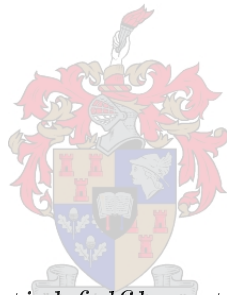


Reducing energy costs within schools in South Africa using solar and intelligent hot water interventions

by

Stefan Gerber



*Thesis presented in partial fulfilment of the requirements for
the degree of Master of Science in Engineering in the Faculty
of Engineering at Stellenbosch University*

Supervisor: Prof M.J. Booysen

Co-supervisor: Dr. A.J. Rix

December 2019

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2019

Copyright © 2019 Stellenbosch University
All rights reserved.

Abstract

The educational gap within socio-economic groups in South Africa is immense and learners within no-fee schools are at a disadvantage with most of the available funding being used for personnel salaries. As the energy costs in South Africa rise, the remaining non-personnel funding will decrease, limiting spending on teacher-support materials and school maintenance. This has led to lasting problems within these communities and constrains the ability of the education system to provide learners of these schools with a pathway out of poverty. Schools are currently billed on commercial and industrial tariff structures, and by reducing their energy usage and maximum monthly demand, money can be saved to be better spent improving the quality of education delivered. To address this problem a comprehensive system capable of estimating the potential financial viability of solar and load-shifting interventions within a school environment was developed.

A method capable of determining the energy usage within schools was developed. This process creates a generic energy consumption profile for a building from measured energy usage data of a subset of schools, and is expanded to scale the load profile using only usage data from utility bills as well as seasonal dates to produce a load forecast. The method was validated using five datasets each containing the hourly energy usage measurement data from schools over a period of three years, and was capable of forecasting the yearly energy consumption of the schools to within an averaged accuracy of 5% while estimating the maximum monthly demand to 6% of the measured usage.

A PV optimisation technique was implemented using the forecast to estimate the potential profitability of various solar system sizes by determining the internal rate of return and the utilisation of the system's generating capacity. It was able to identify the optimal system size of the schools with the best return on investment, presenting itself as a valuable tool for reducing the financial burden many schools face.

Three control schemes of intelligent water heater scheduling were researched. Firstly, a priority-based scheduler was configured to heat water using the school's water usage history while diverting any excess solar energy to the water heaters to exploit their energy storage capabilities, increasing the school's energy bill savings to 23.2% per month. Secondly, a bi-thermal control method was added to the priority-based scheduler, employing a temperature delta to

ABSTRACT

iii

increase the amount of solar energy to be stored within the water heater tank while minimising their grid reliance and improving the monthly savings to 24.8% per month. Finally, a demand-limiter control scheme was implemented in conjunction with bi-thermal control resulting in large demand-charge savings and an average energy bill reduction of 26.7% per month, producing the maximum savings while maintaining suitable levels of user comfort.

Uittreksel

Die onderriggaping in sosio-ekonomiese groepe in Suid-Afrika is groot. Leerders in minderbevoorregte skole word benadeel aangesien die meeste van die befondsing aan personeelsalarisse afgestaan word. Soos energiekoste in Suid-Afrika toeneem, sal oorblywende nie-personeelbefondsing kwyn, wat besteding aan onderrigondersteunende material en skoolherstelwerk beperk. Dit het gelei tot probleme in hierdie gemeenskappe en beperk die vermoë van die onderrigstelsel om leerders van hierdie skole te voorsien van 'n weg uit armoede. Skole se elektrisiteitsrekening word tans gehef op grond van kommersiële en industriële tariefstrukture. Deur energieverbruik sowel as die maksimum maandelikse vraag te verminder, kan die kwaliteit van onderrig verbeter word deur die wins sodoende aan te wend. Om hierdie problem aan te spreek, is 'n stelsel ontwikkel om die proses wat vereis word om voordele van 'n sonkragstelsel vir 'n skool te identifiseer, te vereenvoudig. Die stelsel maak gebruik van data-gedrewe vraagvoorskatting en finansiële lewensvatbaarheidsbenaderings vir verskeie intelligente elektriese warmwatersilinderskemas.

'n Metode met die vermoë om energieverbruik in skole te bepaal, is ontwikkel. Hierdie metode skep 'n generiese energieverbruiksprofiel vir 'n gebou vanaf gemete energieverbruiksdata van 'n stel skole. Die proses is uitgebrei om die lasprofiel te skaleer deur slegs verbruiksdata van elektrisiteitsrekening en seisoenale datums in ag te neem om 'n lasvoorspelling te maak. Die metode is bekragtig deur vyf datastelle elk met uurlikse energieverbruiksmetings van skole oor 'n tydperk van drie jaar. Die ontwikkelde tegniek is in staat om die jaarlikse energieverbruik van die skole vooruit te skat met 'n gemiddelde akkuraatheid van 5% terwyl die maksimum maandelikse aanvraag beraam word tot binne 6% van die gemete verbruik.

'n Fotovoltaïese optimeringstegniek is geïmplementeer wat van die voorskatting gebruik maak om die potensiële winsgewendheid van verskeie sonkragstelselgroottes te beraam. Dit word bereik deur die interne opbrengskoers en effektiwiteit van die stelsel se opwekkingskapasiteit te bepaal. Hierdeur is die optimale stelselgrootte van die skole met die beste opbrengs op belegging bepaal, wat die tegniek 'n waardevolle stuk gereedskap vir die vermindering van die finansiële las maak.

Drie intelligente beheerskemas vir elektriese warmwatersilinders is ondersoek. Die eerste beheerskema het 'n prioriteitsskeduleerder geïmplementeer

wat die warmwatersilinders verhit deur gebruik te maak van die historiese verbruiksdata. Die skeduleerder het ook enige oortollige sonenergie na die silinders herlei om hulle kapasiteit as termiese batterye ten volle te gebruik. Dit het die skool wat ondersoek is se maandelikse elektrisiteitrekening met 23.2% verlaag. Die tweede beheerskema het 'n nuwe, bi-termiese verhittingsmetode geïmplimenteer deur die warmwatersilinder na verskillende temperature te verhit vir netwerkenegie en sonkrag met 'n maandelikse besparing van 24.8%. Laastens is 'n beheerskema ontwikkel om te verhoed dat die maksimum maandelikse kragverbruik verhoog word deur die warmwatersilinders. Hierdie skema het ook die bi-termiese metode geïmplimenteer en die grootste maandelikse besparing van 26.7% bereik sonder dat die gemak van die verbruikers ontwrig word.

Patents and publications

The work in this manuscript has been published as follows:

- S. Gerber, M.J. Booysen, A.J. Rix, "Managing peak demand and energy costs through PV and intelligent scheduling of water heaters at two schools in South Africa", REPS, September 2018, Stellenbosch, South Africa.
- Accepted for publication: S. Gerber, M.J. Booysen, A.J. Rix, "Combining grid-tied PV and intelligent water heater control to reduce the energy costs at schools in South Africa", Energy for Sustainable Development, June 2019.

Additionally, the work in this manuscript has been submitted for publication as follows:

- S. Gerber, M.J. Booysen, A.J. Rix, "Toward sustainable developing cities: A simplified forecasting model for sizing grid-tied PV using monthly electricity bills", Sustainable Cities and Society, August 2019.

Finally, the work in the manuscript has been patented as follows:

- S. Gerber, M.J. Booysen, A.J. Rix, "Method and system for controlling temperature in a water heater", Provisional Patent Application No. 2019/02357, 18 April 2019.

The datasets and GUI for this thesis are available at:

- <https://bit.ly/2OXt9gA>

Acknowledgements

I would like to express my sincere gratitude to the following people and organisations for their invaluable support these past two years:

- My supervisor, prof. Thinus Booysen for his support and guidance and for providing me with opportunities for personal and professional growth.
- Dr. Arnold Rix for being a fantastic co-supervisor throughout my masters. Providing insightful feedback and ideas throughout the study.
- My family and friends for their continued love and support.
- MTN South Africa for the funding I received during my studies.

Contents

| | |
|--|-------------|
| Declaration | i |
| Abstract | ii |
| Uittreksel | iv |
| Patents and publications | vi |
| Acknowledgements | vii |
| Contents | viii |
| List of Figures | xi |
| List of Tables | xiii |
| Nomenclature | xiv |
| 1 Introduction | 1 |
| 1.1 Sustainable development through techno-economic assessment | 1 |
| 1.2 The South African context | 2 |
| 1.2.1 An electricity supply crisis, and the falling costs of solar | 2 |
| 1.2.2 Demand side management in South Africa | 4 |
| 1.2.3 The financial struggle within schools in South Africa | 5 |
| 1.3 Thesis problem statement, hypothesis and objectives | 6 |
| 1.3.1 Problem statement | 6 |
| 1.3.2 Proposed solution | 7 |
| 1.3.3 Hypothesis | 7 |
| 1.3.4 Research objectives | 7 |
| 1.4 Contributions | 7 |
| 1.5 Thesis structure | 8 |
| 2 Literature Review | 10 |
| 2.1 Data-driven energy usage forecasting | 10 |
| 2.1.1 Data cleaning | 10 |

| | | |
|----------|---|-----------|
| 2.1.2 | Time series modelling and forecasting | 12 |
| 2.1.3 | Validation | 15 |
| 2.1.4 | Testing | 15 |
| 2.1.5 | Summary of energy usage forecasting | 15 |
| 2.2 | Load matching and grid interaction | 16 |
| 2.2.1 | Grid interaction | 16 |
| 2.2.2 | Optimised on-site generation | 16 |
| 2.2.3 | Summary of load matching and grid interaction | 17 |
| 2.3 | Demand management within buildings | 17 |
| 2.3.1 | DSM within a school environment | 18 |
| 2.3.2 | Load Shifting using EWHs | 18 |
| 2.3.3 | Summary of DSM techniques | 19 |
| 2.4 | Chapter summary | 19 |
| 3 | System development | 20 |
| 3.1 | Energy usage forecasting | 20 |
| 3.1.1 | Electricity demand within schools in South Africa | 20 |
| 3.1.2 | A review of techniques discussed in literature | 23 |
| 3.1.3 | Forecast model development | 24 |
| 3.2 | Solar system optimisation | 28 |
| 3.2.1 | Data gathering | 29 |
| 3.2.2 | Solar system modelling and optimisation | 29 |
| 3.2.3 | Optimised load-matching | 35 |
| 3.3 | Intelligent EWH control | 36 |
| 3.3.1 | Analysis of usage data | 37 |
| 3.3.2 | EWH modelling | 39 |
| 3.3.3 | EWH DSM techniques | 42 |
| 3.3.4 | EWH control schemes | 44 |
| 3.4 | Graphical user interface development | 49 |
| 3.5 | Chapter summary | 50 |
| 4 | Experimental setup | 51 |
| 4.1 | Simulation environment | 51 |
| 4.2 | Measurement data used | 51 |
| 4.3 | Data cleaning | 53 |
| 4.3.1 | Missing data | 53 |
| 4.3.2 | Smoothing and outlier detection | 53 |
| 4.4 | Simulation setup | 54 |
| 4.4.1 | Load profile determination and forecast | 54 |
| 4.4.2 | Solar PV simulation | 56 |
| 4.4.3 | EWH simulation | 56 |
| 4.5 | Evaluation and metrics | 58 |
| 4.5.1 | Forecast model | 58 |
| 4.5.2 | Solar intervention | 59 |

| | |
|--|-----------|
| <i>CONTENTS</i> | x |
| 4.5.3 EWH control schemes | 59 |
| 4.6 Chapter summary | 61 |
| 5 Results | 62 |
| 5.1 Energy usage forecasting | 62 |
| 5.2 Solar intervention | 65 |
| 5.3 Complete system simulation with EWH intervention | 67 |
| 5.3.1 Load forecast | 67 |
| 5.3.2 Solar intervention | 68 |
| 5.3.3 Smart-scheduled EWH intervention | 68 |
| 5.4 Simulation performance | 73 |
| 6 Conclusion | 75 |
| 6.1 Evaluation of work | 75 |
| 6.1.1 Energy usage forecasting | 75 |
| 6.1.2 Solar system optimisation | 76 |
| 6.1.3 Intelligent EWH control | 76 |
| 6.2 Recommendations for future work | 77 |
| Appendices | 78 |
| A System software user guide | 79 |
| A.1 Installation | 79 |
| A.2 Usage | 79 |
| A.3 Question and bugs | 80 |
| List of References | 82 |

List of Figures

| | | |
|------|--|----|
| 1.1 | Normalised comparison of Eskom tariff increases versus South African consumer price index versus international median solar installation costs [14; 15; 16; 17]. | 3 |
| 1.2 | Percentage breakdown of Grade 6 learners who are literate and numerate for each quintile. | 5 |
| 3.1 | Hourly electricity usage of a school (670 students). | 22 |
| 3.2 | A typical school's daily energy usage profile over a 1 year period with seasonal periods highlighted. | 22 |
| 3.3 | Partial correlation of a school's measured usage data showing surges occurring on 24 hour and 7 day (168 hour) intervals. | 23 |
| 3.4 | Simplified diagram of normalised load profile being scaled using monthly utility bill usages. | 28 |
| 3.5 | Solar module orientation relative to the sun (Brownson [66]). | 30 |
| 3.6 | Equivalent circuit of a single-diode solar model. | 32 |
| 3.7 | Simplified diagram of complete solar modelling process | 35 |
| 3.8 | Diagram illustrating yearly cash flows for the evaluated solar system. | 36 |
| 3.9 | Usage profiles obtained from measurement data | 37 |
| 3.10 | Diagram of energy flows within EWH nodal models | 40 |
| 3.11 | The PPDM in action for various power limits. | 43 |
| 3.12 | Daily profile of an EWH with thermostat control. | 45 |
| 3.13 | Daily profile of an EWH with smart schedule control with solar dumping. | 47 |
| 3.14 | Daily profile of an EWH with bi-thermostat control. | 48 |
| 3.15 | Daily profile of an EWH with demand limiting. | 49 |
| 4.1 | A complete diagram of the developed system simulation process. | 52 |
| 4.2 | Forecasting scaler list training process with anomaly boundaries and entries highlighted. | 54 |
| 4.3 | Effects on Legionella bacteria at different water temperatures with simulation bounds highlighted. | 57 |
| 5.1 | Basic energy usage forecast accuracy for a varying number of utility bills | 63 |

| | | |
|-----|--|----|
| 5.2 | Averaged maximum monthly demand forecast versus measured maximum demand for 2018. | 64 |
| 5.3 | Solar generation usage for varied system sizes. | 66 |
| 5.4 | Internal rate of return for various system sizes. | 67 |
| 5.5 | Energy usage forecast versus measured usage for school A in 2018. | 68 |
| 5.6 | Solar system IRR and generation results for school A. | 69 |
| 5.7 | Usage results for the bi-thermostat control scheme with prioritised heating and solar dumping. | 70 |
| 5.8 | Usage results for the demand-limiting control scheme with a bi-thermostat and solar dumping. | 71 |
| 5.9 | Usage results for the evaluated interventions. | 72 |
| A.1 | The graphical user interface after running the main.py command. | 80 |
| A.2 | Azimuth angle of a school building. | 81 |
| A.3 | Interface after a simulation was performed, with results shown in the console window. | 81 |

List of Tables

| | | |
|-----|--|----|
| 3.1 | Weather data file description | 29 |
| 3.2 | Solar system parameters | 32 |
| 3.3 | Feature table of evaluated interventions | 49 |
| 4.1 | Input parameters used for experimental setup | 55 |
| 5.1 | Overview of evaluated schools | 63 |
| 5.2 | Forecast simulation results | 64 |
| 5.3 | Input data for solar simulations | 65 |
| 5.4 | Simulation results for evaluated school for the period 1 January 2018 to 31 December 2018 | 73 |

Nomenclature

Acronyms and Abbreviations

| | |
|---------|--|
| AOI | Angle Of Incidence |
| ANN | Artificial Neural Network |
| AR | Autoregressive |
| ARIMA | Autoregressive Integrated Moving Average |
| ARMA | Autoregressive Moving Average |
| BOS | Balance Of System |
| CEC | California Energy Commission |
| CS | Control Scheme |
| DT | Decision Tree |
| DSM | Demand-Side Management |
| EWH | Electric Water Heater |
| EMS | Energy Management System |
| GUI | Graphical User Interface |
| IRENA | International Renewable Energy Agency |
| IRR | Internal Rate Of Return |
| LCOE | Levelised Cost of Energy |
| LR | Linear regression |
| LCF | Load Cover Factor |
| LM | Load Matching |
| MAPE | Mean Absolute Percentage Error |
| NREL | National Renewable Energy Laboratory |
| NPV | Net Present Value |
| OCGT | Open Cycle Gas Turbine |
| OECD | Organisation for Economic Co-operation and Development |
| PPDM | Peak Power Demand Manager |
| PV | Photovoltaic |
| KNN | K-Nearest Neighbours |
| REIPPPP | Renewable Independent Power Producer Procurement Programme |

SARIMA Seasonal Autoregressive Integrated Moving Average
 SAM System Advisor Model
 SCF Supply Cover Factor
 TOU Time Of Use

Constants

c Specific heat capacity of water [4184 Joule/kg.°C]
 ρ Density of water [1000 kg/m³]

List of Symbols

α_s Solar altitude angle
 β Surface tilt angle
 δ Declination angle
 γ Azimuth angle
 B_{Energy} Total monthly energy consumption
 B_{Scaled} Scaled monthly energy matrix
 C_t Cash flow for period t
 d Thermocline thickness
 D_S Seasonal weekday total usage lists
 G_S Normalised seasonal hot water usage lists
 I_b Incidence beam irradiance
 I_d Incidence diffuse irradiance
 I_r Incidence ground reflected irradiance
 k Thermal conductivity
 M_{basic} Basic energy scaling matrix
 M_{max} Maximum monthly demand scaling matrix
 $M_{offpeak}$ TOU offpeak energy scaling matrix
 M_{peak} TOU peak energy scaling matrix
 $M_{standard}$ TOU standard energy scaling matrix
 p Probability
 P_{elec} Electrical power rating of an EWH
 Q_{elec} Electrically restored enthalpy
 Q_{draw} Enthalpy of water drawn from EWH
 Q_{loss} Enthalpy loss through standing losses
 r Discount rate
 R_{therm} Thermal resistance of an EWH cylinder
 S_n Seasonal daily usage lists

| | |
|-------------|---|
| T_{amb} | Ambient temperature |
| t_{event} | Time until next EWH usage event |
| t_{et} | Minimum time before EWH usage event that grid heating may occur |
| t_{heat} | Time to heat EWH to target temperature |
| U_S | Normalised seasonal energy usage lists |
| V_{cyl} | Volume of the EWH cylinder |
| V_{draw} | Volume of water drawn from EWH |

Subscripts

| | |
|-----|----------------------|
| t | t^{th} time period |
| n | n^{th} position |

Chapter 1

Introduction

Driven by a surge in economic growth, the combined share of total electricity production within developing countries surpassed that of developed Organisation for Economic Co-operation and Development (OECD) countries, and in 2016 the share of global electricity produced by non-OECD countries reached 56% [1]. Increased public awareness and political pressure have resulted in a universal push to move away from carbon emitting energy supplies [2]. As a result, renewable energy is the fastest growing source of electricity production, accounting for 26% of the global energy supply in 2018 [1]. However, in many developing countries electrification necessary for economic development would not be possible without coal-powered plants, leading to an increase in global carbon emissions from 2017 to 2018 [3; 4]. In South Africa, Eskom is the largest electric utility company, generating approximately 95% of the country's electricity demand [5]. Coal accounts for nearly 80% of the country's electricity needs and produces 85% of the country's total carbon dioxide emissions, resulting in South Africa being one of the largest CO₂ emitting countries in the world [3; 6]. This is an unfortunate consequence of developing nations being unable to afford cleaner alternatives without sacrificing spending in other, more essential sectors [7; 8].

1.1 Sustainable development through techno-economic assessment

Optimising the use of available natural resources has become a point of concern globally as many governments aim to reduce their emissions without compromising human development. During recent years the impact of urbanisation has become apparent through rapid increases in the number of carbon-emitting residential and commercial buildings, especially in developing countries [9]. Additionally, many rural communities lack sufficient electrical power sources due to their remoteness, harming the potential for economic growth within those regions [10]. Renewable energy sources have the potential to satisfy the

need for providing low-carbon power within cities and decentralised energy solutions to remote communities. However, this opportunity must be met with careful consideration regarding the effective management of the available supply to reduce unnecessary expenses while providing sufficient generation [11]. The design and development of renewable energy systems is a complex process involving multiple stakeholders with different aims and objectives including a clear understanding of the environmental, technical and economic factors, and if not managed properly, these complexities can lead to needless delays and avoidable capital expenses [12]. Simplifying the implementation through improved project assessment methods is therefore imperative to allow for renewable systems to be installed at a rate suitable for sustainable development, particularly in developing countries without the necessary resources to perform cost versus benefit analyses in rural areas requiring urgent intervention.

1.2 The South African context

In South Africa, a nation dependent on coal for energy generation, sustainability initiatives are often neglected in favour of more economical alternatives. As a member of the International Renewable Energy Agency (IRENA), numerous policies have been created to promote the growth of renewable energies [13]. However, a rise in cleaner forms of energy may not only facilitate the adoption of renewable resources, but alleviate the strain on an already ailing electricity supply.

1.2.1 An electricity supply crisis, and the falling costs of solar

Since the founding of electricity generation within South Africa in 1923 the utility has always been a single state owned and operated entity. Throughout most of Eskom's history, the country's electricity prices were some of the least expensive in the world [18]. This stems from a considerable over-investment in generation capacity in the late 1970's, allowing the real, inflation-adjusted average electricity price to fall by 40% between 1978 and 2004 as the country's demand could not keep up with the oversupply, effectively subsidising the cost of electricity. Due to a poor development strategy, energy spending remained restricted from 1998 until 2004 as national spending was focused on generating economic growth [18; 19; 20]. However, rising energy demand promoted by steady economic growth led to a diminished energy supply, and by 2007 Eskom was unable to reliably provide electricity to South Africa, leading to a national energy crisis and rolling "load shedding". To help fund generational capacity the average cost of electricity across all sectors increased by 357% from 2007 to 2018, greatly outpacing the consumer inflation rate as presented in Figure 1.1 [21]. This placed a large financial strain on the country and its

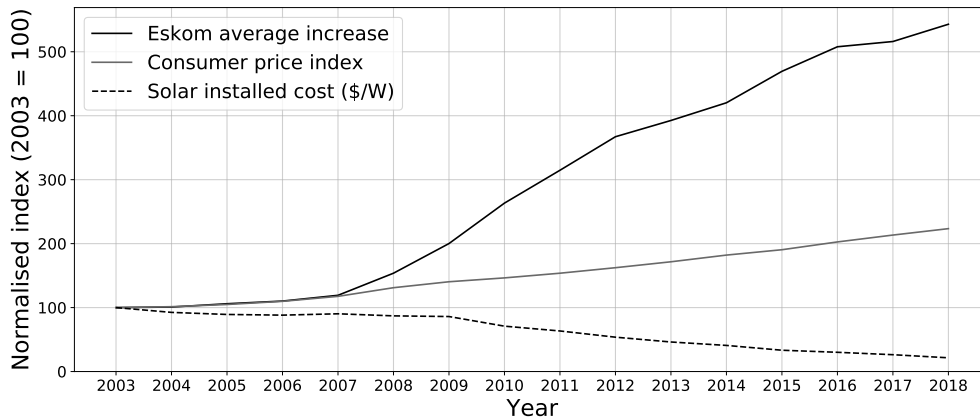


Figure 1.1: Normalised comparison of Eskom tariff increases versus South African consumer price index versus international median solar installation costs [14; 15; 16; 17].

citizens, especially in poorer communities [22; 23]. A notable disadvantage of subsidised energy is that it can place an unnecessary burden on public finances. The national budget for 2019 outlined a 3-year R69 Billion (\$4.8 Billion) support plan to keep Eskom operating [24]. Funds that could have been better spent on health, education and social welfare. Other harmful consequences of subsidizing the cost of energy include:

- Inexpensive electricity costs discourages the use of efficient technologies within industry and households in South Africa, leading to the country producing nearly 1.5% of global carbon emissions in 2017 [18].
- Energy subsidies lead to investments in energy-intensive industries as operating costs are lower. A study has found that this is done at the expense of employment generation [25].
- Energy subsidies discourage private investments in the nation's energy sector as companies cannot compete with electricity sold at below the cost of generation [18].

Since 2010 new policies have allowed for private investment through the Renewable Independent Power Producer Procurement Programme (REIPPPP), whereby several producers provided cleaner energy through wind and solar farms, generating several thousand local employment opportunities [26] and alleviating the need for Eskom to use extremely expensive Open Cycle Gas Turbine (OCGTs) during peak usage hours [27]. Driven by technological advances, economies of scale and deployment subsidies the installation costs of solar power has declined [22; 28]. In January 2018 the International Renewable Energy Agency released figures indicating that renewable energy has emerged

as a sufficient method for providing new power needs and found that the levelised cost of energy (LCOE) from solar photovoltaics (PV) decreased by 69% between 2010 and 2016 [13].

However, many organisations (e.g. Departments of Education and Public Works) still lack the necessary resources required to determine the benefits of renewable energy solutions, particularly due to the complexity involved in determining an accurate cost versus benefit forecast, as it will vary on a case by case basis [29]. Another challenge faced when designing solar systems is that it may provide excess energy during certain times of the day in certain seasons or load conditions. Effectively managing the demand and potential excess supply is therefore necessary to achieve the greatest benefit, and requires thoughtful considerations and planning [30].

1.2.2 Demand side management in South Africa

Consistent and reliable delivery of electricity requires the demand to be met with a sufficient supply, and can vary dramatically depending on weather conditions and the time of day. The electricity utility will then adjust to the changing demand by dispatching the required amount of generation. However, during peak usage hours the generation may be supplied by less efficient and more economically expensive sources. In South Africa, this demand is supplied by OCGTs burning very large amounts of diesel fuel at a great expense [27]. Demand side management (DSM) is a process of managing the electricity demand to optimise the use of available generation, and includes all methods that influence a customer's energy use to reduce the electricity demand while remaining mutually beneficial to the customer and the utility [31]. All DSM techniques result in either peak-shaving, load-shifting or simply a reduction in energy use by modifying the usage profile of consumers [31].

In South Africa the electric water heaters (EWHs) in multiple cities were retrofitted with control mechanisms allowing the utility company to switch off the EWHs during peak times to reduce the stress on the electricity supply by means of direct control [32]. Different forms of centralised EWH control have been proposed, with three key control objectives, namely, cumulative energy usage reduction, efficient load management, and user comfort [32; 33]. Direct control provides the benefit of being able to reduce the strain on the grid during high demand periods, but does not take the comfort of the user into account. Other forms of DSM regularly implemented include frequency regulation to maintain a stable supply during periods of excess or reduced generation. South Africa's electricity tariffs are structured as domestic, commercial, agricultural, industrial and municipal providing indirect demand response through load-limiting and time-of-use pricing schemes, whereby the cost of electricity changes throughout the day depending on the demand [17]. This, in addition to increasing electricity prices, presents an opportunity for large monthly savings through effective load management and energy saving interventions.

