SMART METERING AND ENERGY ACCESS PROGRAMS: AN APPROACH TO ENERGY POVERTY REDUCTION IN SUB-SAHARAN AFRICA

by

BENNOUR BIN AMADE BACAR

Thesis submitted in partial fulfilment of the requirements for the degree

Master of Engineering: Energy

in the Faculty of Engineering & the Built Environment

at the Cape Peninsula University of Technology

Supervisor: Dr Ali Almaktoof

Bellville
October 2022

DECLARATION

I, Bennour Bacar, declare that the contents of this dissertation/thesis represent my own unaided work, and that the dissertation/thesis has not previously been submitted for academic examination towards any qualification. Furthermore, it represents my own opinions and not necessarily those of the Cape Peninsula University of Technology.

Signed 25 / 01 / 2023 **Date**

ABSTRACT

Evolving technologies can provide continuous and accurate energy data to plan, implement, and maintain energy systems for areas where electricity access is a challenge, particularly in sub-Saharan Africa (SSA) where over 53% of the world's energy-poor population resides.

This research aims to analyse the applicability of smart metering data to the sustainable energy access planning (SEAP) framework for energy access programs (EAPs), toward the reduction of energy poverty in SSA.

Household energy data based on energy access criteria from an SSA country was generated using smart metering technologies, then applied to the analysis and calculation of energy access indicators, demand forecasting through machine learning, and energy systems' optimization and cost analysis.

The approach involved five related components. Country-specific data was collected, analysed, and used to define an energy profile. This profile was then applied as input to a smart metering experiment using a variable household electrical load and a smart meter to measure electricity usage, from which data was collected on General Packet Radio Service (GPRS) communications via Meter Data Management (MDMS) software. The resulting energy data was analysed on its applicability to the SEAP framework and explored over three exercises that included the analysis and calculation of energy access indicators, demand forecasting through machine learning, and energy systems' optimization and cost analysis.

The measured household energy data, analysed and explored using tools and platforms that include Python, Azure ML Studio, and Homer Pro, were directly or indirectly applicable to all assessments in the SEAP framework and exposed the possibility of generating additional data for further use on applications that require a specific range of datasets. These capabilities presented the potential for energy planners and policymakers to use improved data to determine the indicators for the implementation and monitoring of an energy access program; furthermore, it unlocked aspects of data forecasting and optimization of energy systems in terms of sizing and cost.

ACKNOWLEDGEMENTS

I wish to thank:

- God Almighty, for giving me health, strength, focus, wisdom, and patience.
- Dr Almaktoof, my supervisor, for his support and guidance.
- My parents, for their endless prayers.
- My wife, for her patience, encouragement, and priceless support.
- My children, as they are the sunlight to my emotional solar panel.
- Ontec Systems, for their support.
- All those who have directly or indirectly contributed to my studies and research.

DEDICATION

This thesis is dedicated to God, the Almighty, the Omniscient. Nothing is without His will, as I am a channel to His plan.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Terms	Definition
AC	Alternating Current
ADB	Asian Development Bank
ADC	Analogue to Digital Conversion
AFF	Affordability Assessment
AMI	Advanced Metering Infrastructure
AT&C	Aggregated Technical and Commercial
BEN	Benefit Assessment
BR	Billing Register
CST	Cost Assessment
CSV	Comma Separated Values
DC	Direct Current
DCU	Data Concentrator Unit
Discounted payback	Payback period considering the applied discount to the costs
EAP	Energy Access Program
ED	Energy Demand Assessment
EP	Energy Poverty Assessment
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
GTF	Global Tracking Framework
HAN	Home Area Network
HH	Household
IHD	In-House Display
Internal rate of return	Discount rate at which the net present cost is the same for both base case and best-case energy systems
IPYNB	Interactive Python Notebook

IR	Infra-red
Levelized Cost of Electricity	Cost per kWh of the energy system's useful electrical energy produced, on average and over the project's lifetime
LP	Load Profile
NAN	Neighbourhood Area Network
Net Present Cost	Present value of the energy system costs (capital, replacement, O&M, fuel, emissions, grid costs) minus the present value of revenue received (salvage, grid sales).
O&M	Operation and maintenance
PBI	Power Business Intelligence
PLC	Power Line Carrier
RF	Radiofrequency
RSC	Resource assessment
SE4ALL	Sustainable Energy for All
SEAP	Sustainable energy access planning
Simple payback	The period required to recover the investment costs on the best-case energy system, compared to the base case system
SUS	Sustainability Assessment
SW	Software
UN	United Nations
WAN	Wide Area Network
XLS	Excel spreadsheet
XLSX	Excel open XML spreadsheet

CHAPTER ONE
INTRODUCTION

1.1 Introduction

The design of energy systems that are clean, affordable, and accessible for both the poor and non-poor is supported by frameworks such as the sustainable energy access planning (SEAP) (ADB, 2018b). Using these methodologies within an energy access program (EAP), household and energy features are identified to estimate future requirements for energy, based on the collection of data (ADB, 2018a), that is both of primary origin (generated directly from interviews, surveys, etc.) and secondary origin (records kept from health and government institutions, etc.) (BUL, 2021). However, this process has its challenges, the main research problems to be addressed:

a) Input data for planning: missing, estimated, non-contextual, and old.
b) Output data following implementation: no monitoring and tracking of the implemented energy systems.

1.2 Background

The lack of access to electricity is a global challenge affecting nearly 761 million people, of which more than half are in sub-Saharan Africa (SSA); moreover, the rural population is the most affected with only 28% being able to use electricity, notwithstanding overall economic growth (WB, 2021).

Electricity access is key for human and economic growth. For this reason, the UN has defined global access to modern energy services as a universal 2030 goal; furthermore, a common outline was set by the Global Tracking Framework (GTF) of the Sustainable Energy for All (SE4ALL) initiative, defining access to electricity as “availability of an electricity connection at home or the use of electricity as the primary source for lighting” (IEA&WB, 2014; Shrestha & Acharya, 2015).

