:16-06:20	15	10	3	39	16	6	57		
06:21-06:25	42	21	2	42	15	1	24		
06:26-06:30	33	31	1	73	32	3	36		
06:31-06:35	59	30	1	59	48	4	55		
06:36-06:40	47	49	0	75	43	3	67		
06:41-06:45	50	23	5	58	84	5	90		
06:46-06:50	45	38	2	79	84	3	110		
06:51-06:55	43	31	3	62	48	4	76		
06:56-07:00	68	38	7	63	45	3	59		
07:01-07:05	63	49	3	66	59	11	60		
07:06-07:10	32	24	3	56	57	9	48		
07:11-07:15	60	31	2	53	41	7	95		
07:16-07:20	65	38	3	75	47	13	75		
07:21-07:25	50	42	3	82	48	16	92		
07:26-07:30	70	26	5	71	44	14	83		
07:31-07:35	78	34	2	76	32	8	98		
07:36-07:40	70	22	6	63	49	24	40		
07:41-07:45	63	42	6	78	56	18	87		
07:46-07:50	82	36	7	83	64	15	76		

DAY 2 SITE 5 (MOI AVENUE-SLIP ROAD JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
07:51-07:55	86	27	4	55	42	12	103		
07:56-08:00	61	20	5	85	46	16	98		
08:01-08:05	75	37	6	44	36	27	112		
08:06-08:10	71	43	5	65	37	22	75		
08:11-08:15	88	52	8	62	44	15	94		
08:16-08:20	76	54	6	60	38	33	107		
08:21-08:25	75	32	7	85	41	37	95		
08:26-08:30	79	38	9	79	32	26	98		
08:31-08:35	75	41	11	68	38	30	83		
08:36-08:40	90	38	6	71	32	27	101		
08:41-08:45	58	36	7	86	40	24	95		
08:46-08:50	76	41	8	75	24	17	88		
08:51-08:55	70	43	15	69	18	30	86		
08:56-09:00	63	52	11	81	22	24	104		
09:01-09:05	89	34	12	56	16	20	89		
09:06-09:10	70	29	10	64	26	31	92		
09:11-09:15	93	38	12	73	27	24	108		
09:16-09:20	68	21	8	47	18	28	78		
09:21-09:25	110	18	7	56	42	55	83		
09:26-09:30	87	25	10	37	14	28	107		
09:31-09:35	72	31	6	40	18	43	111		
09:36-09:40	81	38	16	49	13	44	94		
09:41-09:45	71	20	14	58	16	56	97		
09:46-09:50	100	22	17	65	22	29	84		
09:51-09:55	87	13	8	47	26	37	101		
09:56-10:00	115	23	7	70	18	24	106		
10:01-10:05	81	20	9	76	16	31	185		
10:06-10:10	94	16	6	80	14	23	126		
10:11-10:15	102	24	10	77	22	35	111		
10:16-10:20	86	30	11	59	20	33	176		
10:21-10:25	98	16	13	63	32	51	196		
10:26-10:30	91	18	7	60	36	44	172		

Table 13: Day 2 Site 6 (City Hall Way-Wabera Street T-Roundabout) Traffic Data

DAY 2 SITE 6 (CITY HALL WAY-WABERA STREET T-ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	12			27	0	1	18	24	0
06:06-06:10	10			18	3	0	18	20	1
06:11-06:15	10			20	1	0	22	17	2
06:16-06:20	15			26	3	0	17	24	1
06:21-06:25	18			28	3	0	17	20	0
06:26-06:30	12			25	4	0	8	26	0
06:31-06:35	9			23	4	1	33	22	5
06:36-06:40	21			27	4	5	24	33	0
06:41-06:45	17			25	2	0	35	20	1
06:46-06:50	31			37	2	1	35	37	1
06:51-06:55	21			48	3	1	37	29	2
06:56-07:00	20			54	5	2	50	50	7
07:01-07:05	28			48	1	3	34	22	5
07:06-07:10	14			62	2	4	24	23	5
07:11-07:15	26			57	7	5	41	45	3

DAY 2 SITE 6 (CITY HALL WAY-WABERA STREET T-ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
07:16-07:20	30			49	8	2	90	40	4
07:21-07:25	34			65	5	0	33	53	1
07:26-07:30	22			58	4	2	37	52	4
07:31-07:35	27			42	4	2	20	44	3
07:36-07:40	32			80	15	3	36	44	4
07:41-07:45	45			44	7	6	42	38	8
07:46-07:50	35			36	4	7	35	92	7
07:51-07:55	35			37	7	6	47	54	7
07:56-08:00	20			80	21	2	54	72	8
08:01-08:05	35			68	4	6	55	110	5
08:06-08:10	32			37	22	3	41	103	12
08:11-08:15	28			55	2	9	49	74	7
08:16-08:20	27			54	6	11	44	63	7
08:21-08:25	24			47	9	9	48	78	6
08:26-08:30	33			35	11	8	47	67	8
08:31-08:35	28			39	7	4	42	70	6
08:36-08:40	38			60	13	1	52	60	11
08:41-08:45	42			58	8	3	44	80	5
08:46-08:50	50			46	14	6	54	63	10
08:51-08:55	27			50	11	8	49	67	9
08:56-09:00	37			39	9	10	50	75	9
09:01-09:05	32			58	23	4	36	59	10
09:06-09:10	37			60	10	7	28	60	5
09:11-09:15	43			65	10	4	84	79	8
09:16-09:20	36			53	14	7	45	76	15
09:21-09:25	55			57	7	3	44	72	4
09:26-09:30	45			44	17	8	50	70	9
09:31-09:35	35			78	21	4	49	65	3
09:36-09:40	41			56	16	2	30	72	4
09:41-09:45	45			62	18	8	40	60	6
09:46-09:50	34			42	19	6	51	63	11
09:51-09:55	37			56	16	4	18	62	12
09:56-10:00	44			62	8	10	19	66	11
10:01-10:05	40			78	11	4	53	54	12
10:06-10:10	50			57	8	6	45	88	8
10:11-10:15	45			48	7	10	37	86	9
10:16-10:20	42			52	11	4	34	79	3
10:21-10:25	39			58	18	7	43	75	2
10:26-10:30	40			46	8	6	31	55	11

Table 14: Day 2 Site 7 (Haile Selassie Avenue-Moi Avenue Roundabout) Traffic Data

DAY 2 SITE 7 (HAILE SELASSIE AVENUE-MOI AVENUE ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	60	10	2	50	128	5	52	208	2
06:06-06:10	66	12	1	100	133	4	77	249	2
06:11-06:15	70	13	2	75	192	3	69	285	0
06:16-06:20	68	33	1	102	171	4	92	267	4
06:21-06:25	72	47	2	72	181	3	68	373	2
06:26-06:30	63	84	0	80	200	3	72	356	5
06:31-06:35	90	48	0	95	148	3	76	273	7
06:36-06:40	92	64	1	148	213	4	112	421	1

