A COST-EFFECTIVE APPROACH TO HANDLE MEASUREMENT AND VERIFICATION SAMPLING AND MODELLING UNCERTAINTIES

by

Zadok Olinga

Submitted in partial fulfillment of the requirements for the degree

Master of Engineering (Electrical Engineering)

in the

Department of Electrical, Electronic and Computer Engineering
Faculty of Engineering, Built Environment and Information Technology
UNIVERSITY OF PRETORIA

September 2015

SUMMARY

A COST-EFFECTIVE APPROACH TO HANDLE MEASUREMENT AND VERIFICATION SAMPLING AND MODELLING UNCERTAINTIES

by

Zadok Olinga

Supervisor(s):	Prof. X. Xia
Department:	Electrical, Electronic and Computer Engineering
University:	University of Pretoria
Degree:	Master of Engineering (Electrical Engineering)
Keywords:	M&V, uncertainty, sampling, modelling, energy, efficiency,

In this study, a measurement and verification (M&V) cost minimisation model is proposed to deal with both the M&V modelling and sampling uncertainties. In order to find the optimal solutions in terms of the modelling accuracy level, and the sample size, the M&V cost that includes the modelling cost, sampling cost, and overhead cost is formulated as the objective function, in which the modelling cost is developed as a function of the modelling accuracy in terms of the coefficient of variation of the room mean square error (CV(RMSE)), and the sampling cost, which is directly related to the sample size.

In order to illustrate the effectiveness of the proposed model, an optimal M&V modelling and sampling strategy is designed for a traffic intersection lamp retrofit project. In addition, partial optimal M&V plans designed with optimal sampling but non-optimal modelling solutions, or with optimal modelling but non-optimal sampling solutions are employed as the benchmark. Comparisons between the optimal and non-optimal solutions show advantageous cost savings performance in the execution of sampling and modelling activities for the case study. More precisely, the optimal solutions reduce the sampling cost by 55% and the total M&V cost by 14% against the solutions obtained by optimal modelling but non-optimal sampling solutions.

To test the applicability and flexibility of the proposed model for the cost-effective design of

similar traffic light retrofit projects, simulations have been carried out to evaluate the model performance when applying the model to M&V projects with different characteristics. The simulation results show that the proposed model is able to offer flexible trade-offs between the modelling and sampling uncertainties; namely, using more accurate baseline models and smaller sample sizes or less accurate baseline models but greater sample sizes to accommodate different practical needs in executing M&V projects with different characteristics.

The major contributions of this study can be highlighted as follows: 1) a M&V modelling cost model is developed, which is able to offer a quantitative analysis of the M&V baseline model uncertainty and cost; and, 2) a M&V cost minimisation model is proposed to handle both the M&V modelling and sampling uncertainties cost-effectively. The effectiveness and flexibility of this model are demonstrated by a case study and simulations.

OPSOMMING

N KOSTE-EFFEKTIEWE BENADERING TOT DIE METING EN VERIFIKASIE MONSTERNEMING EN MODELLERING ONSEKERHEDE HANTEER

deur

Zadok Olinga

Studieleier(s):	Prof. X. Xia
Departement:	Elektriese, Elektroniese en Rekenaar-Ingenieurswese
Universiteit:	Universiteit van Pretoria
Graad:	Master of Engineering (Electrical Engineering)
Sleutelwoorde:	M&V, onsekerheid, steekproefneming, modellering, energie, doeltreffendheid,

In hierdie studie word 'n meet-en-verifieer (M&V) kosteverminderingsmodel voorgestel wat onsekerhede in beide die M&V modellering sowel as die steekproefnemingsproses in ag neem. Om die optimale oplossings te vind in terme van die modelleringsakkuraatheid en steekproefgrootte, word die M&V koste geformuleer as 'n objektiewe funksie wat die modelleringskoste, steekproefnemingskoste en oorhoofse koste insluit. Die modelleringskoste is 'n funksie van die modelleringsakkuraatheid in terme van die koëffisiënt van variasie van die wortel-gemiddelde-kwadratiese fout (CV(RMSE)) en die steekproefnemingskoste, wat direk verwant is aan die steekproef grootte.

Om die effektiwiteit van die voorgestelde model te illustreer word 'n optimale M&V modellerings- en steekproefnemingstrategie ontwerp vir die vervanging van gloeilampe by 'n verkeerskruising. Semi-optimale M&V strategieë, naamlik M&V strategieë met optimale steekproefneming maar nie-optimale modellering, en met optimale modellering maar nie-optimale steekproefneming, word as maatstaf gebruik. Vergelykings tussen optimale en nie-optimale oplossings wys voordelige kostebesparings in die steekproefneming- en modelleringsproses vir die gevallestudie. Om presies te wees, die optimale oplossings verminder die steekproefnemingskoste met 55% en die algehele M&V koste met 14% teenoor oplossings met

optimale modellering maar nie-optimale steekproefneming.

Om die toepaslikheid en buigsaamheid van die voorgestelde model te toets vir die koste-effektiewe ontwerp van soortgelyke verkeersligprojekte is simulasies uitgevoer om die prestasie van die model te evalueer met die toepassing van die model op M&V projekte met verskillende eienskappe.

Die resultate van die simulasies wys dat die voorgestelde model in staat is om buigsame kompromieë tussen modellerings- en steekproefnemingsonsekerhede te tref. Dit word bereik deur die gebruik van meer akkurate basislynmodelle en kleiner monstergroottes, of minder akkurate basislynmodelle maar groter monstergroottes om aan die verskillende praktiese vereiestes in die uitvoering van M&V projekte met verskillende eienskappe te voldoen.

Die hoof bydrae van hierdie studie kan as volg uiteengesit word: 1) 'n M&V modelleringskostemodel is ontwikkel, wat in staat is om 'n kwantitatiewe analise van die M&V basislynmodelonsekerheid en -koste aan te bied; 2) 'n M&V kosteverminderingsmodel is voorgestel om beide die M&V modellerings- en steekproefnemingsonsekerhede op 'n koste-effektiewe manier te hanteer. Die effektiwiteit en buigsaamheid van hierdie model is gedemonstreer deur middel van 'n gevallestudie en simulasies.

LIST OF ABBREVIATIONS

ASHRAE	American Society Of Heating, Refrigeration And Air-conditioning Engineers
CDM	clean development mechanism
CFL	compact fluorescent lamp
CV	sampling coefficient of variation
CV(RMSE)	coefficient of variation of the root mean square
DSM	demand side management
ECM	energy conservation measure
EE	energy efficiency
EEDSM	energy efficiency and demand side management
FEMP	federal energy management program
GP-MCEM	Gaussian process modelling and a Monte Carlo expectation maximisation
GUM	guide to the expression of uncertainty in measurement
IPMVP	international performance measurement and verification protocol
ISO	international standards organisation
kW	kilowatt
kWh	kilowatt-hour
LED	light emitting diode
MBE	mean bias error
M&V	measurement and verification
OLS	ordinary least squares
PF	power factor
RMSE	root mean square error
R	South African Rand
SANAS	South African National Accreditation System
SRS	simple random sampling
SSD	sample size determination
SVM	support vector machines
SANS	South African national standard
VSD	variable speed drives
W	watt

NOMENCLATURE

Symbols

λ_0	the search starting point to solve the optimisation model
$\bar{Y}_i$	the sample mean in the ith group
A_{2n}	the quantity *2-aspect* fittings at a traffic intersection
A_{3n}	the quantity *3-aspect* fittings at a traffic intersection
A_{4n}	the quantity *4-aspect* fittings at a traffic intersection
A_{4rn}	the quantity *4-aspect* fittings with turning arrows at a traffic intersection
a_i	the procurement cost per meter in the ith group
b_i	the installation cost per meter in the ith group
C_m	the M&V modelling cost for one model
C_s	the sampling cost for a M&V project
CV_i	the coefficient of variation of the ith group
CV_m	the CV(RMSE) of a M&V baseline model
E_1	the energy consumption at an intersection with individual aspect retrofits
E_2i	the energy consumption at an intersection with complete set retrofits
G_n	the quantity of green signal lamps at a traffic intersection
n_i	the optimum sample size for the ith group
p	the required relative precision
R_n	the quantity of red signal lamps at a traffic intersection
U_m	the total modelling uncertainty
U_s	the total sampling uncertainty
Y_n	the quantity of yellow signal lamps at a traffic intersection
z	the z-value of a normal distribution curve that corresponds to a desired confidence le

Subscripts

i	group counter
j	sample counter
m	modelling
s	Sampling

ACKNOWLEDGEMENTS

I would like to acknowledge the support of Dr. Xianming Ye during the course of this research.

I would also to acknowledge the support of my three sisters, Berna, Peace and Faith, during the course of this masters degree.

This work is supported by the Centre for New Energy Systems and the National Hub for the Post-graduate Programme in Energy Efficiency and Demand Side Management at the University of Pretoria.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 PROBLEM STATEMENT

1.1.1 Context of the problem

Global energy demand is high, and it is projected to increase by 37% by the year 2040 from its 2014 levels [1]. Along with the need to produce more energy is the need to produce and use it efficiently to reduce carbon emissions [2]. Recently in South Africa, the increased energy demand has manifested itself in the form of energy shortages that have led to load shedding [3]. Energy efficiency (EE) and demand side management (DSM)programmes have been proposed as a means to alleviate the energy crisis [4]. Most of the EE projects in South Africa are sponsored by Eskom, and for the purposes of project financing and decision making, measurement and verification (M&V) is used to quantify the savings from EE and DSM projects, and to also track the performance of those projects over the contractual periods [5].

International guidelines for M&V state that no credible M&V savings can be reported without a measure of uncertainty. The main sources of quantifiable uncertainty are measurement uncertainty, modelling uncertainty, and sampling uncertainty [6]. The quantification and alleviation of uncertainty is a contributor to M&V cost because it requires engineering professionals to perform baseline modelling, sampling, and meter installation. The International Performance Measurement and Verification Protocol (IPMVP) and and the Federal Energy Management Program (FEMP) recommend that M&V cost does not exceed 10% of the average annual savings being assessed [6, 7]. Other guidelines give cost limits based on the IPMVP

M&V option being used. The costs range from a minimum of 1% of the annual measured savings for IPMVP Option A to a maximum of 10% for IPMVP Option D [8],[9]. Therefore, researchers, M&V practitioners, and energy efficiency project participants are eager to find cost-effective solutions to handle M&V uncertainties.

1.1.2 Research gap

A lot of research has been done on M&V baseline modelling. Some of these works also focus on modelling uncertainty [10, 11, 12, 13, 14]. More recently there has also been work focused on M&V sampling uncertainty [15, 16]. However, there is no historical research focussed on combined uncertainty analysis.

The research carried out in this dissertation aims to examine sampling, and modelling uncertainties together as a way of minimising M&V cost whilst offering M&V practitioners flexibility in designing an optimal M&V plan, which either has high sampling uncertainty with low modelling uncertainty, or high modelling uncertainty with low sampling uncertainty. This study pays less attention to measurement uncertainty, because the most frequently used high accuracy power meters have gradually become affordable due to the fast developing meter design and manufacturing technologies. However, efforts on dealing with sampling and modelling uncertainties are currently believed to be the most significant contributors to the entire M&V cost, especially when both the modelling and sampling techniques are used during the M&V process.

A traffic light retrofit project is used to illustrate the performance of the method developed during this research, and a sensitivity analysis is carried out to show its applicability to other M&V projects.

1.2 RESEARCH OBJECTIVE AND QUESTIONS

This research focuses on a traffic signal lamp retrofit project. For this project, individual incandescent traffic signal lamps are replaced with individual energy efficient light emitting diode (LED) lights or complete combined incandescent traffic signal lamp sets are replaced with combined LED signal lamp sets.

The objective is to develop a cost effective approach to handling M&V modelling and sampling uncertainties as a way of giving M&V practitioners flexibility in decision making at the M&V plan phase of the project. The decisions involve the sample size to be used and the required model accuracy.

The following are the research questions:

- Can there be a trade-off between sampling accuracy and modelling accuracy?
- Can the modelling and sampling uncertainties in M&V be minimised while meeting a minimal M&V project cost?
- What is the effect of the sampling coefficient of variation (CV) on the model accuracy?
- Can a minimum M&V cost be achieved while meeting uncertainty criteria such as the 90/10 criteria which, is a commonly used criteria in M&V?

1.3 RESEARCH GOALS

The goal of this research is to develop a cost effective approach to handling M&V sampling and modelling uncertainties. This will be achieved by developing and solving a cost minimisation problem that takes into account the modelling cost and the sampling cost inherent in M&V projects in conjunction with sampling and modelling uncertainties.

1.4 OVERVIEW OF STUDY

This dissertation is organised as follows. Chapter 2 provides an overview of existing literature on M&V practice and contributions to the study of uncertainty in M&V. Chapter 3 sets out to develop a cost minimisation model for M&V by examining the sampling and modelling uncertainties. Chapter 4 lays out a case study on traffic signal lamp retrofit that is used to demonstrate the approach developed in chapter 3. The results are presented and discussed in chapter 5. Chapter 6 gives a conclusion to the dissertation.

CHAPTER 2

LITERATURE STUDY

2.1 CHAPTER OVERVIEW

This chapter presents literature related to energy efficiency (EE), how measurement and verification is used in EE projects, the impact of uncertainty in M&V, and it delves into existing contributions to the study of uncertainty in M&V. Furthermore, it describes the research problem, the research approach and the contribution of this research.

2.2 THE ROLE OF ENERGY EFFICIENCY IN MEETING SOUTH AFRICAN ENERGY NEEDS

Globally it is widely accepted that energy used for industrial development, and domestic use is a finite resource that needs to be used efficiently and generated cleanly. Global energy demand is set to grow by 37% by 2040 [1]. Reflecting this increased energy demand, South Africa has recently been experiencing load shedding in an effort to address its energy shortages. EE is one of the most cost effective solutions adopted for alleviating the energy shortage [1]. In South Africa, energy efficiency has been advocated in government white papers and policy documents [4].