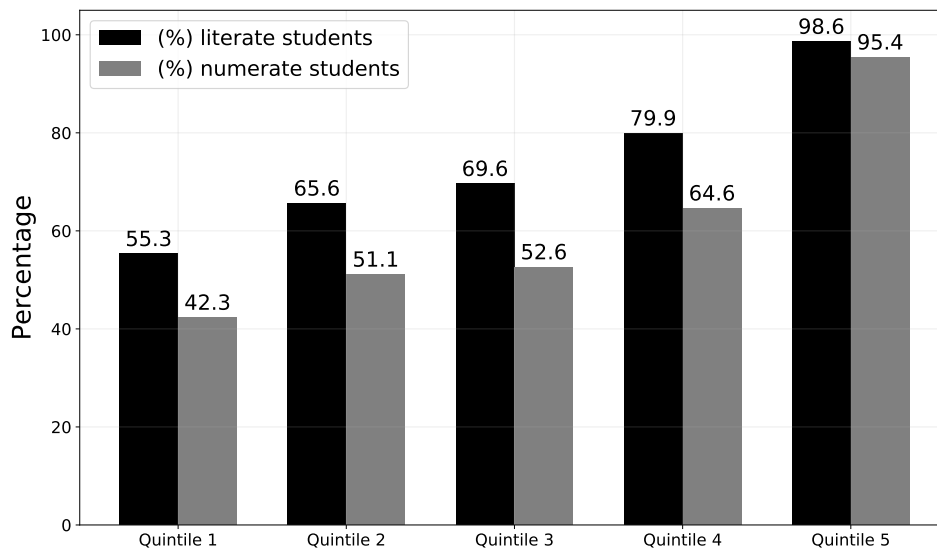


Figure 1.2: Percentage breakdown of Grade 6 learners who are literate and numerate for each quintile [34].

1.2.3 The financial struggle within schools in South Africa

Reducing inequality and poverty in South Africa has been the central concern since the first democratic election in 1994 and despite efforts to restructure the economy the divide between the rich and poor has continued to grow [35; 36]. A study by van der Berg *et al.* [37] reported on how the low quality of education provided in schools within disadvantaged communities can lead to exclusion and marginalisation, limiting the future prospects of learners from a young age, presented in Figure 1.2 [37]. In 2006, 49% of education spending reached the poorest 40% of households [37; 38]. In 2017 this has improved to 54% [39]. However, after personnel spending and conditional grants, only 10% to 20% of the budget remains for non-personnel expenses, which includes textbooks, laboratory equipment, stationary, school-maintenance and utility costs.

Schools are currently billed on a commercial or industrial tariff, and the expected increase in electricity costs will further reduce the already limited non-personnel spending available for schools [17; 21; 39]. This has led to lasting problems within poorer communities that inhibit the possibility of education providing a way out of poverty [37]. Reducing the financial strain placed on schools is therefore imperative. Relief could come by employing alternative energy generation methods such as renewable energy sources, providing a pathway to sustainable development.

Schools provide the potential benefit of having a load-profile suited for solar

intervention with electricity usage coinciding well with solar generation curves. Furthermore, combining the benefits of EWH management and load-matched solar PV generation in a school environment can potentially provide large reductions in monthly demand charges by reducing a school's peak power usage while maintaining user comfort [40]. This will also provide the opportunity to divert excess generated solar energy to be stored in the EWH's, reducing the potential grid load used by the EWH's and maximising the effectiveness of a solar system to allow for the greatest economical benefit [30]. However, achieving this requires a deep understanding of a building's energy usage throughout hours, days, and varying seasons to correctly size a PV system [41], and can be achieved by installing expensive smart meters to measure hourly energy usage over a long representative period.

1.3 Thesis problem statement, hypothesis and objectives

Within the context set out in the previous section, this section presents the problem statement to be investigated during this study, as well as the objectives that need to be met in order to address the problem.

1.3.1 Problem statement

The educational gap within socio-economic groups in South Africa is immense. Learners within no-fee schools (the poorest 60%) are at a disadvantage with most of the available funding being used for personnel salaries. As the energy costs within South Africa rise, the remaining non-personnel funding will decrease, limiting spending on teacher-support materials including textbooks, feeding schemes, stationary and internet access. This places these schools at a disadvantage and constrains the ability of the education system to provide a pathway out of poverty for poor children. Schools are billed on commercial and industrial tariff structures. These tariffs include being billed for basic energy usage, as well as peak electrical usage within a calendar month. Reducing the peak demand can lead to large reductions in monthly utility costs. EWHs usually use large amounts of electricity throughout the day in order to maintain a suitable operating temperature, and their usage spikes can increase the monthly peak demand of a school, leading to unnecessary expenditure. However, since EWHs can store energy as a thermal battery, correctly timed heating schedules will ensure sufficiently warm water while not increasing the monthly demand limit. This, in addition to an optimally sized solar system may give schools an opportunity to reduce their monthly utility spending. Still, many organisations still lack the necessary resources required to determine the benefits of renewable energy solutions due to the complexities of accurately determining the feasibility of a project.

1.3.2 Proposed solution

This study aims to address the above problem statement by developing a (1) computationally inexpensive, accurate forecasting model using easily accessible input parameters, (2) to use the forecast to develop a method capable of sufficient load-matching using a solar model, (3) to combine the thermal-storage capabilities of EWHs with load-matched solar through smart-scheduling techniques to reduce peak monthly demand and basic energy usages, and (4) to provide accurate cost and savings forecasts to determine the viability of the energy-saving intervention at a particular school.

1.3.3 Hypothesis

A system can be developed to simplify the process required to accurately determine the benefits of combining solar PV and load-shifting interventions, and provide accurate financial viability forecasts for schools using easily accessible input parameters.

1.3.4 Research objectives

A set of comprehensive research objectives were selected in order to perform the work necessary to prove the hypothesis. The research objectives were:

Research objective 1:

To develop a computationally inexpensive and accurate model capable of forecasting the energy usage and maximum monthly demand of a school using easily accessible parameters.

Research objective 2:

To develop a method capable of determining the energy usage of a school after solar intervention using a cost optimised, load-matched solar system.

Research objective 3:

To combine the thermal-storage capabilities of EWHs with load-matched solar through smart-scheduling techniques capable of reducing peak monthly demand and basic energy usages.

Research objective 4:

To provide cost and savings forecasts for various system sizes and interventions.

1.4 Contributions

This paper makes the following contributions:

- Presents a comprehensive system capable of estimating the potential financial viability of solar and EWH interventions within a school environment using easily accessible input parameters, removing the need for smart meters and time-intensive analysis.
- Presents a computationally inexpensive method for forecasting time series data with multi-dimensional seasonality components.
- Presents a method capable of using a verified solar model in conjunction with an energy demand forecast to estimate the potential financial return of a solar system.
- Development of a computationally inexpensive method to assess the performance of EWH control schemes. The algorithm considers the needs of both the customer and the utility and is capable of diverting any excess generation to heat EWHs using a prioritised heating mechanism.

1.5 Thesis structure

Chapter 2 presents literature relevant to the development of a system able to support the study hypothesis. Data-driven forecasting techniques are investigated with priority given to methods using time-series data. Additionally, relevant studies pertaining to load-matched renewable energy systems and demand-management within buildings are also investigated. This chapter concludes with identified challenges with regard to the development of a system capable of adhering to all research objectives.

Chapter 3 describes the research methods and design choices implemented during the development of the system, comprised of three separate components namely *energy usage forecasting*, *solar system optimisation* and *intelligent electric water heater control*.

Chapter 4 defines the experimental bounds and parameters used to simulate the developed system. The chapter introduces with data cleaning measures used to ensure that the measurement data sets used to test the system is suitable. Thereafter, the simulation setup is described with all parameters and constants defined. The chapter concludes with metrics used to evaluate each of the three system components.

Chapter 5 presents the results of the system simulation. Firstly the forecasting and solar optimisation techniques are explored with results shown for five schools with real measured datasets. Secondly the EWH control schemes are evaluated by performing a complete system case study, whereby real energy and hot water usage data from a single school is used to complete a full system simulation. The chapter concludes by expanding on the computational efficiency of the system implementation with simulation times for potentially larger datasets.

Chapter 6 concludes the study by evaluating the four research objectives with regard to the system simulation results. Finally, recommendations for future work regarding energy saving interventions and data-driven forecasting techniques are presented.

Chapter 2

Literature Review

Schools in South Africa are expected to pay increasing electricity costs imposed on them by their municipality and governed by the national energy regulator, inhibiting them from obtaining the resources necessary to improve the quality of education delivered. This chapter presents literature that is relevant to the development of a forecasting method capable of providing accurate demand profiles for schools without the need for expensive smart meters.

Additionally, relevant studies pertaining to load-matched renewable energy systems and demand-management within buildings are discussed to facilitate progress towards research objectives defined in Section 1.3, and were evaluated in a critical manner to ensure that the methodology presented later in this study agrees with published research.

2.1 Data-driven energy usage forecasting

The ability to accurately forecast energy usage has become a widely discussed topic during recent years as the interest in energy saving within today's energy conscious society continues to grow. Data-driven forecasting methods use measured data and building characteristics to predict the future usages using specialised statistical approaches. The effects of reliably determining a building's energy usage are far reaching, as arrangements can be made in advance to mitigate demand during peak times, resulting in large monetary savings [42].

2.1.1 Data cleaning

A forecast is dependent on the data quality used to build a demand forecasting model, and without effective processing the model will yield unsatisfactory results. Data cleaning is a subset of data-preprocessing and a necessary step in the data mining process whereby the original data is prepared and transformed into a dataset suitable for use as a training set [43]. The process aims to

find meaningful relationships while reducing the size using several techniques including filling missing data and outlier detection.

2.1.1.1 Missing data

Missing data can be a result of equipment failure, power outages or connectivity issues and is a frequent occurrence when working with large datasets, especially those reliant on electronic measurement devices [43]. A general rule within data science is to invalidate the dataset if more than 20% of attributable values are missing [44]. However, if the measurements are only partly lost, several techniques are available to restore them.

Simple filling techniques: Simple filling techniques include manually filling the missing values. This is a time consuming process but may be effective when working with very small datasets. Another method is to use a global constant as a replacement value. A large drawback to this method is that it may not provide a useful description of the missing value, while adding complexity to deep learning methods that may place extra weight on features associated with the "constant".

Computed filling techniques: These filling methods use available data from the dataset to fill the remaining entries. An example of such a method is to replace the value of a missing entry by the average value of the attributes associated with that specific entry. A more computationally expensive filling method is to use the most probable value to fill the missing entry. This can be done through techniques including Bayesian formalism or by using a decision tree [43].

2.1.1.2 Smoothing and outlier detection

Noisy data presents one of the greatest problems facing data scientists [44]. When outliers are present in a dataset the data will deviate from the expected norm. Removing anomaly data can be done using multiple methods including regression and clustering [43].

Regression and exponential smoothing: This is done by fitting the available data to a function, and through a best-fit technique it can be used to predict missing values [43]. Time series data is smoothed using a lag operator, whereby the weights of past observation decrease exponentially, allowing for more accurate forecasts [45]. The method also detects any anomaly entries by adding buffers above and below the smoothed forecast line. If a data point appears outside of the buffer, it is invalidated.

Clustering: Using this technique, sets of data points are grouped together based on their attributes. For time series data this can be sectioned into seasonal categories to easily detect outlying entries. An ideal clustered set contains many data points with small variations, separated by a large variation from the next clustered set [43].

2.1.2 Time series modelling and forecasting

After preparing the original data the process of data-driven forecasting can commence. First, the model must be trained using an algorithm that receives the pre-processed dataset. Two types of data-driven models exist, single and combined models [46].

2.1.2.1 Data-driven forecasting models

Single models provide data training and analysis using only one predictive algorithm. These algorithms include conventional methods utilising regressive and statistical calculations. Additionally, classification based procedures and deep learning techniques are also used. Combined models focus on the improvement of the forecasting techniques by using multiple single model algorithms, providing the highest forecasting accuracy at the cost of increased computational complexity [46].

2.1.2.2 Conventional modelling and forecasting methods

These methods include linear and autoregressive models and are widely used for forecasting time series data due to their ease of use and forecasting accuracy. These approaches can offer the best results if the training dataset provides linear trends, and does so while remaining computationally inexpensive.

Linear regression (LR): This technique is used to gain a better understanding of the relationship between a scalar response (or dependant variable) and a predictor (or independent variable). If more than one predictor is present, the process is known as multiple regression modelling, and is often used for forecasting from small datasets [46].

Koen and Holloway [47] completed a study investigating South Africa's electricity demand, and developed a modelling methodology using multiple regression to derive forecasts to be used for strategic planning. Changes relating to economic and demographic growth for various scenarios were used as predictor values, and helped filled the gap within the long-term electricity forecasting domain.

Zhu *et al.* [48] developed a data-driven process capable of using smart metering data to create daily load profiles for various building types, and accounts for weather and holiday periods. The constructed profiles built using regression

algorithms are then compared to the observed usages to detect energy usage anomalies. The process required smart-meters to be installed at all times, and was only able to provide reliable detection results for a specific building once a large number of data points have been measured.

Autoregressive modelling (AR): This is the most frequently used time-series forecasting method. Autoregressive Moving Average (ARMA) modelling is an method capable of using past values to forecast future usages. If the training set is not stationary, meaning that the values of the dataset are time-invariant, an initial differentiation (the "integrated" term) is necessary, producing an Autoregressive Integrated Moving Average (ARIMA) model. This method has been proven to provide accurate estimations for trend-based datasets, but does not support time series with a seasonal component [49]. Seasonal Autoregressive Integrated Moving Average (SARIMA) modelling methods solves this by supporting the direct modelling of a seasonal component within a time-series. A drawback of SARIMA based models is their computational complexity with regards to the large datasets required to accurately calculate the trend and seasonal elements within the time series, necessary during each "re-training" of the forecast model.

Heylman *et al.* [45] investigated energy usage trends from over 200 buildings on the University of Virginia's campus to determine the most effective techniques for forecasting building energy usage. The study also examines the clustering of buildings based on their energy usage trends rather than the building's functional use. It was determined that one year's hourly usage data was insufficient to accurately predict long term usage patterns. Short term SARIMA models used for 2-day forecasts accurately determined the overall usage trend, but failed to capture demand fluctuations during the day as the hourly data points were insufficient to produce minute-by-minute forecasts. Clustering was done by grouping buildings based on energy usage with mixed results as some buildings displayed seasonal patterns while others did not, and it was concluded that additional information was required to more precisely group the campus buildings.

2.1.2.3 Classification-based modelling and forecasting methods

Simply put, classification based models perform pattern recognition within a sample dataset. These observations are then returned in a set of user defined, quantifiable properties known as *features* [46].

K-nearest neighbours (K-NN): This is a method often used for classification, and relies on similar patterns being present within time-series data that can then be classified according to their properties.

Zala and Abhyankar [50] examines the ability to design a time-of-use tariff for three major sectors (residential, commercial and industrial) by decompos-

ing the measured load known to the electricity supplier using representative load profiles and the elasticity of different types of loads. The study developed a k-means clustering method to find the usage patterns of consumers without the need for expensive smart-meters. However, the method was required to subdivide each sector into smaller sub-sectors as usage patterns still varied within each tariff type. These variations were apparent but still accurately identified the individual major sector usages, and could possibly be applied to cluster individual rooms and buildings within a larger load measurement.

K-means modelling is relatively simple to implement and shows good forecasting accuracy. However, K-NN algorithms can become exponentially complex as the training dataset grows.

Decision trees (DTs): Decision trees are constructed from successive nodes with various conditions [46]. At each node a test is performed considering a specific condition of the input dataset, and successive tests are performed until the input value is assigned to the most suitable category. DTs consist of a single *root node*, conditional *decision nodes*, and successive *child nodes*.

Yilmaz *et al.* [51] presented a methodology for creating representative electricity demand profiles for households in Switzerland. By examining the energy usage profiles of 656 multi-family apartments with a temporal resolution of 15 minutes, the study determined that five defining features including mean daily usage, morning-, midday-, evening- and night usages were enough to adequately define the households usage patterns. Households were then categorised into three distinct groups through decision tree learning, allowing for improved load forecasting and enhanced demand response targeting. The paper considers electricity use patterns within individual households, and states that daily profiles for a given household significantly differ from the averaged profile, and concludes that more features defining the underlying traits of households need to be explored. These feature requirements were investigated by Zhang *et al.* [52] and introduced three engineering methods including visualisation, selection and extraction. The three methods were tested using measured data from 1000 homes, evaluating 132 features to create lists containing the ten most important features. The study concluded that the three methods discussed resulted in three different feature lists. However, certain features appeared in all rankings, and by increasing the detail with which these elements are trained using machine learning models, the accuracy of the predictions improved.

Decision trees present a major advantage as a forecasting technique due to it being simple to understand and interpret through white box modelling. Additionally, the algorithm is computationally inexpensive and performs well with large datasets [46]. A limitation of DTs is that they aren't very robust, and variations in the training data can result in large changes in the DTs prediction, reducing the accuracy of forecasts.

2.1.2.4 Deep learning modelling and forecasting methods

During recent years there has been a shift towards new modelling techniques in the field of data-driven forecasts [49].

Artificial neural networks (ANNs): This is a machine learning technique inspired by the biological neural network that constitute a human brain. The method is capable of forecasting non-linear patterns through "learning" from repeated examples and the technique has been shown to accurately predict electricity usage [46]. However, ANNs require very large datasets to perform reliable predictions that are only marginally better compared to regression methods while requiring large amounts of processing power [46].

Selecting a suitable algorithm to train a forecast model is a crucial step necessary for accurate predictions as the applicability of each algorithm varies on a case-by-case basis [49]. Once defined and trained, it can be tested using unseen data as to provide an unbiased evaluation of the model.

2.1.3 Validation

When working with large amounts of data, a validation dataset is used. It is created by dividing the measured data into three parts, a test set, training set and a validation set. It is useful for tuning key modelling parameters to improve the model's performance, as well as for selecting the appropriate forecasting model if many are compared [46]. If only small measured datasets are available the validation set will be forfeited and the measurement data will be re-partitioned into a training set and test set [49].

2.1.4 Testing

The last step of data-driven forecasting is the testing phase whereby a test dataset is used to evaluate the performance of the trained model. A general rule of thumb is to divide the measurement data such that 60%, 20% and 20% is allocated for training, validation and testing respectively [46], or in the case of smaller datasets 70% for the training set and 30% for the test set [49]. The performance of a forecasting model is determined by its ability to produce usage estimates that closely relate to the actual measured demand, and is quantified by using accuracy metrics [49].

2.1.5 Summary of energy usage forecasting

Although various load forecasting techniques have been discussed [45; 47; 48; 50; 51], none of the solutions present a method for forecasting individual building (or property) load through easily accessible input parameters, essential for

the South African educational context in which smart-meters are not commonly used. Access to smart meter data (where used) is difficult to obtain, and the requisite skills for detailed simulations and available resources (time and money) at government level are limited. Additionally, some of the existing works [45; 50] performed the forecasting using computationally expensive methods, unsuitable for schools wanting immediate forecasts that are readily available (eg. web interface or mobile phone).

2.2 Load matching and grid interaction

Optimising the size of renewable energy systems has gained interest during recent years as the awareness of energy saving and clean energy continues to grow. Load matching (LM) presents many challenges as it involves the interaction of two dynamic phenomena, building energy demand and on-site generation.

2.2.1 Grid interaction

Salom *et al.* [53] analysed five case studies to determine a set of quantitative indicators enabling grid designers and operators to better understand the impact of net zero energy buildings on the local grid. This was done by determining the import/export relationship between the individual buildings and the surrounding grid, commonly known as grid interaction. The study found that three indicators, the load-cover factor, supply-cover factor and the loss-of-load probability factor, used to determine the likelihood of the on-site generation failing to supply electricity provided the most insight. The load-cover describes the percentage of electrical demand satisfied by the on-site energy system and the supply-cover factor describes the percentage of on-site generation used by the building. These indicators proved useful for providing insight into the daily and seasonal capabilities of various renewable energy systems and their respective grid interaction. The authors concluded that the indicators can prove beneficial to grid designers in the near future, but that further studies are needed to determine reference indicator values for certain building types, topologies and climates.

2.2.2 Optimised on-site generation

Maleki *et al.* [54] presented a simplified renewable energy sizing method for schools by generating demand profile estimations for specific building layouts. During their research various field studies were conducted to gain an understanding of school topologies and energy usages. From this information a series of equations were created relating to school population and building size. These equations were capable of determining the optimal energy system

size without the need for intensive simulations techniques. A drawback to the technique used is the homogeneous nature of South Korean schools where the study was conducted. Information presented in the study confirmed that most of the classrooms were similarly sized and seated a similar number of students, far different from the education standards in a developing country such as South Africa.

Yang *et al.* [55] developed an optimal design method for a hybrid solar-wind system needed to power a continuous-load telecommunication station. The method is based on a genetic algorithm with two goal parameters including minimising the loss-of-load probability and the annualised cost of the system. The method provided favourable results with a high system utilisation rate while the battery's state of charge remained above 50% for most of the one year field test. An advantage of the method described includes taking into account system design characteristics including PV modules specifications, slope angle and the available installation area as these affect the resulting energy production and system installation costs.

2.2.3 Summary of load matching and grid interaction

Optimising on-site generation through load matching allows for reductions in system costs, grid usage and loss-of-load probability. However, the performance of a load-matched system is burdened by a building's energy usage patterns, and can be improved by altering the demand profile to better fit the generation of the renewable energy system. Load matching and grid interaction studies were investigated [53; 54; 55], providing key insight into feature selection and building characteristics required to optimally size an energy system. However, only one of the studies Yang *et al.* [55] present a method for optimally sizing renewable energy systems from a load profile necessary for determining more accurate cost versus benefit evaluations, and only does so using a fixed load, much different from the constantly fluctuating demand present in a school environment.

2.3 Demand management within buildings

Demand-side-management (DSM) is a process during which the behaviour of consumer's demand for energy is modified. Successful DSM reduces the strain placed on the local electricity supplier during peak times, and enables for more reliable usage forecasts that can be used to improve the performance of renewable energy systems to better fit demand [53]. However, effective DSM requires a thorough understanding of the varying needs and usage patterns of consumers that differ dependant on multiple factors including lifestyle, building type and location, and adjusting to the needs of the customers is an essential requirement for any successful DSM program [32].

2.3.1 DSM within a school environment

Schools generally accommodate many people during the daytime and use large amounts of resources from utility companies [56]. Gallego Sanchez Torija *et al.* [56] considered the incorporation of water usage in energy audits within a school environment. The objective of the research was to measure the water usage and potential emissions per user within the school building, as well as the potential improvement due to a series of proposals aimed at reducing water usage while maintaining user comfort. By altering the flow rate of the taps installed within the school the researchers were able to reduce water usage by up to 82% with a project payback period of less than 2 years while maintaining user comfort.

Mauri *et al.* [57] evaluated a school situated in a rural, off-grid area with limited energy generation and storage capabilities. The research focused on optimising the school's power delivery systems to match the building's demand using a smart scheduling energy management system (EMS). The EMS was able to effectively match the building's usage but relied heavily on an expensive backup battery system when the renewable energy system was unable to keep up with demand. This provides a challenge for buildings situated in rural areas, as battery storage systems are expensive [33]. A more convenient solution would be to use existing infrastructure to manage the demand of buildings.

2.3.2 Load Shifting using EWHs

Managing the power demand of buildings can not only reduce their monthly energy costs, but the resulting carbon footprint as well [33]. EWHs provide an opportunity for intervention as they are present in nearly all building types, and can be used for energy storage as a thermal battery [58]. Different forms of centralised EWH control have been proposed, with three key control objectives, namely, cumulative energy usage reduction, efficiently managing the load, and user comfort [32; 33; 58]. Direct control provides the benefit of being able to reduce the strain on the grid during high demand periods, but does not take the comfort of the user into account. Combining the benefits of EWH management and load-matched solar PV generation in a school environment can potentially provide large reductions in monthly demand charges by reducing a school's peak power demand while maintaining user comfort. This will also provide the opportunity to divert excess generated solar energy to be stored in the EWH, reducing the potential grid load used by the EWHs and maximising all three control objectives.

Utilising EWHs for DSM requires physical models of the water heater's dynamic behaviour during heating, cooling and usage periods. Kepplinger *et al.* [33] compared the performance of a computationally inexpensive single-node model and a two-node model with thermocline tracking. The thermocline is the layer mixing the warm and cold water nodes within the EWH and moves

along the height of the EWH tank when hot water leaves and cooler water enters during use [32]. The study found that both models accurately reproduced the real measured results with the single-node model only performing marginally worse.

In South Africa the electric water heaters within households in multiple cities were retrofitted with control systems by Roux *et al.* [32], allowing the energy provider to switch off the EWHs during peak hours to reduce the stress on the supply by means of direct control. The proposed algorithm uses a load spreading approach in which EWHs compete for access to the available power. The method takes grid load limits, real-time temperature, water usage patterns and user comfort into account, and was capable of reducing buildings' peak demand while maintaining user-comfort levels beyond that of the commonly used ripple control.

2.3.3 Summary of DSM techniques

Various DSM techniques within school environments [56; 57] and load shifting methods using EWHs [32; 33; 58] were discussed. Still, none of the solutions present a peak monthly demand-limiter through intelligent EWH control, crucial for reducing demand charges at institutions billed on a demand-based tariff structure, combined with an optimally sized solar system for maximal utility bill reduction. Many of the proposed solutions were created for developed countries, unlike South Africa, where effectively managing the available energy supply is essential. Additionally, none of the existing interventions implement user-specific EWH control schemes using real-time usage data to maximise utility bill and carbon emission savings while maintaining user comfort.

2.4 Chapter summary

Chapter 2 reviewed relevant literature regarding energy usage forecasting, grid interaction and DSM using EWHs. The method and shortfalls of each literature component was discussed and highlights a set of considerable challenges that need to be addressed in order to achieve the research objectives defined in Section 1.3.

Chapter 3

System development

The objective of this study is to develop a system capable of simplifying the process required to accurately determine the benefits of solar and load-shifting interventions, and provide accurate financial viability forecasts for schools using easily accessible input parameters. This chapter discusses the research methods and design choices implemented during the development of the system, comprised of three separate components namely *energy usage forecasting*, *solar system optimisation* and *intelligent electric water heater control*.

3.1 Energy usage forecasting

In this section the development of a load forecasting method that uses monthly utility bills to produce an energy usage profile of a school is assessed.

3.1.1 Electricity demand within schools in South Africa

As of 2017 there are more than 26 000 public schools in South Africa, educating 12 million students [59]. Schools are structured in 5 quintile categories catering to students from the poorest 20 percent (quintile 1), to the most wealthy 20 percent (quintile 5) of the population. This presents an opportunity for intervention as schools use a large portion of national resources, and can be used as a means of communication towards learners and their families promoting the benefits of energy efficiency and sustainable development [60].