DAY 2 SITE 7 (HAILE SELASSIE AVENUE-MOI AVENUE ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:41-06:45	78	42	0	55	161	5	84	290	3
06:46-06:50	85	70	1	60	176	4	109	301	2
06:51-06:55	86	81	5	80	285	4	112	418	5
06:56-07:00	97	84	7	62	223	2	93	419	3
07:01-07:05	80	68	3	65	197	2	132	376	1
07:06-07:10	80	34	4	53	186	3	90	249	9
07:11-07:15	70	111	8	70	342	3	113	597	3
07:16-07:20	90	74	8	72	195	6	103	422	1
07:21-07:25	80	63	3	60	204	3	99	306	4
07:26-07:30	78	87	0	36	206	9	116	256	2
07:31-07:35	90	63	1	70	192	7	113	1007	2
07:36-07:40	100	80	8	90	562	5	38	1426	2
07:41-07:45	61	17	7	106	362	5	76	954	1
07:46-07:50	102	72	9	27	172	10	136	322	4
07:51-07:55	70	67	10	85	289	5	57	669	2
07:56-08:00	82	24	4	71	490	9	93	316	1
08:01-08:05	83	53	5	96	344	8	141	197	3
08:06-08:10	100	49	3	50	285	7	73	172	5
08:11-08:15	50	104	17	70	266	12	132	134	7
08:16-08:20	70	97	13	100	198	10	135	144	3
08:21-08:25	65	58	24	20	194	13	155	161	1
08:26-08:30	62	43	16	21	178	14	147	136	2
08:31-08:35	73	62	10	50	180	9	119	172	1
08:36-08:40	60	70	11	70	140	7	111	155	5
08:41-08:45	110	122	7	60	124	5	103	186	3
08:46-08:50	90	49	14	100	146	10	118	166	1
08:51-08:55	80	94	14	0	165	5	143	138	6
08:56-09:00	76	113	0	62	130	4	125	121	3
09:01-09:05	62	84	14	105	128	3	138	152	2
09:06-09:10	61	104	15	80	129	5	129	134	1
09:11-09:15	52	88	7	70	118	6	137	153	8
09:16-09:20	65	74	12	130	104	6	148	140	2
09:21-09:25	85	67	23	40	102	6	99	176	4
09:26-09:30	72	28	0	72	98	12	89	129	4
09:31-09:35	68	59	18	57	95	3	102	135	2
09:36-09:40	70	104	13	100	146	5	137	189	9
09:41-09:45	76	66	14	63	100	7	104	128	5
09:46-09:50	80	47	15	35	78	5	117	168	4
09:51-09:55	73	43	9	70	83	7	142	159	4
09:56-10:00	70	67	16	38	100	6	108	116	2
10:01-10:05	72	56	6	40	128	10	127	133	7
10:06-10:10	70	85	19	30	113	14	112	139	4
10:11-10:15	81	78	11	42	139	14	114	187	2
10:16-10:20	70	73	24	30	118	12	106	119	2
10:21-10:25	58	44	14	23	120	11	119	145	6
10:26-10:30	55	27	23	40	99	9	102	154	1

Table 15: Day 3 site 1 (Haile Selassie Roundabout) Traffic Data

DAY 3 SITE 1 (HAILE SELASSIE ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	81	6	2	38			42	7	4	23	8	0

DAY 3 SITE 1 (HAILE SELASSIE ROUNDABOUT) TRAFFIC DATA												
DIRECTION	**NORTHWARD**			**SOUTHWARD**			**EASTWARD**			**WESTWARD**		
TIME	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**
06:06-06:10	90	10	5	47			37	12	3	41	10	0
06:11-06:15	79	7	2	41			49	7	5	36	14	2
06:16-06:20	180	7	4	63			51	8	4	64	9	2
06:21-06:25	206	9	3	78			34	14	4	73	6	4
06:26-06:30	212	12	7	103			40	11	3	91	7	0
06:31-06:35	232	14	6	65			75	9	3	79	4	0
06:36-06:40	297	14	7	105			42	5	7	94	12	0
06:41-06:45	221	28	6	119			41	10	8	97	8	4
06:46-06:50	325	26	12	115			57	7	2	78	16	3
06:51-06:55	131	15	12	165			70	7	3	98	14	2
06:56-07:00	260	24	13	119			32	13	4	101	10	0
07:01-07:05	237	24	10	137			78	8	3	68	16	0
07:06-07:10	222	39	11	123			61	10	15	98	18	1
07:11-07:15	228	23	21	139			76	16	5	102	9	0
07:16-07:20	114	33	9	176			64	12	8	112	8	0
07:21-07:25	263	27	13	63			122	18	7	126	14	0
07:26-07:30	162	37	19	187			79	4	7	108	16	0
07:31-07:35	240	21	12	39			96	19	14	92	8	0
07:36-07:40	165	18	24	139			105	15	3	107	11	3
07:41-07:45	280	29	20	85			88	7	9	115	12	0
07:46-07:50	169	20	19	196			82	3	10	52	14	4
07:51-07:55	258	31	33	0			69	6	4	112	8	4
07:56-08:00	190	24	28	134			77	4	4	129	16	5
08:01-08:05	179	33	36	180			87	11	15	140	32	0
08:06-08:10	207	21	35	105			133	15	11	0	19	4
08:11-08:15	121	24	38	99			93	4	23	68	18	2
08:16-08:20	189	14	28	46			68	4	10	74	16	0
08:21-08:25	88	13	38	161			120	9	11	98	17	4
08:26-08:30	192	18	45	72			92	7	15	112	36	6
08:31-08:35	297	17	28	0			112	4	13	126	32	0
08:36-08:40	205	15	45	49			69	2	10	131	30	0
08:41-08:45	97	13	28	168			85	7	13	54	28	0
08:46-08:50	210	27	19	160			49	4	9	68	27	0
08:51-08:55	231	18	27	0			63	4	7	74	16	0
08:56-09:00	308	17	45	88			78	8	12	98	19	3
09:01-09:05	134	15	39	12			115	2	11	89	18	4
09:06-09:10	267	5	23	267			39	2	9	120	0	0
09:11-09:15	224	7	32	0			45	5	16	112	19	3
09:16-09:20	129	13	31	172			108	3	17	74	36	4
09:21-09:25	299	8	37	0			75	4	20	68	46	10
09:26-09:30	207	7	30	140			62	5	10	54	31	8
09:31-09:35	230	16	47	139			17	4	9	93	29	0
09:36-09:40	231	9	46	179			119	3	6	110	18	7
09:41-09:45	206	14	38	0			37	0	13	97	27	4
09:46-09:50	249	13	40	0			119	11	10	68	20	0
09:51-09:55	61	10	36	285			51	8	17	23	18	6
09:56-10:00	252	15	50	1			82	7	21	54	37	18
10:01-10:05	250	26	40	30			68	9	11	34	39	10
10:06-10:10	149	11	44	296			44	8	29	41	48	3
10:11-10:15	146	9	46	0			89	9	14	51	19	6
10:16-10:20	181	14	56	211			78	6	21	0	38	14
10:21-10:25	163	15	35	129			58	8	19	35	44	9
10:26-10:30	333	13	39	44			63	5	13	74	18	0