To provide sustainable access to a minimum amount of energy for the basic needs of both poor and non-poor, while creating an energy supply system that is socially inclusive, the SEAP framework was released by the Asian Development Bank (Shrestha & Acharya, 2015). Included in this methodology are assessments on energy poverty, energy demand, energy resources, cost, affordability, sustainability and benefits.

Key common indicators comprise the SEAP elements specified: electricity consumption, usage patterns, demand profile, with also modern energy services’ components that include availability (can be used), affordability (non-prohibitive cost), reliability (available most of the time), convenience (safe and available when needed), and quality (e.g., level of voltage) (Shrestha & Acharya, 2015). These aspects require energy data, which survey or historical data alone might not provide, or might not be accurate enough if available.

Smart metering technologies monitor, log, collect and provide data for energy planners to accurately measure and get valuable information which, combined with other useful information (such as that acquired from household surveys), will assist in motivating and taking informed decisions toward EAP initiatives. Data flows from energy monitoring translate into insights, forecasting, improved planning, effective implementation, better results, and progress tracking.

1.3 Significance of the research

By coupling the data-acquisition features of the smart metering technology with EAP, this research will aid energy planners in effectively extracting, interpreting, and using the data that derives from all the features the devices and applications can provide. It is a ripple effect that will be valuable to the EAP, implementation, and tracking, but ultimately and most importantly, to those who need access to modern energy in sub-Saharan Africa.

1.4 Objectives of the research

This research aims to analyse and evaluate the applicability of Smart Metering data to the SEAP framework, towards energy access planning.

The objectives are:

a. To analyse and identify the data from the smart meter, required for the SEAP.

b. To establish the applicability of the data to each of the SEAP assessment components.

c. To analyse and identify data from the smart meter for EAP post-implementation monitoring and targeting.

1.5 Research questions

- What data can the smart meter provide as an input to SEAP assessments?
- Is this data inclusive to the full extent of the SEAP framework, i.e., is it applicable to its different assessment components?
- How can monitoring and targeting aid in the sustainability of an EAP implementation?

1.6 Thesis organization

The thesis encompasses 6 chapters, outlined as follows:

Chapter one introduces the SEAP framework for energy access planning and smart metering technologies as contributors to the energy access framework; this applicability defines the thesis scope and subsequent problem definition.

Chapter two outlines each of the elements. The concept of smart metering is explored in its different aspects: the smart meter, communication technologies, meter data collection and management software, security, global deployment figures and the recent COVID-19 impact. On SEAP, the seven assessments of the framework and the energy data relationships are considered. Lastly, the applicability of smart metering technologies to energy access, through the SEAP framework, is discussed.

Chapter three defines the method by which the applicability of smart metering to the SEAP framework will be explored, within the context of an SSA country/region. Processes are presented for data collection and analysis, with an indication of the applicable tools, and ethical considerations. The chapter concludes with an overall description of the methodology.

Chapter four shapes the experiment by collecting the secondary data, and generating primary data through a smart metering experiment, to then explore the applicability of energy data to the SEAP framework. Furthermore, it expands to calculations based on the criteria defined by the applicable assessments of the energy access framework, machine learning prediction to build additional data, and the optimization and cost analysis of energy systems based on the built demand dataset.

Chapter five presents and discusses the results of the experiment, exercises and analyses performed, with relevance to the smart metering technology and resulting household energy data applied to each of the assessments.

Chapter six, lastly, presents a summary of the studies performed in this thesis, outlining the contributions, limitations, and possible areas for future research work.

The following appendices are included:

Appendix A: Types of electronic meters and grid connections

Appendix B: Python commands used for data analysis

CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction

2.2 Smart metering system

2.2.1 The electricity meter

2.2.1.1 Measuring electricity: power, energy, and demand

2.2.1.2 Electricity meter evolution

2.2.1.3 Electricity measurement systems

2.2.1.4 Smart meter

2.2.2 Data exchange: communications

2.2.2.1 Networks and technologies

2.2.2.2 Topologies

2.2.3 Meter data management system

2.2.4 Smart metering and security

2.2.5 Global footprint and impact on COVID-19

2.3 Energy access through the SEAP framework

2.3.1 SEAP framework assessments

2.3.2 Relationship with energy data

2.4 Considerations

2.1 Introduction

The SEAP framework was released by the Asian Development Bank to create an energy supply system that is socially inclusive and provides sustainable access to a minimum amount of energy for the basic needs of both poor and non-poor (Shrestha & Acharya, 2015). SEAPs assessment components of energy poverty, energy demand, energy resources, cost, benefits, sustainability, and affordability, contain indicators that would greatly benefit from smart metering technologies, which would provide useful energy data for energy planners to accurately measure and get valuable information, to assist in informed decisions towards EAP initiatives, as well as the sustainability of the implementation; and as stated by Casals et al. (2020), adding benefit as a tool of not only energy but social analysis.

The sections that follow review the two components of this study in the literature, that is smart metering and energy access program; and discuss the relationship between them.

2.2 Smart metering system

Reinforced by technology growth, smart metering has evolved (Figure 2.1) from just basic manual meter readings to incorporating advanced smart grids, including elements of prosuming, distributed generation, and charging stations for electric vehicles (EVs) while playing a more active role in reducing carbon footprint (Živi et al., 2015).