For this study the measured data of high income (quintile 5) schools were used. These schools are all located in Stellenbosch, South Africa and were chosen as they had municipal smart-meter data available for multiple seasons. From the available measured data the two types of schools were classified as primary and high schools.

3.1.1.1 Tariff structures

Most schools in South Africa are billed on one of three tariff structures that include a commercial fixed (basic) charge, industrial demand and time of use (TOU) tariff structure [17; 61]. They are formulated as follows:

Commercial fixed: This is the most commonly used of the three tariff structures, especially in residential sectors within South Africa. For a commercial fixed tariff a customer is charged a flat energy rate and billed for their usage at the end of each month.

Industrial demand: This tariff structure is unavailable for customers within the residential sector and only suitable for entities consuming a large amount of energy. A reduced basic energy rate and a fixed monthly charge is imposed in addition to a maximum demand charge, during which customers are charged based on their peak monthly demand.

Time of use: This tariff structure promotes customers to use more energy at off-peak times, while reducing their usage during peak times to lessen the strain on the local grid. TOU charges vary during hours of the day, defined by the national energy regulator as standard, off-peak and peak hour rates. The tariff charges also differ for the warmer low-usage season and the colder high-usage season. Unlike the commercial fixed structure TOU is also imposed a monthly reading cost as well as a small demand-charge and is only suitable for customers consuming large amounts of energy.

3.1.1.2 Energy usage profile

A property's electrical usage profile provides great insight into the behaviour of its inhabitants, and by analysing a school's usage, careful considerations can be made to introduce effective DSM through cost-saving interventions.

Schools use the majority of their energy within the daytime when students are present. Figure 3.1 shows the usage profile of a school during a typical school day. A small usage surge appears in the early morning due to bore-hole pumps being used to water the sport's grounds as well as EWH heating elements being powered. A small reduction then appears as the night-time lights and pumps are switched off. Once students and teachers start to arrive a sudden increase appears as computers, classroom- and office lights and air conditioning units are switched on. This remains relatively constant with two intra-day spikes appearing during the break periods when teachers use appliances to cook meals and heat water. Once school ends a sharp decline in energy usage can be seen, and continues to decline as teachers leave the classroom, switching off lights and air-conditioning units or fans. Once all the

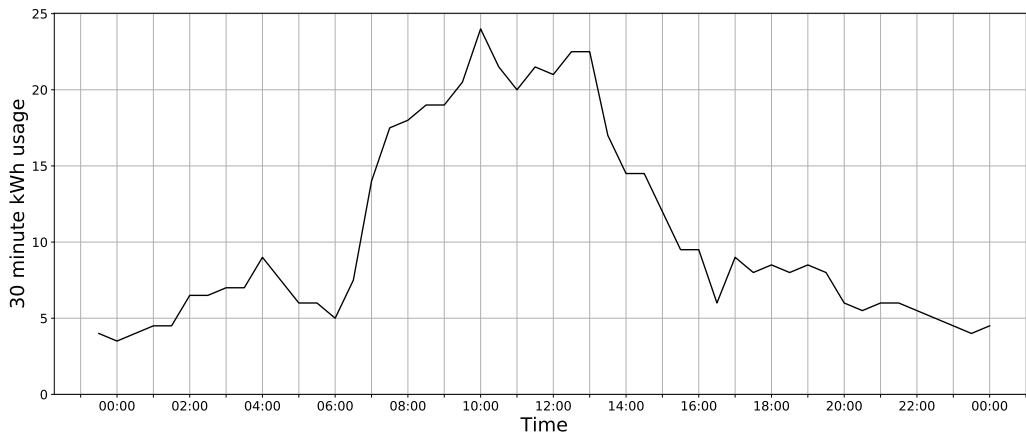


Figure 3.1: Hourly electricity usage of a school (670 students).

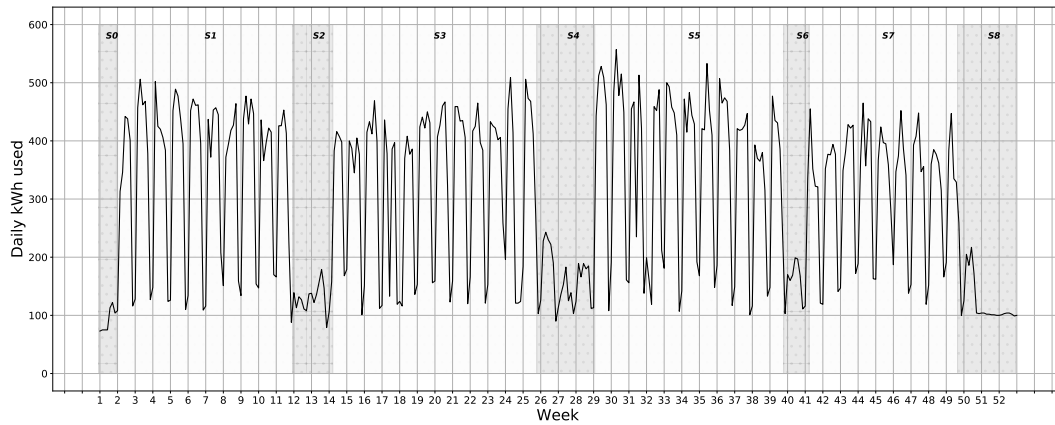


Figure 3.2: A typical school's daily energy usage profile over a 1 year period with seasonal periods highlighted.

teachers and students have left the building the usage profile returns to its base load.

Public schools in South Africa operate on a yearly schedule consisting of four school terms with predetermined dates. Figure 3.2 presents the daily kWh used for a typical high-income public school in South Africa. The highlighted sections labelled S0 to S8 represent the four, "in-season" school terms (S1, S3, S5, S7) and the remaining "off-season" holiday periods. From the figure the weekday spikes during the school terms are clearly visible, with the dips in demand showing the weekend's usage. A clear drop in usage during the holiday periods are visible, giving an indication to the seasonality of a school's electricity usage.

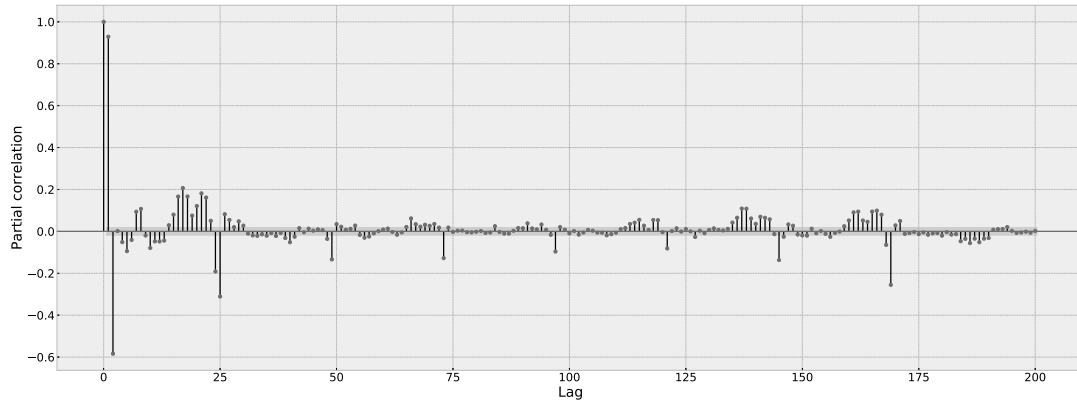


Figure 3.3: Partial correlation of a school's measured usage data showing surges occurring on 24-hour and 7 day (168 hour) intervals.

3.1.2 A review of techniques discussed in literature

Time series analysis encompasses a field of study related to extracting valuable information from time series data. It differs from other data analysis methods due to it having a distinct temporal order, with all data values accompanied by a timestamp [48]. Chapter 2 reviewed various methods employed while performing data-driven time series forecasting, each suitable for a particular case. This section describes the model selection process during which the reviewed methods were scrutinised to obtain the most suitable model for forecasting energy usage within schools while adhering to the research objectives defined in Section 1.3. Figure 3.3 presents the autocorrelation function of a year's hourly energy measurement data from a school, and gives insight into the daily and weekly seasonality patterns within the data with correlation spikes appearing in intervals of 24 hours and a large spike occurring after 7 days (168 hours).

Due to the requirement for the process to remain computationally inexpensive, deep learning method techniques such as artificial neural networks, and K-nearest neighbour classification techniques were not considered. These methods would have been suitable for complex models requiring little to no adjustment, unlike a process where new measurement data is constantly added. Additionally, forecasting a building's energy usage from previous monthly utility bills presents a challenge as most forecasting techniques require historical usage data to provide accurate forecasts.

Linear regression is a technique used to model the relationship between two variables, and can be applied to quantify the strength of the relationship when new data points are added. From this relationship a trend line can be drawn and when it's applied to a school's energy profile it can show the long-term increase or decrease in usage. Applying this to the monthly bills input will allow for trends to become visible on a long-term basis. However, LR models lack the ability to manage the multiple seasonal components within a time series, necessary for energy profiles within schools.

Autoregressive time series modelling is an improved forecasting method

when compared to the linear regressive technique. SARIMA modelling is an autoregressive model supporting seasonal data. However, school usage data contains multiple seasonality components for daily, weekly and term-length periods and is not supported by SARIMA methods. This is solved by using a TBATS model (Exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components) with no seasonality constraints [62]. Applying the historic electricity demand of a school to the model will allow for a base profile to be trained and fit according to a set of characteristics. The usage from schools with these characteristics can then be forecast using the monthly bills. However, an crucial disadvantage to the TBATS method is the inability to account for external variables including term-dates and holidays, meaning that the model will only be fit for a specific calendar year and unable to account for shifting term and holiday dates.

Decision-trees employ a flowchart-like structure containing conditional control statements to separate the original dataset into smaller groups through classification techniques until they are small enough to be described by some label [63]. This allows for shortcomings of previously discussed methods to be mitigated through its *if..then..else* design, as it will be able to adjust to external conditions while remaining computationally inexpensive. As such, decision trees (DTs) will provide the necessary procedural arrangement to distinguish the multiple seasonal elements of a school's energy usage profile.

3.1.3 Forecast model development

Most forecasting models require large datasets to ensure a degree of certainty when predicting future outcomes through commonly used regressive methods. Unfortunately this is not always possible as smart-meters necessary for data capturing are expensive and not common in residential or commercial settings in South Africa. The proposed forecasting method solves this by minimising the computational complexity and required input parameters of the forecast. It does this by converting hourly measured smart meter data into easily usable scalar lists using classification. The input parameters needed to decompose the time series are: the measured energy usage data, the school type (primary or high), school term dates and public holidays. The proposed 3-stage approach is summarised as follows:

1. Generate normalised seasonal scaler lists: This step uses smart electricity meter data from a limited subset of schools in combination with the specified input parameters to perform a time-series decomposition, which is used to determine a normalised seasonal scaler lists that can be used to construct a base profile for a school type.

2. Construct normalised load profile: This step creates a normalised load profile for a school type from the specified term and holiday dates for a certain year consisting of normalised hourly demand values.

3. Scale normalised load profile to monthly bills: The normalised, *base* profile is scaled to a representative hourly usage using the school's monthly bills, which capture the cumulative totals and maximum peak loads.

3.1.3.1 Generate normalised seasonal scaler lists

Firstly, the days of the measured data year is divided into nine seasonal periods, presented by the date-ranges S_0 through S_8 in Figure 3.2, indicating the four school terms and five vacation periods throughout the year.

The vacation periods succeed each school term, with the other period, S_0 , occurring between New Year's day and the first day of school. From the figure, distinctive usage patterns can clearly be seen on a typical daily usage profile over a year, giving an indication to the strong multi-dimensional seasonality within weekly and term-length time frames. Energy usages during individual public holidays are not considered when constructing the base profiles as the process can be simplified by averaging the energy usages during the weekend as their demands were determined to be similar.

Next, the hourly usages for each of the seasonal periods are separated for each day of the week. If a weekday occurs more than once during a seasonal period the usages are averaged for that specific day and hour. New lists now contain the measured hourly usages for each weekday of each seasonal period, containing 24 hourly entries for each of the 7 weekdays.

These usages are normalised to contain the energy usage relative to the entire day's, resulting in a normalised profile, U_S , describing the usage for each hour of each day of the week, for each of the nine seasonal periods, S .

$$U_S = \begin{bmatrix} \frac{E_{0-0}}{(\sum_{i=0}^{23} E_{0-i})} & \cdots & \frac{E_{0-23}}{(\sum_{i=0}^{23} E_{0-i})} \\ \vdots & \ddots & \vdots \\ \frac{E_{6-0}}{(\sum_{i=0}^{23} E_{6-i})} & \cdots & \frac{E_{6-23}}{(\sum_{i=0}^{23} E_{6-i})} \end{bmatrix} \quad (3.1)$$

The process can easily be applied to large datasets. If new data is added the same steps outlined previously are followed and averaged with the existing hourly usage ratios, in effect, re-scaling the seasonal proportions and smoothing the usage profile to be better suited for a greater number of schools. If enough data is available these profiles can be expanded based on other characteristics including building layout and location. The training process did not take into account the proportional relationship of each day in month, as this would require a very large dataset to account for the minor daily variations to avoid being over fit for a measured profile. Instead, the average total usage for each day of the week for every seasonal period is calculated and presented as:

$$D_S = [d_0 \ d_1 \ d_2 \ d_3 \ d_4 \ d_5 \ d_6] \quad (3.2)$$

with d_0 to d_6 being the averaged total daily energy usage for Monday to Sunday of each each seasonal period from D_0 to D_8 .

3.1.3.2 Construct normalised load profile

Once the seasonal lists containing scaler values for the expected normalised hourly usage of each day of the week has been obtained, a normalised load profile for a specific type of school can be built using a *YearConstructor* function. The method uses the nine U_S lists to construct an unscaled version of the specific year's forecast for individual days, using the term dates and public holiday input parameters. For public holidays, the preceding weekend's usages are averaged as stated. Once completed, a standard year would represent all U_S lists summing to 365 entries, and is presented in Algorithm 1. The constructed usage profile consisting of normalised hourly usage values ranging from 0 to 1 can then scaled using the aggregated usages listed on the monthly utility bills.

| | |
|----|--|
| | Result: An hourly normalised load profile for a specific school type |
| 1 | Inputs: year, term dates, holiday dates, U_s ; |
| 2 | begin |
| 3 | if not leapyear then |
| 4 | Get school type; |
| 5 | for 8760 hours do |
| 6 | Determine day of the year (1-365); |
| 7 | Determine day of the week (1-7); |
| 8 | Determine hour of day using (0-23); |
| 9 | Determine season from specified term dates (S0-S8); |
| 10 | Get U_s for that season; |
| 11 | if Holiday then |
| 12 | if Weekday then |
| 13 | Assign U_s scaler to hour of day from an averaged Saturday and Sunday value; |
| 14 | else |
| 15 | Assign U_s scaler to hour of day for the specific weekend day; |
| 16 | end |
| 17 | else |
| 18 | Assign U_s scaler to hour of day for the specific day of the week; |
| 19 | end |
| 20 | end |
| 21 | else |
| 22 | Repeat entire process for 8784 hours; |
| 23 | end |
| 24 | end |

Algorithm 1: *YearConstructor* function.

3.1.3.3 Scale normalised load profile to bills

Schools in South Africa are billed according to one of three tariff structures as discussed in Section 3.1.1. To forecast the usage of a school its monthly bills are used to scale the hourly normalised profile.

Scaling is done by using matrix multiplication to determine a year's load profile with estimated peak- and basic usages using only the relevant monthly

bill components and the school's term dates as input parameters. However, the scaling method does not depend on all twelve bills being available. It requires at least one monthly bill, and scales the value to each of the remaining months using the months for which bills are available. The factors by which the months are scaled are calculated from the relational proportion of each month's energy to every other month's energy usage obtained from the measurement data. This results in 12x12 matrices containing scalars between 0 and 1 for basic energy usage, maximum monthly demand and the three time of use periods structured as follows:

$$M_r = \begin{bmatrix} r_{0,0} & \dots & r_{11,0} \\ \vdots & \ddots & \vdots \\ r_{0,11} & \dots & r_{11,11} \end{bmatrix} \quad (3.3)$$

where M_r represents a scaling matrix and $r_{x,y}$ is calculated as the total energy usage for month x divided by the total energy usage for month y . The total monthly energy usage required to scale the normalised profile is shown as:

$$B_{Energy} = [b_0 \quad \dots \quad b_{11}] \quad (3.4)$$

with b_E being the total energy consumed within each calendar month from January until December (month's 0 to 11). If no monthly bill is available for a certain month b_E will be 0. Performing matrix multiplication and dividing the result by the total monthly bills will result in a 1x12 matrix containing scaled usages for all 12 months.

$$B_{Scaled} = \frac{B_{Energy} \cdot M_r}{n} \quad (3.5)$$

with n being the number of monthly bills used.

In total there are five such scaling matrices. For a basic commercial tariff structure the scaling is done by estimating the energy usage for all months using M_{basic} to scale the given monthly usages to the remaining months. The same process is followed for a demand based tariff with the addition of M_{max} scaling maximum monthly demand. For a time-of-use tariff the off peak, standard and peak times are scaled using the $M_{offpeak}$, $M_{standard}$ and M_{peak} matrices respectively, resulting in a scaled hourly energy usage forecast as the method output.

After the monthly bill usages have been extended to all months of the year the normalised profile can be scaled. The process is shown in Figure 3.4.

First, the averaged energy usage for each day of the week of the applicable seasonal period is obtained from D_S and used to construct a *primary* monthly demand profile for every month. It is assumed that each individual week day (Monday through Sunday) of every seasonal period (S_0 to S_8) will have the

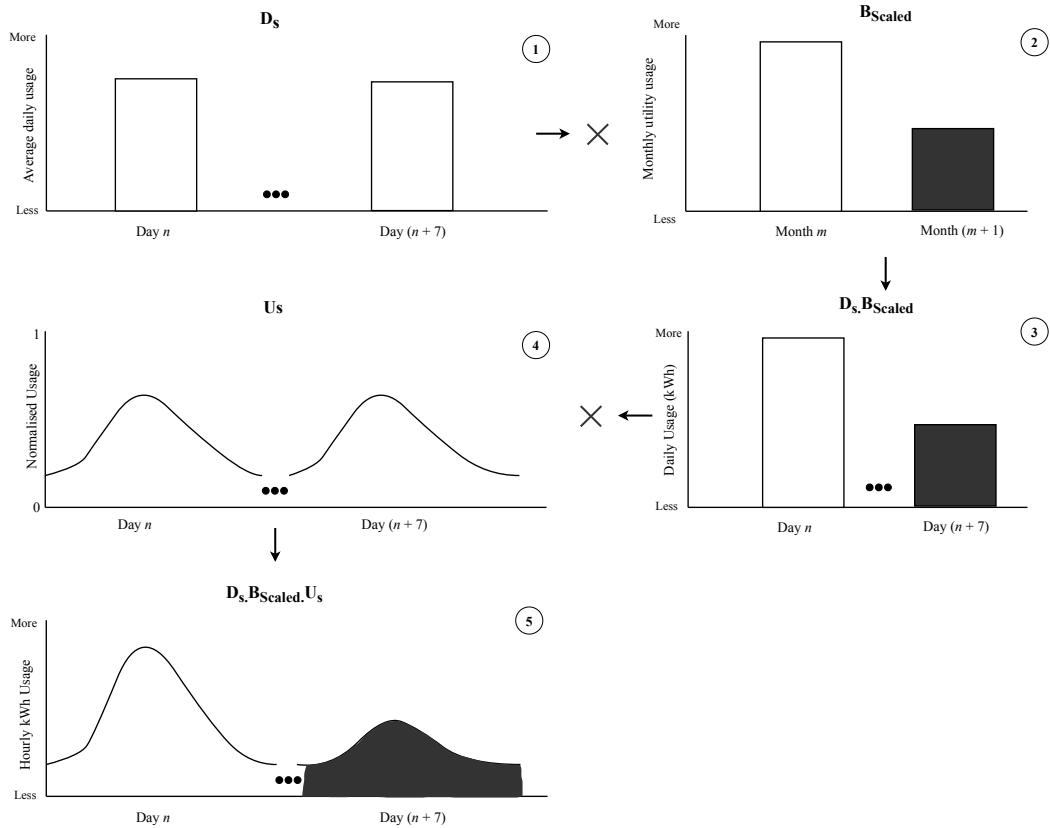


Figure 3.4: Simplified diagram of normalised load profile being scaled using monthly utility bill usages.

same normalised usage profile as determined by U_S and same total usage as determined by D_S . Simply put, if two successive weeks are of the same seasonal period their forecasted demand profile will be identical.

Next, the individual daily usage within a month is divided with the total usage of the monthly *primary* profiles to obtain the proportional usage of each day of the month. This proportion is then multiplied with the appropriate B_{Scaled} scaling usage to effectively split the total monthly usage amongst the month's days.

Finally, each day of every month can now be divided into it's hourly demand profile by multiplying the day's total usage with the appropriate U_S entry dependant on the day of the week and the seasonal period.

3.2 Solar system optimisation

This section details a solar system optimisation technique to overcome the challenges in determining a grid-tied PV solution for a school without the benefit of high resolution smart meter data. In conjunction with the forecasting method discussed in the previous section this procedure is capable of determining the

Table 3.1: Weather data file description

| Symbol | Description | Unit |
|------------|---|-------------------------|
| yr | Year (e.g. 2019) | - |
| mo | Month of the year (1-12) | - |
| day | Day of the year (1-365) | - |
| hr | Hour of the day (0-23) in local time | - |
| T_a | Ambient dry-bulb temperature | $^{\circ}\text{C}$ |
| ϕ | Ground reflectance factor | - |
| D_{snow} | Snow cover depth | cm |
| E_b | Direct radiation received by a surface normal to the sun | W/m^2 |
| E_d | Radiation diffused by the atmosphere that is horizontal to the ground | $\text{W h}/\text{m}^2$ |
| E_g | Total radiation on a horizontal surface | $\text{W h}/\text{m}^2$ |

economic viability of a PV system based on the representative demand profile, geography, solar prediction and solar system model to perform a cost benefit analysis over various PV sizes.

The goal of the solar optimisation method is to maximise the profitability and utilisation of the system using an iterative technique. The scaled load profile obtained from the forecasting model contains hourly energy usage values with estimated demand peaks.

3.2.1 Data gathering

The input parameters required for the simulation is the PV module and inverter specifications, school location, allocated budget, estimated available roof area, roof tilt and azimuth angles. Once received, solar resource and temperature data contained within a weather file is automatically obtained from the European Commission's photo-voltaic geographical information system database for the specified location [64]. If the database is unreachable, a fallback method containing weather files of towns and cities in South Africa is used, selecting the nearest available location. The data fields obtained from the weather file are shown in Table 3.1 and used by the performance model to obtain location-specific forecasts [65].

3.2.2 Solar system modelling and optimisation

The solar PV generation modelling is performed using System Advisor Model (SAM) [65] open-source libraries that employ a photo-voltaic performance model, and calculates the AC electrical output for each hour over a one year period.

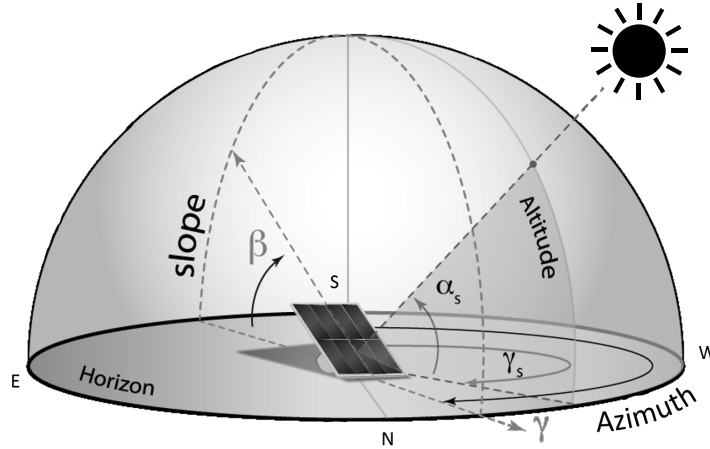


Figure 3.5: Solar module orientation relative to the sun (Brownson [66]).

3.2.2.1 Supplementary sub-model calculations

Sun position

From the building specifications and data contained in the weather file shown in Table 3.1 the orientation of the solar module surface in relation to the sun can be calculated using the following two formulas [67]:

$$\sin \alpha_s = \cos L \cos \delta \cos H + \sin L \sin \delta \quad (3.6)$$

$$\sin \gamma_s = \frac{\cos \delta \sin H}{\cos \alpha_s} \quad (3.7)$$

With α_s and γ_s representing the altitude and azimuth angles of the sun as presented in Figure 3.5. L is the latitude of the array in degrees and is gained from the weather file data. The δ is the declination angle of the sun, described as the angle between the equator and a line drawn from the centre of the earth to the centre of the sun. Approximated by [67]:

$$\delta = 23.45 \sin \left[\frac{360}{365} (\text{day} - 81) \right] \quad (3.8)$$

H in Equations 3.6 and 3.7 is the local hour angle, presenting the number of degrees that the earth must rotate before the sun will be directly over the local meridian. SAM's hour angle calculations account for right ascension and is determined by calculating the location's local mean sidereal time, a time scale based on earth's location measured relative to fixed celestial objects [65].

Surface angles and irradiance determination

Once the sun's altitude and azimuth angles have been determined the angle of incidence (AOI), defined as the angle between direct radiation from the sun and a line normal to the solar module surface can be calculated as:

$$AOI = \arccos \left(\sin (90^\circ - \alpha_s) \cos (\gamma_s - \gamma) \sin \beta + \cos (90^\circ - \alpha_s) \cos \beta \right) \quad (3.9)$$

With β and γ being the tilt and azimuth angles of the solar modules as defined by the input parameters and shown in Figure 3.5.

For this study a fixed tilt tracking option was chosen as it is the most economical and suited for permanent installation on a building roof with minimal maintenance.

The nominal global incidence irradiance I_g varies greatly dependant on location and weather and is the sum of the beam, diffuse and ground-reflected irradiance:

$$I_g = I_b + I_d + I_r \quad (3.10)$$

with:

$$I_b = E_b \cos AOI \quad (3.11a)$$

$$I_d = E_d \frac{1 + \cos \beta}{2} \quad (3.11b)$$

$$I_r = \psi \left(E_b \cos (90^\circ - \alpha_s) + E_d \right) \frac{(1 - \cos \beta)}{2} \quad (3.11c)$$

with ψ representing the ground reflectance ratio and E_b , E_d being the beam irradiance and diffuse irradiance.

Panel degradation and surface soiling

Depending on the environment, the modules may become covered in dust and sediment, introducing DC losses. For this study the losses were calculated at a linearly increasing rate. Starting at 0% at the beginning of the month, the hourly loss slowly increased until reaching 3% by the end of every month. It was assumed that rainfall and regular cleaning with water would occur at least once per month [68].