Table 16: Day 3 Site 2 (Kenyatta Avenue Uhuru Highway Roundabout) Traffic Data

DAY 3 SITE 2 (KENYATTA AVENUE UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	**NORTHWARD**			**SOUTHWARD**			**EASTWARD**			**WESTWARD**		
TIME	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**
06:00-06:05	91	27	8	88	22	9	97	4	3	39	11	3
06:06-06:10	107	41	4	148	30	2	95	8	1	48	6	5
06:11-06:15	115	23	6	123	53	3	82	2	2	35	7	1
06:16-06:20	124	57	6	164	68	4	87	0	1	30	16	3
06:21-06:25	111	34	4	132	54	6	91	1	0	35	7	4
06:26-06:30	96	53	1	163	35	2	97	3	2	70	12	5
06:31-06:35	120	30	5	123	63	4	102	6	3	45	3	10
06:36-06:40	97	70	5	77	53	0	78	3	3	74	12	4
06:41-06:45	186	46	6	143	62	2	67	3	2	86	10	5
06:46-06:50	77	41	3	193	59	2	143	0	5	75	17	6
06:51-06:55	218	66	9	203	56	7	80	9	1	79	11	8
06:56-07:00	0	66	5	127	48	16	64	2	5	106	9	10
07:01-07:05	135	62	6	149	50	19	68	3	8	49	12	4
07:06-07:10	100	56	7	167	43	22	38	0	2	84	9	3
07:11-07:15	159	89	2	173	28	11	48	2	7	110	3	10
07:16-07:20	0	62	8	153	37	19	62	8	3	100	3	9
07:21-07:25	146	72	7	169	74	7	88	1	4	85	9	8
07:26-07:30	0	69	6	217	33	8	95	7	3	96	3	11
07:31-07:35	167	83	7	176	41	13	104	6	2	0	13	7
07:36-07:40	217	91	13	154	38	17	114	6	3	66	6	9
07:41-07:45	0	97	15	0	20	15	124	1	4	168	4	10
07:46-07:50	241	58	5	214	31	10	109	8	4	40	10	7
07:51-07:55	80	77	22	193	34	18	122	12	5	73	5	15
07:56-08:00	123	111	22	135	26	14	97	9	5	131	13	17
08:01-08:05	61	95	18	68	31	9	104	7	4	112	1	7
08:06-08:10	0	90	13	139	25	4	96	2	6	123	7	10
08:11-08:15	100	59	23	113	21	13	92	8	8	15	13	15
08:16-08:20	0	54	21	223	32	11	124	6	7	55	6	8
08:21-08:25	96	32	20	93	37	14	127	3	9	177	11	15
08:26-08:30	201	75	18	0	29	18	132	3	8	80	7	19
08:31-08:35	147	45	14	246	21	16	114	6	6	49	14	12
08:36-08:40	177	65	18	39	48	24	118	6	10	116	15	25
08:41-08:45	31	65	18	97	43	28	107	3	7	42	4	13
08:46-08:50	126	41	9	103	55	31	111	7	10	28	10	12
08:51-08:55	145	53	26	46	47	34	132	7	14	102	6	20
08:56-09:00	0	47	16	69	50	28	129	4	3	69	5	18
09:01-09:05	147	37	29	156	41	19	113	4	8	27	4	13
09:06-09:10	115	75	30	188	38	26	118	6	8	31	10	18
09:11-09:15	0	66	17	139	48	21	121	1	9	78	18	22
09:16-09:20	256	47	14	48	32	11	114	4	10	148	4	18
09:21-09:25	71	56	29	161	44	17	117	3	7	28	7	23
09:26-09:30	119	47	21	62	29	31	112	4	9	120	13	24
09:31-09:35	201	58	17	192	26	21	127	6	10	56	3	23
09:36-09:40	113	38	21	97	38	16	102	8	15	108	10	25
09:41-09:45	128	80	28	116	27	31	113	4	8	40	10	13
09:46-09:50	224	46	23	140	29	23	124	6	7	99	12	25
09:51-09:55	0	34	16	189	36	14	112	7	7	60	11	16
09:56-10:00	100	31	35	61	14	21	108	3	12	92	5	26
10:01-10:05	22	53	26	113	19	24	97	24	4	61	10	1
10:06-10:10	193	64	19	144	16	16	121	16	8	101	7	2
10:11-10:15	51	56	21	183	21	11	118	11	9	20	8	6
10:16-10:20	105	80	27	155	24	27	123	27	4	109	9	6

DAY 3 SITE 2 (KENYATTA AVENUE UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
10:21-10:25	90	59	22	122	31	24	114	24	10	107	3	1
10:26-10:30	0	81	25	257	27	16	122	16	9	55	4	8