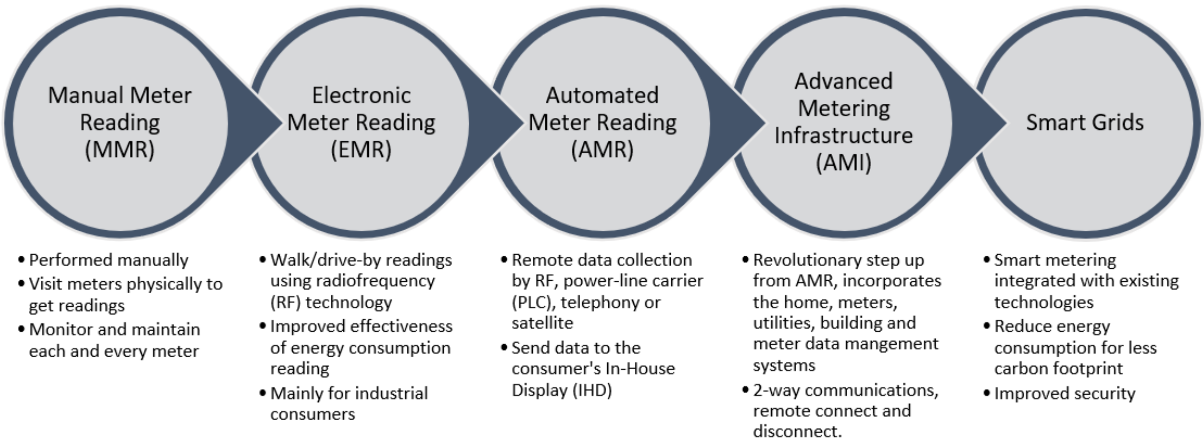


Figure 2.1: Smart metering evolution (Živi et al., 2015)

A typical smart metering system is described by Weranga et al. (2014) as one consisting of smart meters (metering gateway), control devices, appliances, a communication link, and a control centre; furthermore, it may include energy measurement taken at a point of distribution for an area (Sun et al., 2016). These elements are shown in Figure 2.2.

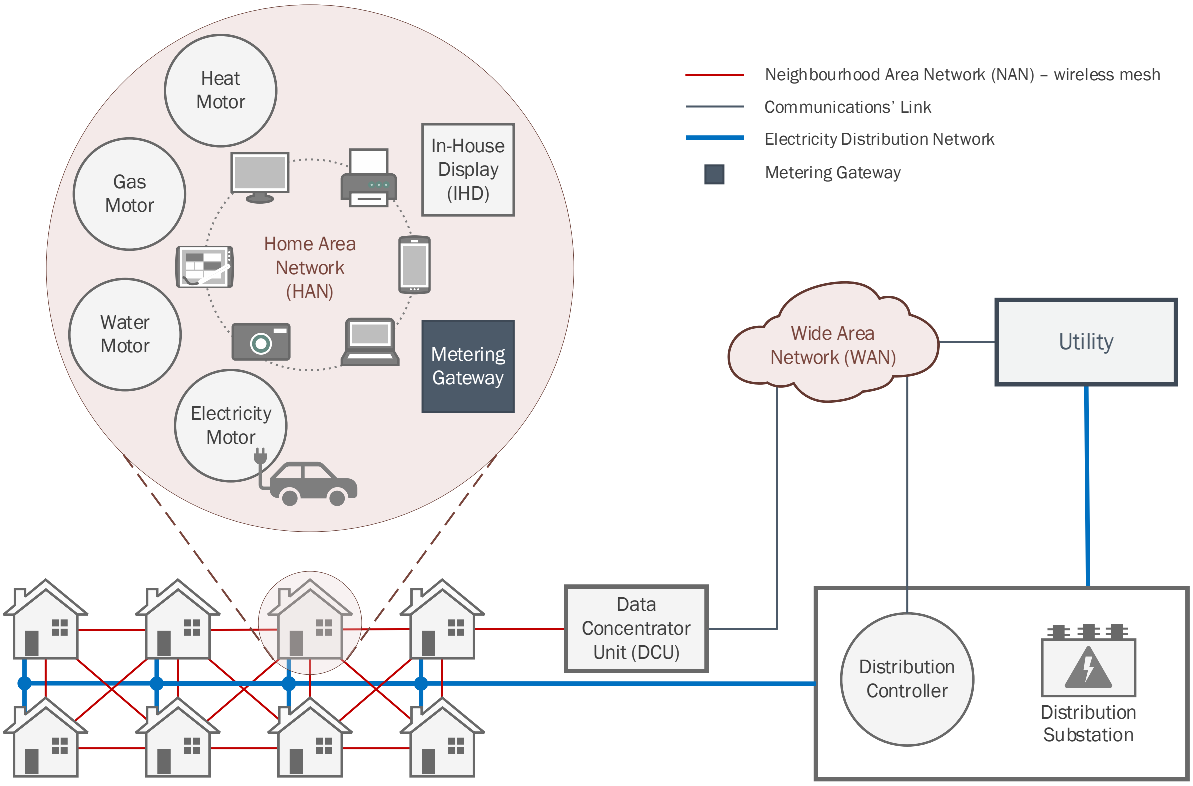


Figure 2.2: Smart metering system components (based on Živi et al., 2015)

According to ADB (2018a), besides improving utilities' operation efficiency and quality of supply (QoS), smart metering systems lead to increased consumer satisfaction and participation; moreover, key benefits include:

- Utilities
    - Reduction in aggregated technical and commercial (AT&C) losses.
    - Lower costs of labour, purchased power and connection/disconnection.
    - Improved power quality, asset management and generation cash flow.
    - Enhanced anomaly detection and overall grid visibility.
    - Integration with renewable energies.
    - Demand management.
- Consumers
    - Accurate bills.
    - Ability to not only monitor electricity consumption but also control and manage home/office appliances.
    - Electricity bill savings through the implementation of time-of-use (ToU) tariffs.

- General
  - Ability to introduce value-added services.
  - Reduced carbon footprint through reduced physical presence for monitoring and meter reading.

A key component of a smart metering system is the device with energy metering and intelligence combined (Weranga et al., 2014), to interact with the energy endpoint, collecting and providing data: the smart meter.