In south Africa, energy efficiency is implemented through the DSM initiative. DSM refers to collaborative programmes aimed at reducing electricity demand by encouraging efficient energy use, particularly at peak periods (7:00 - 10:00 and 18:00-20:00) [4]. This can be achieved by shifting loads to off-peak periods, and by reducing the overall energy consumption through the installation of EE devices like LED lights, variable speed drives (VSD). Energy consumption can also be lowered by more efficient use of heating, ventilation and air-conditioning (systems), and optimising processes [4, 17].

2.3 MEASUREMENT AND VERIFICATION EXPLAINED

M&V provides an unbiased and repeatable process to quantify energy and demand savings in EE [6] and DSM projects [6] in more than 15 countries in Europe [18], North America [18], and Asia [19]; including South Africa [5].

In energy efficiency projects M&V is important because it reduces the risk of poor project performance according to a standard accepted by project stake holders [7], which encourages investment in EE projects [5]. It is also used to track and evaluate the performance of EE project activities, and it provides a level of confidence in EE efforts, which is critical to participation in global initiatives like the United Nations Clean Development Mechanism (CDM).

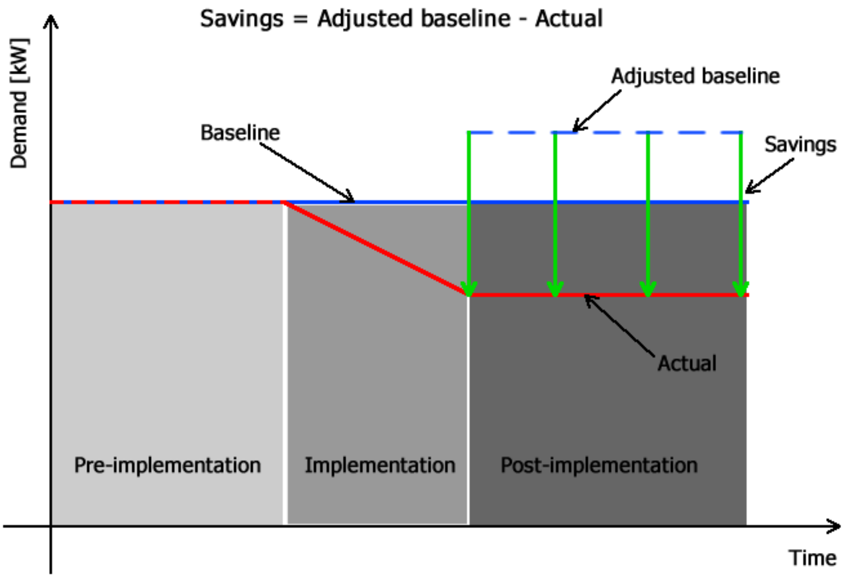

Figure 2.1: The M&V stages in the implementation of EE and DSM activities according to Eskom guidelines [5]

.

Figure 2.1 shows a typical time line for the implementation of EE projects, and the general equation for obtaining energy savings using M&V methodologies. Energy savings cannot be measured directly therefore they are obtained through the deduction of the post implementation energy consumption from the baseline energy consumption. This assumes that the conditions under which this comparison is achieved are the same. To ensure that this is true, an adjustment of the baseline might be necessary in the post implementation phase. Therefore the savings can be more accurately obtained by deducting the post implementation energy consumption from the adjusted baseline energy consumption[6].

M&V guidelines are provided in the IPMVP, which gives common practice recommendations and principles for quantifying energy and water savings [6]. Similar and widely reference guidelines based on the IPMVP are the M&V guideline of the American Society of Heating, Refrigeration and Air Conditioning Engineers (ASHRAE) [20] and the FEMP for federal energy projects in the USA [7]. Due to the wide implementation of EE initiatives, other guidelines have been drawn up for specific countries such as the South African M&V guideline for Demand Side Management projects [5] and the Australian best practice guideline [21]. The International Standards Organisation (ISO) has also released a standard for M&V, which establishes general principles and guidelines for the M&V process [22]. Furthermore, the South African National Standard (SANS) 50010 has been released to provide M&V guidelines for projects eligible under the 12I and 12L tax incentive program in South Africa [23].

The above guidelines provide M&V methodologies, and examples to guide M&V practitioners in calculating energy savings for EE and DSM projects. According to the IPMVP, M&V methodologies can be grouped into four M&V methods; namely, Option A, Option B, Option C, and Option D [6], which are listed in Table 2.1. IPMVP Options A, and B, are suited to energy subsystems that can be segregated from the complete energy consuming system [24]. An example of this is an energy efficiency project where only the lights in a building are being retrofitted. The lights form a subsystem that can be isolated for M&V, while the whole building with HVAC, computers and other electronic devices forms the energy system. When it is necessary to M&V the whole facility energy consumption, IPMVP Option C, and Option B are more suitable. They do not account for sub-facility energy consumption separately [24]. Typically, where HVAC system or building envelop retrofits are being carried out, Option C, and D are more applicable.

Option A is defined in the IPMVP text as a partially measured isolated retrofit with only key parameters being measured [24]. An example of this approach is measuring the power consumption of one fluorescent tube in a factory with the same tubes installed on the same floor, and those tubes are on a set lighting schedule.

When all the parameters of an energy consuming system are measured for M&V this is called IPMVP Option B [24]. An example of this is a retrofit of a chemical boiler in a paper mill where the input energy, the input paper pulp, the output pulp are all measured. In this instance, the whole paper mill is the energy system, and the chemical boiler the subsystem.

Option C is used to M&V whole facilities, and the interactions within the facility energy system are ignored [24]. Case in point are the HVAC and building envelop retrofits as mentioned above.

Option D requires the use of calibrated computer simulations to model the system energy performance, and then to calculate the energy savings once certain parameters are altered in that system [24].

Table 2.1: IPMVP options and their descriptions[6]

IPMVP Option	Description
A	Partially measured isolated retrofit with only key parameters being measured
B	Retrofit isolation with full parameter measurement
C	Complete Facility Measurement
D	Calibrated Simulation (whole facility or sub facility)

2.4 M&V IN SOUTH AFRICA

In South Africa M&V has been used for various EE projects under the DSM programme [5]. This has allowed the development of methodologies, and policies for M&V. In [25], a M&V methodology is developed for the mass roll-out of solar water heaters for domestic houses in South Africa. Through an assessment of the performance of some of the Solar water heating programmes, it is shown that compliance with the existing South African standards in the field of EE is affecting the confidence, and reliability of the reported energy savings according to currently existing M&V standards [25].

Within the industrial sector, the mining sector is one of the primary and obvious targets for

EE schemes due to the high electricity demand in this sector. Typical schemes involve the optimisation of compressed air, ventilation and, water pumping systems. An example of this is given in [26], which involves optimising the pressure set point profile of a pumping system, and fixing leaks that appear in the piping system [26]. M&V has been applied to calculate the energy savings that have been achieved from these projects.

Manufacturing plants are also major consumers of electricity in South Africa, and M&V methodologies have been developed for EE projects that involve lighting retrofits [27], motor controllers [10], to name a few. In [27], a methodology is developed for a lighting retrofit project where energy savings are achieved by replacing the old and inefficient lighting fixtures with new, and more efficient lighting fixtures. And in [10], a motor sequencing controller is installed to reduce the energy consumption of a conveyor belt in a plant. The M&V methodology in [10] involves developing a baseline adjustment model that correlates energy consumption to the production rate of the plant.

Beyond DSM projects such as those mentioned above, supply side energy conservation projects have also been implemented in South Africa, and M&V methodologies have been developed for them. In particular, this has been done with the country's utility Eskom. In [28], a M&V methodology is developed for a load shifting programme implemented in the utility's power plants. Beside energy savings, the project is also implemented with a view to carbon emissions reductions, which are also measured and verified [28].

As mentioned previously the M&V process is pivotal in assisting stakeholders in making decisions about investing in EE projects. The role of M&V in performance contracting projects is outlined in [29]. This study develops a M&V methodology for performance contracting that is aimed at convincing private developers to engage in the performance contracting program [29]. Further work that develops methodologies that promote EE projects via the guarantees of M&V is given in [30]. This work outlines M&V methodologies developed for the utility Eskom for various DSM activities.

All the projects, and methodologies are part of the M&V environment in South Africa but not an exhaustive account of all the EE measures implemented. These measures are widely implemented under the DSM program [5] as part of the country's strategy to reduce energy consumption [4].

2.5 UNCERTAINTY IN M&V

Any statement of reported savings by the M&V process includes some degree of uncertainty since no measurement, model or sample can be 100% accurate [7]. Stating the uncertainty in measured savings lends the savings report more credibility, and it provides the financial backers of a energy conserving measure (ECM) project with more confidence in decision making [20]. In general, the level of uncertainty in any given project is proportional to the complexity of the ECM. The aim of M&V projects should be to limit the uncertainty in the reported savings. This can be achieved by minimising the errors in measurement, and modelling processes [7].

Uncertainties can be grouped into quantifiable uncertainties, and unquantifiable uncertainties. When IPMVP Option A is used, inexact estimates of parameter values, and inadequate positioning of meters can lead to unquantifiable uncertainties. Furthermore, poor estimation of the interactive effects inherent in some IPMVP Option A or Option B projects can also results in unquantifiable uncertainties. There are three types of quantifiable uncertainty; namely, measurement uncertainty, modelling uncertainty and sampling uncertainty [6].

2.5.1 Measurement uncertainty

Measurement uncertainty results from instrumentation error caused by poor measurement equipment calibration, data tracking errors, and human error in data capturing. Measurement uncertainty is unavoidable although it can be mitigated with proper data handling protocols and the use of high accuracy, calibrated metering equipment such as Class 1, Class 2, and Class 3 meters that have precision of +/-1.5%, +/-2.5%, and +/-4% respectively [31, 32, 33, 34]. For three phase energy metering, it is recommended that Class 0.5 meters be used for active power, and Class 2 meters for reactive power for M&V in South Africa [35].

Input data uncertainty has been tackled in [36], where Gaussian Process Modelling and a Monte Carlo Expectation Maximisation (GP-MCEM) framework is used to develop baseline energy models that take measurement uncertainty into account. These models have the benefit of reducing M&V cost by reducing the amount of M&V data [36]. Additionally, popular guidelines for evaluating uncertainty in measured data are provided in the Guide to the Expression of Uncertainty in Measurement (GUM) [31]. The general approach to handling measurement uncertainty advocated in the GUM is describing a measurement using a measurement model in the form of a functional relationship between input and output quantities such as current, and voltage and resistance [37].

2.5.2 Modeling uncertainty

Calculating energy savings includes comparing actual energy consumption at the post-implementation phase of a project to the modelled baseline energy consumption. But baseline energy consumption during the post-implementation period cannot be measured [24, 6]. Therefore, a baseline model needs to be established to calculate the energy that would have been consumed in the post-implementation phase, which leads to modelling uncertainty [6]. Modelling techniques are widely used to characterise the relationship between energy consumption, and a number of energy driving factors such as temperature, production, facility occupancy rate, etc. These techniques include linear regression [13], support vector machines [38], Gaussian Process Modelling [36], cross-validation [39], and the use of neural networks [40]. Modelling uncertainty is unavoidable when using the whole facility and the calibrated simulation approach but not a factor when using IPMVP Options A and B because all the factors relevant to measuring energy consumption (namely power) are fully measured. However, modelling uncertainty can be mitigated by ensuring right function form is used for models, and key parameters are included in the models [6].

A number of existing M&V research articles typically focus on baseline model development. In [10], linear regression modelling is used to develop a baseline adjustment model for a conveyor belt retrofit project. The model accuracy is judged using mean bias error (MBE) and the coefficient of variation of the root mean square error (CVRMSE). In the same manner, baseline modelling for building retrofits carried out using the whole facility approach employs multivariate linear regression with key energy governing factors, such as the outdoor dry bulb temperature [13, 12]. Baseline modelling has also been done using the calibrated simulation approach, which is IPMVP option D. This approach uses energy consumption modelling software such as 'Quest' to model and simulate building energy consumption [11].

In [41], simple linear regression is extended using cross-validation so that the amount of uncertainty in the baseline model can be better estimated. The approach is also used in deciding how much data is needed for baseline estimation. Cross-validation is also used in

[39] for the M&V of a whole building using IPMVP Option C. It is shown that normalised root-mean-square-error (RMSE), and median absolute relative total error are critical to the consideration of modelling uncertainty in determining energy savings [39].

Other regression based modelling approaches include support vector machines (SVM), which has been used to forecast building energy consumption in a tropical region [38], and ordinary least squares (OLS) used in [42] to evaluate empirically-based energy consumption models for centrifugal water chillers.

Unlike prevailing modelling approaches that focus on uncertainty in the baseline model, [43] proposes that baseline models should be evaluated according the ratio of the expected uncertainty in the savings against the total savings. The study focuses on handling modelling uncertainty on project savings rather than the baseline model itself since baseline development is not the ultimate interest of M&V [43].

2.5.3 Sampling uncertainty

Practically, there are projects with large quantities of EE devices spread over large geographical areas, such as large scale lighting retrofit [15], solar water heater roll out [44], and residential rebate programs [45]. Due to budgetary constraints not all devices can be metered, and so sampling is used, which introduces sampling uncertainty into the reported savings [46]. Sampling uncertainty is avoidable for a small population project where meters are applied to all involved EE units, or when the whole facility IPMVP option is used where the energy usage of an entire facility is captured by a single measurement point [6]. Sampling uncertainty is also mitigated through the use of a sufficient sample size and appropriate sampling methods. These methods include simple random sampling, stratified random sampling, and cluster sampling to name a few [47, 48].