Table 17: Day 3 Site 3 (University Way Uhuru Highway Roundabout) Traffic Data

DAY 3 SITE 3 (UNIVERSITY WAY UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	72	10	13	0	2	4	14	21	2	53	20	10
06:06-06:10	81	9	12	147	15	6	17	19	7	71	10	15
06:11-06:15	58	11	10	128	1	7	12	23	1	68	14	12
06:16-06:20	105	16	9	133	0	8	19	28	2	43	11	20
06:21-06:25	63	15	14	141	3	5	24	17	4	110	12	10
06:26-06:30	74	11	12	156	4	6	16	25	1	94	16	7
06:31-06:35	90	20	11	113	5	12	37	19	3	72	18	9
06:36-06:40	94	14	5	119	6	7	49	27	6	100	21	12
06:41-06:45	69	16	3	137	10	13	89	31	2	102	19	11
06:46-06:50	105	35	6	153	30	10	96	24	5	80	17	8
06:51-06:55	110	19	6	176	22	15	74	38	7	113	15	8
06:56-07:00	78	54	9	122	15	7	112	32	4	145	18	10
07:01-07:05	113	49	11	126	29	8	89	43	11	97	16	9
07:06-07:10	128	37	18	129	30	12	43	36	9	108	14	12
07:11-07:15	131	63	19	134	16	6	64	27	6	129	19	10
07:16-07:20	149	33	12	74	15	12	91	22	5	77	13	8
07:21-07:25	65	20	7	107	19	8	0	29	6	85	12	8
07:26-07:30	77	42	17	172	16	19	0	35	12	108	10	11
07:31-07:35	110	52	18	109	15	13	69	43	9	95	15	9
07:36-07:40	152	24	11	156	22	19	114	37	8	0	13	10
07:41-07:45	98	28	14	131	17	22	98	45	12	165	15	9
07:46-07:50	24	33	9	173	16	14	113	32	10	76	10	12
07:51-07:55	138	24	15	167	19	23	109	24	7	0	12	15
07:56-08:00	95	19	13	115	24	27	86	37	12	195	9	13
08:01-08:05	139	25	20	155	17	22	91	43	9	0	11	13
08:06-08:10	40	28	13	227	21	25	76	31	14	219	8	16
08:11-08:15	51	41	24	141	23	19	68	26	11	79	13	18
08:16-08:20	87	27	11	163	25	23	94	35	8	0	10	18
08:21-08:25	45	35	14	155	21	14	87	47	13	155	10	14
08:26-08:30	171	18	19	163	20	32	39	32	17	72	7	15
08:31-08:35	140	50	14	11	12	15	47	35	19	43	9	17
08:36-08:40	104	3	21	123	9	24	54	27	11	29	11	18
08:41-08:45	162	43	19	139	22	26	0	16	8	81	13	21
08:46-08:50	106	24	14	144	15	24	0	28	24	48	8	16
08:51-08:55	139	31	20	212	17	27	69	32	18	73	12	21
08:56-09:00	115	27	9	188	12	32	117	27	21	39	15	27
09:01-09:05	82	34	21	117	9	28	102	34	15	48	10	23
09:06-09:10	127	41	18	151	13	24	89	29	12	0	14	26
09:11-09:15	107	27	14	137	11	27	74	38	27	213	9	29
09:16-09:20	32	13	18	145	3	30	69	24	22	38	7	25
09:21-09:25	111	21	21	152	17	35	81	32	18	47	11	28
09:26-09:30	54	14	4	112	15	33	54	26	25	73	13	31
09:31-09:35	65	34	18	93	19	40	46	24	21	52	15	27
09:36-09:40	109	60	41	144	18	32	57	18	26	38	10	27
09:41-09:45	86	41	33	126	21	35	81	23	18	0	8	32

DAY 3 SITE 3 (UNIVERSITY WAY UHURU HIGHWAY ROUNDABOUT) TRAFFIC DATA												
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
09:46-09:50	42	19	20	139	17	32	63	34	27	123	14	30
09:51-09:55	98	27	14	125	20	28	92	27	15	79	14	24
09:56-10:00	57	41	19	134	21	16	87	28	14	98	8	19
10:01-10:05	119	18	13	148	17	23	0	19	26	69	11	23
10:06-10:10	123	24	7	153	19	30	0	14	19	0	13	27
10:11-10:15	103	19	24	172	31	27	18	27	12	107	10	31
10:16-10:20	75	29	18	144	35	26	37	39	19	18	12	28
10:21-10:25	96	16	17	169	28	22	49	24	13	0	10	25
10:26-10:30	59	35	20	129	15	27	54	43	22	149	15	23

Table 18: Day 3 Site 4 (Kenyatta Avenue-Moi Avenue-Mondlane Street Junction) Traffic Data

DAY 3 SITE 4 (KENYATTA AVENUE-MOI AVENUE-MONDLANE STREET JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	26	96	2	24	39	2	4	38	4
06:06-06:10	35	84	2	27	33	3	2	42	11
06:11-06:15	41	107	1	32	46	3	8	55	7
06:16-06:20	39	100	4	25	42	5	10	46	9
06:21-06:25	40	93	6	26	44	2	17	40	3
06:26-06:30	52	110	1	31	39	4	12	63	3
06:31-06:35	48	89	5	30	41	0	10	68	6
06:36-06:40	44	107	4	23	52	2	14	59	5
06:41-06:45	56	85	5	22	46	3	8	83	7
06:46-06:50	65	150	4	26	48	5	17	92	12
06:51-06:55	49	147	4	34	51	3	19	76	13
06:56-07:00	38	126	6	22	54	4	13	67	3
07:01-07:05	42	129	4	24	58	7	10	72	7
07:06-07:10	54	146	13	29	62	9	18	83	7
07:11-07:15	67	123	8	39	60	10	7	64	9
07:16-07:20	58	192	13	34	79	13	21	59	5
07:21-07:25	71	156	9	43	94	6	23	67	13
07:26-07:30	64	98	12	44	88	9	16	56	12
07:31-07:35	52	121	18	41	72	12	13	63	5
07:36-07:40	68	203	12	34	90	13	10	148	6
07:41-07:45	71	192	8	36	84	11	17	70	9
07:46-07:50	84	189	11	32	98	9	11	68	10
07:51-07:55	92	143	11	36	102	10	21	67	13
07:56-08:00	75	127	15	31	99	7	18	48	6
08:01-08:05	63	174	15	35	106	11	14	118	16
08:06-08:10	88	125	14	38	96	11	13	92	10
08:11-08:15	76	136	13	41	88	13	13	71	3
08:16-08:20	95	110	17	34	89	15	23	154	10
08:21-08:25	82	141	12	40	109	10	24	145	14
08:26-08:30	74	148	14	37	93	13	16	60	22
08:31-08:35	58	148	16	35	80	15	20	72	24
08:36-08:40	93	161	16	34	94	18	26	69	17
08:41-08:45	102	103	10	31	74	13	25	64	17
08:46-08:50	88	110	14	32	98	20	17	72	24
08:51-08:55	79	79	9	40	86	18	21	87	26
08:56-09:00	94	78	13	26	96	13	23	149	10

DAY 3 SITE 4 (KENYATTA AVENUE-MOI AVENUE-MONDLANE STREET JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			SOUTHWARD			EASTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
09:01-09:05	82	96	15	31	94	16	26	155	17
09:06-09:10	73	99	14	33	67	12	18	67	11
09:11-09:15	68	104	19	22	72	18	28	138	9
09:16-09:20	57	89	16	41	82	8	22	62	23
09:21-09:25	72	109	16	32	76	10	23	93	30
09:26-09:30	61	121	19	28	86	12	15	89	22
09:31-09:35	86	111	11	39	79	25	22	83	8
09:36-09:40	77	122	30	35	110	13	25	89	11
09:41-09:45	69	110	24	36	90	19	24	71	13
09:46-09:50	92	98	16	33	87	10	19	84	19
09:51-09:55	78	113	18	35	95	13	26	157	17
09:56-10:00	64	132	11	42	58	12	30	88	18
10:01-10:05	72	106	27	26	65	17	21	90	12
10:06-10:10	83	99	14	30	61	7	25	61	20
10:11-10:15	59	87	13	27	57	13	22	77	10
10:16-10:20	48	104	33	31	62	21	27	75	27
10:21-10:25	56	96	19	34	55	21	28	61	13
10:26-10:30	60	75	32	29	63	19	26	55	16