The panel degradation rate was chosen at 0.8% per year [13], culminating in a 21.2% decrease in efficiency by the end of the module's expected 25 year life cycle.

Wiring, module mismatch and connection losses amounted to 4.4% and was estimated from previous literature [42; 67].

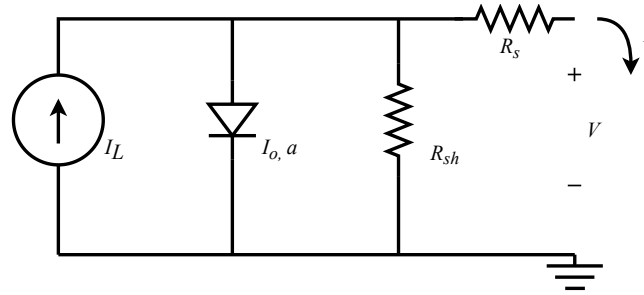


Figure 3.6: Equivalent circuit of a single-diode solar model.

Table 3.2: Solar system parameters

| Description | Variable | Unit |
|---|-----------------|-----------------|
| Solar input parameters | | |
| Maximum power point voltage | V_{mp} | V |
| Maximum power point current | I_{mp} | A |
| Open circuit voltage | V_{oc} | V |
| Short circuit current | I_{sc} | A |
| Temperature dependence of V_{oc} | β_{oc} | V/ $^{\circ}$ C |
| Temperature dependence of I_{sc} | α_{sc} | A/ $^{\circ}$ C |
| Temperature dependence of maximum power point | γ_{spec} | %/ $^{\circ}$ C |
| Number of cells in series | N_{ser} | - |
| Calculated solar coefficients | | |
| Modified nonideality factor | a | - |
| Light current | I_L | - |
| Diode reverse saturation current | I_o | - |
| Series resistance | R_s | - |
| Shunt (parallel) resistance | R_{sh} | - |
| Temperature coefficient adjustment factor | $adjust$ | - |
| Inverter input parameters | | |
| Maximum AC power | $P_{ac,0}$ | W |
| Maximum DC power | $P_{dc,0}$ | W |
| Power usage during operation | $P_{s,0}$ | W |
| Maximum power usage at night | $P_{nt,0}$ | W |

No external or self-shading losses were taken into account as the modules were assumed to be mounted flat on the roof of a school building without interference from nearby buildings or trees. Additionally, no snow losses were taken into account as it is a rare occurrence in the Southern part of South Africa where the study was performed.

3.2.2.2 Solar circuit model and coefficient generator

Once the effective global irradiance for a location is determined the generation potential for solar modules can be estimated. For this study a 6-parameter solar module model is used. The specific module sub-model implementation by the California energy commission (CEC) by Dobos [69] is an improvement to the five parameter model originally defined by De Soto *et al.* [70], and allows for

customised modules to be modelled from manufacturers specifications. This is done by converting the specifications to circuit coefficients described in Table 3.2 to create an equivalent single-diode circuit model as shown in Figure 3.6.

The six coefficients are calculated from the manufacturer specifications shown in Table 3.2 by simultaneously solving six non-linear constraint equations. This is done using Newton's method, whereby a root x is found through successive iterations such that $f(x) = 0$.

The coefficients are defined as [69]:

$$\mathbf{x} = \begin{bmatrix} a \\ I_L \\ I_o \\ R_s \\ R_{sh} \\ adjust \end{bmatrix} \quad (3.12)$$

And the six constraint equations are shown as:

$$\mathbf{F} = \begin{bmatrix} I_L - I_o(\exp[\frac{I_{sc}R_s}{a}] - 1) - \frac{I_{sc}R_s}{R_{sh}} - I_{sc} \\ I_L - I_o(\exp[\frac{V_{oc}}{a}] - 1) - \frac{V_{oc}}{R_{sh}} \\ I_L - I_o(\exp[\frac{V_{mp}+I_{mp}R_s}{a}] - 1) - \frac{V_{mp}+I_{mp}R_s}{R_{sh}} - I_{mp} \\ I_{mp} - V_{mp} \left[\frac{\frac{I_o}{a} \exp[\frac{V_{mp}+I_{mp}R_s}{a}] + \frac{1}{R_{sh}}}{1 + \frac{I_o R_{sh}}{a} \exp[\frac{V_{mp}+I_{mp}R_s}{a}] + \frac{R_s}{R_{sh}}} \right] \\ (I_L + \alpha_{sc} \cdot (1 - adjust/100) \cdot \Delta T) - I_{o,T'} \exp[\frac{V_{oc,T'}}{a_{T'}} - 1] - \frac{V_{oc,T'}}{R_{sh}} \\ \gamma_{spec} - \gamma_{model} \end{bmatrix} \quad (3.13)$$

with $T' = \Delta T + 25^\circ\text{C}$ (STC) and γ_{model} is determined by:

$$\gamma_{model} = \left. \frac{\partial P}{\partial T} \right|_{mp} \quad (3.14)$$

Equation 3.14 forms part of the to the sixth constraint equation and relates a maximum power output to a specific operating cell temperature, and the other parameters are adjusted accordingly until $\gamma_{spec} - \gamma_{model} = 0$. Iteratively solving for $\mathbf{F}(\mathbf{x})=0$ using Newton's method will then determine the 6 unknown coefficients.

The electrical behaviour of the single-diode equivalent model can now be described using:

$$I = I_L - I_o \left(\exp \left[\frac{V + IR_s}{a} \right] - 1 \right) - \frac{V + IR_s}{R_{sh}} \quad (3.15)$$

Through iterative simulations the maximum power point can be calculated for the user-defined module specification and determined coefficients to obtain the module DC output at maximum power defined as:

$$P_{dc,m} = V_{mp} I_{mp} \quad (3.16)$$

The total array power can now be calculated by multiplying the number of modules with the DC output of a single panel.

$$P_{dc,total} = P_{dc,m} N_{modules} \quad (3.17)$$

3.2.2.3 Inverter model description

Converting the DC output of the solar array to an AC output necessary for use within buildings in South Africa requires an inverter. During the modelling process it was assumed that a system would use identical inverters, and that the DC power from the solar array would be divided equally among the total number of inverters $N_{inverters}$:

$$P_{dc,inverter} = \frac{P_{dc,total}}{N_{inverters}} \quad (3.18)$$

SAM's inverter model is an implementation of an empirical model described by King *et al.* [71] capable of simulating the AC output power for a specific DC input [65]. From the manufacturer specifications shown in Table 3.2 the inverter's AC power output can be calculated as [65]:

$$P_{ac,inverter} = \frac{P_{ac,0}}{P_{dc,0} - P_{s,0}} (P_{dc,inverter} - P_{s,0})^2 \quad (3.19)$$

The system's AC power is then determined by multiplying a single inverter's output power with the total number of inverters. During the evening when no solar power is generated the inverters will still use a small amount of power $P_{nt,0}$.

3.2.2.4 Complete solar performance model description

SAM's solar model performs a yearly generation forecast consisting of 8760 hours from the input parameters described. A diagram of the solar modelling process can be found in Figure 3.7. The process is repeated for every hour and performed as follows:

1. Determine the sun's location and relative panel orientation using the input parameters and weather data.

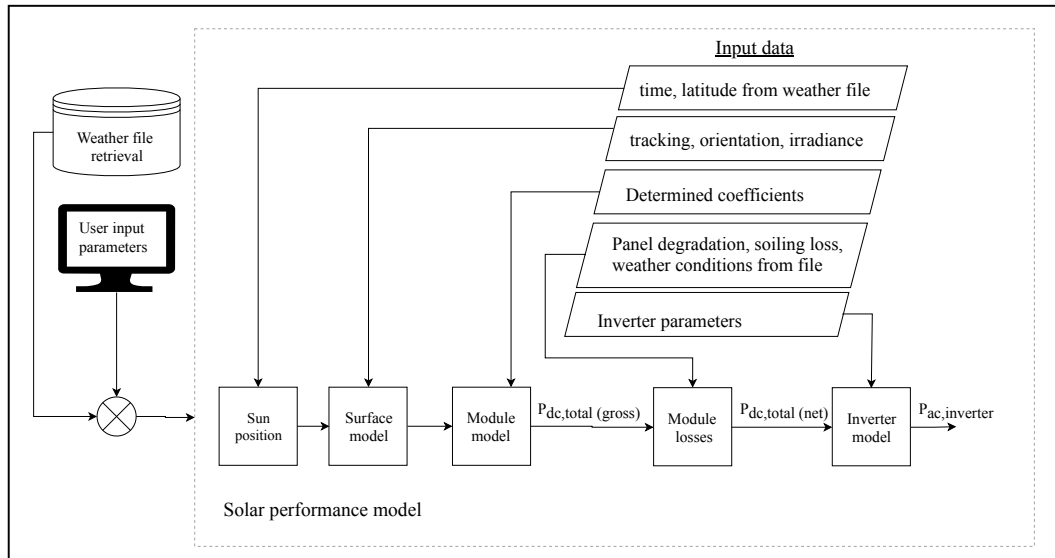


Figure 3.7: Simplified diagram of complete solar modelling process

2. Calculate the nominal global incident irradiance.
3. Apply standard yearly soiling and degradation factors to the PV array model.
4. Get the array DC output power using the determined coefficients.
5. Use the inverter model to calculate the net AC power generated by the system.

Taking the hourly usage forecast and subtracting the hourly generated solar now presents a new hourly load profile.

3.2.3 Optimised load-matching

In Chapter 1 a research objective stating the requirement for a cost-optimised solution was made. As such, an economical approach is selected. The approach is based on the concept of annualised system cost and focuses on accurately defining the yearly financial in-flows and out-flows.

To effectively describe the system cash flow the complete system arrangement must be described including accounting for local municipal electricity rates. Figure 3.8 illustrates the cash flows associated with the solar system. The yearly savings are calculated by determining the years' energy cost without a solar system, then subtracting the total of all monthly energy bills and finally adding all income gained from exporting excess energy to the grid. The system life cycle is estimated to be 25 years with inverter replacements every 10 years, consistent with standard manufacturer specifications [72]. A small yearly maintenance cost (1% of system cost) to reduce soiling losses, and

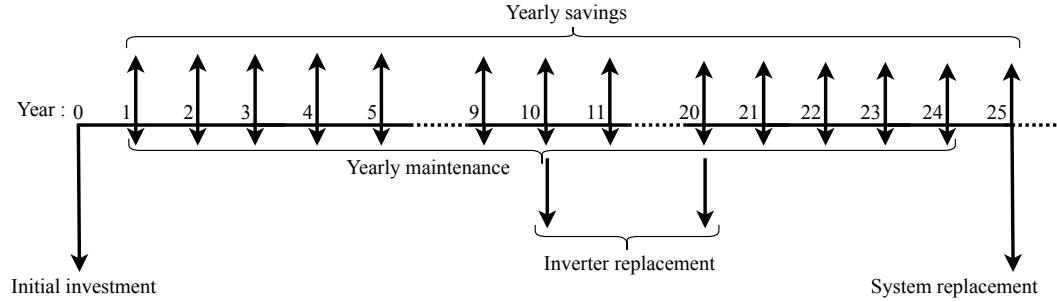


Figure 3.8: Diagram illustrating yearly cash flows for the evaluated solar system.

module degradation rates of 0.8% per year was chosen for the simulation [13]. Further assumptions made include electricity prices rising by 7% per year, considered modest within the South African context [21].

Solar system size optimisation is done by estimating the profitability of the investment using the internal rate of return (IRR) metric shown below:

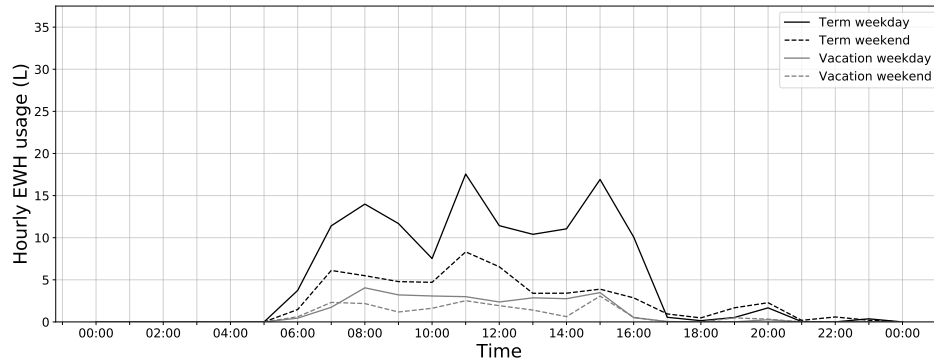
$$NPV = \sum_{t=1}^T \frac{C_t}{(1+r)^t} - C_0 = 0 \quad (3.20)$$

where C_t is the net cash inflow during the period t , C_0 is the initial investment cost, r is the discount rate and T is the number of time periods representing the total years of the system's lifetime.

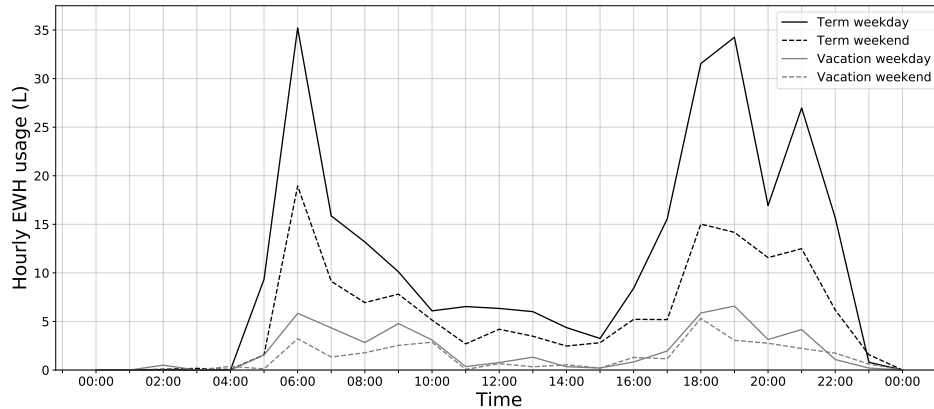
The IRR is calculated as the discount rate r that equates to a net present value (NPV) of zero. The higher the IRR, the more financially viable the solar system becomes. The IRR metric quantifies the economic effect of a solar system over its life cycle, and concerns multiple external and internal factors. It should be noted that the first year's cash flow does not consider the benefits obtained through tax incentives, even though South African law permits a full deduction of the system's value as depreciation during the first year of operation for commercial installations [73]. Additionally, the initial installation cost is dependant on system size and was determined using results obtained from a recent solar markets report [13].

3.3 Intelligent EWH control

This section describes the development of intelligent water heater control schemes (CSs). The development will assess the use of a grid-tied PV solution in combination with load shifting through smart scheduling of energy storing electric water heaters to reduce both energy usage from the grid, peak load, and the electricity bills of a school, while ensuring the hot water demand is met.



(a) Typical daily demand profiles of an EWH installed within a school.



(b) Typical daily demand profiles of an EWH installed within a residence on the school premises.

Figure 3.9: Usage profiles obtained from measurement data

3.3.1 Analysis of usage data

For this study the measured hot-water usage data from a high income (quintile 5) school was used as the measurement data was available for multiple seasons. The data had a temporal resolution of 1 minute and was available for an entire year. The data fields provided contained the volume of hot water used and the time of usage.

3.3.1.1 EWH usage within schools

Figure 3.9a shows a typical school's daily hot water demand during summer months for term- and vacation periods, presenting a steady usage rate during the daytime, with a large reduction present outside of school hours. Figure 3.9b presents the demand for EWHs installed within staff residences located on a school premises with surges appearing before and after school hours when students are waking up and returning from school. It was found that usage

patterns remained similar during colder months (May to September), but that the average hot-water usage increased by 40%. These figures paint a clear picture to the flow of people in and out of a building and how it relates to the usage of hot water [56].

3.3.1.2 Water usage decomposition

In order to perform scheduling of EWHs prior knowledge of their usage patterns must be known. As such, EWH intervention requires the installation of smart meters or manually entering likely usage times to perform well. For this study a vectorisation technique was implemented to determine the likelihood of EWH usage from historic measurement data by converting it into an array form. Due to its nature as a high-level programming technique, vectorisation drastically reduces the computational complexity of a process by reducing large datasets into more concise vector lists. The measurement data was converted into seasonal lists containing the probability (ranging from 0 to 1) of an EWH being accessed. These lists were created for two types of EWHs according to their usage patterns, school EWHs and residence EWHs. An EWH access was defined as a hot-water event during which more than 2 litres of water was used. A total of 9 seasonal lists were created and is shown below:

$$G_S = \begin{bmatrix} w_{0-0} & \dots & w_{0-47} \\ \vdots & \ddots & \vdots \\ w_{6-0} & \dots & w_{6-47} \end{bmatrix} \quad (3.21)$$

Each of the seasonal lists contain 48 entries for each day of the week representing a temporal resolution of thirty minutes for the EWH usage probability. Additionally the vectorisation function added a backshift operator, whereby the importance of each of the preceding weeks' access probability would decrease, meaning that newer measurement data would have a larger effect on determining the new probabilities than older usages. The backshift operator was chosen to decrease the weight of the preceding weeks' by 10% every week. As such, usage probabilities older than 10 weeks would have no effect. This was done to adjust to the changing usage patterns within seasonal periods spanning multiple weeks, as temperature changes during these periods lead to varying usage volumes and access probabilities.

$$Probability(\%) = \frac{\sum_{i=0}^n p(i) * (1 - 0.1 * i)}{n - \sum_{i=0}^n i * 0.1} * 100 \quad (3.22)$$

Equation 3.22 is repeated for every entry of G_S with n being the total number of weeks i up to a maximum of 10, and p being the probability of a specified week. It should be noted that the probability of use within a 30-minute timeslot is either a yes or no (1 or 0) for a specific week, but becomes more defined once multiple weeks of data is available.

For example, if probability vectors (chosen as 0%, 100% and 100%) for the first 3 weeks are available, the projected probability of usage would be determined as:

$$\begin{aligned} \text{Probability}(\%) &= \frac{(0 * 0.8) + (1 * 0.9) + (1 * 1)}{3 - (0.2 + 0.1 + 0)} * 100 \\ &= 70.37\% \end{aligned} \quad (3.23)$$

This provides improved forecasting reliability compared to the non-backshift method result of 66.66%.

3.3.2 EWH modelling

Digital representations of EWH behaviour through simulation requires for specialised thermal models to be defined. This section provides an overview of the thermal properties of EWHs and a description of the nodal models used during this study.

3.3.2.1 Thermal properties of EWHs

The thermal properties of EWHs are governed by a set of heat and mass transfer principles that need to be accounted for to accurately model EWH behaviour. This section provides a brief overview of work further expanded upon by Cloete [74].

Water enthalpy and enthalpy reduction: Water enthalpy refers to the heat energy stored within a body of water combined with the product of the pressure and volume of the water. The enthalpy contained within an EWH may decrease in one of two ways. The first mechanism for enthalpy reduction occurs when hot water is drawn off at the EWH outlet and replaced by cold water at the inlet during usage. The second method occurs when heat is radiated into the environment through the EWH surface as standing losses. Thermal radiation occurs across a temperature boundary of any two masses, equalising the energy states between them. By properly insulating the temperature boundary the rate of energy loss can be reduced.

Enthalpy restoration: For the purpose of this study the thermal energy contained within the body of water in the EWH is increased by an electrical heating element. Other forms of heating not represented within the nodal models include burning natural energy sources such as gas, coal or wood.

3.3.2.2 EWH nodal models

EWH nodal models are used extensively for research purposes, especially pertaining to DSM [32] and smart-grid [33] applications as described in Chapter

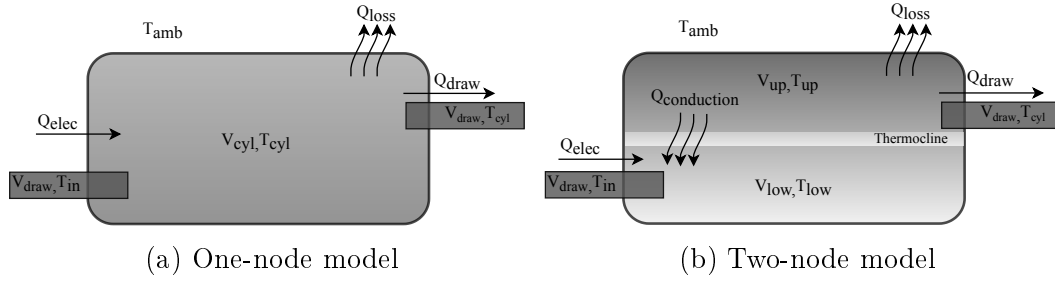


Figure 3.10: Diagram of energy flows within EWH nodal models

2. This section describes the two nodal models developed by Nel [75] and used during this study to model EWH behaviour by simulating the energy flows inside an EWH as presented in Figure 3.10.

One-node model:

The first, *one-node* model regards the water within the EWH as a uniform mass. This model assumes that water leaving the outlet is at the same temperature as the average of the body of water in the EWH, and that water entering at the inlet will instantaneously mix with the water in the tank to create a new average temperature.

Enthalpy reduction: For the one-node model the enthalpy reduction due to drawn off water being displaced by cooler inlet water is determined as:

$$Q_{draw}[n] = c\rho V_{draw}[n] \left(T_{cyl}[n] - T_{in}[n] \right) \quad (3.24)$$

With c being the specific heat capacity of water (4184 J/kg, °C), ρ is the density of water (1000 kg/m³), $T_{cyl}[n]$ is the temperature of V_{draw} litres of water leaving the tank at sample n and $T_{in}[n]$ is the temperature of V_{draw} litres of water entering the tank.

Standing losses occur as energy lost to the environment due to heat dissipation over a sampling interval n and are calculated as follows:

$$Q_{loss}[n] = c\rho V_{cyl} \left(T_{cyl}[n] - T_{cyl}[n+1] \right) \quad (3.25)$$

With V_{cyl} being the volume of water stored in the EWH and T_{cyl} being the average temperature of water within the tank. The change in temperature due to thermal radiation is determined by:

$$T_{cyl}[n+1] = T_{cyl}[n] - \frac{T_{cyl}[n] - T_{amb}[n]}{c\rho V_{cyl} R} \quad (3.26)$$

R is the thermal resistance of the EWH (determined to be 0.46 from previous studies [74; 75]), defined as a measurement of how much an object rejects

heat flow when presented with a temperature differential between the EWH and the environment T_{amb} .

Enthalpy restoration: The enthalpy contained within the tank can be increased by an electric heating element. The increase in energy over a sampling period is determine as:

$$Q_{elec}[n] = P_{elec}[n]\Delta n \quad (3.27)$$

With $P_{elec}[n]$ being the power of the heating element in Watts, and Δn the fraction of the interval for which it is on, chosen as 1 (the entire interval) for all performed simulations. The increase in temperature within the tank is described by:

$$T_{cyl}[n + 1] = T_{cyl}[n] + \frac{Q_{elec}[n]}{c\rho V_{cyl}} \quad (3.28)$$

Two-node model:

The *two-node* model is used to model EWH behaviour when large amounts of water greater than the threshold volume $V_{threshold}$ (chosen as 2 Litres) is used within a single interval. This is done to account for the stratification process, whereby lower-density warm water rises above the colder, denser water to create a *thermocline* that separates the two masses [74]. The temperatures of the two masses are governed by the following differential equations:

$$cm_{up}T_{up} = -\frac{1}{R_{up}}\left[T_{up} - T_{amb}\right] - \frac{1}{R_{therm}}\left[T_{up} - T_{low}\right] \quad (3.29a)$$

$$cm_{low}T_{low} = -\frac{1}{R_{low}}\left[T_{low} - T_{amb}\right] - \frac{1}{R_{therm}}\left[T_{low} - T_{up}\right] + Q_{elec} \quad (3.29b)$$

With m_{up} and m_{low} being the mass (in kg) of the upper and lower bodies of water. The thermal resistance R is divided amongst the two masses based on their tank area coverage. When in the two-node state, heat from the two masses will transfer in a process known as de-stratification. The thermocline between the two parts is assumed to have a thickness of 10mm with no mass [75], and is determined using:

$$R_{therm} = \frac{d}{kA_{cond}} \quad (3.30)$$

Where k is the thermal conductivity of water and d is the thickness of the thermocline. A_{cond} is the area where the two masses meet. The thermal conductivity of k varies as the temperature changes, and is estimated by [76]:

$$k = (-8.354 \times 10^{-6})T^2 + (6.53 \times 10^{-3})T - 0.5981 \quad (3.31)$$

The EWH will remain in the two-node model state until the volume used within an interval period is less than $V_{threshold}$ or $T_{up} = T_{low}$ as a result of the heating element increasing the temperature of the lower body of water, or the upper mass heat lowering from thermal radiation.

3.3.2.3 EWH orientation

During the two-node state the orientation of the EWH is important. This is due to the area A_{cond} changing depending on the arrangement. For this study all EWHs were assumed to be in a horizontal orientation as this is the norm in South Africa.

3.3.3 EWH DSM techniques

The intelligent water heater CSs developed during this study will each implement various DSM techniques to reduce the cost of energy used. Five techniques are used and expanded upon in this section including:

- Smart scheduling
- Peak power limiting
- Prioritised heating
- Excess solar dumping
- Bi-thermostat control
- Demand limiting

3.3.3.1 Smart scheduling

User-based scheduling is a technique commonly employed in many household and businesses to reduce the energy used for water heating when compared to the manufacturer standard *thermostat control* [77]. Smart-scheduling performs EWH heating by means of a bi-directional communication link between a schedule controller and an EWH controller. The schedule controller is equipped with a centralised algorithm capable of managing multiple EWHs. It will turn on the EWH heating elements within a certain period, t_{et} , of the expected usage time based on prior usage patterns to increase user comfort and decrease overall energy usage.

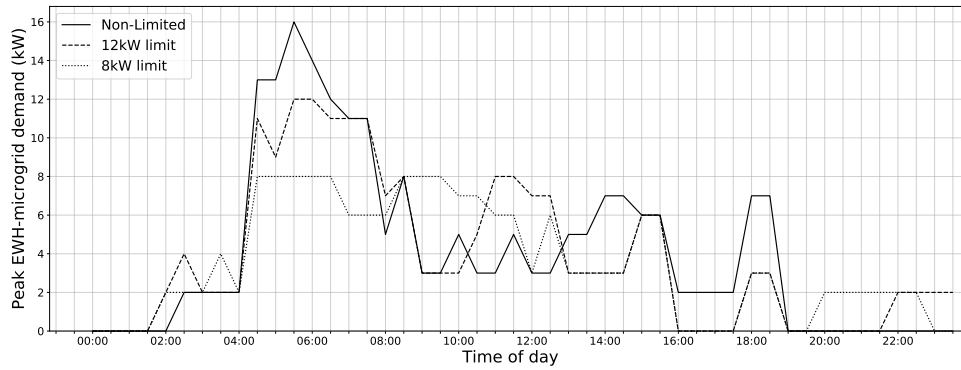


Figure 3.11: The PPDM in action for various power limits.