In order to satisfy the 90/10 criterion for CDM projects with minimal cost, an optimal sampling plan is developed in [15] that balances the sampling uncertainties across different lighting groups with different level of uncertainties. A further study [16] proposes improvements to longitudinal CDM sampling designs for lighting retrofit projects. Both ensure cost savings by ensuring an optimal number of samples is used to handle sampling uncertainty.

2.6 M&V COST

Achieving a higher level of M&V accuracy by reducing M&V uncertainties usually implies greater cost [20]. However M&V budgets tend to be limited. IPMVP and FEMP recommend that M&V cost does not exceed 10% of the average annual savings being assessed [6, 7]. Other guidelines give cost limits based on the IPMVP M&V option being used. Among these the measurement and verification handbook developed by the Systems Engineering and Management Corporation recommends that the M&V cost be limited to 5% for IPMVP Option A and to 10% for Option B and C of the average annual savings being assessed[8]. The California Evaluation Framework goes further in breaking down M&V cost according to the IPMVP option selected. It recommends a cost limit of 1% - 3% for Option A, 3% - 5% for Option B, 1% - 10% for Option C and 3% - 10% for Option D [9]. Typically M&V costs increase as the complexity of the ECM increases [7]. Therefore, researchers, M&V practitioners, and energy efficiency project participants are eager to find cost-effective solutions to handle M&V uncertainties.

According to the IPMVP, the cost of measuring and verifying energy savings depends on many factors as given in [6]. Among those factors are certain key factors that are related to uncertainty mitigation. These are the following [6]:

- The IPMVP M&V option selected; typically this is decided by the nature and complexity of the EE project;
- The quantity and complexity of meters and other measurement equipment;
- The sample size of EE devices to be measured;
- The level of engineering skill needed to develop the M&V plan;
- The complexity of baseline models used to describe the energy performance of a system;
- M&V accuracy requirement.

The above-mentioned factors contribute to the cost of meter procurement, installation, data collection, and the M&V labour cost associated with the level of M&V practitioners. These costs can be grouped into sampling cost and modelling cost. The sampling cost includes the cost of meter procurement and installation. The meter procurement cost depends on detailed meter specifications. The modelling cost includes the cost of M&V expertise in estimation and handling of uncertainty that is related to the number of independent variables, the complexity of these variables, the complexity of the system being modelled, and the required confidence and precision.

2.7 SAMPLING IN M&V

According to the Eskom M&V guidelines [5], established statistical methods for sampling can be employed to limit the quantity of measurement while meeting predetermined accuracy levels, and encouraging project sponsorship [5].

It is seldom necessary to consider an entire population in order to make a some fairly strong statistical inferences about it. It is possible to make inferences using just a random sample [49]. According to the central limit theory, regardless of the shape of the distribution of the population, the shape of the sample distribution of the mean approximates a normal distribution with sample mean $\bar{y}$ and sample standard deviation $s/\sqrt{n}$ where the sample size n is sufficiently large. Therefore the sample mean is an unbiased estimate for the population mean μ. This allows for the use of sample data to make statistical inferences [49]. Most M&V projects assume a normal distribution [6], and for normally distributed populations, any size sample n is considered sufficiently large [50].

There are numerous sampling methods used in M&V as mentioned in [48]. The following subsections will expand on these methods.

2.7.1 Simple random sampling

Simple random sampling is the most commonly used method of sampling [15]. It is equivalent to “drawing n names from a hat". The defining feature is that the final sample could be any set of n distinct names, and all such sets are equally likely [46].

Because simple random sampling is the simplest way of sampling, it provides the basis for the development of sampling theory. For this reason it is used as a basis for comparison with other sampling methods [51]. It is most suitable for use when sampling a relatively large, homogenous population of elements [51] and when more efficient sampling techniques are not viable[48]. Furthermore, because of its simplicity, simple random sampling is practical when there is limited information about the population to be sampled and also in situations where data collection can be performed efficiently [48].

However, because of its methodology, simple random sampling has some inherent disadvantages. These disadvantages are as follows [51]:

- Because simple random sampling requires that all elements be identified and labeled before sampling, it becomes expensive and impractical where large populations of elements are involved,
- Since each sample has an equal probability of being chosen, simple random sampling may result in samples spread over a large geographical area, which again becomes impractical because of cost,
- Simple random sampling is not useful in situations where it is necessary to focus on subgroups of elements within a population.

Its simplicity means that simple random sampling is often chosen as the method for sampling within M&V projects [15]. However, as mentioned above, when the population to be sampled is large, non-homogenous and spread over a large geographical location it makes simple random sampling impractical, and therefore, other methods of sampling need to be considered.

2.7.2 Stratified random sampling

As mentioned in the previous subsection, simple random sampling is not suitable when the population is non-homogenous, and it is necessary to focus on subgroups of elements within a population. A sampling approach that can be used in those instances is stratified random sampling. Stratified random sampling is suitable in situations where there is significant variation within elements of the population to be sampled but not within sub-groups of that population. It is then necessary to group elements into relatively homogenous subgroups called ‘strata’ [48]. Each stratum can then be sampled more efficiently using the simple random sampling approach.

For stratified random sampling to be used the population must be divided into strata that are mutually exclusive and exhaustive [51]. These two terms can be defined as follows [48]:

- Mutually exclusive means that every element in the population should be put into only one group,
- Exhaustive means that no population element can be excluded from stratification.

Because population elements grouped into strata are more homogenous than when the entire population is considered, samples taken from the strata will have less variation and consequently lead to more accurate estimates being made [48]. Where the population density varies across a region, stratification can ensure that accurate estimates can be made with equal accuracy for different sub-regions [48].

2.7.3 Cluster sampling

Cluster sampling is unlike simple random sampling and stratified random sampling where the sampling is done on the population elements. Instead the population elements are divided up into more 'natural' sub-groupings called 'clusters' and those clusters are then sampled instead of the individual elements within the clusters . An example of this is grouping CFL lights according to different geographical areas and sampling those geographical areas [48].

The following are key features of cluster sampling [51]:

- The process of selecting elements for clustering might be stepwise. For example, having city blocks as clusters and households as units within the clusters might involve first selecting a sample of city blocks then selecting a sample of households from the selected city blocks. These sampling steps are referred to as 'stages' and a sampling plan might involve many stages,
- Different sampling techniques such as simple random sampling or stratified random sampling can be used to select clusters,
- More than one sampling frame might be involved in the process of cluster sampling. An example can be a list of countries at the first stage of sampling, followed by a list

of townships within those counties at the second stage, then a list of schools in each township selected, etc,

- After the first stage of sampling, only those clusters chosen are used to compile the sampling frame.

The main advantages of cluster sampling are its feasibility and its cost-effectiveness depending on the spread of the population to be sampled [51]. A case in point is when elements to be sampled are spread over a wide area such as a continent, the only feasible and cost-effective solution will be to use cluster sampling [48].

The key disadvantage of cluster sampling is that the standard errors obtained from it tend to be larger than those obtained from other sampling methods. This is because units within each cluster are usually homogenous with respect to multiple criteria. Therefore, selecting multiple units from a cluster, as is typical in cluster sampling, is redundant. The result of this redundancy is high standard errors of estimates [51].

2.7.4 Multi-stage sampling

Multistage sampling is a complex form of cluster sampling, which involves the use of multiple stages of cluster sampling to achieve greater efficiency in the sampling [51]. In some situations the clusters selected in the first stage of sampling might be too large that it becomes prohibitively expensive to sample each element in each cluster. Furthermore, as mentioned in the previous subsection, the units within each cluster might be too homogenous, which leads to redundancy [51]. To reduce the sampling cost and reduce redundancy, it might be better to sample the units within the clusters as another stage of sampling. In effect, data will only be collected for the second stage, or other proceeding stages of sampling other than the original clusters after the first stage of sampling [48].

2.7.5 Sample accuracy (confidence and precision)

The IPMVP recommends that sampling should be done in a "statistically meaningful way". To this end it cites confidence and accuracy criteria such as the 90/10 criteria for CDM or the common 80/20 criteria as recommended in [52], [6].

The accuracy of the estimate involves constructing a confidence interval within which one is sufficiently sure that the true population value lies or equivalently, placing a bound on the probable error of the estimate [49]. According to [49], for any chosen sample, the confidence interval z is described as

$$z = 100(1 - \alpha), \tag{2.1}$$

where α is the chosen precision, say 10% or 5% which would give a confidence of 90% and 95% respectively [49]. An interpretation of the above equation is as follows. Say if the confidence is given as 90% under simple random sampling, then for 90% of the possible samples of size n, the interval covers the true value of the population mean μ [49]. Note that the confidence interval cannot be given without an implied precision value [6].

2.7.6 Sample size determination

The point of sample size determination (SSD) is to make sure a representative sample of EE measures is taken without unnecessary cost. Sample size determination methods have been grouped into two, namely; the frequentist methods, and the Bayesian methods [47].

2.7.6.1 Frequentist methods

Frequentist methods of SSD deal with problems that have a normal distribution with a known variance. The sample size can then be determined based on a given confidence and precision interval [47].

If the confidence interval is z and the precision p then the initial sample size for simple random sampling is calculated as [49]

$$n_0 = \frac{z^2 CV^2}{p^2} \tag{2.2}$$

Where the CV is the coefficient of variation defined as $CV = s/\mu$, that is the standard deviation s divided by the mean μ.

If the overall population N is not large relative to the sample size n, then a finite correction factor needs to be applied to the above equation. Which is given as [49]

$$n = \frac{n_0 N}{n_0 + N} = \frac{CV^2 z^2 N}{CV^2 z^2 + N p^2}, \tag{2.3}$$

2.7.6.2 Bayesian methods

Bayesian methods on the other hand have been developed from the realization that it is often impossible to take more than one sample and that a simple sample may not be sufficiently large to allow clear cut inferences to be made. Bayesian methods are said to employ the optimizing of utility functions to find the right sample size. A key advantage of Bayesian sampling methods over the frequentist methods is that they can allow for uncertainty that is inherent in any estimate [47].

2.8 MODELLING

Energy savings in EE projects cannot be measured directly. To measure them it is necessary to compare a modelled baseline energy consumption profile to a measured post implementation energy profile under similar conditions. However at the post implementation phase the baseline profile cannot be obtained. For this reason, baseline service level adjustment models have to be developed to bring both time periods under the same set of conditions [6].

M&V baseline adjustment models typically compare energy consumption of EE measures with criteria such as ambient temperature, quantity or rate of production or occupancy rates. Regression analysis is typically used to develop these models [6] although Bayesian modelling, and Gaussian modelling have also been used [43], [53].

In regression modelling several metrics are recommended by the IPMVP for quantifying model quality. They are the coefficient of determination (R^2), the coefficient of variation of the root mean square error (CVRMSE) and the coefficient of variation of the standard deviation (CVSTD), which is used for mean models [43].

The coefficient of determination found in most statistical texts is commonly used in engineering to describe the strength of the relationship between a dependent variable and a set of independent predictor variables. A R^2 value closer to unity indicates a stronger relationship and a value closer to 0 indicates a poor relationship [54]. For linear models R^2, the CVRMSE, and CVSTD are given by the following equations [43]:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}, \tag{2.4}$$

$$CVRMSE = \frac{1}{\bar{y}}\sqrt{[\frac{\sum(y_i - \hat{y}_i)^2}{n-p}]}, \tag{2.5}$$

$$CVSTD = \frac{1}{\bar{y}}\sqrt{[\frac{\sum(y_i - \bar{y})^2}{n-p}]}, \tag{2.6}$$

where:

y_i	The actual value i,
$\bar{y}$	The mean value,
$\hat{y}$	The predicted value,
n	The number of observations,
p	The number of parameters in the regression model.

R^2 is bounded between 0 and 1, while the CVRMSE and the CVSTD are both bounded between 0, and infinity. For a model to be considered a good model, the CVRMSE or the CVSTD must be closer to 0 and R^2 closer to 1. R^2 represents that variation of the actual data y_i about the modelled value $\hat{y}$ compared to the variation of the actual data about the average value $\bar{y}$. However, the CVRMSE or CVSTD are normalized differently to the R^2 value. They represent the variation of the actual data y_i about the predicted value $\hat{y}$, in case of the CVRMSE or the variation of the actual data about the mean value $\bar{y}$, in case of the CVSTD, all normalised by the data mean. This difference dictates the suitability of each metric as goodness-of-fit measures. R^2 tends to be dependent on the slope of the data while the CV depends on the spread of the data. Consequently if the aim of a M&V study is to evaluate how well a baseline energy model captures the variation in data then R^2 is the useful metric, while if the aim is to obtain the uncertainty in the savings prediction then the CVRMSE is more useful [43].

2.9 RATIONAL FOR THIS STUDY

It is evident that a number of existing M&V studies focus on handling M&V modelling uncertainty, especially during the process of baseline model development. There are also studies on dealing with sampling uncertainty, and guidelines to reduce measurement uncertainty. However, no specific study has considered a combined uncertainty analysis among the M&V measurement, sampling, and modelling uncertainties towards an optimal M&V plan. Combined analysis would involve examining measurement and sampling uncertainties, measurement and modelling uncertainties, sampling and modelling uncertainties, or a combination of all three types of uncertainties, which often exist in M&V practice.

This work aims to examine sampling and modelling uncertainties together whilst minimising M&V cost, which offers M&V practitioners flexibility in designing an optimal M&V plan that either has high sampling uncertainty with low modelling uncertainty or high modelling uncertainty with low sampling uncertainty. This study pays less attention to measurement uncertainty as it is commented that measurement uncertainty makes a negligible contribution to the overall uncertainty for electricity metering cases where population variance is not unusually low [55]. However, efforts on dealing with sampling and modelling uncertainties are believed to be the most significant contributors to the entire M&V cost, especially when both the modelling and sampling techniques are used during the M&V process.