2.2.1 The electricity meter

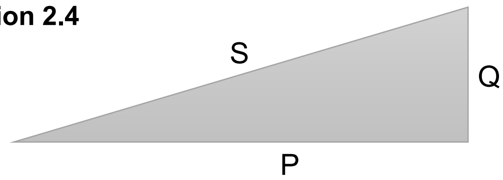
A smart meter is an energy device that has three main functions: (i) to measure electricity usage (or electricity generated); (ii) to switch the consumer on/off; (iii) remotely control the electricity consumption (Halder, 2014). Barai et al. (2016) further describe the smart meter as a device beyond just a traditional meter, capable of:

- Monitoring the distribution grid, to detect issues and send notifications on outages or restoration.
- Measuring and recording daily, interval-based power usage, and wirelessly sending collected high-resolution data over a network to a data management system, when part of AMI.
- Allowing utilities to better communicate with the consumers, with pertinent information such as outages and their causes, expected restoration, and others.
- Improving grid planning and management by improved response against the occurrence of outages, their frequency and duration.

2.2.1.1 Measuring electricity: power, energy, and demand

The measurement of electricity by meters relies on basic electricity principles, explained in Table 2.1.

Table 2.1: Electricity measurements overview (Toledo, 2013)

Electric Measurement	How it is calculated
Power The rate of electricity transferred in a circuit per unit of time; it is the basis for measuring electricity. For example, a water heater would consume more electricity than an electric shaver over the same period because the former uses more watts than the latter. **Demand** Power, but measured within a predefined time interval (typically of 15 or 30 minutes) and expressed in kilo-units; utilities use this measurement to assess the consumer's highest demand from the network (maximum demand), to invest, build and maintain it. Example: two consumers can use the same energy amount for a certain month, but with one reaching a maximum demand of 100 kW, while the other just 40 kW; the first would have used high-power equipment, leading to higher peaks of consumption. **Power factor** The ratio between real power used and apparent power supplied to a load, indicating how efficient the power usage is.	In an alternating current (AC) circuit, the electric power has 3 components, calculated from the measured voltage *V* and current *I*, expressed in *volt (V)* and *ampere (A)* respectively, as well as the phase angle between them, *φ*: ▪ Active power *P*, as the real power consumed via caloric dissipation in a circuit, expressed in *watt (W)*: $P = V \times I \times \cos\varphi$ **Equation 2.1** ▪ Reactive power *Q (Var)*, as the imaginary, unusable power used only for magnetic (e.g., motors) or electric (e.g., capacitors) fields, expressed in *volt-ampere reactive (VAr)*: $Q = V \times I \times \sin\varphi$ **Equation 2.2** ▪ Apparent power *S*, as the vector sum of active and reactive powers, expressed in *volt-ampere (VA)*: $S = V \times I$ **Equation 2.3** Based on formulas 2.1, 2.2 and 2.3, the relationship between the 3 power components is expressed by equation 2.4, and the corresponding power triangle in Figure 2.3: $S^2 = P^2 + Q^2$ **Equation 2.4**  Figure 2.3: The power triangle The component $\cos\varphi$ from formula 2.1 defines the power factor PF, which combined with formula 2.3 can be expressed as: $\boldsymbol{PF = \frac{P}{S}}$**Equation 2.5**
Energy Power used over time results in energy consumed; this energy is the measurement generally used by utilities around the world for residential electricity billing.	Calculated from electric power, electric energy is also classified as active, reactive, or apparent, typically expressed in *kilowatt-hour (kWh)*, *kilovolt-ampere reactive hour (kVArh)* or *kilovolt-ampere hour (kVAh)* respectively. If *X* is expressed as either P, Q or S on a given period *t* (hours), then electric energy consumed *E* can be calculated as: $E = X \times t$ **Equation 2.6**

2.2.1.2 Electricity meter evolution

The levels of intelligence and energy measurement complexity that are currently present on the smart meters passed through several steps of evolution and technology improvement. Figure 2.4 illustrates the early history of the electricity meters, back to the 19th century when gas was used for urban street lighting.

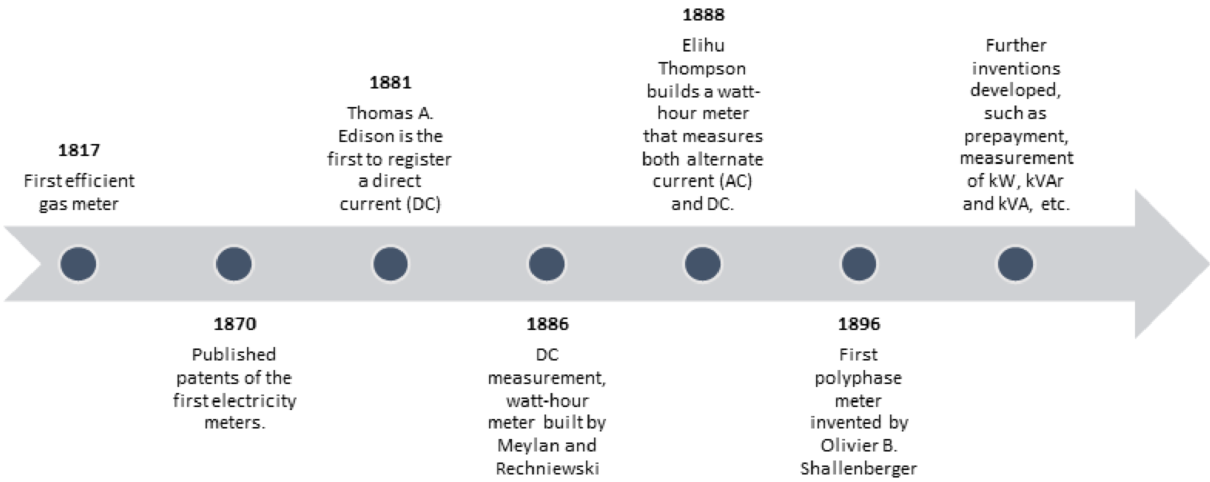


Figure 2.4: History of early electricity meters (based on Toledo, 2013)

Many meter technologies have been on the market, provided by different manufacturers, but overall the induction electromechanical and static electronic are the two types available, which perform power and energy measurements used over a time range; furthermore, whilst billing and monitoring were the main use for meters, currently the applications are more diverse, ranging from real-time energy management, load shifting, demand forecasting, cogeneration monitoring, in-home smart plugs, among others (Toledo, 2013).