3.3.3.2 Peak power limiting

In addition to managing the heating schedules of the EWHs, the schedule controller also implements a peak power demand manager (PPDM). The PPDM ensures that the total EWH power usage does not exceed a specified limit, reducing overall demand spikes due to multiple EWH elements switching on at the same time. An example of this process is shown in Figure 3.11, with peak power limiting done for various power limits and the maximum power usage for the EWH-microgrid shown for 30-minute segments of the day.

3.3.3.3 Prioritised heating

Priority control calculates the urgency with which each EWH connected to a schedule controller needs to be heated using a cost function similar to the one proposed by Roux *et al.* [32]. The cost function is based on two key factors: (1) the time until the next expected water usage event, and (2) time required to heat the water from its real-time measured temperature to a target internal temperature, T_{set} . Assuming both factors carry the same weight the cost function can be defined as:

$$k_{gn} = t_{heat} - t_{event} \quad (3.32)$$

The cost function is calculated for each simulated interval n for every EWH, gn , where t_{event} is the time before the next expected usage and t_{heat} is the required heating time from the present EWH outlet temperature to the target temperature T_{set} . The time before the next expected usage is based on historical usage patterns discussed in Section 3.3.1.2, and the time to heat the EWH is determined by its measured internal temperature, volume and element rating. Only the EWHs with highest priority, k_{gn} , are heated such that a minimal amount of energy is used while maintaining user comfort and limiting the power draw of the EWH elements depending on the selected PPDM control configuration.

3.3.3.4 Excess solar dumping

EWHs are used as a thermal battery, storing hot water for later use. When a solar PV system generates excess energy, it can be used to augment water heating instead of exporting the energy to the grid at a reduced rate [61]. The mechanism uses the previously discussed priority list to assign water heaters to excess solar supply, but only as much as is available and based on the prioritised water heaters' element rating. For example, if 5 kW were available, and the top three water heaters in decreasing priority had respective element ratings of 3 kW, 3 kW and 2 kW, the simulation algorithm would use solar power to heat the former and the latter. Grid power would be used for the middle one, but only if the PPDM allows for 3 kW grid draw at the time.

3.3.3.5 Bi-thermostat control

This technique introduces a bi-thermal heating mechanism. For this approach, the EWHs are heated to different target temperatures for grid heating (T_{setG}) and solar heating (T_{setS}), allowing solar energy to be diverted to the EWHs if excess solar is available even if the water temperature is higher than the grid target temperature. By employing a higher target temperature for solar heating, the EWHs will remain suitably hot for usage for longer periods before needing to use energy from the grid.

3.3.3.6 Micro-grid demand limiting

Demand limiting is done by the schedule controller. The algorithm retrieves historical- and current building energy usage data, comparing the current power usage P_{cur} to the recorded maximum P_{max} . The approach then controls the EWHs such that the buildings' peak monthly power demand is not affected by the water heaters. This would imply that EWH heating does not contribute to any of the school's peak demand charges when billed on an *industrial demand* or *time of use* tariff. However, if the user wishes to do so an optional demand threshold parameter allows the maximum demand to be exceeded by a percentage $D_{threshold}$.

3.3.4 EWH control schemes

This section describes the EWH control schemes that were evaluated during this study. Three approaches are compared with the normally used *always on* thermostat control. All three use excess solar dumping to augment supply and a centralised demand-drive priority water heating. The three incremental approaches are (1) with a single target-temperature thermostat and modulated schedule employed, (2) with a bi-thermostat and modulated schedule employed, and (3) a demand-limiting controller with the bi-thermostat. The first two aim to manage the total energy through schedule control, while the

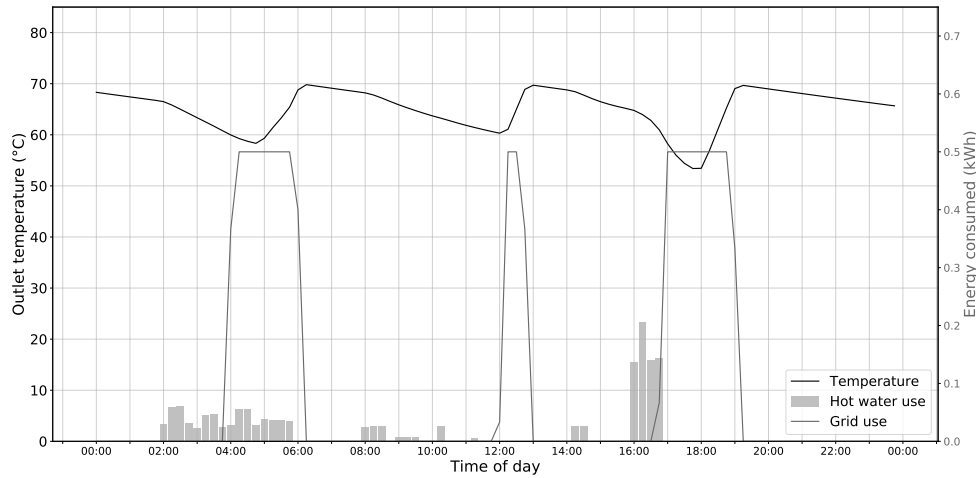


Figure 3.12: Daily profile of an EWH with thermostat control.

third manages the peak demand. The heating state and source of the EWH is evaluated for each simulated interval, and determined by a decision tree with various operational nodes dependant on the selected control scheme CS_{set} . A pseudocode implementation is shown in Algorithm 2.

3.3.4.1 Thermostat control

Thermostat- or "*always-on*" control is the standard control scheme for most household EWHs and serves as a baseline comparison for the other developed control schemes. The mechanical thermostat within the water heater is set to a desired temperature T_{set} and always supplied with power from the grid. No other scheduling techniques are used for this control scheme, and it will simply switch on the heating element once the measured internal temperature is below the set point. Most mechanical thermostats are prone to measurement inaccuracies about T_{set} that need to be accounted for during simulations known as *hysteresis*, and was chosen as 7.4% based on field tests by previous studies [74].

Figure 3.12 presents the usage profile of a thermostat controlled EWH over 15 minute periods with accompanying temperature and energy usage measurements. The EWH is installed in a school building and the shown usage profile measurements were taken on the 20th of February 2018. From the figure the switching of the EWH element is visible when the water temperature is below or above the target value. The control scheme is seen to be reactive towards usage events, with usage temperatures dropping to as low as 54°C from the chosen target of 65°C, as would be the case on a normal water heater Eskom [78]. Additionally, the heating elements are powered unnecessarily for a large part of the day, resulting in unnecessary energy use.

```

Result: Heat EWH by means of selected control scheme
1 Inputs:  $CS_{set}$ ,  $t_{et}$ ,  $T_{set}$ ,  $T_{setG}$ ,  $T_{setS}$ ,  $G_s$ ,  $P_{cur}$ ,  $P_{max}$ ,  $D_{threshold}$ ;
2 begin
3   for each sampling interval do
4     Get selected control scheme from  $CS_{set}$ ;
5     if  $CS == Always\ On$  then
6       if  $OutletTemp < (T_{set} - hysteresis)$  then
7          $HEATSTATE = TRUE$ 
8       end
9       if  $OutletTemp > (T_{set} + hysteresis)$  then
10         $HEATSTATE = FALSE$ 
11      end
12      if  $HEATSTATE == TRUE$  then
13        Switch on EWH element
14      end
15    end
16    if  $CS == Smart\ schedule$  then
17      Get expected usage time from  $G_s$ ;
18      Calculate costfunction;
19      Get EWH priority ranking;
20      if ( $EWH$  is 1st on priority list) and (Expected usage is within  $t_{et}$  hours) then
21        if ( $SolarExcess > EWHrating$ ) and ( $OutletTemp < T_{set} - hysteresis$ ) then
22          Divert excess solar and switch on EWH element
23        end
24        if ( $costfunction < 0$ ) and ( $SolarExcess < EWHrating$ ) and ( $OutletTemp < T_{set} - hysteresis$ ) then
25          Heat EWH from grid
26        end
27      end
28    end
29    if  $CS == Bi-thermostat$  then
30      Get expected usage time from  $G_s$ ;
31      Calculate cost function;
32      Get EWH priority ranking;
33      if ( $EWH$  is 1st on priority list) and (Expected usage is within  $t_{et}$  hours) then
34        if ( $SolarExcess > EWHrating$ ) and ( $OutletTemp < T_{setS} - hysteresis$ ) then
35          Divert excess solar and switch on EWH element
36        end
37        if ( $costfunction < 0$ ) and ( $SolarExcess > EWHrating$ ) and ( $OutletTemp < T_{setG} - hysteresis$ ) then
38          Heat EWH from grid
39        end
40      end
41    end
42    if  $CS == \mu Grid\ demand\ limiting$  then
43      Get expected usage time from  $G_s$ ;
44      Calculate cost function;
45      Get EWH priority ranking;
46      if  $EWH$  is 1st on priority list then
47        if ( $SolarExcess > EWHrating$ ) and ( $OutletTemp < T_{setS} - hysteresis$ ) then
48          Divert excess solar and switch on EWH element
49        end
50        if ( $P_{cur} + EWHrating < P_{max} * (100\% + D_{threshold})$ ) then
51          if ( $costfunction < 0$ ) and ( $SolarExcess > EWHrating$ ) and ( $OutletTemp < T_{setG} - hysteresis$ ) then
52            Heat EWH from grid
53          end
54        end
55      end
56    end
57  end
58 end

```

Algorithm 2: Pseudocode implementation of the evaluated EWH control schemes.

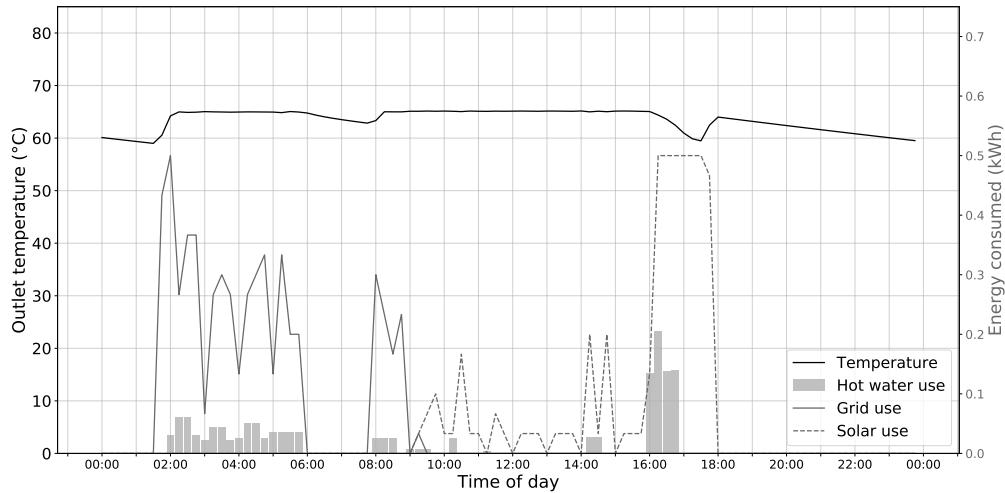


Figure 3.13: Daily profile of an EWH with smart schedule control with solar dumping, showing the heightened temperature achieved through excess energy dumping.

3.3.4.2 Smart Schedule control with solar dumping

The first energy-saving intervention uses a smart schedule that overrides the prioritisation control to only heat for a set time before hot water use to reduce standing losses. This method uses only a single target temperature on the thermostat, namely 65 °C as seen in Figure 3.13. The schedule has to balance the need to limit the standing losses of the water heater and the need to deliver hot water when needed, while dumping the excess energy from the solar supply into the EWHs. Unless solar power is available, the scheduler will only allow switching the element on if the time until the next event, t_{event} is within a specified time t_{et} .

When comparing the daily usage profile of the smart scheduling approach the lack of sustained power to the EWH element is noticeable as a result of the scheduler only prioritising the EWH heating when usage is expected. During the afternoon the excess available solar (presented as a dashed line) is used to heat the EWH, further reducing grid usage.

3.3.4.3 Bi-thermostat control with prioritisation and solar dumping

The second energy-saving intervention, based on the first, introduces a bi-thermal heating mechanism in conjunction with excess available solar. The bi-thermostat employs a temperature delta for grid (T_{setG}) and solar (T_{setS}) heating. This allows for additional thermal energy to be stored in the EWH tank while further reducing grid usage with a lower grid-use target temperature. It should be noted that the bi-thermal approach includes all the features

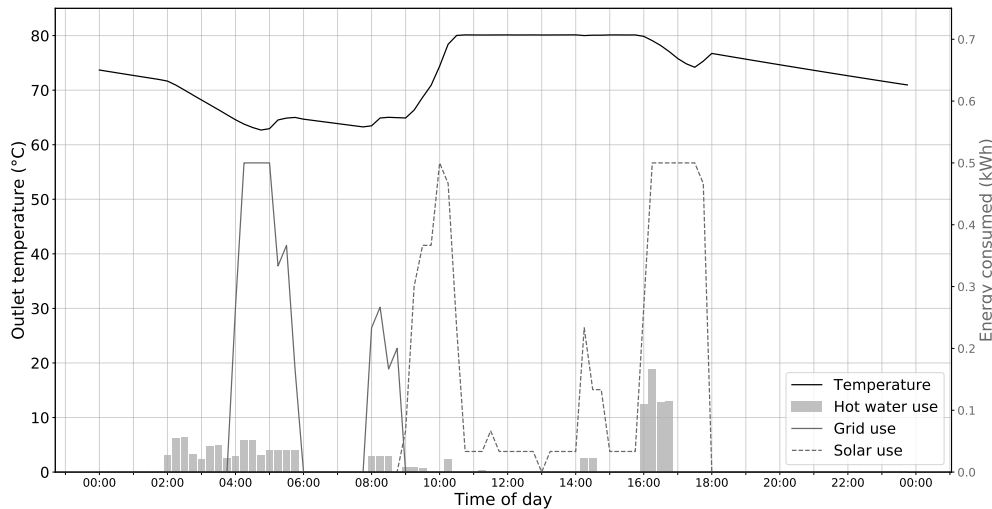


Figure 3.14: Daily profile of an EWH with bi-thermostat control.

of the prioritisation scheduler, in addition to having multiple heating thresholds.

The control scheme implementation is shown in Figure 3.14 with a temperature delta of 30°C . The grid- and solar heating threshold were set 50°C and 80°C respectively, well within the operating ranges of standard EWHs [79]. Due to the higher solar threshold the water temperature at the start of the day is higher than the previously discussed interventions, further reducing the need to use the grid.

3.3.4.4 Demand-limiting control with a bi-thermostat and solar dumping

The final CS is a demand-limited heating method in conjunction with the bi-thermal heating mechanism with excess solar dumping, which controls the EWHs such that the buildings' peak monthly power demand is not affected by the water heaters. This is done by heating the EWHs according to priority without exceeding a calculated power limit in order to not increase the recorded peak monthly demand unless a $D_{threshold}$ percentage is specified. Since the intention of this scheme was to perform demand limiting, the scheduled heating time limitation t_{et} was removed, allowing the EWHs to be heated to their target temperatures and remain hot enough during peak-demand times to minimise the number of cold events at the cost of increased standing losses.

As an example Figure 3.15 is presented during which the maximum monthly demand would have been exceeded if the EWH was switched on at 08:00 AM as was done for the previous intervention. However, due to the demand-limit being established from an existing monthly maximum and present measured data the EWH element was not heated using the grid. Once excess solar

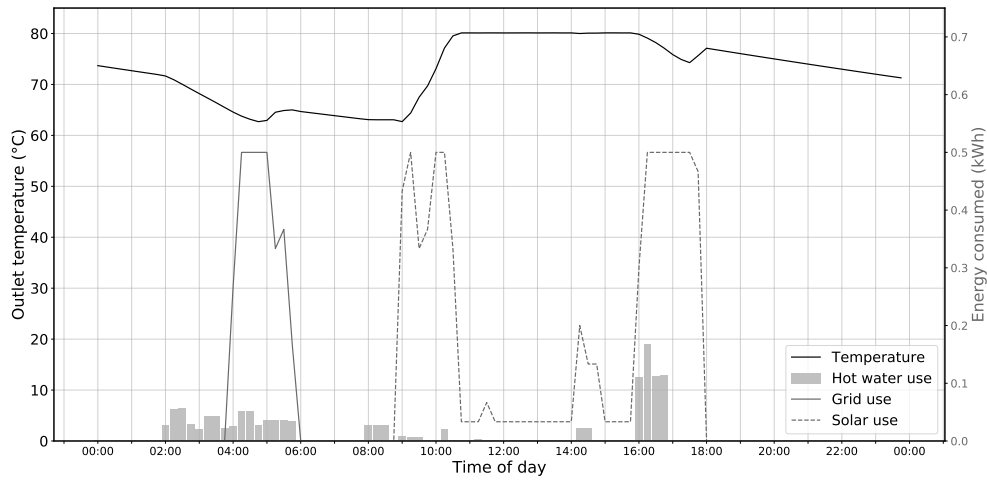


Figure 3.15: Daily profile of an EWH with demand limiting.

Table 3.3: Feature table of evaluated interventions

| Energy saving intervention | Schedule Control | Prioritised heating | Excess solar dumping | Bi-thermostat | μ Grid demand limiting |
|--------------------------------------|------------------|---------------------|----------------------|---------------|----------------------------|
| Thermostat-control | ✗ | ✗ | ✗ | ✗ | ✗ |
| Smart-schedule control and solar PV | ✓ | ✓ | ✓ | ✗ | ✗ |
| Bi-thermal control and solar PV | ✓ | ✓ | ✓ | ✓ | ✗ |
| Demand-limiting control and solar PV | ✗ | ✓ | ✓ | ✓ | ✓ |

became available it was once again heated to the solar heating threshold of 80°C. Table 3.3 presents a summary of all the DSM techniques implemented in each of the evaluated CSs.

3.4 Graphical user interface development

A graphical user interface (GUI) was developed using *PyQt*, a Python binding of the cross-platform GUI toolkit Qt. The GUI simplifies the process for schools to size a solar PV system by determining an optimal installation size. The software does not take the savings from smart-scheduled EWHs into account and only performs usage forecasting with optimised solar functionality. The output of the GUI presents the estimated yearly kWh usage of the school, the optimal solar installation size with an estimated payback period, the maximised budget installation with an estimated payback period, and two graphical plots with the IRR and solar generation used and exported by the

school. The user manual for the system software can be found in Appendix A.

3.5 Chapter summary

Chapter 3 expanded on each of the three system components. The energy usage forecasting section expanded on the tariff structures imposed on schools in South Africa, and the energy usage patterns seen within schools. The section concluded with the development of the forecasting model. The solar system optimisation section explained the workings of the SAM performance model and presented a cash flow based approach for determining the optimal solar system size. The chapter concluded with a DSM section pertaining to hot water usage within schools as well as load management and load shifting using an EWH controller.

Chapter 4

Experimental setup

This chapter defines the experimental bounds and parameters used to simulate the developed system. Data cleaning measures performed on the raw data are also explained and all assumptions made are clearly noted. Performance metrics necessary to determine the system's ability to fulfil the stated research objectives are also discussed. A diagram of the developed system's simulation process is shown in Figure 4.1.

4.1 Simulation environment

All models and processes simulated during this study were implemented in the Python programming language. The use of Python within the data-science community has grown during recent years due to its ease of use as a high-level programming language with many data analysis libraries. Libraries of particular note include Pandas for handling large datasets, NumPy for optimised numerical processing, SciPy for differential equation solving and Matplotlib for data plotting and visual analysis.

4.2 Measurement data used

The data used for the forecasting simulations was obtained from municipal smart meters installed within each school and used to determine their monthly energy bill. The simulation uses measured energy usage data of 5 high income (quintile 5) schools (three primary schools, two high schools) in South Africa. The temporal resolution of the data is 60 minutes and is available from January 2016 until December 2018. The schools were selected due to having measured energy data available for multiple seasons.

For the EWHs, the measured energy and hot water usage data from a high school was used. The building's EWH water usage and energy usage was captured with smart water heater controllers that were retrofitted to seven EWHs installed within the school. Of the seven water heaters, three were

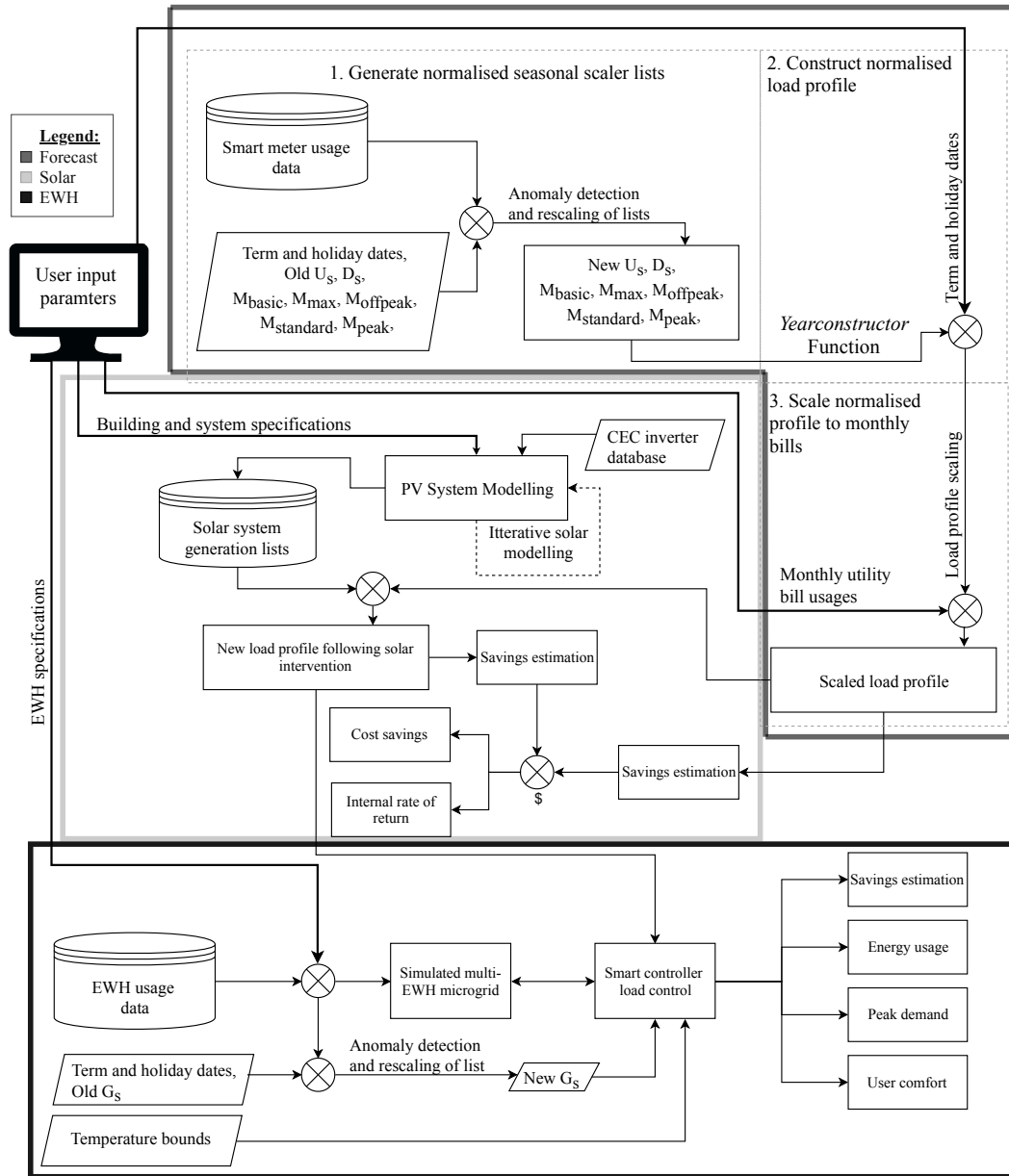


Figure 4.1: A complete diagram of the developed system simulation process.

installed within the school building and 4 were installed in student residences on the school premises. The data was captured over a 12-month period with a temporal resolution of 1 minute.

4.3 Data cleaning

In order to generate reliable simulation results the measurement data was cleaned to ensure that no missing values were present, and outlier detection measures were performed for the constructed seasonal lists.

4.3.1 Missing data

Energy measurements from the five schools totalled 131 520 entries, of which 2.97% (3912 entries) of values were missing from the raw data. Missing data can be a result of equipment failure, power outages or connectivity issues. Missing data was filled by taking into account the seasonality of the school's energy usage patterns. The computed filling technique was implemented as follows:

If no data was present at a specific entry, the data-entry from exactly 7 days prior would be selected as a filling value granted that it was during the same season and of the same day-type (normal day or holiday). If this was not possible the value 7 days later would be selected. Lastly, if a value could still not be retrieved the previous or next day's hour would be selected if it was of the same weekday-type (weekday or weekend day). This process was then repeated for all empty data entries.

For the water measurement data a total of 125 679 data points were retrieved of which 5.09% (6392) was missing. It should be noted that although the temporal resolution was recorded at 1 minute, the smart meters would only return a value with a timestamp if a usage event occurred or if the meters were pinged by the central server. Missing data was filled by using the average usage volume of the specific EWH.

4.3.2 Smoothing and outlier detection

If too much noise is present while training the seasonal scalar lists the forecasts will deviate from the expected norm. Figure 4.2 shows the rescaling process during which new data samples are added to calculate a new normalised list U_s . In the example the hourly usage for a Monday of the second school term is shown ($U_3[0]$). The upper and lower detection boundaries are set at 1 standard deviation (34% above the mean). If a sample is detected outside of the boundary (as seen for 14:00), the outlier is disregarded during the rescaling process, and not taken into account to generate the new hourly usage list for

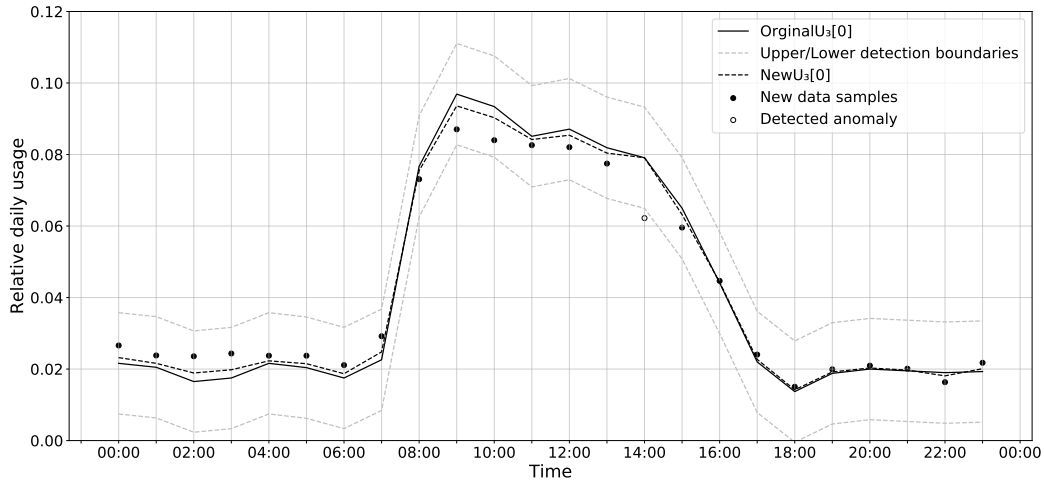


Figure 4.2: Forecasting scaler list training process with anomaly boundaries and entries highlighted.

that particular day and season. The same process is repeated when creating the EWH usage lists G_s .