2.10 CONTRIBUTION OF THIS STUDY

This study proposes an M&V cost minimisation model to deal with both the M&V modeling and sampling uncertainties with minimal cost. In order to find the optimal solutions in terms of the modelling accuracy level and the sample size, the M&V cost that includes the modelling cost, sampling cost, and overhead cost is formulated as the objective function, in which the modelling cost is developed as a function of the modelling accuracy in terms of CV(RMSE) and the sampling cost is directly related to the sample size.

In order to illustrate the effectiveness of the proposed model, an optimal M&V modelling and sampling strategy is designed for a traffic lighting retrofit project. In addition, partially optimal M&V plans designed with optimal sampling but non-optimal modelling solutions, or with optimal modelling but non-optimal sampling solutions are employed as the benchmark.

To test the applicability and flexibility of the proposed model for the cost-effective design of similar traffic lighting projects, simulations have been carried out to evaluate the model performance when applying the model to M&V projects with different characteristics.

The major contributions of this study can be highlighted as follows: 1) an M&V modelling cost model is developed, which is able to offer quantitative analysis of the M&V baseline model uncertainty and cost; 2) an M&V cost minimisation model is proposed to handle both the M&V modelling and sampling uncertainties cost-effectively. The effectiveness and flexibility of this model are demonstrated by a case study and the simulation results.

CHAPTER 3

M&V COST MINIMISATION MODEL DEVELOPMENT

3.1 CHAPTER OVERVIEW

In this section, a M&V cost minimisation model is developed to design optimal M&V plans that handle both the M&V modelling and sampling uncertainties cost-effectively. For this purpose, typical M&V cost factors are reviewed, and a model is developed to characterise the relationship between the M&V modelling cost, and accuracy. In addition, the formulation of combined M&V uncertainties including both the modelling, and sampling uncertainties are introduced under different practical scenarios. Based on the M&V cost, and modelling analysis, a general M&V cost minimisation model is developed, in which the sum of sampling, and modelling costs are considered as the objective function, and the M&V accuracy requirements are formulated as the constraints of the optimisation model.

3.1.1 M&V modelling cost analysis

There are numerous factors affecting M&V cost as mentioned in various guidelines, and protocol documents [6, 7]. Most of these M&V cost factors are related to the handling of M&V uncertainties; namely, the M&V measurement, sampling, and modelling uncertainties, in order to achieve the desired M&V accuracy. Thus, the M&V project cost can also be categorised into the metering cost, sampling cost, modelling cost, and the overhead cost. The metering cost normally includes the meter procurement, calibration, installation, and commissioning cost. The amount of the metering budget is decided by the required meter

Table 3.1: CVRMSE values and the estimated modelling cost

CVRMSE (%)	Modelling Cost (R)
5	223 750
10	155 500
15	130 500
20	103 000
25	92 000
30	74 750

device specifications, and the service level of calibration, installation from different suppliers [56]. Sampling cost is directly related to the sample size; a greater sample size implies a higher sampling cost. The modelling cost includes the level of M&V professionals involved in improving the model accuracy. The modelling cost also depends on the modelling techniques to be used, and the complexity of the system being modelled. The overhead cost should be a constant including reporting, communication, documentation, and management.

The metering cost can be obtained from various measurement service suppliers and the sampling cost is proportional to the required sample size. However, the modelling cost cannot be easily estimated as the required expertise from M&V professionals, and the modelling complexities vary in different projects. In this study, the quoted and approved modelling costs for more than 300 M&V projects from a South African National Accreditation System (SANAS) accredited M&V inspection body under the South African National EEDSM programme are assessed and analysed [23]. Although the M&V modelling cost is complicated, and different for different projects, the modelling accuracy plays an important role in deciding the M&V modelling cost. In this study, the CV(RMSE) values are taken as the key indicator to evaluate the modelling accuracy. Models with a lower CV(RMSE) indicate a higher M&V modelling accuracy.

To obtain a modelling cost model, regression analysis is applied to the data in Table 3.1. Linear regression being a common form of regression analysis is compared to exponential regression to fit the data. The comparisons are made based on the coefficient of determination R^2 and the $CVRMSE$. A high R^2 value and a low $CVRMSE$ value indicate a model with a good fit and a high accuracy. Table 3.2 shows the results of the regression analysis on the M&V cost data. Figure 3.1 shows the results of linear fitting and exponential fitting.

Table 3.2: Modelling cost fitting data

Fitting criteria	Linear model	Exponential model
R^2	0.903	0.975
$CVRMSE$	0.130	0.082

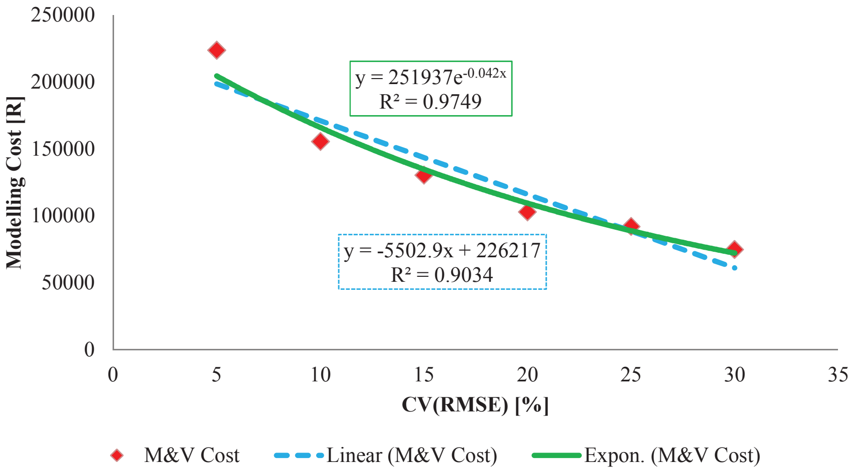

Figure 3.1: M&V modeling cost vs CV(RMSE)

It is found that the exponential M&V modelling cost model has a coefficient of determination R^2=0.975 with $CV(RMSE)$ = 0.082 while the linear modelling cost model has R^2=0.903 and $CV(RMSE)$=0.130. Therefore, the exponential model is believed to be more accurate to represent M&V modelling cost against the CV(RMSE). The M&V modelling cost to be used for the M&V cost minimisation is given in Equation (3.1).

$$C_m = 251937e^{-0.042CV_m}, \tag{3.1}$$

where CV_m denotes the CV(RMSE) of an M&V baseline model while C_m denotes the M&V modelling cost for one baseline model.

3.1.2 M&V sampling cost

In addition to the M&V modelling cost, the M&V metering, sampling and the overhead cost are given as follows. Let a_i denote the procurement and b_i represent the installation, calibration, and commission cost for the ith type of metering device to be used for M&V, n_i

denote the sample size of the ith type of metering device. Then the metering and sampling cost can be denoted by Equation (3.2) [15],

$$C_s = (a_i + b_i)n_i, \tag{3.2}$$

where:

C_s	The metering and sample cost of a M&V project.
a_i	The procurement cost per meter
b_i	The installation cost per meter
n_i	The sample size for which meters are installed
i	M&V project.

Let C_o denote the overhead cost of an M&V project, then the overall M&V project cost is a combination of the modeling cost, metering and sampling cost, and the overhead cost.

3.2 COMBINED M&V UNCERTAINTY ANALYSIS

In this study, both the modelling and sampling uncertainties will be considered and handled in designing a cost-effective M&V plan. As introduced in the IPMVP, the uncertainties in M&V can be combined in either an additive or multiplicative way provided that they are independent. Independence means that the random errors affecting one uncertainty component do not affect the other [6]. In order to quantify both the sampling and modelling uncertainty together, the combined uncertainty U is calculated by

$$U = \sqrt{Um^2 + Us^2}, \tag{3.3}$$

where Um and Us denotes the combined modelling and sampling uncertainties, respectively.

Practically, different combinations of uncertainties may exist in various M&V projects. As illustrated in Figure 3.2, energy efficiency activities may range from simple EE lighting replacement actions to a holistic EE strategy that improves all lighting, water heating, and space heating/cooling systems. For the M&V process of various EE projects, when measurement uncertainties are negligible, the combined sampling and modelling uncertainties can be categorised into different scenarios as follows:

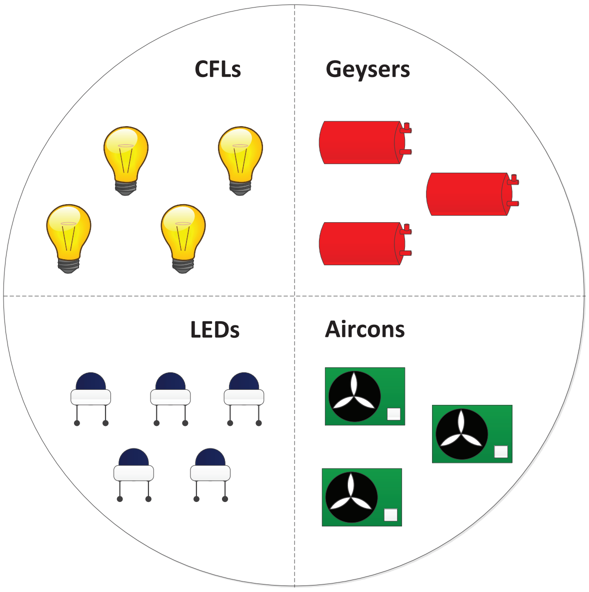

Figure 3.2: Illustration of combined uncertainty analysis.

1) **Combined sampling uncertainty analysis**. Given a large-scale lighting retrofit project that includes both CFL and LED technologies, modelling uncertainty is negligible as the lighting energy consumption is well characterised by the product of wattage and usage time. However, sampling uncertainty needs to be handled for M&V in this case. As introduced in [15], sampling uncertainties can be cost-effectively handled by applying a stratified sampling approach. To apply the optimisation approach in [15], the lighting population is firstly stratified into different strata by the CV of the energy usage of individual lamps. The optimal sample size can then be assigned to each stratum for sampling. In this scenario, combined sampling analysis will be applied across different lighting strata.

The sampling uncertainty in the ith lighting stratum is described as by the sample standard error and is defined as

$$Us_i = \frac{cv_i \bar{Y}_i}{\sqrt{n_i}}, \tag{3.4}$$

where:

cv_i the sampling CV of the ith stratum,

n_i the required sample size in the ith lighting stratum,

$\bar{Y}_i$ the sample mean.

The combined sampling uncertainty across each lighting stratum is expressed as

$$U_s = \sqrt{\sum_{i=1}^{I} \left(\frac{N_i}{N}\right)^2 \cdot \frac{(cv_i \bar{Y}_i)^2}{n_i}}, \tag{3.5}$$

where:

I the total number of lighting strata,

N_i the population size of the ith stratum,

N the combined project population.

2) **Combined modelling uncertainty analysis.** Given a small scale energy conservation project that aims to improve the energy efficiency of several water heaters, and air conditioners, sampling uncertainties do not exist when each device is measured for M&V. However, energy baseline models need to be established to adjust the baseline under post-retrofit conditions for savings determination. For projects with both water heaters, and air conditioners involved, the baseline model may be designed as a function that characterises the relationship between the total energy usage between the heating or cooling degree days over the reporting period [12]. Alternatively, separate models can be designed for both the water heaters, and the air conditioners in order to improve the modelling accuracy. When one baseline model is developed, no combined modelling uncertainty analysis is required. But when two or more models are applied in one M&V project, the combined modelling uncertainty must be performed to evaluate the total modelling uncertainty.

Given an M&V project with J baseline models, uncertainty for each model is formulated as

$$Um_j = CVm_j \bar{Y}_j, \tag{3.6}$$

where $\bar{Y}_j$ is the average baseline energy consumption, and CVm_j is the CV(RMSE) of the jth model. The combined modelling uncertainty is given by

$$Um = \sqrt{\sum_{j=1}^{J} \left(\frac{N_i}{N}\right)^2 \cdot Um_j^2}. \tag{3.7}$$

where:

J	is the total number of baseline models,
N_i	the population size of the ith stratum,
N	the combined project population.

3) **Combined modelling and sampling uncertainty analysis.** In some cases, an energy conservation project could have various types of technologies or devices involved; in this instance, both the modelling and sampling efforts need to be made for M&V. In this case, combined sampling uncertainty includes the sampling uncertainties across all the strata while the combined modelling uncertainty includes the modelling uncertainties of each individual baseline model. The total uncertainty includes both the combined sampling, and modelling uncertainties, which is calculated by Eq. (3.3).

3.3 M&V OPTIMISATION MODEL

In this subsection, an M&V optimisation model is developed to handle both the M&V sampling, and modelling uncertainties cost-effectively. The aim of the optimisation is to achieve the desired M&V accuracy with minimal M&V cost. As introduced in Subsection 3.1.1, the M&V cost includes metering, and sampling cost, modelling cost, and overhead cost. The M&V accuracy is defined in terms of the combined modelling and sampling accuracy, which is set to meet the 90/10 criterion in this study. As introduced in [6], the relationship between the 90/10 criterion accuracy, and the combined sampling and modelling uncertainty is characterised by

$$p = \frac{z \times U}{\bar{Y}}, \tag{3.8}$$

where p is the relative precision and z is the z score related to a confidence level [6].

Let an M&V project have I sampling strata and J models for the baseline adjustment, it is expected to find the optimal sample size n_i in each sampling stratum, and the optimal accuracy level CVm_i for each baseline model that achieve the desired M&V accuracy with minimal M&V cost. This is an optimisation problem that aims to find the optimal solutions

λ=$(CVm_1, \ldots, CVm_J, n_1, \ldots, n_I)$, which minimises the overall M&V cost $f(\lambda)$

$$f(\lambda) = \sum_{j=1}^{J} 251937e^{-0.042CVm_j} + \sum_{i=1}^{I}(a_i + b_i)n_i + C_0, \tag{3.9}$$

subject to the constraints

$$p = \frac{z \times U}{\bar{Y}} \leq 10\%, \tag{3.10}$$

where U is the total uncertainty that is calculated by the Eqs. (3.3)-(4.6).