2.2.1.3 Electricity measurement systems

Of the different subsystems that comprise the metering system, the measurement system is the main subsystem, as illustrated in Figure 2.5.

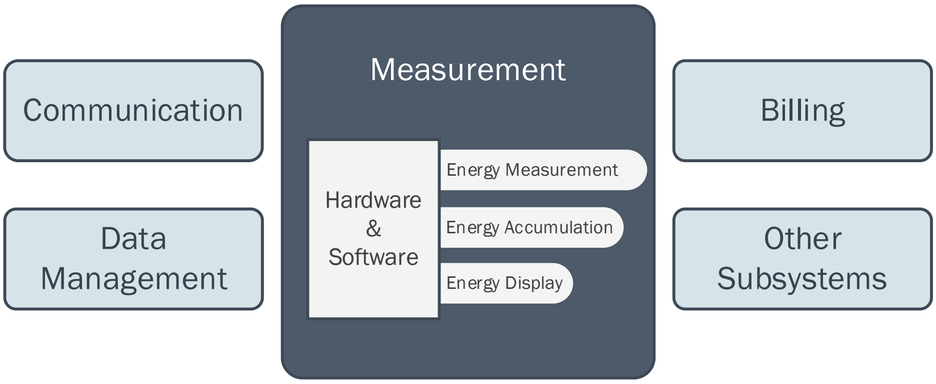


Figure 2.5: A metering system and its subsystems (based on Toledo, 2013)

Electromechanical induction meters

This meter is based on 1888's Thompson design, with energy consumption recorded using electrical and mechanical components, with the following characteristics (Toledo, 2013):

- Can be single- or poly-phase
- Operation principle (Figure 2.6):
  - The voltage circuit **(1)** and current circuit **(2)** have coils, in parallel with the main circuit and in series with the load, respectively
  - The voltage supply and the current flow generate electromagnetic fields via coils, which rotate the disk **(4)**
  - When the disk rotates, so do the pivot and the spindle **(6)**, causing the register display to rotate proportionally to the energy usage
  - The magnetic brake rotor **(5)** controls rotor speed relative to the energy used
  - Other elements include the display dials (7), the stator (3), meter cover and base, the cover and terminal connections, the ID plate and the stator **(4)**

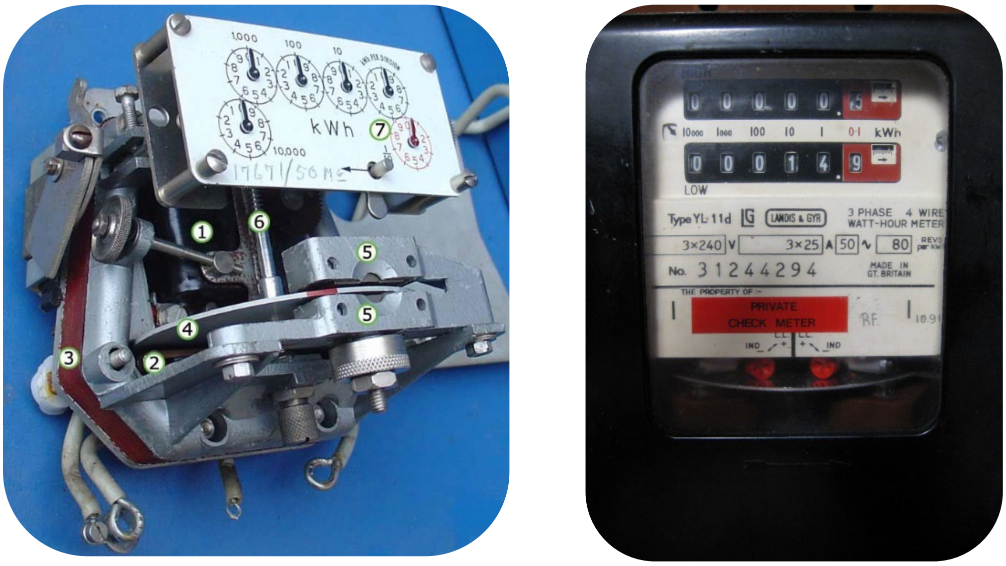


Figure 2.6: Electromechanical meter interior and exterior (Ali@gwc.org.uk, 2005; RODALCO, 2012b).

Drawbacks of the electromechanical meter include kWh or kVArh measured separately (Toledo, 2013), as well as moving parts that wear over time and the limitation for manual readings only (Weranga et al., 2014).

Programmable electronic registers

A transition that adds an electronic element to the mechanical, where the electromechanical meter counts its disk turns via an infra-red (IR) device, then transmits the resulting pulses to an embedded (Figure 2.7) or external element (programmable electronic register); moreover, these devices are still commonly used by utilities globally, even with electronic meters' availability (Toledo, 2013).

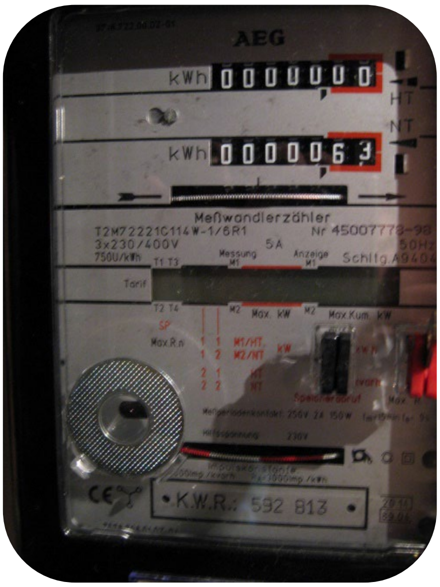


Figure 2.7: Programmable electronic meter (RODALCO, 2012a).