Table 19: Day 3 Site 5 (Moi Avenue-Slip Road Junction) Traffic Data

DAY 3 SITE 5 (MOI AVENUE-SLIP ROAD JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	16	13	0	48	23	2	78		
06:06-06:10	39	15	1	60	39	4	63		
06:11-06:15	42	20	4	54	46	6	84		
06:16-06:20	51	25	3	63	58	7	120		
06:21-06:25	60	19	3	58	37	11	91		
06:26-06:30	43	16	1	71	20	3	153		
06:31-06:35	67	11	5	68	23	2	135		
06:36-06:40	48	29	2	76	45	10	147		
06:41-06:45	44	12	1	92	38	5	146		
06:46-06:50	66	10	3	63	59	4	154		
06:51-06:55	52	17	2	77	35	6	107		
06:56-07:00	59	22	6	85	41	7	123		
07:01-07:05	47	37	10	98	27	8	167		
07:06-07:10	62	42	12	121	51	7	143		
07:11-07:15	57	24	7	101	53	6	107		
07:16-07:20	71	32	12	99	29	6	94		
07:21-07:25	58	40	18	134	37	8	145		
07:26-07:30	54	33	11	61	38	7	132		
07:31-07:35	76	21	8	119	31	21	121		
07:36-07:40	52	19	7	74	27	8	94		
07:41-07:45	61	19	3	94	27	14	134		
07:46-07:50	75	15	4	87	33	12	98		
07:51-07:55	69	22	13	80	21	18	74		
07:56-08:00	90	19	6	102	32	21	98		
08:01-08:05	67	12	5	96	32	22	109		
08:06-08:10	73	19	4	82	20	13	131		
08:11-08:15	67	10	7	87	26	9	74		
08:16-08:20	92	13	8	66	16	16	84		
08:21-08:25	87	15	13	74	27	14	126		

DAY 3 SITE 5 (MOI AVENUE-SLIP ROAD JUNCTION) TRAFFIC DATA									
DIRECTION	NORTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
08:26-08:30	102	25	19	69	23	21	87		
08:31-08:35	81	21	12	72	27	18	75		
08:36-08:40	79	29	20	44	29	30	102		
08:41-08:45	124	26	12	66	18	20	79		
08:46-08:50	61	33	19	81	21	24	65		
08:51-08:55	111	42	19	88	16	37	107		
08:56-09:00	73	30	21	96	27	42	83		
09:01-09:05	59	12	25	78	15	14	65		
09:06-09:10	95	21	19	90	18	27	70		
09:11-09:15	72	17	9	113	17	20	97		
09:16-09:20	89	12	20	83	19	11	104		
09:21-09:25	134	19	23	92	12	21	84		
09:26-09:30	53	26	30	84	5	8	128		
09:31-09:35	42	33	12	73	12	9	79		
09:36-09:40	75	40	13	85	17	22	85		
09:41-09:45	52	19	12	110	12	10	97		
09:46-09:50	63	20	16	64	5	17	100		
09:51-09:55	41	23	19	73	11	19	89		
09:56-10:00	110	33	29	63	14	20	77		
10:01-10:05	85	34	17	83	26	18	121		
10:06-10:10	63	21	11	70	12	19	97		
10:11-10:15	56	36	19	67	32	18	74		
10:16-10:20	65	40	23	61	17	16	81		
10:21-10:25	88	29	24	68	21	20	79		
10:26-10:30	72	21	17	56	18	12	96		

Table 20: Day 3 site 6 (City Hall Way-Wabera Street T-Roundabout) Traffic Data

DAY 3 SITE 6 (CITY HALL WAY-WABERA STREET T-ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	8			17	2	0	22	19	3
06:06-06:10	9			15	3	0	17	19	1
06:11-06:15	5			19	2	0	16	35	3
06:16-06:20	16			23	4	0	27	21	0
06:21-06:25	14			39	2	0	28	18	0
06:26-06:30	13			37	1	0	24	35	1
06:31-06:35	18			36	6	1	28	29	4
06:36-06:40	9			42	4	4	23	51	3
06:41-06:45	20			33	1	1	35	8	4
06:46-06:50	23			37	6	3	25	43	4
06:51-06:55	22			41	4	1	40	32	3
06:56-07:00	17			36	1	2	30	33	3
07:01-07:05	25			38	2	0	40	25	3
07:06-07:10	24			40	6	2	33	36	4
07:11-07:15	21			42	4	3	36	41	3
07:16-07:20	34			32	4	2	41	38	4
07:21-07:25	27			38	3	1	32	22	5
07:26-07:30	22			49	5	3	39	43	3
07:31-07:35	23			50	4	3	48	55	5
07:36-07:40	27			32	3	1	39	66	3
07:41-07:45	37			48	2	5	18	76	5
07:46-07:50	28			56	5	3	44	105	3

DAY 3 SITE 6 (CITY HALL WAY-WABERA STREET T-ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
07:51-07:55	23			47	6	2	37	190	14
07:56-08:00	27			27	6	4	42	96	5
08:01-08:05	27			52	5	7	44	84	6
08:06-08:10	31			58	8	2	50	70	10
08:11-08:15	34			42	4	10	65	82	8
08:16-08:20	32			52	6	5	53	63	12
08:21-08:25	25			62	3	1	49	60	8
08:26-08:30	28			48	3	7	43	67	3
08:31-08:35	26			56	9	8	30	58	4
08:36-08:40	31			32	7	6	26	47	1
08:41-08:45	43			38	10	7	44	58	4
08:46-08:50	31			52	5	3	40	99	6
08:51-08:55	32			43	6	3	37	81	12
08:56-09:00	35			36	7	7	41	79	17
09:01-09:05	48			59	9	17	28	56	5
09:06-09:10	22			62	5	12	30	87	12
09:11-09:15	25			42	4	7	50	72	8
09:16-09:20	28			65	1	3	41	76	7
09:21-09:25	49			47	2	7	39	59	12
09:26-09:30	42			56	6	3	37	92	5
09:31-09:35	33			58	10	9	35	66	9
09:36-09:40	34			40	11	2	42	100	5
09:41-09:45	34			64	8	5	43	78	11
09:46-09:50	26			79	18	2	41	61	10
09:51-09:55	53			53	7	14	56	93	8
09:56-10:00	38			49	9	5	43	49	7
10:01-10:05	46			53	11	8	45	102	13
10:06-10:10	48			59	15	10	46	91	11
10:11-10:15	40			45	17	12	21	27	13
10:16-10:20	40			47	20	14	36	103	13
10:21-10:25	33			73	7	5	39	120	14
10:26-10:30	46			62	16	8	45	109	13