3.4 CHAPTER SUMMARY

A M&V cost minimisation model that takes into account the M&V modelling and sampling cost has been developed. The M&V modelling cost has been formulated as a function of the CV(RMSE) by exponentially fitting practical M&V cost data to the estimated CV(RMSE) values for multiple M&V projects. Through the use of stratification on EE device populations with varying sample CV values and modelling requirements, it is shown that it is possible to have a combined uncertainty analysis for any given M&V project. The overall uncertainty is set to be less than or equal to the 90/10 confidence and precision requirements; this forms the constraint for the optimisation problem. Solutions to the optimisation problem provide CV(RMSE) values for each model in a M&V project, and sample sizes for each sampling group within a M&V project. The next chapter presents a case study used to demonstrate the effectiveness of the optimisation approach developed in this chapter.

CHAPTER 4

CASE STUDY: TRAFFIC INTERSECTION SIGNAL LAMP RETROFIT

4.1 CHAPTER OVERVIEW

In this chapter a case study is presented to illustrate the cost effective handling approach developed in this thesis. The focus is a traffic lamp (aspect) retrofit project implemented in South Africa. Background on the project is given, this includes technical details about the functioning of the intersections, how the intersections are grouped for sampling is explained, and finally a M&V cost minimisation problem is framed in the context of this case study, and solved in the next chapter. Discussions are also provided to solve the proposed M&V cost minimisation model.

4.2 TRAFFIC INTERSECTIONS PROJECT BACKGROUND

A traffic light retrofit project that replaces 56 W incandescent signal lamps with an equal number of 15 W LED signal lights has been implemented in several municipalities in South Africa. A number of 2 200 traffic intersections have been retrofitted by more than 125 000 LED signal lamps. Due to different conditions of existing traffic light systems, the lamp retrofits have been done in two ways. One solution is to replace 56W signal lamps, which include red, amber, and green coloured ones within a traffic light sets with 15W LED lamps. The other is to change the whole traffic light sets, which include 4-aspect, 3-aspect, and 2-aspect ones with new whole sets fitted with LEDs. Since this project is financially sponsored by the local government, the energy savings of this project need to be accurately quantified by an M&V process. Figures 4.1-4.3 illustrate the scope of the traffic lamp retrofits, the configuration of typical intersections and a timing diagram that corresponds to that configuration.

Figure 4.1 is an illustration of what single aspects (lamps) looks like compared to the combined aspects. The replacement of a single aspect involves removing a particular aspect (Red, Yellow and Green) and replacing it with another single LED aspect. While the combined aspect retrofit involves removing the whole lamp fitting and replacing it with a new LED fitted lamp fitting.

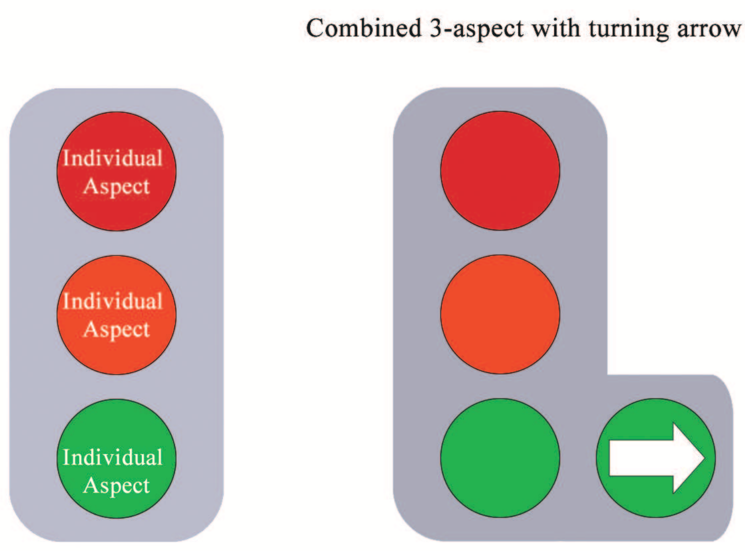

Figure 4.1: Image showing individual aspects and a combined aspects

Figure 4.2 shows a planning diagram of a traffic intersection in Pretoria, South Africa. In this diagram, signal lamps are grouped for the sake of control. A particular set of signal lamps is switched onto Red, Yellow or Green depending on the phase of operating. This is done via a timing schedule. An example of this timing schedule is shown in Figure 4.3.

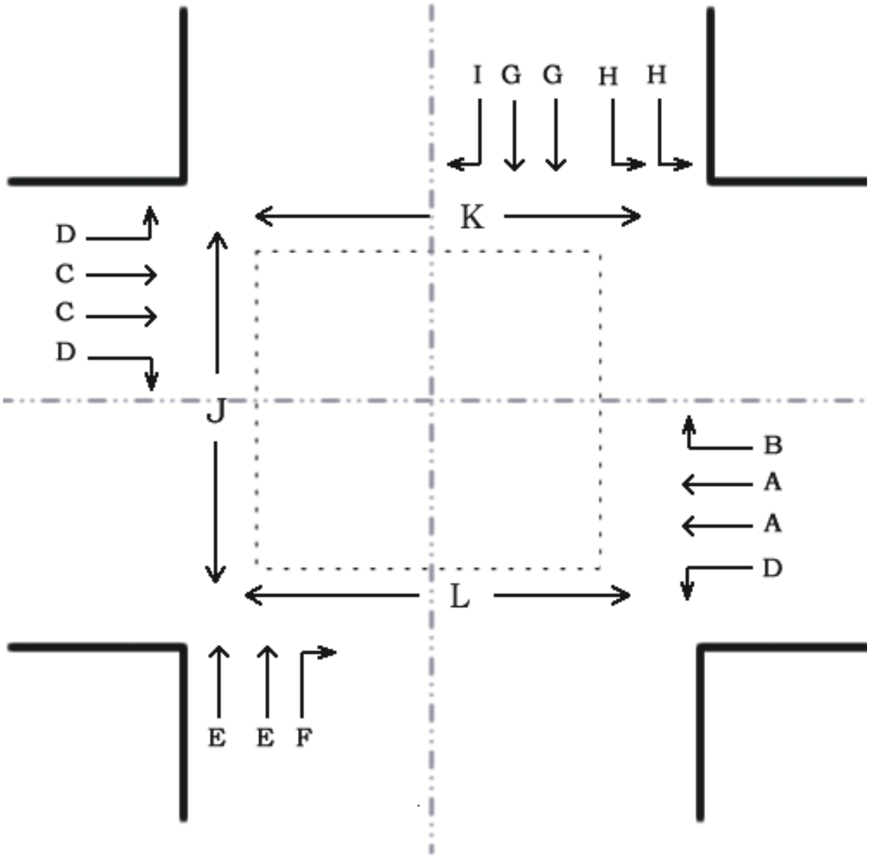

Figure 4.2: A planning image of a intersection showing signal groups

Figure 4.3 shows a timing diagram for a traffic intersection in Pretoria. In this figure the solid green line indicates when the signal lamp groups are set to a steady green state, the dashed green line shows when the signal lamp groups are set to a flashing green state. The blank spaces are all when the signal lamp groups are set to Red. A single cycle of operation lasts 85 seconds, and repeats throughout the day. Throughout all these changes in operational phase, the energy consumption vacillates since the quantity of lights varies across signal lamp groupings.

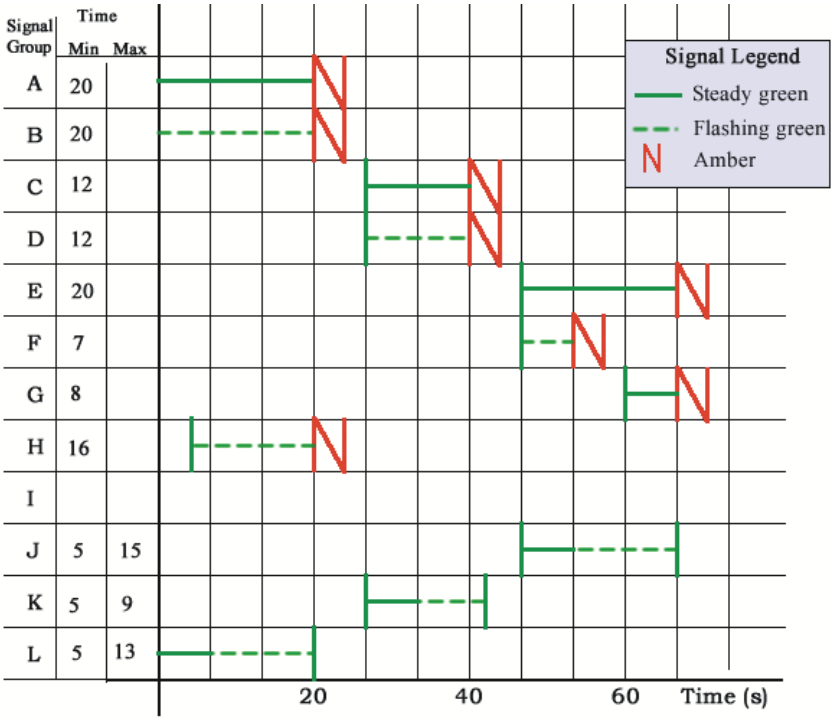

Figure 4.3: Timing diagram for the intersection in figure 4.2

4.3 INTERSECTION CLASSIFICATION

As shown in Figures 4.2-4.3 the energy consumption at any given intersections varies throughout a day of operation. This variation is dependent on the number of phases of operation, and the quantities of lamps. Different phases of operation, require different configurations of the traffic lamps to implement them, and these different configurations require varying quantities of lamps to implement. Because of this, the traffic intersections can be grouped into large intersections, which includes intersections with 3 or more phases of operations, and small intersections that includes all intersections with no more than 2 phases of operation. The large intersections have been retrofitted with complete sets of lamp fittings, while the smaller intersections have been retrofitted with individual lamp fittings as illustrated in Figure 4.1.

Given the two different light retrofit scopes, the energy consumptions of a few intersections have been recorded over a short period for both complete-set retrofit and individual retrofit traffic intersections. Based on metered data gathered during the test sampling, it is found that the average energy consumption of the intersections with individual signal lamp retrofits is 1.91 kWh with a maximum standard deviation of 0.382, while the average energy consumption of intersections with combined traffic set retrofits is 1.415 kWh with a standard deviation of 0.708. According to the installation database, there are 1320 intersections with individual signal lamps retrofitted and 880 intersections with traffic lamp set retrofits. According to the test sampling, sampling CV of the daily energy consumption of the 1320 intersections is less than 0.2 while the CV of the 880 traffic lamp set retrofit intersections is taken as 0.5. The traffic intersections are classified into two groups for stratified sampling according to the different sampling CV values of each stratum. Strata 1, henceforth known as Group I contains the complete-set retrofit intersections and strata 2 henceforth known as Group II contains all the individual retrofit intersections.

Optimal sample sizes for each strata will be decided by the model (3.9)-(4.4). The same sample size will be used at both the baseline, and post-retrofit periods. Since each stratum of will have varying levels of uncertainty in their daily energy consumption, meters with different accuracies, and prices will be installed randomly at selected intersections from each stratum.

4.4 M&V APPROACH

In order to reliably quantify the energy savings for this traffic light retrofit project, the project boundary, metering, sampling plan, baseline calculation, and baseline adjustment approaches need to be specifically designed.

The project boundary includes all the 2 200 traffic intersections, and all the LED signal lamps. For the M&V purpose, it is applicable to measure the daily energy consumption either in terms of the traffic intersections or individual LED signal lamps. In order to reduce the sampling population, which will consequently reduce the sample size, and sampling cost, it is decided that the energy consumption per traffic intersection will be measured. Therefore, the IPMVP Option C: the whole facility measurement approach is applied to this M&V case study.

4.5 BASELINE METHODOLOGY

The following section presents the M&V plan, as well as the sampling and metering plan:

1. The electricity use of each intersection is considered as independent from that of the city and of other traffic intersections concerned in the project,

2. Each traffic intersection is fed electricity through a single mains at an adjacent circuit board. It is therefore taken as a whole facility; for this reason IPMVP option C is used to M&V this project,

3. The electrical parameters relevant to this project will be the daily power consumption for each intersection, which will be measured using metering equipment,

4. Once a required sample size has been obtained, the daily average power consumption will be obtained by installing meters at randomly selected intersections,

5. The same sample size will be used at the baseline period as well as the post implementation period.

There are two traffic light baseline modelling approaches introduced in [57]. The daily energy consumption per intersection can be formulated as the quantity of each type of signal lamps in terms of different lamp colours. For instance, the daily energy consumption per intersection E_1 can be denoted by

$$E_1 = \beta_0 + \beta_1 R_n + \beta_2 Y_n + \beta_3 G_n, \tag{4.1}$$

where β_0, β_1, β_2, and β_3 are regression coefficients; R_n, Y_n, and G_n are the quantity of Red, Yellow, and Green signal lamps, respectively. Alternatively, the energy consumption for the traffic set retrofit is given as

$$E_2 = \alpha_0 + \alpha_1 A_{2n} + \alpha_2 A_{3n} + \alpha_3 A_{4n} + \alpha_4 A_{rn}. \tag{4.2}$$

where α_0, α_1, α_2, α_3 and α_4 are regression coefficients; A_{2n} and A_{3n} denote the quantity of *2-aspect* and *3-aspect* traffic light sets, respectively; A_{3n} and A_{rn} denotes the quantity of the *4-aspect* fittings with pedestrian signals, and *4-aspect* fittings with turning arrows, respectively.