Static electronic meters

Used for most smart metering installations globally (in contrast to only high-consumption users in the past), static electronic meters merge functions of both the electromechanical meter and programmable electronic registers, allowing them to (Toledo, 2013):

- Measure various parameters: power, energy, voltage, current, etc.
- Monitor events and send alarms.
- Calculate tariff rates.
- Operate contacts and switches from programmed events and functions.
- Store profiles for energy and other parameters.
- Perform forecasts on consumption, telemetry functions, analysis of energy measurements, load monitoring and monetary conversions.
- Manage single or multiple tariff rates.
- Provide switching events by external sources or internal clock, for ripple control, clock synchronization, RF broadcast, local DCUs, remote servers, load shifting, etc.

Similarly to Thompson's initial design of the electromechanical meter, but with less mechanical and more electronic components, a static electronic meter has the following features, as indicated by Toledo (2013):

- Single-phase or polyphase
- Comprises electronic and electrical components functioning together to register energy consumption

- The measurement of voltages and currents for energy consumption is achieved via internal sensors of high precision, and measurement algorithms

Electronic meters have different classifications, based on different aspects such as form factor, billing, load, etc. Some of these features are summarized in Table 2.2 and illustrated in Appendix A.

Table 2.2: Different types of electronic meters (Toledo, 2013)

Classification	Category	Description
Design	Single module	▪ Components in a single, usually seal-protected unit. ▪ Replacements are only possible in a laboratory.
	Modular	▪ Components in separate, detachable modules. ▪ Replacements and upgrades are possible in the field. ▪ Plug-in meters are mainly implemented in South Africa for prepaid electricity, and socket meters in North America.
Connection to the grid (load requirement)	Direct measurement	▪ Typically applied to residential customers. ▪ Circuit load flows through the meter's current circuit. ▪ Typical load requirement in some countries in Europe is up to 230/400 V and 100 A.
	Indirect measurement	▪ Typically applied to commercial and industrial customers (bigger loads). ▪ Circuit load flows through instrument transformers (current and/or voltage), to decrease the values to a level that the meters can support. ▪ The typical load requirement in some countries in Europe is higher than 230/400 V and 100 A.
Payment mode	Credit or post-paid	▪ Over a defined cycle, meter readings are collected and used for energy bills that are sent to and paid by the customer. ▪ The meter register count is progressive (cumulative). ▪ Constraints associated with no customer visibility on energy consumption, late or non-payments, disconnections/reconnections, etc.
	Debit or prepaid	▪ The customer purchases energy units and loads them onto the meter, which works on a regressive register count. ▪ The customer has more control over energy consumption and can purchase based on budget and need, without the constraints of reconnection fees. ▪ Recent developments now allow prepayment to be managed from the prepaid vending system, using smart meters: the system reads the consumption and subtracts it from the energy balance; once the balance reaches zero, it sends a command to the meter to disconnect the load.

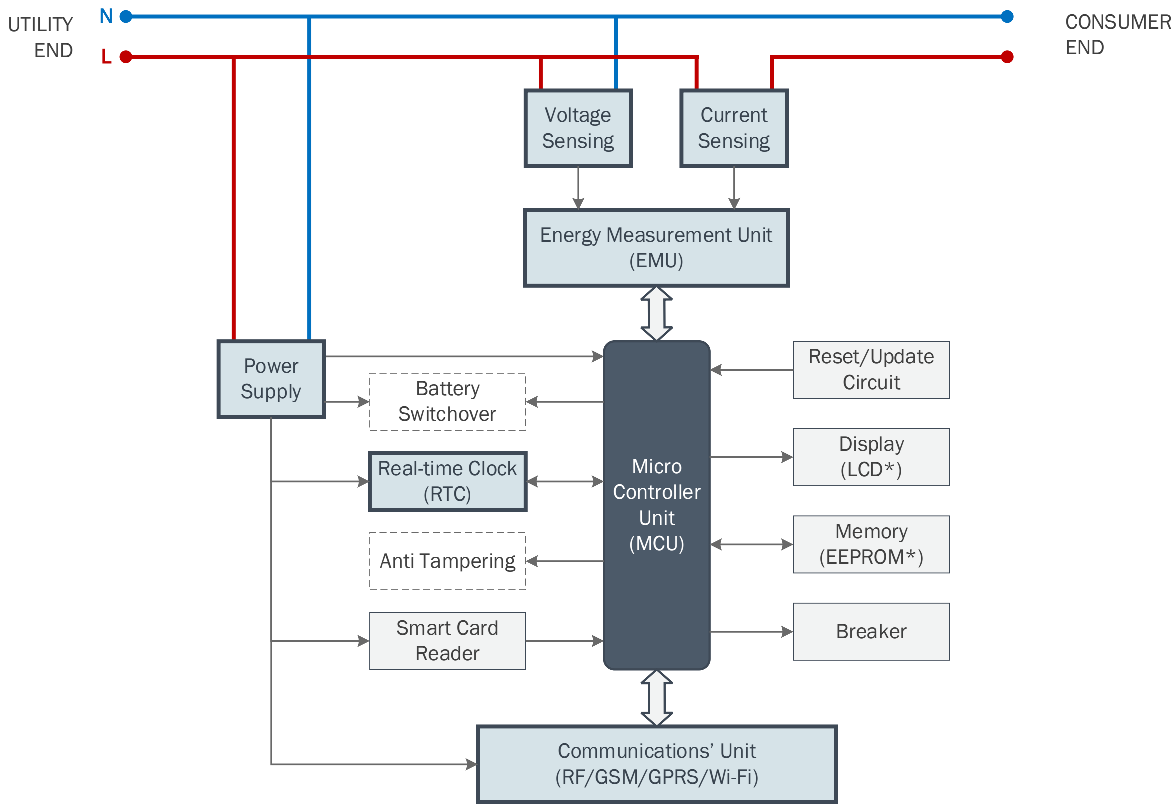
2.2.1.4 Smart meter

A step-up from electronic meters, smart meters (Figure 2.8) bring in features such as two-way communication and remote disconnection and reconnection to the existing electricity measurements and AMR, playing a critical role in today's AMI (Figure 2.1); moreover, functions like pre-payment, load profiling, tamper detection, power outage notification and multi-tariffs contribute to the enhancement of QoS and customer service (Weranga et al., 2014).