4.4 Simulation setup

This section details the input parameters and their resultant outputs for each of the components of the simulated system. All components were simulated for a period of 365 days. All input parameters used during the experimental setup is listed in Table 4.1.

4.4.1 Load profile determination and forecast

The school data was divided amongst the two types of schools, three primary schools with learners from Grade 1 to Grade 7, and two high schools with learners from Grade 8 to Grade 12. This was done as enough data was available, and due to noticeable changes regarding their energy usage characteristics with high schools presenting larger demand surges as a result of more classrooms, on premises residences and after school curricular activities.

Each year was individually added to calculate new normalised seasonal lists for the two types of schools. The yearly trend from the usage data was not taken into account due to newly constructed buildings on the premises only affecting a single school, and carrying over the trend from only one building would be nonsensical. The input parameters for the load profile determination is the energy usage data from a specific school type for a year and the school term and holiday dates for that specific year.

Table 4.1: Input parameters used for experimental setup

| Parameters | Description |
|---|--|
| Time series decomposition | |
| Measured usage data | 8760 hours |
| School type | Primary school or High school |
| School term dates | Start & end dates for each of the 4 terms |
| Public holidays | A list of dates |
| Load profile construction | |
| School type | Primary school or High school |
| Tariff structure | Basic, demand-based or time-of-use |
| Monthly energy usages | kWh used, max monthly kVA or TOU usages (minimum one month) |
| School term dates | Start & end dates for each of the 4 terms |
| Public holidays | A list of dates |
| Solar PV simulation | |
| School location | Latitude and longitude |
| Available budget | ZAR(R) |
| Estimated available roof area | m ² |
| Roof tilt | degrees |
| Azimuth | degrees |
| Custom module specifications* (Required for 6-paramater model) | Max power voltage, max power current, open circuit voltage, short circuit current, number of cells per module, module area |
| Defult module: | Max power: 320W; V_{mp} : 37.5V; I_{mp} : 8.53A; V_{oc} : 46V; I_{sc} : 9.5A |
| EWH modelling | |
| Number of EWHs on premises | School EWHs and residence EWHs |
| Element power | kW |
| Tank volume | Liter |
| Standard target temperature T_{set} | 65 °C |
| Bi-thermostat temperature settings [T_{setG}, T_{setS}] | [50,80]°C |
| Cold event threshold $T_{threshold}$ | 40 °C |
| EWH orientation | Horizontal |
| EWH thermal resistance | 0.46°C/W |
| Thermostat hysteresis | 7.4% |

*Note:** These are optional simulation input parameters.

Once the base load profile for a school type has been created it can be scaled to perform a yearly forecast. The input parameters for the scaling procedure includes the monthly energy usages obtained from utility bills, the school type and its tariff structure as well as the term and holiday dates for that specific year. The simulation output is a list containing the kWh usages for every hour over the 1 year period.

4.4.2 Solar PV simulation

The solar PV generation modelling is performed using SAM's open source libraries that employ a photo-voltaic performance model developed by the National Renewable Energy Laboratory (NREL) NREL [65], and calculates the AC electrical output for each hour over a one year period. The input parameters required for the simulation is the school location, allocated budget, estimated available roof area, roof tilt and azimuth angles with the option for any module to be specified from manufacturers specifications by using the discussed 6-parameter sub-model implementation. If no module is specified the 320W module defined in Table 4.1 is used. During the simulation the inverter best suited for the solar generation capacity is selected from a California Energy Commission (CEC) database within SAM's libraries [65]. The initial investment cost for all system sizes was chosen from recent estimates for module, inverter, balance of system (BOS) and labour costs [13].

Weather data is automatically obtained from the European Commission's photo-voltaic geographical information system database for the specified location [64], or collected using the discussed fallback method if the online database is unreachable. The simulation output is a list containing the kWh generation estimates for each hour of the simulated year. The method iterates through installed capacity steps of 1 kWp, storing one year's hourly generation estimates for each iteration within a database until the maximum system size determined by the investment budget or available roof area is reached.

4.4.3 EWH simulation

The EWH usage data was divided amongst the two types of buildings they were installed in namely the school building and the residences on the school premises. This division was chosen due to their differences in usage profiles as presented by Figure 3.9. The input parameters for the EWH usage profile determination is the measured hot water usage with their specific timestamp, the type of building the EWH is installed in and the school term and holiday dates for that specific year. It should be noted that the creation of the seasonal lists G_s is done on a rolling basis, such that the smart-scheduling will only start to take affect after the first week of measurements have been completed. During the first week of operation the control scheme will default to thermostat control with peak limiting and solar excess dumping.

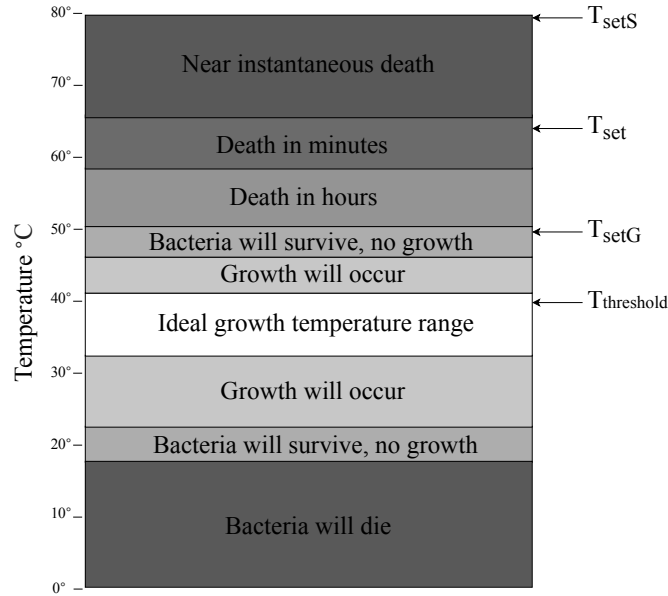


Figure 4.3: Effects on Legionella bacteria at different water temperatures with simulation bounds highlighted.

The three EWH control schemes are simulated at a temporal resolution of 1 minute over a 365 day period. The parameters required to model an EWH are: the orientation, tank volume, element power rating, thermostat hysteresis and thermal resistance. The following measurements are required for each sample period (chosen as 1 minute): volume water used, inlet water temperature, and ambient temperature. The temperature bounds for the simulation was selected with careful attention given to Legionnaires' disease, a severe form of pneumonia induced by the Legionella bacteria.

Figure 4.3 presents the effects of different temperatures on the bacteria, giving an indication to the dangers faced when not thoroughly heating the EWH and a necessary compromise at the expense of slightly increasing grid usage or solar dumping to ensure that the water temperature is high enough to kill of any remaining bacteria. The upper bound for solar dumping was selected as it remained within the operational ranges of most standard EWHs [79], and the threshold temperature for ideal user-comfort was selected from results obtained during a previous study where a field test was performed with a student group [80].

As a result of the varying temperature bounds the amount of hot water consumed will differ for the same overall usage volume as presented by the following energy balance equation:

$$V_{draw} = V_{total} \frac{(T_{use} - T_{in})}{(T_{warm} - T_{in})} \quad (4.1)$$

with V_{draw} being the volume of hot water leaving the EWH with a tem-

perature of T_{warm} , V_{total} being the combined volume of hot and cold used at the tap at T_{use} and T_{in} being the inlet temperature of 20°C. Jacobs *et al.* [80] found that the desired temperature for water at the shower outlet to be 40.2°C, chosen as the T_{use} temperature.

However, the measured data obtained for this study only provided the volume of hot water leaving the EWH, and was assumed to have a target temperature of 65°C as it is within the standard range of a typical South African EWH [78; 80]. Therefore, all V_{draw} volumes have to be scaled from the original measurement data such that:

$$V_{warm(New)} = V_{warm(measured)} * \frac{T_{measured}}{T_{warm}} \quad (4.2)$$

To ensure that the energy balance is maintained throughout all simulated periods. The method output includes the energy used, peak power usage, cold event count and mean event temperature for each EWH.

4.5 Evaluation and metrics

This section describes the metrics used to evaluate the capabilities of the developed system to fulfil the research objectives of the study. In addition to the individual component metrics the system will also be assessed as a whole to determine its computational efficiency and to determine its financial viability when combining multiple interventions.

4.5.1 Forecast model

A primary objective of this study is to accurately forecast the potential monthly bills of a school. As such, the reliability of the forecast is determined by comparing the predicted usage and monthly bills to the real, measured usages:

$$\text{Average forecast accuracy (\%)} = \left| \frac{\sum_{i=1}^n y_{forecast,i} - \sum_{i=1}^n y_{observed,i}}{\sum_{i=1}^n y_{observed,i}} \right| * 100 \quad (4.3)$$

With $y_{forecast,i}$ and $y_{observed,i}$ being the energy usage at hour i . The usages are summed for all n hours within a specified day, month or year and then used to determine the overall percentage error for that time frame.

Further comparisons are made using the mean absolute percentage error (MAPE) metric to observe the severity of hourly fluctuations when comparing the measured and forecast data, and is described as:

$$MAPE(\%) = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_{forecast,i} - y_{observed,i}}{y_{observed,i}} \right| * 100 \quad (4.4)$$

With $y_{forecast,i}$ and $y_{observed,i}$ being the energy usage at hour i . The usage errors are summed for all hours within a specified day, month or year and then divided by the total number of hours n .

The effectiveness of the forecasting technique is determined by training the non-scaled load profile for each school type, and then testing it using an independent school as to avoid over-fitting. In the case of the three primary schools: Two schools were used to train the profile, and the third school's monthly usage data obtained from utility bills was used to test it.

4.5.2 Solar intervention

The performance of the simulated solar system is determined by its ability to reduce the effective cost of solar within a specified building. The economic viability of the system is described by the IRR rate in Equation 3.20. The system grid interaction is defined using two indicators. The first is the load-cover factor, and describes the percentage of electrical demand satisfied by the on-site energy system over n simulation intervals:

$$LCF(\%) = \frac{\sum_{i=1}^n UsedGeneration_i}{\sum_{i=1}^n BuildingDemand_i} * 100 \quad (4.5)$$

The second grid interaction indicator is the supply-cover factor, specifying the percentage of on-site generation used by the building:

$$SCF(\%) = \frac{\sum_{i=1}^n ExcessGeneration_i}{\sum_{i=1}^n TotalGeneration_i} * 100 \quad (4.6)$$

These indicators proved useful for providing insight into the daily and seasonal capabilities of various renewable energy systems and their respective grid interaction. A third indicator, the loss-of-load probability factor, used to determine the likelihood of the on-site generation failing to supply electricity is often employed for system sizing purposes in net-zero building design [53]. However, solar optimisation performed during this study was done in a *cost-focused* manner, and the possibility of net-zero or zero loss-of-load probabilities were not considered.

4.5.3 EWH control schemes

The performance of the three developed intelligent water heater control schemes is measured in terms of their ability to maximise three control objectives, namely, cumulative energy usage reduction, efficiently managing the load, and user comfort. All approaches will be compared to the commonly employed thermostat control approach serving as a baseline, as this is the most common within the South African context [77].

4.5.3.1 Energy usage

A major research objective of this study is to reduce energy costs within schools by means of energy saving interventions. The monthly energy costs of a building is directly related to the basic energy usage and as such, it follows that reducing the overall energy usage of the EWHs is a high priority. This metric evaluates the total energy usage for each of the developed control schemes by comparing them to the baseline usage over the simulated period.

4.5.3.2 Peak demand reduction

A key objective of the control schemes is their ability to sufficiently manage the load of EWHs through optimal power delivery scheduling, while ensuring that the EWH demand does not exceed the specified limit.

Peak power limiting:

All control schemes are subject to grid demand limits by means of a specified peak power usage managed by a peak power demand manager. The PPDM is evaluated to ensure the limit is not exceeded, and the new demand profile will give insight into the load shifting that occurred as a result of smoothing the original non-limited power surges.

Maximum demand reduction:

The demand limiting approach is evaluated to ensure that the grid usage of the EWHs did not affect the building's peak power demand, allowing for a potential reduction in monthly demand charges. This would confirm that the schedule controller correctly referenced the recorded monthly maximum with present building usage before heating the EWHs.

4.5.3.3 User comfort

From a user's perspective, the water temperature during a usage event is their highest priority. If the outlet temperature of the EWH is too cold, it will likely discourage the user from continuing with the employed energy saving intervention, and is therefore a necessary control objective that has to be met.

Mean event temperature:

To gain a perspective of user-comfort during the simulated intervention period the mean hot water temperature during a usage event is measured. This metric will then be able to provide insight with regards to the control scheme's ability to maintain a set target temperature. Additionally, the ability

of the control schemes to mitigate the threat of Legionnaires is evaluated by its ability to regularly achieve temperatures within the bounds deadly for the bacteria as presented in Figure 4.3. Chosen as a minimum of 50°C, and required to be heated to the target temperature within 48 hours as to prevent Legionella growth (with an incubation period of 2-10 days [81])

Cold events:

The second measure for user comfort is the number of cold events experienced. A cold event is defined as a hot water usage event with a recorded temperature below a set threshold $T_{threshold}$. Recording the number of cold events compared to overall usage events will determine the viability of an EWH approach as too many cold events will diminish the enthusiasm users to continue with the energy saving intervention.

4.6 Chapter summary

This chapter discussed the experimental setup used during the simulation of the developed system, with data cleaning measures and an outlier detection algorithm is explained. The simulation setup section defined the input and output variables of each system component as well as assumptions made regarding system parameters. The chapter concludes by defining the metrics used to evaluate the system such that it fulfils the requirements set out in Chapter 1 of the study.

Chapter 5

Results

This chapter expands on the results obtained from the discussed experimental setup.

Section 5.1 is concerned with research objective 1. Forecasting results are provided with accuracy metrics quantifying the performance of the model. The forecasting results are shown for all five schools and also presented per school type. Section 5.2 relates to the combination of the yearly demand forecast and the solar modelling process. The solar generation results are shown for both school types and concludes with a system viability forecast. Section 5.3 pertains to the addition of EWH intervention to the forecast and solar components of the system. This section describes a full system analysis relating to all four research objectives. A single school is chosen and results are section-alised for each component with the complete system evaluated for utility bill reduction, basic energy and peak-demand savings and user-comfort for each of the different energy-saving interventions. The chapter concludes with the simulation performance metrics in Section 5.4. These metrics give an indication to the scalability of the developed system for large scale interventions.

5.1 Energy usage forecasting

This section details the forecasting results obtained from the experimental setup. Table 5.1 provides information about the five tested schools.

For each evaluated school the normalised load profile was trained using measurement data (January 2016 until December 2018) from the remaining schools of the same type. The monthly utility bills for the evaluated school from 2017 was then used to scale the normalised profile and compared to the real, measured, 2018 usage profile. Figure 5.1 shows the averaged basic energy forecasting accuracy for both school types (grey for primary schools and black for high schools) depending on the number of utility bills used to scale the normalised profile. As can be seen in the figure, the increase in overall forecast accuracy coincides with an increase in prediction accuracy. The maximum

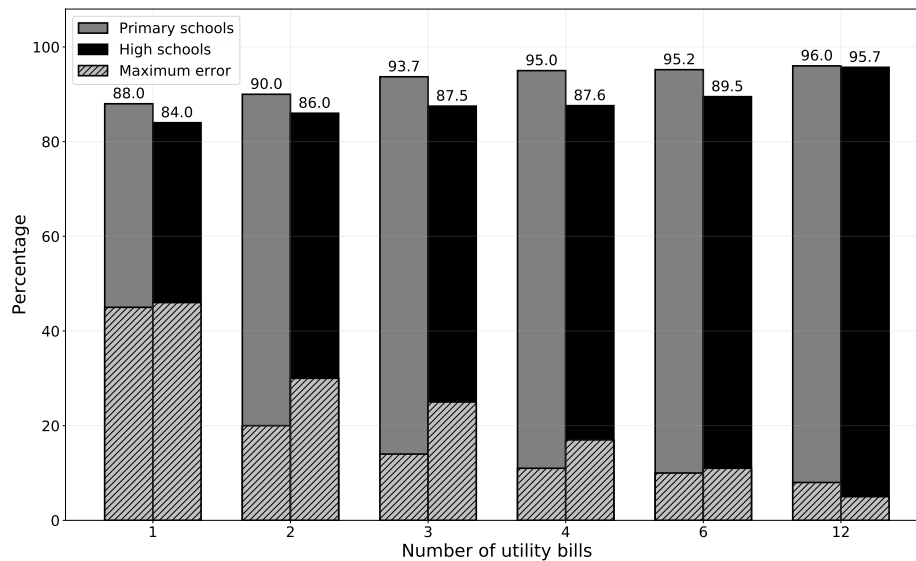


Figure 5.1: Basic energy usage forecast accuracy for a varying number of utility bills

Table 5.1: Overview of evaluated schools

| School | School type | Tariff structure | Number of learners | Average yearly energy usage* (MWh) |
|--------|-------------|------------------|--------------------|------------------------------------|
| A | Primary | Demand | 910 | 165.6 |
| B | Primary | TOU | 671 | 150.4 |
| C | Primary | TOU | 820 | 112.6 |
| D | High | TOU | 550 | 182.6 |
| E | High | TOU | 718 | 334.3 |

Note:* Average usage per year from 2016 to 2018

error metric (dashed sections) indicates the largest total usage error for a 1 year forecast for each school type and number of utility bills. This gives an indication to the deceiving nature of averages, and highlights the importance of using a large number of utility bills. For the remaining performance evaluations made in this study all forecasts were made using a full year's utility bills.

Figure 5.2 presents the average maximum monthly demand forecast for primary and high schools versus the measured maximum monthly demand during 2018 using 12 month's demand data. Overall the results are favourable with the demand forecast accuracy averaging 97% for primary schools and 93% for high schools.

All usage and demand forecasting results are presented in Table 5.2. The maximum monthly forecast metric returns the percentage difference of a fore-

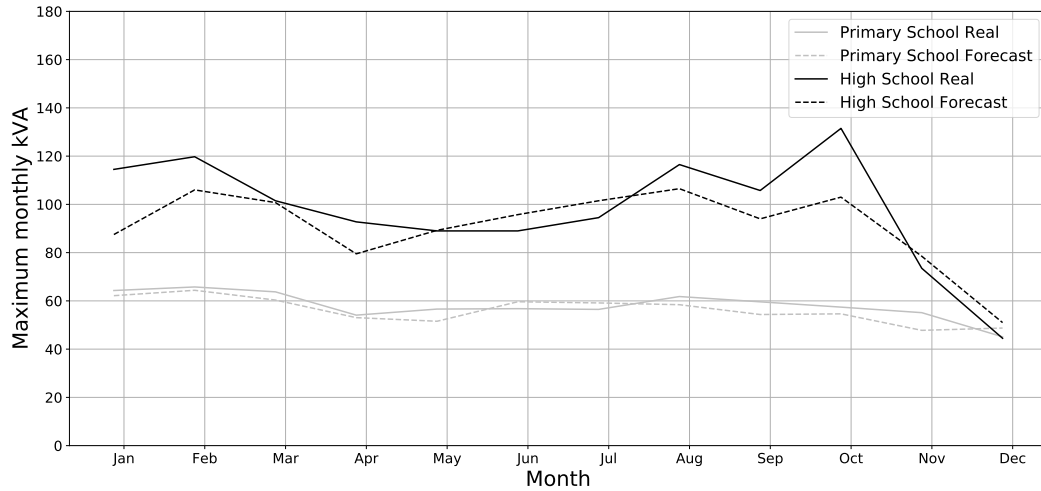


Figure 5.2: Averaged maximum monthly demand forecast versus measured maximum demand for 2018.

Table 5.2: Forecast simulation results

| School | Maximum monthly forecast error | | Average yearly forecast accuracy | | MAPE | | |
|----------|--------------------------------|------------|----------------------------------|------------|-------------------|------------------|-----------------|
| | Energy (%) | Demand (%) | Energy (%) | Demand (%) | Hourly energy (%) | Daily energy (%) | Peak demand (%) |
| A | 15.0 | 26.8 | 98.6 | 99.6 | 39.8 | 22.1 | 8.07 |
| B | 23.6 | 18.7 | 97.2 | 96.9 | 36.9 | 21.7 | 9.10 |
| C | 36.1 | 15.5 | 90.7 | 93.4 | 44.6 | 26.8 | 6.15 |
| D | 26.7 | 36.9 | 95.4 | 97.3 | 37.8 | 25.1 | 17.6 |
| E | 29.9 | 46.6 | 96.9 | 88.1 | 40.4 | 21.3 | 23.5 |

casted month's basic energy usage and maximum monthly demand versus the comparative measured month, and returns the largest value. With respect to basic energy usage it is seen that all schools performed similarly, and for monthly demand the primary schools fared better than the high schools. This is mostly due to more training data being available for the primary schools (Schools A, B and C versus D and E), as the process of training the seasonal usage lists and scaling matrices results in a *smoothing* effect, reducing the impact of usage surges only seen at a specific school. This *smoothing* will generate a "best fit" profile for a school dependant on a specific feature set defining a building type (such as location or quintile). Due to nature of demand forecasting all evaluated schools will differ from the normalised profile, but to varying degrees.

The average yearly forecast accuracy provides the percentage difference for the whole year's energy usage and summed monthly demands. A higher accuracy indicates that the schools' overall usage and monthly demands remained similar to the input data year, and reaffirms the models ability to effectively

Table 5.3: Input data for solar simulations

| School | Est. available roof area (m ²) | Roof tilt (°) | Azimuth (°) | Latitude | Longitude |
|----------|--|---------------------|----------------|------------|------------|
| A | 2100 | 28 | 350 | 33°93'31"S | 18°88'16"E |
| B | 1500 | 24 | 340 | 33°94'69"S | 18°85'57"E |
| C | 1800 | 24 | 340 | 33°94'51"S | 18°85'44"E |
| D | 1300 | 25 | 350 | 33°93'11"S | 18°87'97"E |
| E | 2000 | 35 | 340 | 33°94'57"S | 18°85'94"E |

scale the normalised lists to create hourly usage figures from single monthly usage inputs.

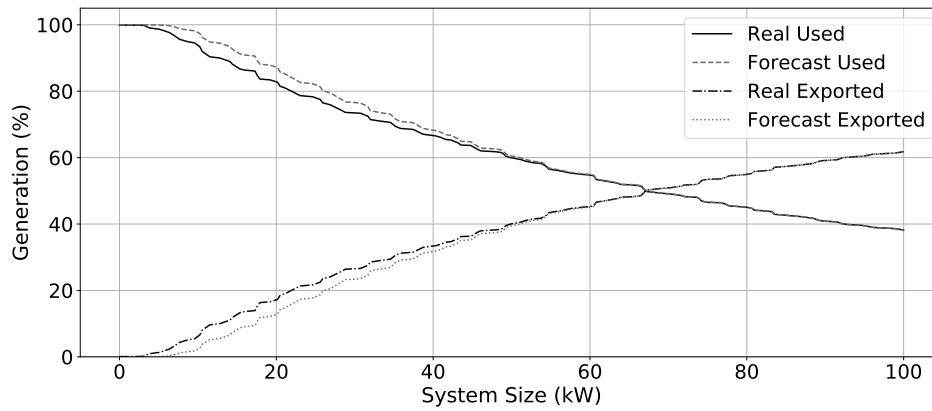
The MAPE metric quantifies the consistency of a forecast, and penalises outlying results. Furthermore, large scale changes also contribute to larger MAPE scores with similar usage errors during daytime and night time resulting in different error rates. The hourly and daily energy usage metrics show the MAPE scores for 8760 and 365 simulation intervals respectively. The worse performing hourly scores are largely due to the intra-day usage scale changes. The school's performed well for the monthly peak demand MAPE scores, indicating a preferable mean monthly demand forecast.

5.2 Solar intervention

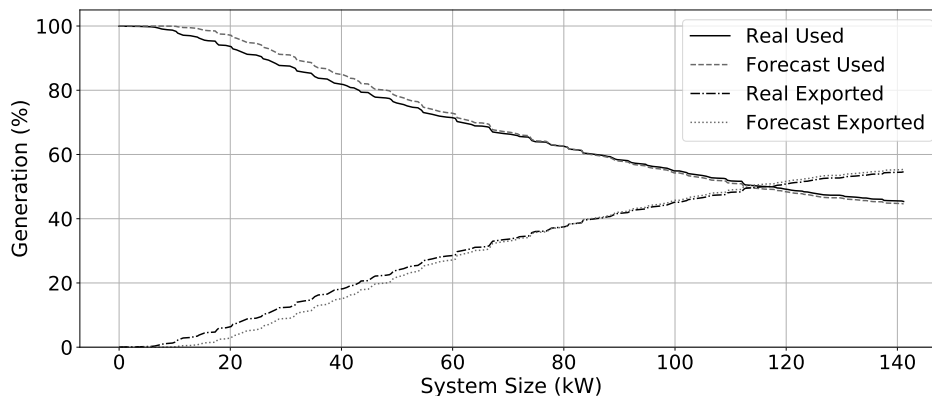
After the scaled load profile is determined a series of solar systems with varying generation capacities are simulated. The hourly generation estimates are then subtracted from the hourly load forecast to calculate the energy and demand savings, the system utilisation and if the results justify an investment using an IRR metric. Table 5.3 shows the input data used to perform the solar simulations.

Figures 5.3a and 5.3b present the averaged percentage of solar generation capacity used versus the generated energy exported to the electricity supplier for primary and high schools respectively. These figures give an insight into the grid-interaction and load cover the system provides. Both figures show similarly shaped graphs, indicating similar inter-day usage profiles for the forecast and measured loads. However, the forecast usage profile is slightly better suited for solar intervention due to a larger percentage of generation capacity being used. Also seen is the difference in the export mid-point, whereby the solar system utilisation drops below 50% with high schools requiring systems nearly 60% larger than primary schools.