The baseline energy consumption will be the aggregated energy consumption of each intersection in the project multiplied by the number of days in the baseline measurement period. The two baseline models (4.1)-(4.2) will be applied for baseline adjustments under the post-retrofit period.

4.6 OPTIMISATION MODEL FOR THE CASE STUDY

The optimisation model developed in section 3 is applied to the case study presented in this section to solve the cost minimisation and uncertainty handling problem. The initial values presented in Table 4.2. The optimisation model is presented below for the two groups of intersection retrofits.

$$f(\lambda) = 251937e^{-0.042CVm_1} + 251937e^{-0.042CVm_2} + (a_1 + b_1)n_1 + (a_2 + b_2)n_2 + C_0, \quad (4.3)$$

subject to the constraints

$$p = \frac{z \times U}{\bar{Y}} \leq 10\%, \quad (4.4)$$

where U is the total uncertainty that is calculated by the Eqs. (3.3)-(4.6).

where,

$$U = \sqrt{U_m^2 + U_s^2}, \quad (4.5)$$

$$Um = \sqrt{\sum_{j=1}^{J} \left(\frac{N_i}{N}\right)^2 \cdot Um_j^2}. \quad (4.6)$$

$$U_s = \sqrt{\sum_{i=1}^{I} \left(\frac{N_i}{N}\right)^2 \cdot \frac{(cv_i\bar{Y}_i)^2}{n_i}}, \quad (4.7)$$

$$\bar{Y} = \frac{N_1 Y_1 + N_2 Y_2}{N_1 + N_2}. \tag{4.8}$$

where:

CV_{m1}	the required modelling CVRMSE of Group I
CV_{m2}	the required modelling CVRMSE of Group II
CV_1	the sampling CV for Group I
CV_2	the sampling CV for Group II
p	the required precision (10%)
z	the z-value associated with the required confidence (90%)
n_1	the required sample size of Group I
n_2	the required sample size of Group II
N	the overall project population
U_m	the overall modelling uncertainty
U_s	the overall sampling uncertainty
U	the overall project uncertainty

To obtain optimal sample sizes and model accuracies, the cost minimisation model in Eqs. (3.9)-(4.4) is solved using the case study specific information given in Table 4.2. The optimisation problem in Eqs. (3.9)-(4.4) is a non-linear problem and it is solved using “MATLAB" simulation software, specifically the *fmincon* optimisation function. The following settings are employed for the optimisation function are given in Table 4.1. They are the tolerance on the function value, *tolfun*, the tolerance on the constraints, *tolcon*, and the termination tolerance on the design variables, and *tolx*.

Table 4.1: The optimisation settings

Parameter	Value
tolfun	10^{-45}
tolcon	10^{-45}
tolx	10^{-45}

The optimal sample sizes are integers which are obtained using integer programming algorithms. The topic of this thesis deals with the practical problem of minimising M&V project cost, therefore, integer sample sizes are obtained from the optimisation. Once the optimal solution has been found, the sample sizes are rounded up using the MATLAB 'ciel' function. Mathematically, the sample sizes are sub-optimal solutions. The starting point of the optimisation is arbitrarily chosen as $\lambda_0 = (0.3,0.10,20,50)$. With lower bounds $lb = (0,0,0,0)$ and upper bounds $ub = (1, 1, \infty, \infty)$. The optimal solutions to the case study are given in Table 5.1.

In Table 4.2, the meter unit price for *Group I* is R500 and R1500 for *Group II*. The disparity in meter price is due to the fact that group two has more uncertainty associated with it, reflected in its higher CV of sampling value. Therefore, it requires much more sophisticated metering equipment, which is more expensive. The same is reflected in the meter installation cost. The simulations are done at 90% confidence and 10% precision; the same criteria required by the CDM methodology.

The rest of the initial values for the optimisation are drawn from field data and calculations such as those for the sampling CV value and the estimated sample means $\bar{Y}_i$.

Table 4.2: Initial values for the case study.

Parameter	Group I	Group II
Meter unit purchase price	a_1 = R 500	a_2 = R 1500
Installation cost per meter	b_1 = R 195	b_2 = R 320
sample CV values	$cv_1 = 0.20$	$cv_2 = 0.5$
Estimated $\bar{Y}_i$	$\bar{Y}_1$ = 1.91 kWh	$\bar{Y}_2$ = 1.415 kWh
Population	$N_1 = 1320$	$N_2 = 880$

4.7 BENCHMARK FOR COMPARISON

Before solving the case study, solutions without optimisation are calculated as a benchmark for comparison. As mentioned in the chapter 2, there is no existing study that has a cost analysis in dealing with M&V modelling uncertainties. Therefore, the mathematic modelling of the relationship between M&V baseline modelling cost, and the model accuracy is one of the major contribution in providing quantitative cost analysis for the M&V baseline modelling process.

Though there is no benchmark in handling M&V modelling uncertainty with detailed cost analysis, optimal solutions have been provided in [15] to dealing with the M&V sampling uncertainties cost-effectively. In the absence of a direct benchmark to the study, it is proposed that the optimal solutions be compared with the partially optimised solutions, in order to highlight the effectiveness of the proposed M&V cost minimisation model. The partial optimal solutions (POS) are obtained by

POS1: Optimal modelling but non-optimal sampling approach. In this approach, the optimal modelling accuracy is assigned but the sample sizes are not optimised and calculated by the sample size determiantion formula as given in [49],

$$n_0 = \frac{z^2 cv^2}{p^2}.$$

POS2: Optimal sampling but non-optimal modelling approach. In this approach, the optimal sample sizes are assigned but the model accuracy is not optimised. As the ASHRAE M&V guidelines [58] recommends that the IPMVP: Option C baseline models should have a poorest CV(RMSE) of 25%, the CV(RMSE) of 25% is chosen in this approach to establish the benchmark.

The initial values in Table 4.2 are used to calculate both the optimal and partial optimal solutions. The results for POS1 and POS2 are presented in Tables 5.2-5.3.

4.8 CHAPTER SUMMARY

A traffic light retrofit project implemented across multiple municipalities in South Africa is used as a case study to illustrate the applicability of the optimisation approach developed in this thesis. Each traffic intersection is considered as a whole facility, and all the intersections are stratified into two groups based on the type of model applied to them for baseline adjustment. Because there is no historically existing M&V modelling cost model, two partially optimal solutions are proposed as a basis for comparison with the optimal solutions that will be calculated in the next chapter.

CHAPTER 5

RESULTS

5.1 CHAPTER OVERVIEW

In this chapter the optimal results to the M&V cost minimisation problem presented in the case study are given. Partial optimal solutions, which are used as a benchmark for comparison to the optimal solutions are also given. A discussion on the optimal results is given, and a comparison is carried out with the partially optimal results. Furthermore, to demonstrate the applicability of the optimal approach developed in this thesis, simulations on the sampling CV have been carried out, and those results are also presented, and discussed in this chapter.

5.2 OPTIMAL SOLUTIONS FOR THE CASE STUDY

The optimal solutions to the case study are given in Table 5.1. These solutions show that optimal model accuracies of 3.46% and 10.53% can be expected for Group I and Group II respectively. Furthermore, they show that the optimal sample sizes are 28 and 22 for Group I and Group II respectively. These are the optimal solutions necessary to meet a minimum overall M&V project cost for the traffic light retrofit project while meeting the 90/10 criteria for confidence and precision.

In Tables 5.1-5.3, the overall modelling CV(RMSE) is calculated by dividing the total modelling uncertainty U_m as given in Eq. (4.6) over the weighted average daily energy consumption $\bar{Y}$ (Eq. (4.8))

$$CV(RMSE) = \frac{U_m}{\bar{Y}}$$

Table 5.1: Optimal solutions to the case study.

Parameters	Group I	Group II	Overall
Optimal CV(RMSE)	3.458%	10.532%	4.181%
Optimal Sample Size	28	22	50
Sampling cost	R19 460	R40 040	R59 500
Modelling cost	R217 880	R161 880	R379 760
Project Cost	R237 340	R201 920	R439 260

5.3 PARTIALLY OPTIMAL SOLUTIONS TO THE CASE STUDY

The optimal model accuracies in Table 5.1 are used to generate partially optimal solutions with optimal modelling accuracy but non-optimal sample sizes. These results are in Table 5.2. These solutions reveal a greater overall M&V cost than the optimal solutions since non-optimal sample sizes are being used.

Table 5.2: Partially optimal solutions to the case study: optimal modelling

Parameters	Group I	Group II	Overall
Optimal CV(RMSE)	3.458%	10.532%	4.181%
non-optimal Sample Size	11	68	79
Sampling cost	R7 645	R123 760	R131 405
Modelling cost	R217 880	R161 880	R379 760
Project Cost	R225 525	R285 637	R511 161

Table 5.3 presents the partially optimal solutions calculated by taking the optimal sample sizes, and using the non-optimal CV(RMSE) fo 25% for both groups. These solutions reveal a decreased overall M&V cost compared to the optimal solutions in Table 5.1 but with a much poorer overall model accuracy of 18.67%. The low cost is due to the poor CV(RMSE), which entails a much lower M&V modelling cost.

Table 5.3: Partially optimal solutions to the case study: optimal sampling

Parameters	Group I	Group II	Overall
non-optimal CV(RMSE)	25%	25%	18.67%
Optimal Sample Size	28	22	50
Sampling cost	R19 460	R40 040	R59 500
Modelling cost	R88 162	R88 162	R176 325
Project Cost	R107 622	R128 202	R235 825

5.4 COMPARISON OF THE OPTIMAL SOLUTIONS TO THE PARTIALLY OPTIMAL SOLUTIONS

When comparing the results given in Tables 5.2 and 5.1, the optimal solutions reduce the sampling cost by 55% and the total M&V cost by 14% against the solutions obtained by the POS1. The results given in Table 5.3 offer a lower M&V cost than the optimal solution. However, as the model accuracy in Table 5.3 is much lower than the optimal model accuracy, the solutions in Table 5.3 cannot satisfy the required 90/10 criterion for the M&V reporting.

These results show that the optimisation approached developed in this thesis is able provide M&V practitioners a tool to handle M&V modelling, and sampling uncertainties cost-effectively. They show that it is possible to have high modelling accuracies with low sample sizes while meeting the required confidence and precision criteria for M&V baseline models.

5.5 SIMULATION RESULTS

The optimal solutions to the case study in Section 5.2 illustrate the advantageous performance of the proposed M&V cost optimisation model in designing an optimal M&V plan for a specific traffic light retrofit M&V project. In order to test the applicability, and flexibility of the proposed model for the cost-effective design of similar traffic light projects, simulations have been carried out to evaluate the model performance when applying the model to M&V projects with different characteristics. Through the simulations, it is expected that the capability of the proposed M&V cost minimisation model to offer flexible solutions will be identified, which will provide multiple optimal solutions to M&V practitioners to mitigate practical constraints. For instance, some possible solutions may require very high modelling accuracy with a very small sample size, which is not easily implementable. In this case, it is expected that more easily implementable optimal solutions that shift the modelling accuracy to the sample accuracy will be found; such that a lower modelling accuracy with greater sample sizes will be required to satisfy the M&V accuracy.

In the case study, the estimated sampling uncertainty is CV_{Ref}={0.2, 0.5}, which represents cv_1=0.2 and cv_2=0.5 in the two traffic light strata. In order to investigate the flexibility, and the model performance against different sampling uncertainties, two simulations are carried out as follows: the optimal modelling accuracy in terms of the CV(RMSE), and sample sizes are obtained by the optimisation approach with CV_{Ref}={0.2, 0.5} as a reference. In the two simulations, CV_{Ref} is changed by ±10%, ±20%, and ±50%. The settings for the optimisation are kept the same as given in the case study. The search starting point for the simulations is λ_0=(30,25,50,50). The results of these simulations are presented in the following subsections.

5.5.1 Simulations on the CV(RMSE)

In the first simulation, the sample sizes; namely, n_1=28 and n_2=22, are assigned to the two traffic light strata. When the sampling uncertainties change, the optimal accuracy levels of the baseline models are obtained and presented in Figure 5.1. It shows that when sampling uncertainty increases, more accurate models are required if the sampling efforts are limited. The consequence of this is an increasing M&V project cost shown in Fig. 5.2.

5.5.2 Simulations on the sample sizes

In the second simulation, the model accuracy; namely, CV_{m1}=3.46% and CV_{m2}=10.53% are assigned to the two traffic light strata. When the sampling uncertainties change, the optimal sample sizes are obtained and presented in Figure 5.3. It shows that when sampling uncertainty increases, more sample sizes are required if the modelling efforts are limited. The result of this is also reflected in Fig. 5.4, which shows an increasing M&V cost as the required sample sizes increase, and the model accuracies also increase to compensate for the increased sampling uncertainty.