Table 21: Day 3 Site 7 (Haile Selassie Avenue-Moi Avenue Roundabout) Traffic Data

DAY 3 SITE 7 (HAILE SELASSIE AVENUE-MOI AVENUE ROUNDABOUT) TRAFFIC DATA									
DIRECTION	SOUTHWARD			EASTWARD			WESTWARD		
TIME	VEH	PED	MOT	VEH	PED	MOT	VEH	PED	MOT
06:00-06:05	57	12	2	90	10	0	87	120	4
06:06-06:10	72	10	1	88	150	4	91	193	0
06:11-06:15	67	13	0	94	120	4	93	187	3
06:16-06:20	70	51	1	76	120	1	102	117	1
06:21-06:25	73	50	1	94	66	0	92	197	2
06:26-06:30	80	71	0	92	122	2	96	172	3
06:31-06:35	85	78	3	82	145	5	73	184	5
06:36-06:40	52	67	2	95	58	2	83	170	3
06:41-06:45	102	103	0	70	170	0	92	137	1
06:46-06:50	80	49	0	78	123	1	119	148	4
06:51-06:55	87	81	3	102	145	0	97	192	7
06:56-07:00	70	118	7	98	181	1	108	167	5
07:01-07:05	85	97	3	69	94	3	105	182	3
07:06-07:10	80	72	5	80	90	1	124	208	2
07:11-07:15	72	81	6	71	123	3	83	136	3

DAY 3 SITE 7 (HAILE SELASSIE AVENUE-MOI AVENUE ROUNDABOUT) TRAFFIC DATA									
DIRECTION	**SOUTHWARD**			**EASTWARD**			**WESTWARD**		
TIME	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**	**VEH**	**PED**	**MOT**
07:16-07:20	71	79	5	72	130	2	102	147	6
07:21-07:25	83	81	3	67	70	1	116	127	4
07:26-07:30	92	67	2	76	83	3	132	148	2
07:31-07:35	100	79	10	65	46	1	128	184	6
07:36-07:40	110	51	7	80	245	2	76	1050	2
07:41-07:45	101	48	3	93	200	5	125	980	7
07:46-07:50	90	109	4	88	250	3	116	420	3
07:51-07:55	70	43	2	70	500	7	129	1320	6
07:56-08:00	76	79	6	72	337	5	133	510	2
08:01-08:05	56	107	4	92	213	9	117	280	4
08:06-08:10	71	72	12	102	67	15	114	176	3
08:11-08:15	64	57	5	105	53	4	106	156	8
08:16-08:20	52	112	3	55	51	3	104	223	2
08:21-08:25	64	60	13	64	23	10	97	180	3
08:26-08:30	55	48	2	63	37	7	88	173	0
08:31-08:35	60	64	9	102	72	4	121	184	2
08:36-08:40	64	72	9	84	103	6	96	157	3
08:41-08:45	61	43	0	69	38	11	73	180	3
08:46-08:50	53	61	10	73	59	3	97	152	2
08:51-08:55	60	39	3	64	39	5	118	126	1
08:56-09:00	80	112	7	64	54	12	123	130	1
09:01-09:05	85	38	1	73	27	6	96	145	3
09:06-09:10	73	64	21	60	42	13	112	117	2
09:11-09:15	60	27	3	68	36	9	133	128	3
09:16-09:20	59	71	7	89	44	7	77	125	2
09:21-09:25	68	99	14	14	15	9	103	158	6
09:26-09:30	64	81	12	87	67	0	88	118	4
09:31-09:35	45	94	17	108	61	3	93	136	7
09:36-09:40	81	61	20	102	24	2	74	95	5
09:41-09:45	105	87	13	90	66	3	91	127	1
09:46-09:50	90	117	7	103	68	4	119	176	2
09:51-09:55	99	22	4	83	52	14	104	112	4
09:56-10:00	86	99	13	92	45	8	122	84	2
10:01-10:05	65	65	10	81	41	3	117	207	7
10:06-10:10	80	43	19	43	44	1	97	178	8
10:11-10:15	78	46	13	88	71	2	102	184	5
10:16-10:20	83	35	6	70	63	5	129	218	4
10:21-10:25	77	97	29	87	61	4	134	187	5
10:26-10:30	40	81	19	110	73	7	131	540	4