Figure 2.8: A single-phase smart meter (Mapondera, 2015). CC BY-SA 3.0

Hardware structure

As illustrated in Figure 2.9, the smart meter acquires input signals using the voltage and current sensors, to then condition, convert (analogue to digital) and compute them within the microcontroller; further operations are performed by other hardware (HW) elements such as the communications' system, real-time clock, and power supply.



*LCD = Liquid Crystal Display, EEPROM = Electrically Erasable Programmable Read-Only Memory

Figure 2.9: A smart meter's hardware structure (based on Weranga et al., 2014)

Typical components of a modern smart meter (highlighted in Figure 2.9) as described by Weranga et al. (2014), include the:

- **Voltage sensing unit**, acquiring the input voltage signal generally through resistor dividers (Figure 2.10), given their low cost; with the output voltage V_{out} (to the analogue-to-digital conversion process) determined by the input voltage V_{in} and the resistors R_1 and R_2 ($R_1 >> R_2$), as per Equation 2.7.

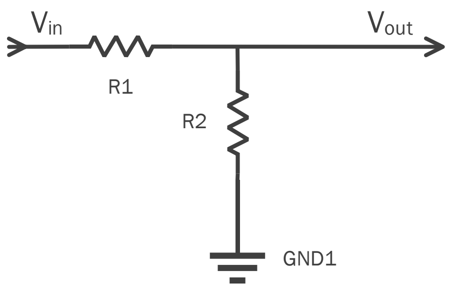


Figure 2.10: Resistor divider

$$V_0 = \frac{R_2}{R_1 + R_2} V_{in}$$

Equation 2.7

- **Current sensing unit**, acquiring the input current signal via current sensors and anti-aliasing filters. Different types include Current transformers, Rogowski coils, shunt resistors, and linear current sensors (Hall effect).
- **Power supply**, to the HW elements in the smart meter, such as the MCU, battery, energy chip, LCD, RTC, and communications unit. Schematics vary by meter designer but generally include step-down transformers, rectifiers (diode bridge), converters from alternating current (AC) to direct current (DC) and vice versa, regulators, and filters.
- **Energy measurement unit (EMU)**, performing signal conditioning, analogue-to-digital conversion (ADC), and computation. The unit's chips provide data such as kWh, kVArh and kVAh energy, or pulse (frequency) outputs; some smart meters additionally include measurements of Root Mean Square (RMS) voltage and current, temperature and frequency, and events/alarms on QoS, tampering detection, Total Harmonic Distortion (THD), and communications. The EMU can be a separate chip or the MCU itself, single- or multi-phase, and can operate in 2 or 4 quadrants.
- **Microcontroller (MCU)**, the smart meter's core unit that runs all the functions, which may include all the EMU functions added to data calculations, display (e.g., via LCD) and smartcard reading. The MCU can act as an internal, single multi-tasking unit, performing all tasks, such as measurement of energy and routine calculations, or as an external unit to handle meters with stepper motor counters instead of an LCD.
- **Real-time clock (RTC)**, keeping the current date and time on the smart meter; time drift (about 60 min/year) is addressed by either periodic synchronization via smart network, manual correction at regular intervals, or by having a highly accurate RTC. The RTC can be built-in to the EMU, or work as a separate RTC, accessed by the MCU, the latter the most common.
- **Communications systems**, allowing the smart meter access to a communications network to send data and events/alarms to a server, receive commands (e.g., remote disconnection or reconnection of supply), and communicate to other devices such as other meters, appliances, and IHDs. Communication networks and protocols include:
    - HAN: Zig-bee, Power Line Carrier (PLC), Wi-Fi, etc.
    - NAN: Zig-bee, etc.
    - WAN: Global Systems for Mobile Communication (GSM), General Packet Radio Service (GPRS), 3G, etc.

2.2.2 Data exchange: communications

Communication networks, interfaces, and technologies are important components of a smart metering system (Figure 2.2). While smart meters are the main elements in an AMI, they also need to communicate to other smart meters, appliances, other types of meters (usually gas and water), and the energy supplier (Weranga et al., 2014).

2.2.2.1 Networks and technologies

The architecture in Figure 2.11 shows the different levels on which the networks exist. These networks comprise technologies and devices (Table 2.3) that enable the data exchange link between the smart meter, local devices and the middleware (Toledo, 2013). The middleware comprises the software (SW) system, such as the Meter Data Management System (MDMS), databases and other applications, managing the smart metering devices, data, functionalities and upstream integration with platforms within, typically, the utility environment.