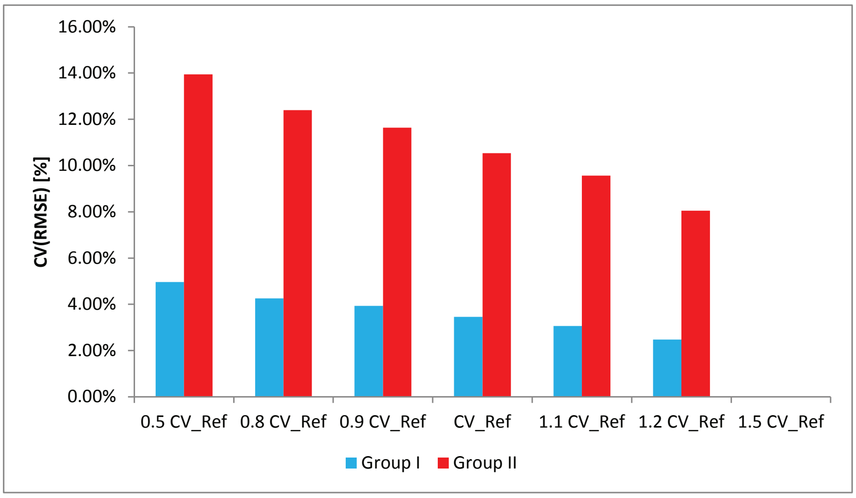

Figure 5.1: CV(RMSE) when sampling uncertainties change

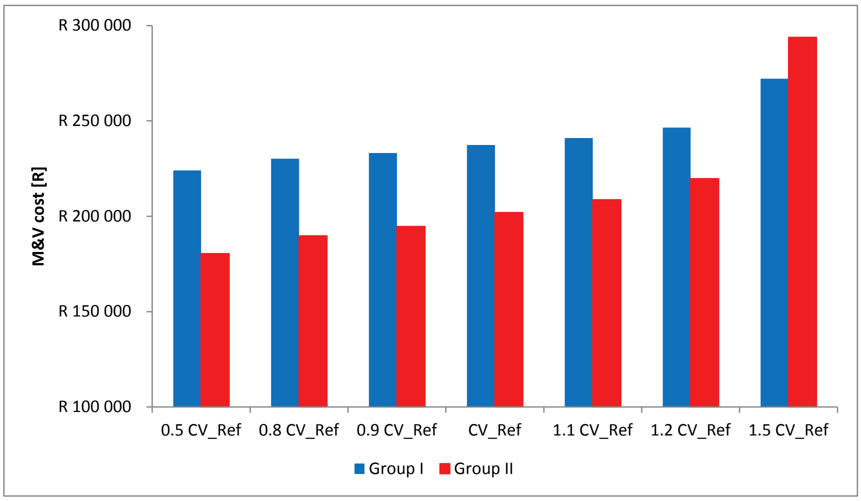

Figure 5.2: M&V cost when sampling uncertainties change

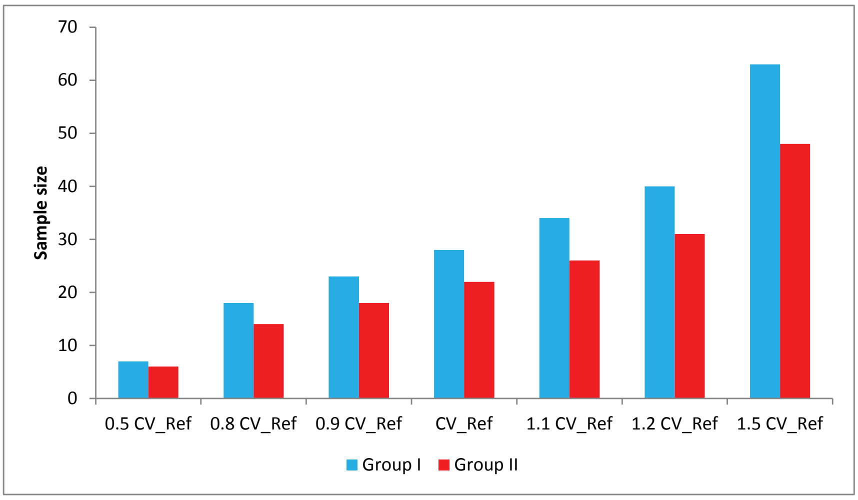

Figure 5.3: Sample size when CV(RMSE) changes.

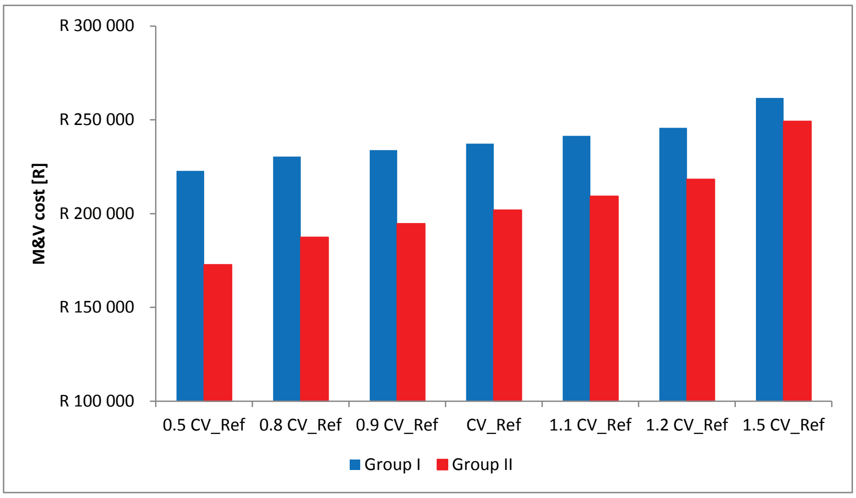

Figure 5.4: M&V cost when CV(RMSE) changes.

5.6 DISCUSSION OF THE SIMULATION RESULTS

Varying the sample CV for the same sample sizes causes the sampling uncertainty to also vary. To mitigate an increase in uncertainty caused by increasing the sampling CV, the modelling uncertainty has to drop. This means lower CV(RSME) values are proposed by the algorithm for prospective baseline models. When the sampling CV decreases, the sampling uncertainty also decreases, which means that the same accuracy criteria can be met by employing less accurate models. To satisfy this scenario the algorithm proposes higher CV(RMSE) values for the respective groups of retrofits. Because mitigating modelling uncertainty contributes significantly to M&V cost, employing less accurate models while meeting the accuracy constraints lowers the overall M&V cost.

When the required modelling accuracy is constrained, only the sampling accuracy can be changed. In this instance the simulations show that as the sampling uncertainty increases, more samples are needed to meet the overall M&V accuracy criteria. The reverse is true as the sampling uncertainty decreases; less samples are needed to meet the overall M&V uncertainty requirements for the same modelling accuracy.

From the simulation results, it is clear that a trade-off is possible between the modelling accuracy, and the sampling accuracy. The proposed optimisation model is able to provide the M&V practitioner the choice of having a more accurate baseline model with fewer sample sizes or a less accurate baseline model with greater sample sizes to achieve the same M&V accuracy requirements.

5.7 CHAPTER SUMMARY

The results of the optimisation of the M&V cost minimisation problem using the traffic light retrofit project as a case study show that CV(RMSE) values of 3.46% and 10.58% are needed for Group I and Group II respectively. They also show that sample sizes of 28 and 22 are required for the two groups in order to meet the 90/10 criteria. By comparing those results with the partially optimal results obtained by using optimal modelling with non-optimal sample sizes shows that optimal solutions reduce the sampling cost by 55% and the overall M&V cost by 14%. Simulations have been performed by varying the sampling CV of both groups of retrofits for fixed sample sizes, and fixed CV(RMSE) values respectively.

They show that for limited sampling efforts, more accurate modelling is required to meet the accuracy criteria (90/10 in this case), and that with limited modelling efforts, more samples are required to meet the same criteria. These results show that it is possible to have a trade-off between modelling and sampling uncertainties for particular M&V projects. This offers the M&V practitioner flexibility in designing cost-effective M&V plans that either have more modelling efforts or more sampling efforts.

CHAPTER 6

CONCLUSION

A cost-effective approach to handling both sampling, and modelling uncertainties in M&V has been developed. By developing a cost minimisation problem that takes into account the M&V modelling cost, and the cost of sampling, it has been shown that it is possible to have a trade-off between modelling accuracy, and sampling accuracy when either the modelling or sampling efforts are limited.

To illustrate the effectiveness of the developed model, an optimal M&V modelling, and sampling strategy has been designed for a traffic intersection lamp retrofit project. In addition, partially optimal M&V plans designed with optimal sampling but non-optimal modelling solutions, or with optimal modelling but non-optimal sampling solutions are employed as the benchmark. Comparisons between the optimal and non-optimal solutions show advantageous cost savings performance in the execution of sampling, and modelling activities for the case study. More precisely, the optimal solutions reduce the sampling cost by 55%, and the total M&V cost by 14% against the solutions obtained by optimal modelling but non-optimal sampling solutions.

A simulation analysis that evaluates the effect of sample CV, and modelling accuracy was carried out to show the applicability, and flexibility of the proposed model for the cost-effective design of similar traffic light projects. The simulation results show that the proposed model is able to offer flexible trade-offs between between the modelling and sampling uncertainties, namely; using more accurate baseline models, and fewer sample sizes or less accurate baseline models but greater sample sizes to accommodate different practical needs in executing M&V projects with different characteristics.

The simulations are done for the 90/10 criteria but it is possible to apply this approach to other popular M&V uncertainty criteria such as the 80% confidence and 20% precision criteria. And it is shown that a higher sampling CV, that leads to poor sampling accuracy will require more stringent modelling in a EE retrofit project, which leads to better modelling accuracy, and achieves the required M&V accuracy criteria.

Due to all the above, it has been shown that it is possible to apply this approach to other M&V projects. Specifically projects where the EE measures can be grouped into sub-groups of homogenous energy consumption. The cost handling approach can be used in the planning phase of M&V projects to allow flexibility in decision making about M&V cost. Possible decisions that can be made are whether to focus on more rigorous modelling or applying greater sampling.

6.1 RECOMMENDATIONS

The work done in this thesis focuses on modelling and sampling uncertainties in M&V. This is done because it is assumed measurement uncertainty can be ignored due to the low cost of measurement equipment compared to the cost of modelling, and the high precision of existing measurement equipment. However, to gain more control on the M&V process, it is possible to include measurement uncertainty in future work.

Another recommendation is that the algorithm and models developed in this masters thesis should be further developed into a user-friendly software tool to assist M&V practitioners and professionals in giving a more accurate indication of M&V cost to project developers and other clients. In South Africa, this is particularly relevant to the Integrated Demand Management (IDM) department at Eskom for its EEDSM programs. This potential tool would also be useful for the 12I and 12L tax incentive program by helping M&V teams bring down M&V cost thus further improving the financial feasibility of EE projects.

REFERENCES

[1] "World energy outlook 2014," Internation Energy Agency, https://www.iea.org/.

[2] "Global greenhouse gas emissions data," United States Environmental Protection Agency, http://www.epa.gov/climatechange/ghgemissions/global.html.

[3] "Eskom load-shedding avoids total blackout," Business Day Live, http://www.bdlive.co.za/business/energy/2015/04/16/eskom-load-shedding-avoids-total-blackout.

[4] Department of Minerals and Energy, "Energy efficiency strategy of South Africa," Technical Report, 2005.

[5] Eskom, "The measurement and verification guideline for demand-side management proejcts," Technical Report, 2011.

[6] Efficiency Valuation Organization (EVO), "International performance measurement and verification protocol: concepts and options for determining energy and water savings," Technical Report, 2012.

[7] Department of Energy (USA), "M&V guidelines: measurement and verification for federal energy projects," Technical Report, 2008.

[8] Air Force Civil Engineer Support Agency, "Energy measurement and verification reference handbook," System Engineering and Management Corporation, Technical Report, 2011.

[9] TecMarket Works, "The California evaluation framework," California Public Utilities Commission, Oregon, Technical Report, 2004.

[10] A. Z. Dalgleish and L. J. Grobler, "Measurement and verification of a motor sequencing controller on a conveyor belt," *Energy*, vol. 28, no. 9, pp. 913–927, 2003.

[11] M. Ke, C. Yeh, and J. Jian, "Analysis of building energy consumption parameters and energy savings measurement and verification by applying equest software," *Energy and Buildings*, vol. 61, pp. 100–107, 2013.

[12] B. Dong, S. E. Lee, and M. H. Sapar, "A holistic utility bill analysis method for baselining whole commercial building energy consumption in singapore," *Energy and Buildings*, vol. 37, no. 2, pp. 167–174, 2005.

[13] G. Escrivá-Escrivá, C. Roldán-Blay, and C. Álvarez Bel, "Electrical consumption forecast using actual data of building end-use decomposition," *Energy and Buildings*, vol. 82, pp. 73–81, 2014.

[14] Z. Ma, P. Cooper, D. Daly, and L. Ledo, "Existing building retrofits: Methodology and state-of-the-art," *Energy and Buildings*, vol. 55, pp. 889–902, 2012.

[15] X. Ye, X. Xia, and J. Zhang, "Optimal sampling plan for clean development mechanism energy efficiency projects," *Applied Energy*, vol. 112, pp. 1006–1015, 2013.

[16] H. Carstens, X. Xia, and X. Ye, "Improvements to longitudinal clean development mechanism sampling designs for lighting retrofit projects," *Applied Energy*, vol. 126, pp. 256–265, 2014.

[17] R. Pelzer, E. H. Mathews, D. F. L. Roux, and M. Kleingeld, "A new approach to ensure successful implementation of sustainable demand side management (dsm) in south african mines," *Energy*, vol. 33, no. 4, p. 1254 Ű 1263, 2008.

[18] F. Stern and D. Vantzis, "Protocols for evaluating energy efficiency - both sides of the atlantic," IEPEC 2014 International Energy Policy and Programme Evaluation Conference, 2004.

[19] The Regulatory Assistance Project (RAP), "Energy efficiency evaluation, measurement, and verification: A regional review of practices in China, the European Union, India, and the United States," Technical Report, 2014.

[20] J. S. Haberl, C. Culp, and D. E. Claridge, "ASHRAE's guideline 14-2002 for measurement of energy and demand savings: how to determine what was really saved by the retrofit," in *Proceedings of the Fifth International Conference for Enhanced Building Operations, Pittsburgh, Pennsylvania*, Pittsburgh, Pennsylvania, 2005, pp. 1–13.

[21] Australian Energy Performance Contracting Association, "A best practice guide to measurement and verification of energy savings," Commonwealth of Australia, Technical Report, 2004.

[22] ISO, "ISO 50015:2014 measurement and verification of energy performance of organizations - general principles and guidelines," Technical Report.