Appendix V: MATLAB Code for Day 3 Site 2 Traffic Flow GGM(1,1) Training

```
%The Nairobi CBD "Day 3 Site 2 Traffic Flow GGM(1,1) Training" MATLAB
code sample file

    % 1. Northbound file

tic

DN=[91  107 115 124 111 96  120 97  186 77  218 0 135 100 159 0 146 0
167 217 0  241 80  123 61  0  100];          % First 27 vehicle flow
data points

AN=[DN(1:4); DN(2:5); DN(3:6); DN(4:7); DN(5:8); DN(6:9); DN(7:10);
DN(8:11);  DN(9:12);  DN(10:13);  DN(11:14);  DN(12:15);  DN(13:16);
DN(14:17);  DN(15:18);  DN(16:19);  DN(17:20);  DN(18:21);  DN(19:22);
DN(20:23); DN(21:24); DN(22:25); DN(23:26); DN(24:27)];             %
24 groups formed by DGT in 4s.

    VN= zeros(24,4);

for j=1:24;

    XN=[AN(j,:)];                                          % Selecting
a group

    XNa=cumsum(XN);                                       % AGO

    ZN=0.5*[XN(1)+XNa(2), XNa(2)+XNa(3), XNa(3)+XNa(4)]  % Background
value Z

    YN=[XN(2:4)];                                          % Measured
vector Y

    BN=[-ZN;1,1,1]';                                     % Data matrix
B

    QN=inv(BN'*BN)*(BN'*YN');                              % Parameter
vector

    for i=1:4;

        t=i-1;

        VN(j,i)=(XN(1)-QN(2)/QN(1))*exp(-QN(1)*t)+QN(2)/QN(1);

    end

end

    VN;
```

```
    WN=[VN(:,1)  VN(:,2)-VN(:,1)  VN(:,3)-VN(:,2)  VN(:,4)-VN(:,3)];  %
IAGO

    VSN=fliplr(WN);

    H0N=sum(diag(VSN,0));

    H1N=zeros(1,3);

    for i=1:3;

        j=1;

        H1N(j,i)=sum(diag(VSN,i));

    H2N=zeros(1,23);

    for i=1:23;

        J=1;

        H2N(j,i)=sum(diag(VSN,-i));

    end

    end

    H1SN=fliplr(H1N);

    HN=[H1SN H0N H2N];

    PN=[HN(1)  HN(2)/2  HN(3)/3  HN(4:24)/4  HN(25)/3  HN(26)/2  HN(27)]  %
Fitted

    %                                                              values

    NRMSE=sqrt(mean((PN-DN).^2))                          % Northbound
RMSE

    NRMSPEe=sqrt(mean((PN-DN).^2/(DN).^2)*100^2)          % Northbound
RMSPE

    NMAE=mean(abs(DN-PN))                                 % Northbound
MAE

    NMAPD=100*sum(abs(DN-PN))/sum(abs(DN))                % Northbound
MAPD

    % 2. Southbound file

DS=[88 148 123 164 132 163 123 77 143 193 203 127 149 167 173 153 169
217 176 154 0 214 193 135 68 139 113];       % First 27 vehicle flow
data points.
```

```
AS=[DS(1:4); DS(2:5); DS(3:6); DS(4:7); DS(5:8); DS(6:9); DS(7:10);
DS(8:11);  DS(9:12);  DS(10:13);  DS(11:14);  DS(12:15);  DS(13:16);
DS(14:17);  DS(15:18);  DS(16:19);  DS(17:20);  DS(18:21);  DS(19:22);
DS(20:23); DS(21:24); DS(22:25); DS(23:26); DS(24:27)];            %
24 groups formed by DGT in 4s.

    VS= zeros(24,4);

for j=1:24;

    XS=[AS(j,:)];                                          % Selecting
a group

    XSa=cumsum(XS);                                       % AGO

    ZS=0.5*[XS(1)+XSa(2), XSa(2)+XSa(3), XSa(3)+XSa(4)]; % Background
value Z

    YS=[XS(2:4)];                                          % Measured
vector Y

    BS=[-ZS;1,1,1]';                                     % Data matrix
B

    QS=inv(BS'*BS)*(BS'*YS');                              % Parameter
vector

    for i=1:4;

        t=i-1;

        VS(j,i)=(XS(1)-QS(2)/QS(1))*exp(-QS(1)*t)+QS(2)/QS(1);

    end

end

    VS;

    WS=[VS(:,1)  VS(:,2)-VS(:,1)  VS(:,3)-VS(:,2)  VS(:,4)-VS(:,3)];  %
IAGO

    VSS=fliplr(WS);

    H0S=sum(diag(VSS,0));

    H1S=zeros(1,3);

    for i=1:3;

        j=1;

        H1S(j,i)=sum(diag(VSS,i));
```

```
    H2S=zeros(1,23);

    for i=1:23;

        J=1;

        H2S(j,i)=sum(diag(VSS,-i));

    end

    end

    H1SS=fliplr(H1S);

    HS=[H1SS H0S H2S];

    PS=[HS(1) HS(2)/2 HS(3)/3 HS(4:24)/4 HS(25)/3 HS(26)/2 HS(27)] %
Fitted

    %                                                          values

    SRMSE=sqrt(mean((PS-DS).^2))                          % Southbound
RMSE

    SRMSPEe=sqrt(mean((PS-DS).^2/(DS).^2)*100^2)          % Southbound
RMSPE

    SMAE=mean(abs(DS-PS))                                 % Southbound
MAE

    SMAPD=100*sum(abs(DS-PS))/sum(abs(DS))                % Southbound
MAPD

    % 3. Eastbound file

DE=[97   95   82   87   91   97   102 78   67   143 80   64   68   38   48   62
88   95   104 114 124 109 122 97   104 96   92];     % First 27 vehicle
flow data points.

AE=[DE(1:4); DE(2:5); DE(3:6); DE(4:7); DE(5:8); DE(6:9); DE(7:10);
DE(8:11);  DE(9:12);  DE(10:13);  DE(11:14);  DE(12:15);  DE(13:16);
DE(14:17);  DE(15:18);  DE(16:19);  DE(17:20);  DE(18:21);  DE(19:22);
DE(20:23); DE(21:24); DE(22:25); DE(23:26); DE(24:27)];             %
24 groups formed by DGT in 4s.

    VE= zeros(24,4);

for j=1:24;

    XE=[AE(j,:)];                                          % Selecting
a group

    XEa=cumsum(XE);                                       % AGO
```

```
    ZE=0.5*[XE(1)+XEa(2), XEa(2)+XEa(3), XEa(3)+XEa(4)]; % Background
value Z

    YE=[XE(2:4)];                                          % Measured
vector Y

    BE=[-ZE;1,1,1]';                                    % Data matrix
B

    QE=inv(BE'*BE)*(BE'*YE');                             % Parameter
vector

    for i=1:4;

        t=i-1;

        VE(j,i)=(XE(1)-QE(2)/QE(1))*exp(-QE(1)*t)+QE(2)/QE(1);

    end

end

    VE;

    WE=[VE(:,1)  VE(:,2)-VE(:,1)  VE(:,3)-VE(:,2)  VE(:,4)-VE(:,3)];  %
IAGO

    VSE=fliplr(WE);

    H0E=sum(diag(VSE,0));

    H1E=zeros(1,3);

    for i=1:3;

        j=1;

        H1E(j,i)=sum(diag(VSE,i));

    H2E=zeros(1,23);

    for i=1:23;

        J=1;

        H2E(j,i)=sum(diag(VSE,-i));

    end

    end

    H1SE=fliplr(H1E);
```

```
    HE=[H1SE H0E H2E];

    PE=[HE(1) HE(2)/2 HE(3)/3 HE(4:24)/4 HE(25)/3 HE(26)/2 HE(27)] %
Fitted

    %                                                          values

    ERMSE=sqrt(mean((PE-DE).^2))                          % Eastbound
RMSE

    ERMSPEe=sqrt(mean((PE-DE).^2/(DE).^2)*100^2)          % Eastbound
RMSPE

    EMAE=mean(abs(DE-PE))                                 % Eastbound
MAE

    EMAPD=100*sum(abs(DE-PE))/sum(abs(DE))                % Eastbound
MAPD

    % 4. Westbound file

DW=[39  48  35  30  35  70  45  74  86  75  79  106 49  84  110 100
85  96  0   66  168 40  73  131 112 123 15];    % First 27 vehicle
flow data points.

AW=[DW(1:4); DW(2:5); DW(3:6); DW(4:7); DW(5:8); DW(6:9); DW(7:10);
DW(8:11);  DW(9:12);  DW(10:13);  DW(11:14);  DW(12:15);  DW(13:16);
DW(14:17);  DW(15:18);  DW(16:19);  DW(17:20);  DW(18:21);  DW(19:22);
DW(20:23); DW(21:24); DW(22:25); DW(23:26); DW(24:27)];            %
24 groups formed by DGT in 4s.

    VW= zeros(24,4);

for j=1:24;

    XW=[AW(j,:)];                                          % Selecting
a group

    XWa=cumsum(XW);                                       % AGO

    ZW=0.5*[XW(1)+XWa(2), XWa(2)+XWa(3), XWa(3)+XWa(4)]; % Background
value Z

    YW=[XW(2:4)];                                          % Measured
vector Y

    BW=[-ZW;1,1,1]';                                      % Data matrix
B

    QW=inv(BW'*BW)*(BW'*YW');                              % Parameter
vector

    for i=1:4;
```

```
        t=i-1;

        VW(j,i)=(XW(1)-QW(2)/QW(1))*exp(-QW(1)*t)+QW(2)/QW(1);

    end

end

    VW;

    WW=[VW(:,1)  VW(:,2)-VW(:,1)  VW(:,3)-VW(:,2)  VW(:,4)-VW(:,3)];  %
IAGO

    VSW=fliplr(WW);

    H0W=sum(diag(VSW,0));

    H1W=zeros(1,3);

    for i=1:3;

        j=1;

        H1W(j,i)=sum(diag(VSW,i));

    H2W=zeros(1,23);

    for i=1:23;

        J=1;

        H2W(j,i)=sum(diag(VSW,-i));

    end

    end

    H1SW=fliplr(H1W);

    HW=[H1SW H0W H2W];

    PW=[HW(1)  HW(2)/2  HW(3)/3  HW(4:24)/4  HW(25)/3  HW(26)/2  HW(27)] %
Fitted

    %                                                              values

    WRMSE=sqrt(mean((PW-DW).^2))                              % Westbound
RMSE

    WRMSPEe=sqrt(mean((PW-DW).^2/(DW).^2)*100^2)              % Westbound
RMSPE
```

```
    WMAE=mean(abs(DW-PW))                                    % Westbound
MAE

    WMAPD=100*sum(abs(DW-PW))/sum(abs(DW))                   % Westbound
MAPD

    % PLOTS

    % Northbound PLOT

    t=[0:1:26];                              % Time sample (t)

    NDt=minus(DN,PN)                        % Northbound de-trended data
series

    SDt=minus(DS,PS)                        % Southbound de-trended data
series

    EDt=minus(DE,PE)                         % Eastbound de-trended data
series

    WDt=minus(DW,PW)                         % Westbound de-trended data
series

    figure

subplot(4,2,1)

plot(t,DN,'-*b',t,PN,'-ro');               % Graph of real and fitted
values

%legend('Real            data','Simulated           data            by
GGM(1,1)','location','northwest');

title('NORTHWARD');                         % Graph title

%ylabel('Vehicle Flow');                    % y-axis label

grid on;

subplot(4,2,3)

plot(t,NDt,'-+g');                         % plotting de-trended graph

%legend('Simulation Error','location','northwest')

grid on;

    % Southbound PLOT

subplot(4,2,2)

plot(t,DS,'-*b',t,PS,'-ro');               % Graph of real and fitted
values
```

```
%legend('Real            data','Simulated            data            by
GGM(1,1)','location','northwest');

title('SOUTHWARD');                          % Graph title

grid on;

subplot(4,2,4)

plot(t,SDt,'-+g');                          % plotting de-trended graph

%legend('Simulation Error','location','northwest')

    grid on;

    % Eastbound PLOT

subplot(4,2,5)

plot(t,DE,'-*b',t,PE,'-ro');                % Graph of real and fitted
values

legend('Real            data','Simulated            data            by
GGM(1,1)','location','northwest');

title('EASTWARD');                           % Graph title

%xlabel('Time Sample (t)');                  % x-axis label

ylabel('Vehicle Flow');                      % y-axis label

grid on;

subplot(4,2,7)

plot(t,EDt,'-+g');                          % plotting de-trended graph

legend('Simulation Error','location','northwest')

%title('EAST');                              % Graph title

xlabel('Time Sample (t)');                   % x-axis label

%ylabel('Vehicle Flow');                     % y-axis label

grid on;

    % Westbound PLOT

subplot(4,2,6)

plot(t,DW,'-*b',t,PW,'-ro');                % Graph of real and fitted
values
```

```
title('WESTWARD');                          % Graph title
suptitle('DAY 3 SITE 2 TRAFFIC FLOW GGM(1,1) TRAINING');
grid on;
subplot(4,2,8)
plot(t,WDt,'-+g');                          % plotting de-trended graph
%xlabel('Time Sample (t)');                 % x-axis label
grid on;
toc
```