Figures 5.4a and 5.4b present the averaged internal rate of return for primary and high schools over a 25 year system lifespan. The figures show the averaged IRR as well as the maximum and minimum IRR boundaries obtained for the different schools. The assumptions made with regards to yearly cash



(a) Primary school solar generation results

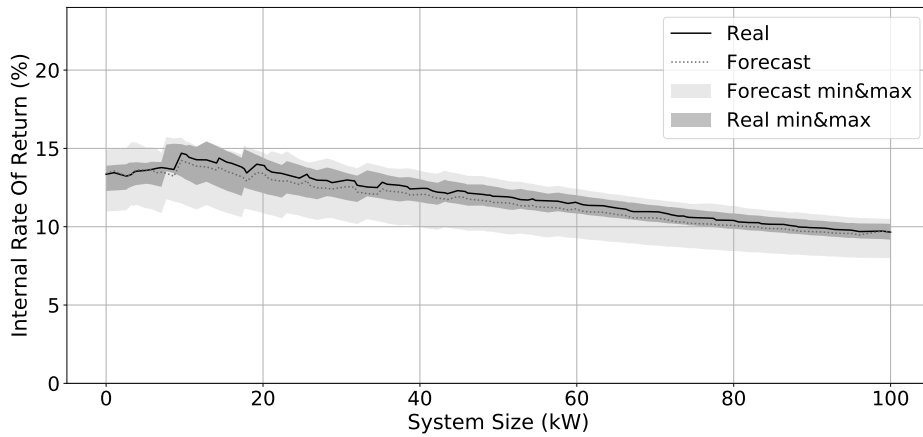


(b) High school solar generation results

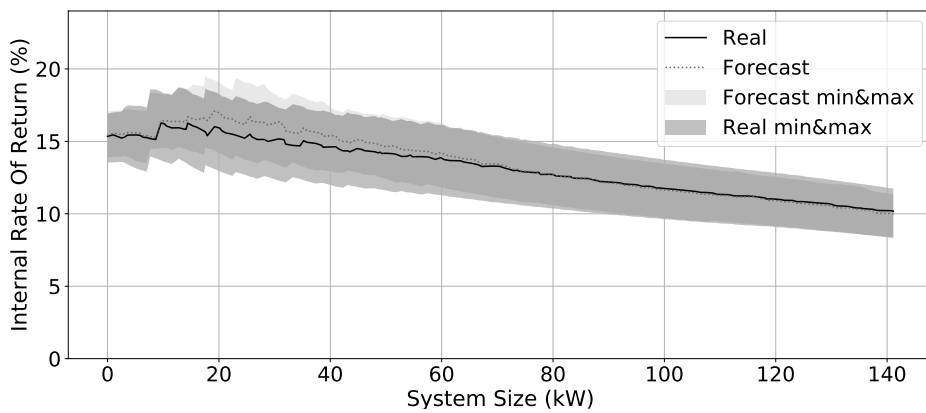
Figure 5.3: Solar generation usage for varied system sizes.

flow, necessary to calculate the IRR, was discussed in Section 3.2.3. The IRR gives an indication of the profitability of an investment, with a larger percentage equating to a greater return. The point of peak savings is achieved for system sizes of 12kW DC and 17kW DC for primary and high schools respectively with expected payback times of 6.6 and 5.8 years accounting for the time value of money. If the school wishes to install a different sized system the payback times can still be determined using the IRR metric.

The shaded areas present the minimum and maximum IRR graph paths for the real and forecast simulation results, indicating a greater discrepancy within the primary school results as the minimum forecast IRR was 1.0% to 1.5% lower than the real value. Differences in the forecasted and real IRR rate are due to variations in their daily load profile. This, in addition to dissimilar monthly demand estimations led to mismatched yearly cash-flows being magnified over the 25 year life cycle of the system.



(a) Primary school internal rate of return.



(b) High school internal rate of return.

Figure 5.4: Internal rate of return for various system sizes.

5.3 Complete system simulation with EWH intervention

This section presents a complete overview of the developed system for a school. Input parameters used are those presented in the simulation setup section and the evaluated school data is that of **School A** unless explicitly stated otherwise. School A is billed on a demand based tariff, and reducing the maximum monthly demand is a key priority.

5.3.1 Load forecast

All forecasting results of school A are presented in Table 5.2. The yearly usage profile is shown in Figure 5.5. The forecast successfully captures the seasonal

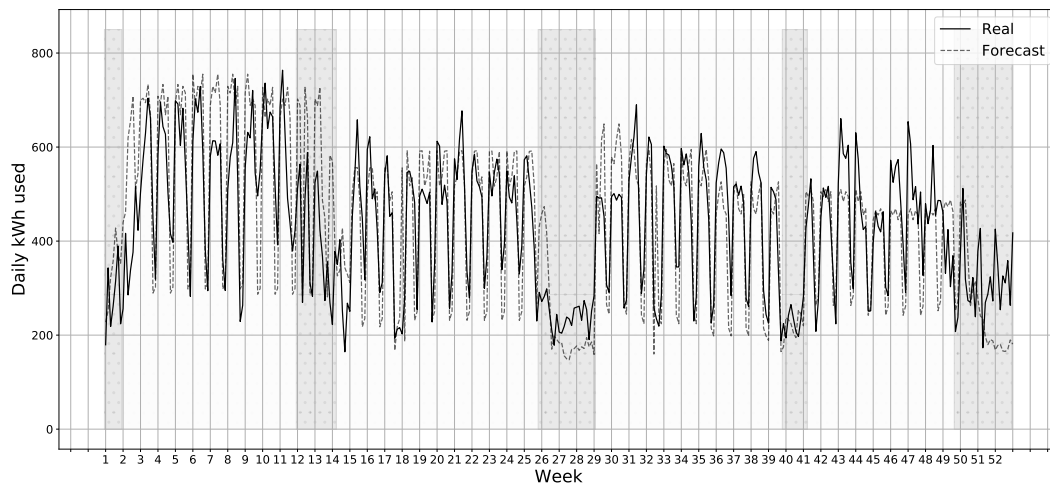


Figure 5.5: Energy usage forecast versus measured usage for school A in 2018.

trend of the school with weekday fluctuations present. The profile is seen as "best-fit" and does not cope well with large intra-day surges as seen in the measured profile but still manages to provide a demand profile suitable for solar optimisation.

5.3.2 Solar intervention

For the case presented in this study it was assumed school A had R500 000 available for the solar system. Figure 5.6 shows the IRR and generation results for varied system sizes. The simulation limit was set by the available budget, resulting in a maximum available system size of 37440 Watt DC (at a cost of R13.35 per installed Watt). The optimal system size for the school is determined to be 8kW DC, with an IRR of 16% and a system utilisation rate of 100%. However, for the purpose of this study excess generation is preferred with regards to dumping solar energy into the EWHs, and it was assumed school A selected the largest possible system consisting of 117 solar modules. The load cover factor of the chosen system was calculated as 0.31, meaning 31% of the building's total energy usage was supplied by the solar system. The complementary index, the supply cover factor was 0.74, meaning 74% of the system's total generation capacity was consumed by the building, as shown in Figure 5.6.

5.3.3 Smart-scheduled EWH intervention

School A is equipped with seven EWHs on its premises, three within the school building (2kW, 100L) and four within residences on the school grounds (3kW, 150L) totalling a potential peak power usage of 18kW if used simultaneously. For the smart-scheduling and bi-thermal control schemes the peak power man-

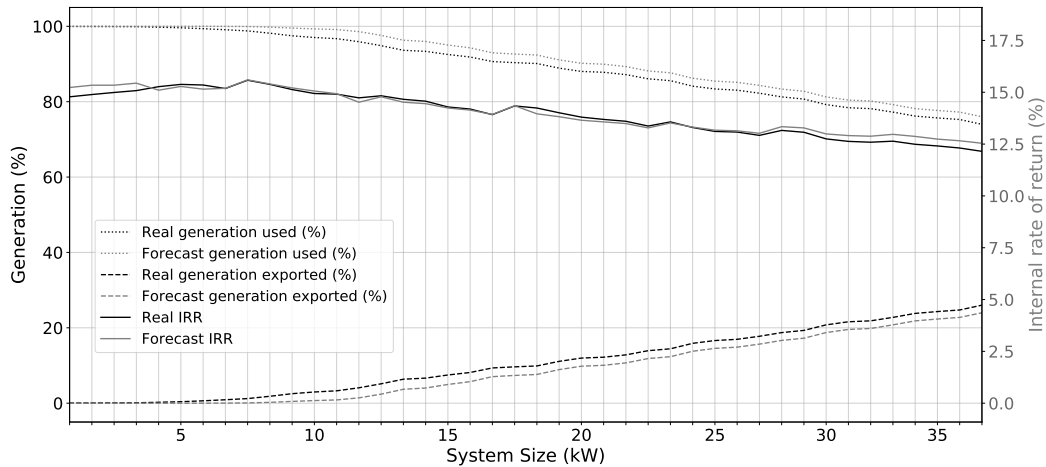
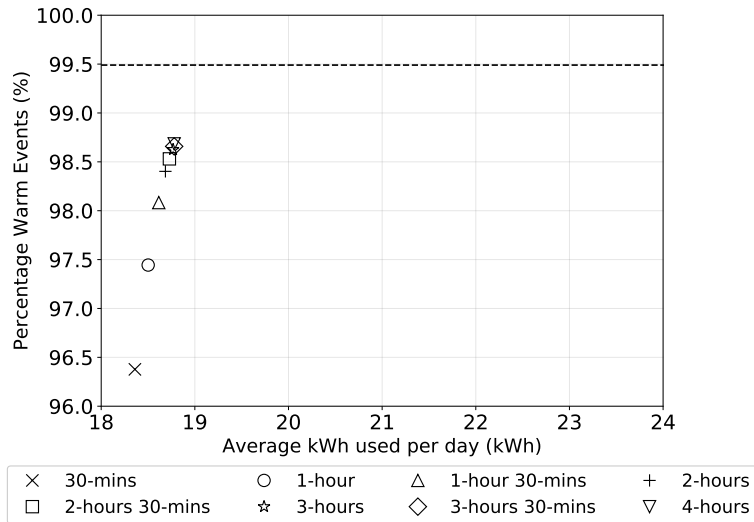


Figure 5.6: Solar system IRR and generation results for school A.

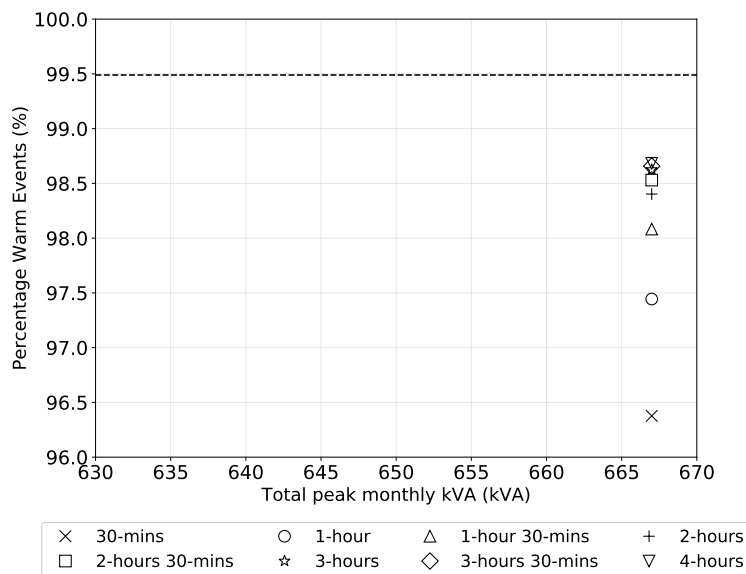
aged by the PPDM was chosen to be 10kW to reduce the potential for large monthly demand surges while allowing enough EWHs to be heated simultaneously to maintain a suitable level of user-comfort.

Figures 5.7a and 5.7b show the average daily kWh grid-usage and total peak kVA for the year for the bi-thermal heating mechanism with a prioritisation scheduler and excess solar dumping. The plotted results show the usages for various scheduling threshold periods t_{et} , whereby an EWH will not be allowed to be heated if the expected usage time t_{event} is greater than t_{et} . A longer scheduling period results in larger standing losses as the EWHs are heated earlier, and reduces the number of cold events by allowing the heater elements to be switched on for a longer period. From Figure 5.7b it is seen that the maximum monthly kVA usages for the year, defined as the sum of the building's peak monthly demand throughout the evaluation period remain the same for all scheduling periods. This is due to the water heaters being scheduled to heat for the first usages very early in the morning. By the time the peak-demand period arrives late morning during winter months or early afternoon during summer months the water has already been heated to a suitable temperature regardless of the scheduling period. From the results the cold event percentages reached a steady state for scheduling periods of greater than 3 hours, and was selected as the optimal scheduling period providing maximal user comfort for the energy used. The dashed line represents the warm event percentage of the thermostat or baseline EWH control scheme. The decreased user comfort compared to thermostat control for scheduling periods are due to the CS's lower grid target temperatures (50°C versus 65°C).

Figures 5.8a and 5.8b present the average daily kWh grid usage and maximum monthly kVA for the year for the demand limiting control scheme with a bi-thermostat and solar dumping. The plotted results show the usages as the demand limit is raised by a specified percentage $D_{threshold}$. A higher demand limit increases the peak monthly kVA for the year while reducing cold events



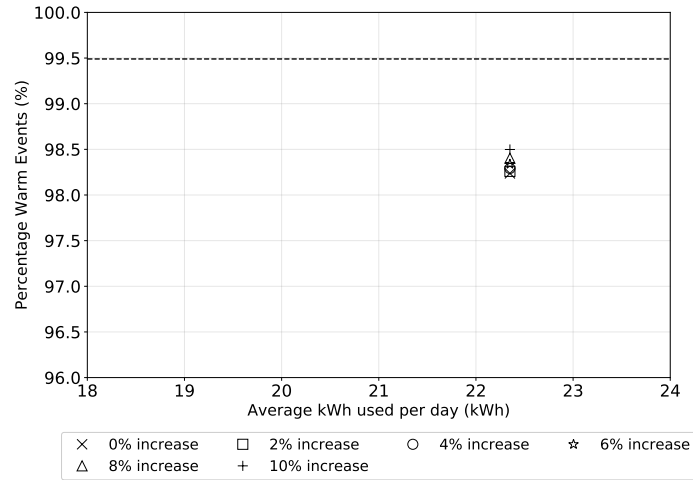
(a) kWh used from grid



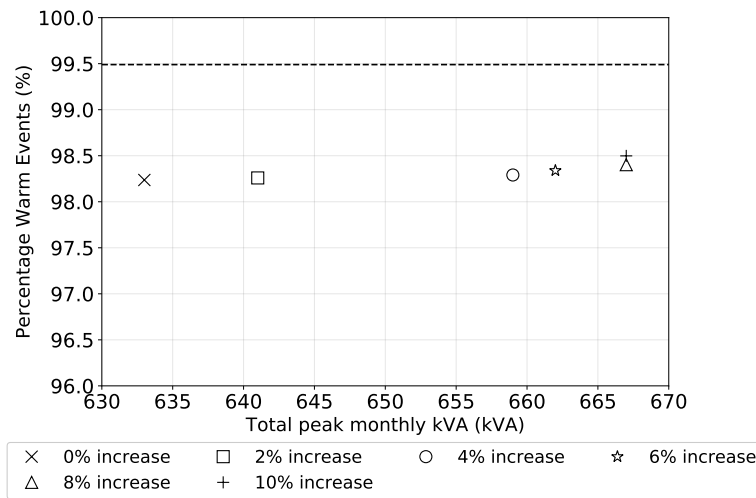
(b) Peak monthly kVA for the year

Figure 5.7: Usage results for the bi-thermostat control scheme with prioritised heating and solar dumping.

by allowing more EWHs to be heated simultaneously. The large increase in kWh usage compared to the bi-thermostat CS is due to the 3 hour scheduling period, t_{et} , being removed. As a result, the EWHs enter a thermostat control state, ensuring that the water heaters are always above the minimum threshold temperature of 50°C without taking the estimated time before the next water usage event into account. This was done to allow the water sufficient time to heat, as no EWH elements are allowed to be active during the demand-limited period with a 0% increase in peak monthly demand.



(a) kWh used from grid

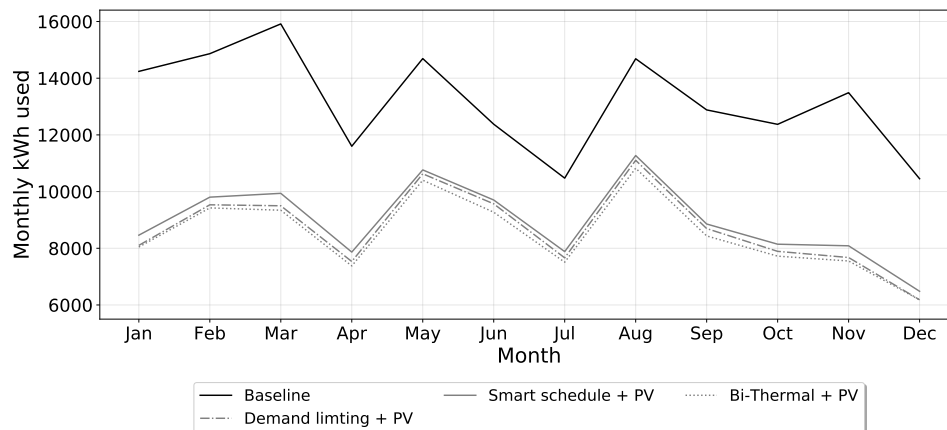


(b) Peak monthly kVA for the year

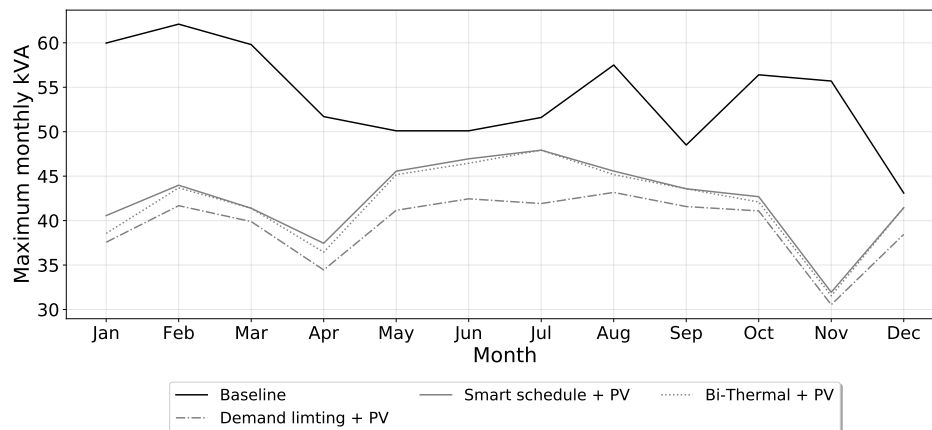
Figure 5.8: Usage results for the demand-limiting control scheme with a bi-thermostat and solar dumping.

From the maximum monthly kVA graph the cold events for the 4% and 6% as well as the 8% and 10% increases are exactly the same. It should be noted that this is due to the percentage increase in allowable demand not being greater than the EWH element rating of (2kW and 3kW). From the results a demand-limiter with 0% demand increase was selected as a result of providing the greatest demand-charge savings while maintaining a suitable level of user comfort.

Figures 5.9a and 5.9b present the kWh used and the peak monthly power demand for each intervention for each month of the year. From the graphs it is apparent that all three interventions provide a noticeable improvement compared to the measured baseline. From the kWh per month graph it is



(a) kWh used from grid



(b) Peak monthly kVA for the year

Figure 5.9: Usage results for the evaluated interventions.

clear that the introduction of bi-thermal temperature thresholds allowed for a sizeable reduction in energy usage compared to the smart schedule CS by exploiting the EWHs thermal capacity through an increased solar dumping threshold.

The peak demand savings graph shows large improvements for the simulated interventions when compared to the building's baseline. The relatively small improvements during the winter months can be attributed to the fact the the school's monthly peak demand occurs earlier in the morning when solar PV generation is minimal. During the summer months the school's energy usage profile is better suited for solar intervention, with the peak monthly

demand occurring early afternoon when solar PV generation is at its highest. Moreover, from the graph the improvements seen from the demand limiting CS is clearly visible compared to the other two simulated configurations, resulting in large monthly demand savings. The complete system results are presented in Table 5.4.

Table 5.4: Simulation results for evaluated school for the period 1 January 2018 to 31 December 2018

| Parameter | School baseline | Solar & prioritisation scheduling | Solar & bi-thermostat | Solar & demand-limited EWHs |
|---|-----------------|-----------------------------------|-----------------------|-----------------------------|
| Yearly energy usage (MWh) | 158.1 | 106.6 | 101.5 | 103.4 |
| Daily energy usage [min,mean,max] (kWh) | [165,433,764] | [94,292,673] | [92,278,661] | [94,283,665] |
| Daily max kVA [min,median,max] | [11,35,62] | [10,27,54] | [10,27,53] | [10,25,52] |
| Warm event percentage (%) | 99.45 | 99.24 | 98.52 | 98.21 |
| Maximum hours under 50 °C | 2.5 | 28.2 | 32.5 | 5.4 |
| Average utility bill saving (%) | 0 | 23.2 | 24.8 | 26.7 |

5.4 Simulation performance

The system was designed with computational simplicity in mind, allowing for viability assessments to be determined speedily. The simulation performance was evaluated on a PC with a i5 6600k with a clock speed of 4.6 GHz and 16 GB of DDR4 RAM. The python simulation time was recorded with the *timeit* module and the ran as a single thread.

The load load profile construction and scaling was completed in 1.34 seconds per school. For the solar simulation a weather file was retrieved from an online database (taking 5-10 seconds dependant on the download rate). The solar simulations were completed in 93 milliseconds per iteration. If a building is simulated with a 100kW DC solar system in 1kW increments the simulation will take 9.3 seconds to iterate through all configurations. The EWH model simulations were completed in 3.18 seconds per EWH per year, roughly 150 times faster than existing two-node models [82].

The evaluated full system simulation described in Section 5.3 took 31.1 seconds to complete. A larger system with a 200kW DC system and 50 EWHs was simulated as well and took 3 minutes and 12 seconds to finish.

Chapter 6

Conclusion

The educational gap within socio-economic groups in South Africa is immense, and with the recent increases in the cost of Electricity many schools have been forced to divert a greater portion of their non-personnel funds to cover the expense. This places these school at a disadvantage, and constrains the ability of the education system to provide learners of these schools with a pathway out of poverty.

To address this problem a system was developed to simplify the process required to determine the benefits of a solar system for a school by providing financial viability forecasts using easily accessible input parameters. The method removes the need for investing in smart meters, overcomes the dependence on external contractors or suppliers, and negates a technical understanding and time-intensive analysis of solar systems. Additionally, research into the benefits of load shifting and demand reduction using intelligent geyser scheduling was completed.

6.1 Evaluation of work

This section presents a summary of the work and results presented in Chapters 2 to 5, with references to the research objectives stated in Chapter 1.

6.1.1 Energy usage forecasting

A method capable of determining the energy usage within schools was developed in Section 3.1 with results presented in Section 5.1. The process creates a generic energy consumption profile for a building from measured energy usage data of a subset of schools, and is expanded to scale the load profile using only monthly usage data from utility bills and seasonal dates to produce a load forecast. The method was validated using five datasets each containing the hourly energy usage measurement data from schools over a period of three years, and was capable of forecasting the yearly energy consumption of the

schools to within an averaged accuracy of 5% while estimating the maximum monthly demand to 6% of the measured usage. These simulations were performed within 1.34 seconds on a conventional computer and confirm the first research objective:

Research objective 1:

To develop a computationally inexpensive and accurate model capable of forecasting the energy usage and maximum monthly demand of a school using easily accessible parameters.

6.1.2 Solar system optimisation

A PV optimisation technique is developed in Section 3.2 and uses the energy usage forecast to determine the potential profitability of the solar system through metrics including the internal rate of return and the utilisation of the system's generating capacity. The results are shown in Section 5.2 and it was found that the potential profitability of a solar system tailored to the school's budget could be determined, presenting itself as a valuable tool for reducing the financial burden many schools face. These results confirm the second research objective and partially confirms the fourth research objective:

Research objective 2:

To develop a method capable of determining the energy usage of a school after solar intervention using a cost optimised, load-matched solar system.

Research objective 4:

To provide cost and savings forecasts for various system sizes and interventions.

6.1.3 Intelligent EWH control

Three control schemes of intelligent water heater scheduling were researched in Section 3.3 with results shown in Section 5.3. Firstly, a priority based scheduler was configured to heat water using the school's water usage history, while diverting any excess solar energy to the water heaters to exploit their energy storage capabilities. This reduced the evaluated school's monthly energy bill by 23.2% compared to the school's always-on control baseline. Secondly, a bi-thermal control method was added to the priority based scheduler, employing a temperature delta to increase the amount of solar energy to be stored within the water heater tank while minimising their grid reliance and improving the monthly savings to 24.8% per month. Finally, a demand-limiter control scheme was implemented in conjunction with bi-thermal control resulting in large demand-charge savings and an average energy bill reduction of 26.7% per month, producing the maximum savings while maintaining suitable levels of user comfort. These results confirm the third and fourth research objectives:

Research objective 3:

To combine the thermal-storage capabilities of EWHs with load-matched solar through smart-scheduling techniques capable of reducing peak monthly demand and basic energy usages.

Research objective 4:

To provide cost and savings forecasts for various system sizes and interventions.

6.2 Recommendations for future work

Research done during this study has shown there to be limitations with the methods used regarding data collection using physical devices installed within a small number of similarly structured (quintile 5) schools. Because of these limitations, more research should be done by implementing the interventions in a larger number of schools to provide more conclusive results.

Monthly demand reduction due to PV and EWH interventions is difficult to estimate from hourly measurement data, and increasing the temporal resolution of the measurement data will allow for more accurate estimations to be made.

An improvement to the demand-limiting approach presented in this study is to employ optimal battery sizing methods to achieve further demand reduction while increasing the solar system's supply cover factor. This will reduce peak grid power usage during the early mornings, particularly during winter months.

Future work can also be expanded to include other forms of intervention, including air-conditioning units and heat pumps, broadening the applicability of the performed research.

Appendices

Appendix A

System software user guide

A.1 Installation

Open a command window in the main directory containing the main.py file.

Note: This software only works with python versions 3.6.0 and 3.7.0.

First, change the console directory to the main software folder (containing the main.py file). Then create a virtual environment using the "**python -m venv myvenv**" command. Then change directory to the "Scripts" folder of the "myvenv" virtual environment using commands: "**cd myvenv**" followed by "**cd Scripts**". Now activate the virtual environment using command "**activate myvenv**" and change the directory back to the main program folder using the "**cd..**" command twice.

Once the virtual environment has been activated the program can be run using the command: "**python main.py**". (An internet connection is necessary to install the packages)(Required packages are: *matplotlib*, *numpy*, *scipy* and *PyQt4*). Once the packages have been installed the interface will appear as shown in Figure A.1.

A.2 Usage

The program requires the following input data:

1. The School type (Primary or High).
2. The School's tariff structure (Basic, Demand or Time-of-use).
3. The monthly usage data for the specific tariff type, at least one month's data is required.
4. The available budget for the installed system.