[23] South African bureau of standards, "SANS 50010:2011 measurement and verification of energy savings," Technical Report.

[24] X. Xia and J. Zhang, "Mathematical description for the measurement and verification of energy efficiency improvement," *Applied Energy*, vol. 111, no. 4, pp. 247–256, 2013.

[25] O. Dintchev, "Overview of the performance of some of the south african solar water heating programmes and related measurement and verification interventions," in *World Energy Engineering Congress, WEEC 2014*.

[26] F. Barnard, A. D. Bod, and L. Grobler, "Energy efficiency measures on mine pumping systems - an m&v methodology," in *Proceedings of the Conference on the Industrial and Commercial Use of Energy, ICUE*.

[27] A. Mavhungu, C. V. D. Merwe, and L. Grobler, "Measurement and verification of a lighting efficiency project implemented in a food and cleaning materials manufacturing plant," in *Energy Engineering: Journal of the Association of Energy Engineering*, vol. 104, no. 6, 2007, pp. 51–62.

[28] H. Brand, J. Vosloo, and G. Bolt, "Carbon dioxide pollution reduction through load shift projects," in *Proceedings of the Conference on the Industrial and Commercial Use of Energy, ICUE*.

[29] W. D. Heijer and L. Grobler, "The role of measurement and verification in eskom inte-

grated demand management's performance contracting initiative," in *2011 Proceedings of the 8th Conference on the Industrial and Commercial Use of Energy, ICUE 2011*.

[30] L. Grobler, W. D. Heijer, and B. Dalgleish, "Measurement and verification of a utility's energy management program," *Energy Engineering: Journal of the Association of Energy Engineering*, vol. 100, no. 5, pp. 40–60, 2003.

[31] JCGM, "Evaluation of measurement data – guide to the expression of uncertainty in measurement," Int. Organ. Stand., Geneva, Technical Report ISBN. 50, 2008.

[32] International Electrochemical Commission, "International Standard 62053 part 21: Static meters for reactive energy (Classes 1 and 2)," Technical Report, 2003.

[33] International Electrochemical Commission, "International Standard 62053 part 23: Static meters for reactive energy (Classes 2 and 3)," Technical Report, 2003.

[34] UNFCCC, "CDM methodology booklet," Technical Report, 2014.

[35] Corporate Technical Audit Department, "Measurement and verification, specification for three phase energy metering for use with measurement and verificaiton projects," Eskom, Tech. Rep. 1, 2005.

[36] M. C. Burkhart, Y. Heo, and V. M. Zavala, "Measurement and verification of building systems under uncertain data: a Gaussian process modeling approach," *Energy and Buildings*, vol. 75, pp. 189–198, 2014.

[37] N. Ridler, B. Lee, J. Martens, and K. Wong, "Measurement uncertainty, traceability, and the GUM," *Microwave Magazine, IEEE*, vol. 8, no. 4, pp. 44–53, 2007.

[38] B. Dong, C. Cao, and S. E. Lee, "Applying support vector machines to predict building energy consumption in tropical region," *Energy and Buildings*, vol. 37, no. 5, pp. 545–553, 2005.

[39] J. Granderson and P. N. Price, "Development and application of a statistical methodology to evaluate the predictive accuracy of building energy baseline models," *Energy*, vol. 66, no. 5, pp. 981–990, 2014.

[40] M. Aydinalp-Koksal and V. I. Ugursal, "Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector," *Applied Energy*, vol. 85, no. 4, pp. 271–296, 2008.

[41] T. Walter, P. N. Price, and M. D. Sohn, "Uncertainty estimation improves energy measurement and verification procedures," *Applied Energy*, vol. 30, pp. 230–236, 2014.

[42] T. Lee, K. Liao, and W. Lu, "Evaluation of the suitability of empirically-based models for predicting energy performance of centrifugal water chillers with variable chilled water flow," *Applied Energy*, vol. 93, pp. 583–595, 2012.

[43] T. A. Reddy and D. E. Claridge, "Uncertainty of "Measured" energy savings from statistical baseline models," *HVAC R Research*, vol. 6, no. 1, pp. 3–20, 2000.

[44] T. Giglio, R. Lamberts, M. Barbosa, and M. Urbano, "A procedure for analysing energy savings in multiple small solar water heaters installed in low-income housing in brazil," *Energy Policy*, vol. 72, no. 4, pp. 43–55, 2014.

[45] M. J. Kaiser and A. G. Pulsipher, "Preliminary assessment of the louisiana home energy rebate offer program using IPMVP guidelines," *Applied Energy*, vol. 87, no. 2, pp. 691–702, 2010.

[46] T. Jayaweera, H. Haeri, J. Keeling, M. S. Khawaja, and J. Rushton, "The uniform methods project: Methods for determining energy efficiency savings for specific measures," The Cadmus Group, Portland, Oregon, Tech. Rep., 2013.

[47] C. J. Adcock, "Sample size determination: a review," *Journal of the Royal Statistical Society: Series D (The Statistician)*, vol. 46, no. 2, pp. 261–283, 1997.

[48] UNFCCC, "General guidelines for sampling and surveys for small-scale CDM project activities. version 2.0," Technical Report, 2009.

[49] S. K. Thompson, *Sampling*. New York: Wiley, 2002.

[50] R. S. Witte and J. S. Witte, *Statistics*. New York: Wiley, 2010.

[51] P. S. Levey and S. Lemeshow, *Sampling of populations: methods and applications*. New

York: Wiley, 2008.

[52] Clean Energy Ministerial, "Energy performance measurement and verification: Guidance on data quality," International Partnership for Energy Efficiency Cooperation, Paris, France, Technical Report, 2014.

[53] Y. Heo and V. M. Zavala, "Gaussian process modeling for measurement and verification of building energy savings," *Energy and Buildings*, vol. 53, pp. 7–18, 2012.

[54] D. C. Montgomery, *Engineering statistics.* New York: Wiley, 2004.

[55] H. Carstens, X. Xia, and S. Yadavalli, "Measurement uncertainty and risk in measurement and verification projects," in *International Enery Program Evaluation Conference*, Long Beach, CA, USA, 10-13 September 2015.

[56] Pacific Northwest National Laboratory, "Metering best practices, a guide to achieving utility resource efficeincy," U.S. Department of Energy, Oregon, Tech. Rep. 2, 2011.

[57] A. Sabapathy and J. Whittaker, "An assessment of energy efficiency and carbon reductions from replacing halogen lamps with light emitting diodes in traffic lights Ű a case study of Gloucestershire countyŠs initiative," *Energy and Environment*, vol. 22, no. 4, pp. 361–374, 2011.

[58] ASHRAE, "ASHRAE Guideline 14: measurement of energy and demand savings," Technical Report, 2002.

APPENDIX A

MATLAB CODE FOR THE OPTIMISATION

A.1 OBJECTIVE FUNCTION

```
%24/04/2013
%M&V Cost Objective function
function y = obj(x,a1,a2,b1,b2)

            %y =251937.*exp(-0.042.*(x(1)./100)) + 251937.*exp(-0.042.*(x(2)./100)) + (a1+b1)*x(3)+(a2+b2).*x(4);
            %y =150000.*exp(-0.042.*(10.*x(1)./100)) + 150000.*exp(-0.042.*(10.*x(2)./100)) + (a1+b1)*x(3)+(a2+b2).*x(4);
            y =(251937.*exp(-0.042.*(x(1))) + 251937.*exp(-0.042.*(x(2))) + (a1+b1)*x(3)+(a2+b2).*x(4))/0.95;
            %y =251937.*exp(-0.042.*(x(1))) + 251937.*exp(-0.042.*(x(2))) + (a1+b1)*x(3)+(a2+b2).*x(4);
            %y =503875.*exp(-0.042.*(x(1)./100)) + 503875.*exp(-0.042.*(x(2)./100)) + (a1+b1)*x(3)+(a2+b2).*x(4);
%
end
```

A.2 CONSTRAINT

```
%29/04/2014
%constraints for "projectCost" The constraints take into account the
%overall confidence and precision
function [c, ceq] = cons(x,N1,N2,E1, E2,cv1,cv2)

S1=cv1.*E1;
S2=cv2.*E2;
M1=(x(1)/100).*E1;
M2=(x(2)/100).*E2;

E=(N1.*E1+N2.*E2)./(N1+N2);


se=sqrt(S1.^2./x(3).*N1.^2+S2.^2./x(4).*N2.^2+M1.^2.*N1.^2+M2.^2.*N2.^2)/(N1+N2);

ceq=[

%x(1)-3.4579;
%x(2)-10.5321;
%x(3)-28;
%x(4)-22;
];

c=[
1.645.*se-0.1.*E;
];

end
```

A.3 SOLUTION

```
%24/04/2014
%optimization of modelling and sampling
%%Initial values
clc;
clear;
close all;


N1=1320;
N2=880;

% % N1=2500;
% % N2=500;

cv1= 0.2;
cv2= 0.5;

%cv1= 0.20*1.5;
%cv2= 0.50*1.5;

E1=1.91; %kW
E2=1.415; %kW

% E1=2.91; %kW
% E2=1.415; %kW

% E1=0.45; %kW
% E2=1.415; %kW


a1 =500;            %value per meter
```

```
a2 = 1500;

b1 = 195;            %installation cost
b2 = 320;

lb =[0;0;0;0;];
ub=[100;100;+inf;+inf;];

%x0: cvr1, cvr2, n1, n2
%x0=[0.3156;0.001;1.8387;0.3734;];

x0=[35;100;50;60;];
% %{
%fmincon
options = optimset('Algorithm','interior-point','Tolcon',1e-45,'tolfun',1e-45,'Tolx',1e-45,'Hessian', {'lbfgs',20},'MaxFunEvals',5000,'MaxIter',5000);

x= fmincon(@(x)obj(x,a1,a2,b1,b2),x0,[],[],[],[],lb,ub,@(x)cons(x,N1,N2,E1, E2,cv1,cv2),options)
%}

%{
%ga
options = gaoptimset('Tolcon',1e-45,'tolfun',1e-45,'PopulationSize',500);
x = ga(@(x)obj(x,cv1,a1,a2,b1,b2,N1,cv2,N2),6,[],[],[],[],lb,ub,@(x)cons(x,N1,E1,cv1,N2,E2,cv2),options)
%}

CVR1 = x(1)
CVR2 = x(2)
s1=ceil(x(3))
```

```
s2=ceil(x(4))

TotalSamples = s1 + s2;
%%calculating the size of the samples

%%calculating the overall cost



scost1 = (a1+b1).*s1;
scost2 = (a2+b2).*s2;
Sample_cost = scost1 + scost2;

mcost1 = 251937.*exp(-0.042.*(x(1)));
mcost2 = 251937.*exp(-0.042.*(x(2)));
mod_cost_tot = mcost1 + mcost2;

tot_cost_group1 = scost1 + mcost1 %total cost group 1
tot_cost_group2 = scost2 + mcost2 %total cost group 2
tot_cost = tot_cost_group1 + tot_cost_group2;

%calculating overall CVRMSE
E=(N1.*E1+N2.*E2)./(N1+N2);

Mod1 = x(1)/100.*E1; %modelling uncertainty group 1
Mod2 = x(2)/100.*E2; %modelling uncertainty group 2
modUn = sqrt(N1.^2*Mod1.^2 + N2.^2*Mod2.^2)./(N1+N2);

Sa1=cv1.*E1./sqrt(x(3)); %sampling uncertainty group 1
Sa2=cv2.*E2./sqrt(x(4)); %sampling uncertainty group 2
samUn = sqrt(N1.^2*Sa1.^2 + N2.^2*Sa2.^2)./(N1+N2);

Usg1 =sqrt(Mod1^2 + Sa1^2); %combined uncertainty for group 1
```

```
Usg2 =sqrt(Mod2^2 + Sa2^2); %combined uncertainty for group 2

U = sqrt(N1.^2*Usg1.^2 + N2.^2*Usg2.^2)./(N1+N2);

Usm = sqrt(modUn^2 + samUn^2);

CVRMSE = (modUn./E)*100;

Combined_CVRMSE = sqrt(N1.^2*CVR1^2 + N2.^2*CVR2^2)./(N1+N2);
%%added to calculate overall standard error
```

A.4 SAMPLE SIZE SIMULATIONS CONSTRAINT

```
%29/04/2014
%constraints for "projectCost" The constraints take into account
    the
%overall confidence and precision
function [c, ceq] = cons(x,N1,N2,E1, E2,cv1,cv2)

S1=cv1.*E1;
S2=cv2.*E2;
M1=(x(1)/100).*E1;
M2=(x(2)/100).*E2;

E=(N1.*E1+N2.*E2)./(N1+N2);


se=sqrt(S1.^2./x(3).*N1.^2+S2.^2./x(4).*N2.^2+M1.^2.*N1.^2+M2.^2.*
    N2.^2)/(N1+N2);

ceq=[

%x(1)-3.4579;
%x(2)-10.5321;
x(3)-28;
x(4)-22;
];

c=[
1.645.*se-0.1.*E;
];

end
```

A.5 CV(RMSE) SIMULATION CONSTRAINT

```
%29/04/2014
%constraints for "projectCost" The constraints take into account the
%overall confidence and precision
function [c, ceq] = cons(x,N1,N2,E1, E2,cv1,cv2)

S1=cv1.*E1;
S2=cv2.*E2;
M1=(x(1)/100).*E1;
M2=(x(2)/100).*E2;

E=(N1.*E1+N2.*E2)./(N1+N2);


se=sqrt(S1.^2./x(3).*N1.^2+S2.^2./x(4).*N2.^2+M1.^2.*N1.^2+M2.^2.*N2.^2)/(N1+N2);

ceq=[

x(1)-3.4579;
x(2)-10.5321;
%x(3)-28;
%x(4)-22;
];

c=[
1.645.*se-0.1.*E;
];

end
```