Appendix VI: Publication List

Getanda, V. B., Kihato, P. K., Hinga, P. K., & Oya, H. (2021). Grey and linear regression models in load forecasting for enhanced smart grid management- Smart grid modelling. *2021 IEEE PES/IAS PowerAfrica*, Nairobi, Kenya, 1-5, doi: 10.1109/PowerAfrica52236. 2021.9543425.

Getanda, V. B., Kihato, P. K., Hinga, P. K., & Oya, H. (2020). Improved grey models in vehicular CO_2 emission modelling and forecasting to decarbonize urban areas – Sustainable environment. In *Proceedings of the 9th International Conference on Appropriate Technology (9th ICAT)*, Pretoria, South Africa, November 23-27, 2020, 88-100. https://www.scribd.com/document/485488919/9th-ICAT-Technology-Exchange-and-Employment-Creation-for-Community-Empowerment-Cross-Pollinating-Innovative-Models.

Getanda, V. B., Kihato, P. K., Hinga, P. K., & Oya, H. (2020). Electricity consumption modeling and medium-term forecasting based on grouped grey model, GGM(1,1). *2020 IEEE PES/IAS PowerAfrica*, Nairobi, Kenya, 2020, pp. 1-5. doi: 10.1109/PowerAfrica49420. 2020. 9219919.

Getanda, V. B., Kihato, P. K., Hinga, P. K., & Oya, H. (2019). Improved grey model, GM (1, 1), in short-term traffic flow forecasting - Smart transportation systems. *2019 IEEE AFRICON*, Accra, Ghana, 2019, 1-9. doi: 10.1109/AFRICON46755.2019.9133801.

Getanda, V. B., Oya, H., Kubo, T., & Sato, Y. (2020). Data grouping techniques' performance analysis in GM(1,1)'s prediction accuracy improvement for forecasting traffic parameters. *Mechatronic Systems and Control, 48*(1), 25-34.

Getanda, V. B., Kihato, P. K., Hinga, P. K., & Oya, H. (2023) Data grouping and modified initial condition in grey model improvement for short-

term traffic flow forecasting. *Automatika*, *64*(1), 178-188. doi: 10.1080/00051144.2022.2119500.