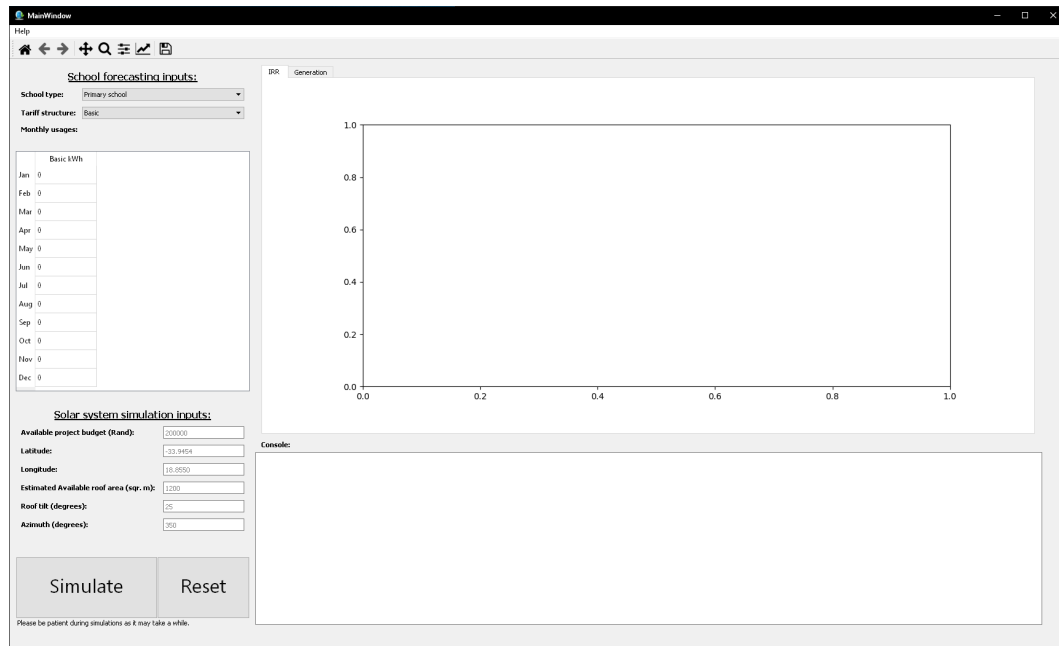


Figure A.1: The graphical user interface after running the main.py command.

5. The latitude and longitude of the school.
6. The roof area, roof tilt and azimuth of the school's roof. (Azimuth is the angular direction of the school's roof in relation to north on a horizontal plane as shown in Figure A.2.

Once the simulation button is clicked all inputs are verified. The simulation outputs are as follows:

1. The estimated yearly kWh usage of the school.
2. The optimal solar installation size with an estimated payback period.
3. The maximised budget installation with an estimated payback period.
4. Two graphical plots with the IRR (internal rate of return) and solar generation used by the school and returned to the grid..

An example output is shown in Figure A.3.

A.3 Question and bugs

Any questions regarding the program's usage, or bugs found can be forwarded to Stefan Gerber at 18195342@sun.ac.za.

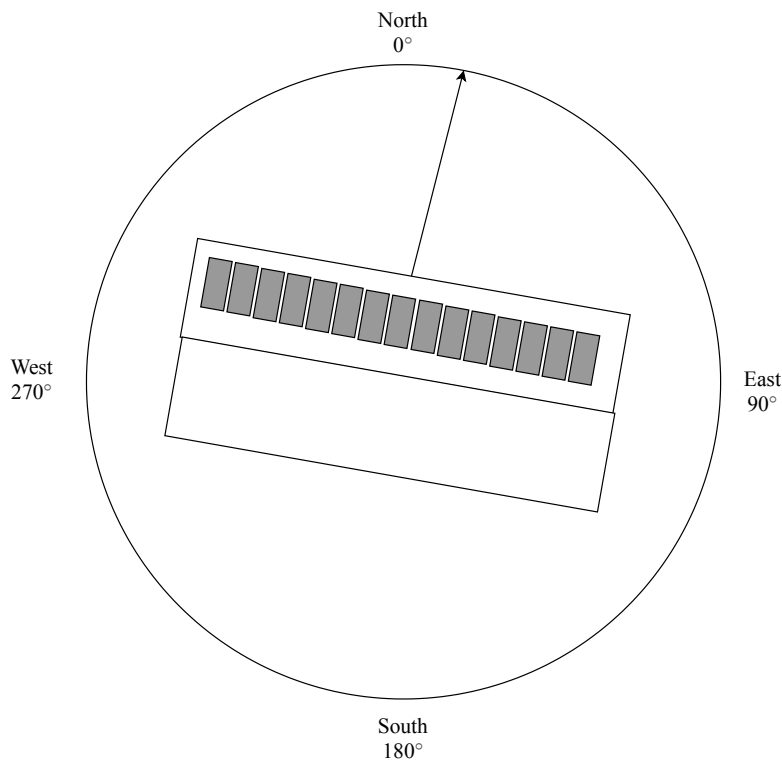


Figure A.2: Azimuth angle of a school building.

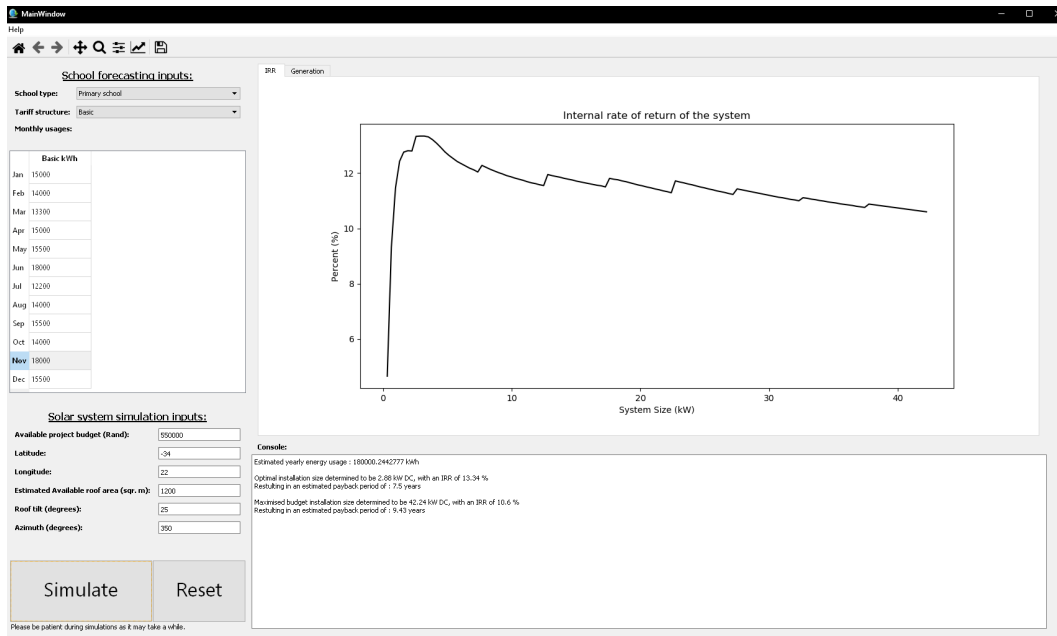


Figure A.3: Interface after a simulation was performed, with results shown in the console window.

List of References

- [1] International Energy Agency, F.: Electricity statistics: Detailed, comprehensive annual data on electricity and heat. <https://www.iea.org/statistics/electricity/>, 2018. Accessed on 2019-06-24.
- [2] Dietz, T., Frank, K.A., Whitley, C.T., Kelly, J. and Kelly, R.: Political influences on greenhouse gas emissions from us states. *Proceedings of the National Academy of Sciences*, vol. 112, no. 27, pp. 8254–8259, 2015. ISSN 0027-8424. <https://www.pnas.org/content/112/27/8254.full.pdf>. Available at: <https://www.pnas.org/content/112/27/8254>
- [3] World Energy Council, U.K.: World energy resources: 2016. <https://www.worldenergy.org/wp-content/uploads/2016/10/World-Energy-Resources-Full-report-2016.10.03.pdf>, 2016. Accessed on 2019-06-25.
- [4] Jackson, R.B., Quéré, C.L., Andrew, R.M., Canadell, J.G., Korsbakken, J.I., Liu, Z., Peters, G.P. and Zheng, B.: Global energy growth is outpacing decarbonization. *Environmental Research Letters*, vol. 13, no. 12, p. 120401, dec 2018. Available at: <https://iopscience.iop.org/article/10.1088/1748-9326/aaf303/pdf>
- [5] Eskom: Eskom yearly integrated report for the year 2017. http://www.eskom.co.za/IR2017/Documents/Eskom_integrated_report_2017.pdf, 2018. Accessed on 2018-07-15.
- [6] Department of Energy, S.A.: Coal resources overview. http://www.energy.gov.za/files/coal_frame.html, 2018. Accessed on 2018-08-22.
- [7] Future of Life Institute, U.S.: Developing countries can't afford climate change. <https://futureoflife.org/2016/08/05/developing-countries-cant-afford-climate-change/?cn-reloaded=1>, 2016. Accessed on 2019-06-23.
- [8] KoÅşak, E., Ulucak, R., Dedeoglu, M. and Ulucak, Z.S.: Is there a trade-off between sustainable society targets in sub-saharan africa? *Sustainable Cities and Society*, vol. 51, p. 101705, 2019. ISSN 2210-6707. Available at: <http://www.sciencedirect.com/science/article/pii/S2210670719311588>

- [9] WSP: Sustainable cities and the role of regional green building programs. <https://gresb.com/sustainable-cities-role-regional-green-building-programs/>, 2018. Accessed on 2019-07-24.
- [10] Jimenez-Estevez, G.A., Palma-Behnke, R., Ortiz-Villalba, D., Nunez Mata, O. and Silva Montes, C.: It takes a village: Social scada and approaches to community engagement in isolated microgrids. *IEEE Power and Energy Magazine*, vol. 12, no. 4, pp. 60–69, July 2014. ISSN 1540-7977.
- [11] Akinyele, D.: Techno-economic design and performance analysis of nanogrid systems for households in energy-poor villages. *Sustainable Cities and Society*, vol. 34, pp. 335 – 357, 2017. ISSN 2210-6707.
Available at: <http://www.sciencedirect.com/science/article/pii/S221067071730358X>
- [12] Wijeratne, W.P.U., Yang, R.J., Too, E. and Wakefield, R.: Design and development of distributed solar pv systems: Do the current tools work? *Sustainable Cities and Society*, vol. 45, pp. 553 – 578, 2019. ISSN 2210-6707.
Available at: <http://www.sciencedirect.com/science/article/pii/S2210670718313404>
- [13] International, Renewable Energy Agency: Renewable Power Generation Costs in 2018. <https://www.irena.org/publications/2019/May/Renewable-power-generation-costs-in-2018>, 2019. Accessed on 2019-06-26.
- [14] StatsSA: Cpi history south africa. <http://www.statssa.gov.za/publications/P0141/CPIHistory.pdf?>, 2019. Accessed on 2019-08-24.
- [15] Laboratory, L.B.N.: An historical summary of the installed price of photovoltaics in the united states from 1998 to 2012. <https://emp.lbl.gov/sites/all/files/lbnl-6350e.pdf>, 2012. Accessed on 2019-08-20.
- [16] Rogers, J.: The price of large-scale solar keeps dropping. <https://blog.ucsusa.org/john-rogers/large-scale-solar-gets-cheaper>, 2018. Accessed on 2019-08-20.
- [17] ESKOM: Tariffs and charges booklet 2019/2020. http://www.eskom.co.za/CustomerCare/TariffsAndCharges/Pages/Tariffs_And_Charges.aspx, 2019. Accessed on 2019-07-02.
- [18] Deloitte: The economic impact of electricity price increases on various sectors of the south african economy. http://www.eskom.co.za/CustomerCare/MYPD3/Documents/Economic_Impact_of_Electrcity_Price_Increases_Document1.pdf, 2018. Accessed on 2019-07-15.

- [19] ESKOM: Eskom 1990 -1999 - african renaissance - powering transformation. <http://www.eskom.co.za/sites/heritage/Pages/1990.aspx>, 1999. Accessed on 2019-07-01.
- [20] ESKOM: Eskom 2000 - 2008 - our recent past -shift performance and grow sustainably. <http://www.eskom.co.za/sites/heritage/Pages/2000.aspx>, 2008. Accessed on 2019-07-01.
- [21] ESKOM: Tariff history. http://www.eskom.co.za/CustomerCare/TariffsAndCharges/Pages/Tariff_History.aspx, 2019. Accessed on 2019-07-04.
- [22] Baker, L. and Sovacool, B.K.: The political economy of technological capabilities and global production networks in south africa's wind and solar photovoltaic (pv) industries. *Political Geography*, vol. 60, pp. 1 – 12, 2017. ISSN 0962-6298.
- [23] Smit, S., Musango, J.K. and Brent, A.C.: Understanding electricity legitimacy dynamics in an urban informal settlement in south africa: A community based system dynamics approach. *Energy for Sustainable Development*, vol. 49, pp. 39 – 52, 2019. ISSN 0973-0826.
- [24] Eskom: Standard & poor's revises eskom's outlook to stable an affirms credit ratings. <http://www.eskom.co.za/news/Pages/2019Mar1B.aspx>, 2019. Accessed on 2019-07-16.
- [25] Inglesi-Lotz, R. and Blignaut, J.N.: Estimating the price elasticity of demand for electricity by sector in South Africa. *South African Journal of Economic and Management Sciences* , vol. 14, pp. 449 – 465, 01 2011. ISSN 2222-3436. Available at: http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S2222-34362011000400007&nrm=iso
- [26] Eberhard, A. and Naude, R.: The South African Renewable Energy Independent Power Producer Procurement Programme: A review and lessons learned. *Journal of Energy in Southern Africa*, vol. 27, pp. 1 – 14, 11 2016. ISSN 1021-447X. Available at: http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S1021-447X2016000400001&nrm=iso
- [27] Donnelly, L.: Eskom burning through diesel again. <https://mg.co.za/article/2018-03-28-eskom-burning-through-diesel-again>, 2018. Accessed on 2019-07-22.
- [28] Hirth, L.: Market value of solar power: Is photovoltaics cost-competitive? *IET Renewable Power Generation*, vol. 9, no. 1, pp. 37–45, 2015. ISSN 1752-1416.
- [29] Sedibe, M.: Inequality of access to resources in previously disadvantaged south african high schools. *Journal of Social Sciences*, vol. 28, no. 2, pp. 129–135, 2011.
- [30] Osuri, S.O., Ngoma, J.L. and Chowdhury, S.P.: Cost-effective renewable energy technologies in universities of south africa: A case study. In: *International Conference on Renewable Power Generation (RPG 2015)*, pp. 1–5. Oct 2015.

- [31] Powerwise: Demand side management. <http://www.powerwise.gov.ae/en/research/programmes-projects/demand-side-management.html>, 2012. Accessed on 2019-07-22.
- [32] Roux, M., Apperley, M. and Booysen, M.: Comfort, peak load and energy: Centralised control of water heaters for demand-driven prioritisation. *Energy for Sustainable Development*, vol. 44, pp. 78 – 86, 2018. ISSN 0973-0826.
- [33] Kepplinger, P., Huber, G., Preißinger, M. and Petrasch, J.: State estimation of resistive domestic hot water heaters in arbitrary operation modes for demand side management. *Thermal Science and Engineering Progress*, vol. 9, pp. 94 – 109, 2019. ISSN 2451-9049.
- [34] SACMEQ: A study of the conditions of schooling and the quality of education. http://www.sacmeq.org/sites/default/files/sacmeq/reports/sacmeq-iii/national-reports/s3_south_africa_final.pdf, 2010. Accessed on 2019-07-16.
- [35] Bank, W.: Overcoming poverty and inequality in south africa. <http://documents.worldbank.org/curated/en/530481521735906534/pdf/124521-REV-0U0-South-Africa-Poverty-and-Inequality-Assessment-Report-2018-FINAL-WEB.pdf>, 2018. Accessed on 2019-07-12.
- [36] Mogues, T. and R. Carter, M.: Social capital and the reproduction of economic inequality in polarized societies. *Journal of Economic Inequality*, vol. 3, pp. 193–219, 02 2005.
- [37] van der Berg, S., Burger, C., Burger, R., de Vos, M., du Rand, G., Gustafsson, M., Moses, E., Shepherd, D., Spaull, N., Taylor, S. and van Broekhuizen and, H.: Low quality education as a poverty trap. Working Papers 25/2011, Stellenbosch University, Department of Economics, 2011. Available at: <https://ideas.repec.org/p/sza/wpaper/wpapers255.html>
- [38] van der Berg, S.: Fiscal incidence of social spending in South Africa, 2006. Working Papers 10/2009, Stellenbosch University, Department of Economics, 2009. Available at: <https://ideas.repec.org/p/sza/wpaper/wpapers82.html>
- [39] McLaren, D.: Funding basic education. <http://section27.org.za/wp-content/uploads/2017/02/Chapter-2.pdf>, 2017. Accessed on 2019-07-04.
- [40] Hohne, P., Kusakana, K. and Numbi, B.: A review of water heating technologies: An application to the South African context. *Energy Reports*, vol. 5, pp. 1 – 19, 2019. ISSN 2352-4847.
- [41] Voss, K., Sartori, I., Napolitano, A., Geier, S., Gonçalves, H., Hall, M., Heiselberg, P., Widén, J., Candanedo, J., Musall, E., Karlsson, B. and Torcellini, P.: Load matching and grid interaction of net zero energy buildings. In: *International Conference on Renewable Power Generation (EuroSun 2010)*, pp. 101–108. 09 2010.

- [42] Raza, M.Q., Nadarajah, M., Quoc Hung, D. and Baharudin, Z.: An intelligent hybrid short term load forecast model for seasonal prediction of smart power grid. *Sustainable Cities and Society*, 12 2016.
- [43] Bhaya, W.: Review of data preprocessing techniques in data mining. *Journal of Engineering and Applied Sciences*, vol. 12, pp. 4102–4107, 09 2017.
- [44] Famili, A., Shen, W.-M., Weber, R. and Simoudis, E.: Data preprocessing and intelligent data analysis. *Intelligent Data Analysis*, vol. 1, no. 1, pp. 3 – 23, 1997. ISSN 1088-467X.
Available at: <http://www.sciencedirect.com/science/article/pii/S1088467X98000079>
- [45] Heylman, C., Kim, Y.G. and Wang, J.: Forecasting energy trends and peak usage at the university of virginia. In: *2015 Systems and Information Engineering Design Symposium*, pp. 362–368. April 2015.
- [46] Bourdeau, M., qiang Zhai, X., Nefzaoui, E., Guo, X. and Chatellier, P.: Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustainable Cities and Society*, vol. 48, p. 101533, 2019. ISSN 2210-6707.
Available at: <http://www.sciencedirect.com/science/article/pii/S2210670718323862>
- [47] Koen, R. and Holloway, J.: Application of multiple regression analysis to forecasting South Africa’s electricity demand. *Journal of Energy in Southern Africa*, vol. 25, pp. 48 – 58, 11 2014. ISSN 1021-447X.
Available at: http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S1021-447X2014000400005&nrm=iso
- [48] Zhu, J., Shen, Y., Song, Z., Zhou, D., Zhang, Z. and Kusiak, A.: Data-driven building load profiling and energy management. *Sustainable Cities and Society*, vol. 49, p. 101587, 2019. ISSN 2210-6707.
Available at: <http://www.sciencedirect.com/science/article/pii/S221067071832362X>
- [49] Kuster, C., Rezgui, Y. and Mourshed, M.: Electrical load forecasting models: A critical systematic review. *Sustainable Cities and Society*, vol. 35, pp. 257 – 270, 2017. ISSN 2210-6707.
Available at: <http://www.sciencedirect.com/science/article/pii/S2210670717305899>
- [50] Zala, H.N. and Abhyankar, A.R.: A novel approach to design time of use tariff using load profiling and decomposition. In: *2014 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, pp. 1–6. Dec 2014.
- [51] Yilmaz, S., Chambers, J. and Patel, M.: Comparison of clustering approaches for domestic electricity load profile characterisation - implications for demand side management. *Energy*, vol. 180, pp. 665 – 677, 2019. ISSN 0360-5442.

- Available at: <http://www.sciencedirect.com/science/article/pii/S0360544219310060>
- [52] Zhang, C., Cao, L. and Romagnoli, A.: On the feature engineering of building energy data mining. *Sustainable Cities and Society*, vol. 39, pp. 508 – 518, 2018. ISSN 2210-6707.
Available at: <http://www.sciencedirect.com/science/article/pii/S2210670717317018>
- [53] Salom, J., Marszal, A.J., WidÅ©n, J., Candanedo, J. and Lindberg, K.B.: Analysis of load match and grid interaction indicators in net zero energy buildings with simulated and monitored data. *Applied Energy*, vol. 136, pp. 119 – 131, 2014. ISSN 0306-2619.
Available at: <http://www.sciencedirect.com/science/article/pii/S0306261914009659>
- [54] Maleki, A., Rosen, M.A. and Pourfayaz, F.: Optimal Operation of a Grid-Connected Hybrid Renewable Energy System for Residential Applications. *Sustainability*, vol. 9, no. 8, pp. 1–20, July 2017.
- [55] Yang, H., Wei, Z. and Chengzhi, L.: Optimal design and techno-economic analysis of a hybrid solar-wind power generation system. *Applied Energy*, vol. 86, no. 2, pp. 163 – 169, 2009. ISSN 0306-2619. IGEC III.
Available at: <http://www.sciencedirect.com/science/article/pii/S0306261908000603>
- [56] Gallego Sanchez Torija, J., Larrumbide, E. and Bedoya, C.: The incorporation of the study into water consumption in energy audits in schools. *Revista de la construccion*, vol. 16, pp. 361–373, 12 2017.
- [57] Mauri, M., Carmeli, M.S., Merlo, M., Brivio, C. and Mbuya, M.: Neural network based load forecasting and fuzzy logic ems for ngarenanyuki school microgrid. In: *2016 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)*, pp. 321–326. 2016.
- [58] Matos, C., Bentes, I., Pereira, S., Faria, D. and Briga-Sa, A.: Energy consumption, co2 emissions and costs related to baths water consumption depending on the temperature and the use of flow reducing valves. *Science of The Total Environment*, vol. 646, pp. 280 – 289, 2019. ISSN 0048-9697.
- [59] schools4sa: Welcome to south african comprehensive schools resource directory. <https://www.schools4sa.co.za>, 2010.
- [60] Dias Pereira, L., Raimondo, D., Corgnati, S. and Gameiro da Silva, M.: Energy consumption in schools - a review paper. *Renewable and Sustainable Energy Reviews*, vol. 40, pp. 911–922, 12 2014.
- [61] Stellenbosch Municipality: Stellenbosch Municipality 2017/2018 tariffs. <https://www.stellenbosch.gov.za/documents/idp-budget/2017-2018-id>

- p-budget/draft-budget-2017-2018/4610-final-tariffs-2017-2018/file, 2018. Accessed on 2018-07-23.
- [62] BroÅłyna, J., Mentel, G., Szetela, B. and Strielkowski, W.: Multi-seasonality in the tbats model using demand for electric energy as a case study. *Economic computation and economic cybernetics studies and research / Academy of Economic Studies*, vol. 52, pp. 229–246, 03 2018.
- [63] Douzal-Chouakria, A. and Amblard, C.: Classification trees for time series. *Pattern Recognition*, vol. 45, no. 3, pp. 1076 – 1091, 2012. ISSN 0031-3203. Available at: <http://www.sciencedirect.com/science/article/pii/S0031320311003578>
- [64] PVGIS, E.C.: Photovoltaic geographical information system. http://re.jrc.ec.europa.eu/pvg_tools/en/tools.html#TMY, 2018. Accessed on 2018-07-12.
- [65] NREL: Solar advisor model technical reference. <https://www.nrel.gov/docs/fy15osti/64102.pdf>, 2018. Accessed on 2018-08-22.
- [66] Brownson, J.R.S.: Collector orientation. <https://www.e-education.psu.edu/eme810/node/576>, 2018.
- [67] Masters, G.M.: *Renewable and efficient electric power systems*. 2nd edn. Wiley-Blackwell, 2013.
- [68] Naeem, M. and Tamizhmani, G.: Cleaning frequency optimization for soiled photovoltaic modules. *2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC)*, pp. 1–5, 2015.
- [69] Dobos, A.P.: Improved coefficient calculator for the california energy commission 6 parameter photovoltaic module model. *Journal of Solar Energy Engineering*, 5 2012.
- [70] De Soto, W., Klein, S. and Beckman, W.: Improvement and validation of a model for photovoltaic array performance. *Solar Energy*, vol. 80, pp. 78–88, 2004.
- [71] King, D., Gonzalez, S., Galbraith, G. and Boyson, W.: Performance model for grid-connected photovoltaic inverters. *Sandia National Laboratories 2007*, vol. 50, p. 44, 2007.
- [72] Rodriguez, C. and Amaratunga, G.: Long-lifetime power inverter for photovoltaic ac modules. *Industrial Electronics, IEEE Transactions on*, vol. 55, pp. 2593 – 2601, 08 2008.
- [73] Dippenaar, M.: The role of tax incentives in encouraging energy efficiency in the largest listed South African businesses. *South African Journal of Economic and Management Sciences* , vol. 21, pp. 1 – 12, 00 2018. ISSN 2222-3436.

- Available at: http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S2222-34362018000100020&nrm=iso
- [74] Cloete, A.H.: *Domestic water heater application for smart grid*. Master's thesis, Stellenbosch University, Stellenbosch University, 3 2017.
- [75] Nel, P.J.C.: *Rethinking electrical water heaters*. Master's thesis, Stellenbosch University, Stellenbosch University, 12 2015.
- [76] Ghiaasiaan, S.M.: *Convective heat and mass transfer*. CRC Press, 2018.
- [77] Cloete, A.H. and Booysen, M.J.: Sustainability through intelligent scheduling of electric water heaters in a smart grid. In: *2016 IEEE 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress*, pp. 848–855. Aug 2016.
- [78] Eskom: Managing your geyser for a more efficient future. <http://www.pcb.org.za/wp-content/uploads/2015/04/EskomGeyserFactSheet.pdf>, 2015.
- [79] CALEFFI: Temperature and pressure relief valve for solar systems. https://www.caleffi.com/sites/default/files/file/01147-07_en.pdf, 2007. Accessed on 2018-10-05.
- [80] Jacobs, H., Botha, B. and Blokker, M.: Household Hot Water Temperature - An Analysis at End-Use Level. In: *International WDSA / CCWI 2018 Joint Conference, Kingston, Ontario, Canada - July 23-25*, pp. 1–14. 2018.
- [81] Mercante, J.W. and Winchell, J.M.: Current and emerging legionella diagnostics for laboratory and outbreak investigations. *Clinical Microbiology Reviews*, vol. 28, no. 1, pp. 95–133, 2015. ISSN 0893-8512. <https://cmr.asm.org/content/28/1/95.full.pdf>. Available at: <https://cmr.asm.org/content/28/1/95>
- [82] Nel, P.J.C., Booysen, M.J. and van der Merwe, B.: A computationally inexpensive energy model for horizontal electric water heaters with scheduling. *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 48–56, Jan 2018.