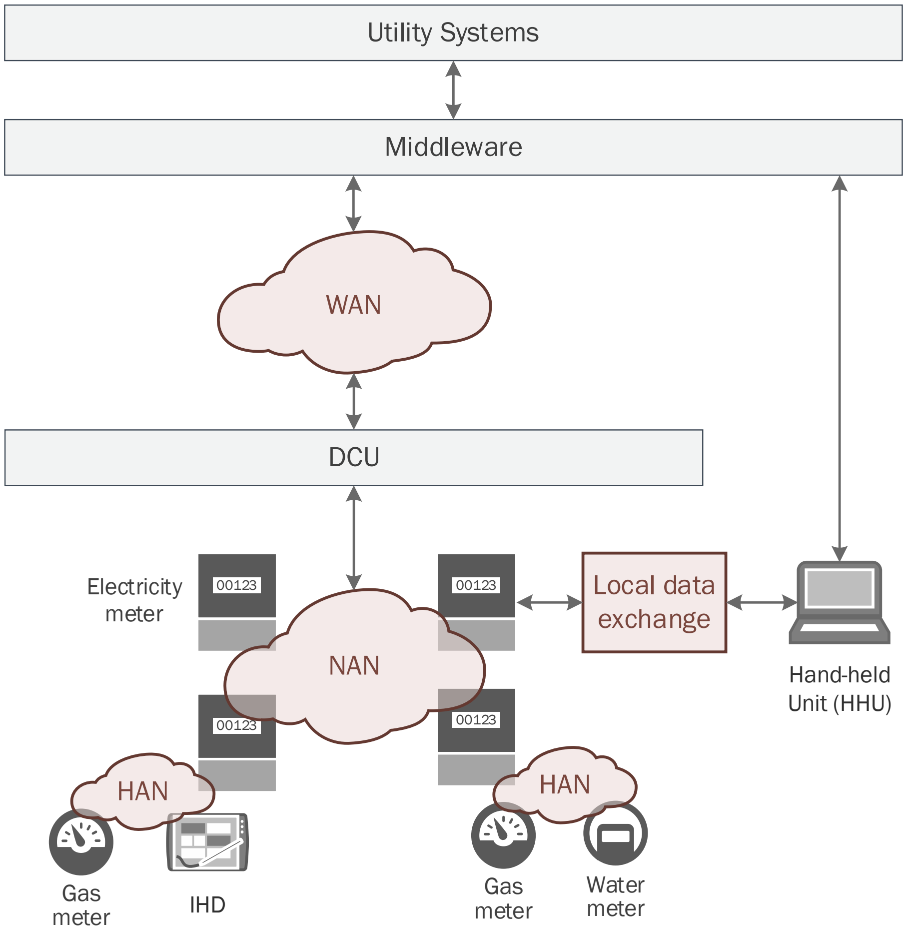


Figure 2.11: Communications on a smart metering system (based on Bajer, 2019; IEC, 2002; Toledo, 2013)

Table 2.3: Communication networks (IEC, 2002; Toledo, 2013; Weranga et al., 2014; Bajer, 2019)

Network/interface	Purpose	Technology
Home Area Network (HAN)	For the communication between the different devices at the customer's home, such as IHDs, sensors, appliances, and meters.	▪ PLC – transmits data over the electricity grid ▪ RF – operates on different frequency ranges of radio operation
Neighbourhood Area Network (NAN)	Transfers data between nearby smart meters and DCUs (WAN access), allowing for data transfer, firmware upgrades, diagnostic and real-time messages.	▪ PLC ▪ RF
Wide Area Network (WAN)	Allows for the interaction between the middleware and the metering system. Includes devices such as DCUs, modems, external gateways, external hotspots, etc.	▪ GSM – transmits via cellular network, using cell-number based, circuit-switched data (CSD) ▪ GPRS - transmits via cellular network, using IP-based, packet-switched data (PSD) ▪ PLC ▪ Public Switched Telephone Network (PSTN) – transmission over a fixed telephone line
Local data exchange	Access the smart meter locally, collecting data to an HHU, to then upload it to the middleware. This access allows for both the reading and programming of the smart meter.	▪ Optical probe – data is collected from the meters via an IR cable attached to the meter's optical interface and an HHU such as a laptop computer.

2.2.2.2 Topologies

Impacting elements such as availability, communications, and robustness, the network topology is a vital component of a smart metering system, allowing its interaction with utility services; available topologies include (Toledo, 2013):

- **Concentrated point-to-point (P2P):** nodes linking to a central device (master, typically a DCU) that collects and routes the data to the WAN, then to middleware. Currently implemented around the world, it is a short-range network (typically from ninety metres to a few kilometres) that uses bus, tree, or star topologies (Figure 2.12), on technologies such as Wi-Fi and PLC. Advantages include easy use, distributed processing, and technology availability; with complexity and single node dependency being some of the disadvantages.

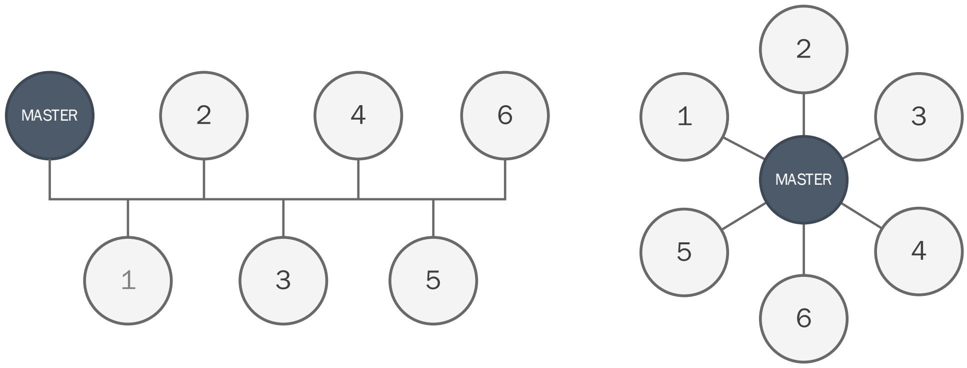


Figure 2.12: A bus (left) and a star (right) network

- **Virtual P2P:** usually applied to metering systems' cellular-based GPRS, GSM and SMS technologies, where infrastructure between devices and middleware is not visible; these are typically owned and managed by 3rd parties (e.g., mobile telecoms providers). This topology (Figure 2.13) allows for simpler middleware access to field devices, however, the dependency on a single comms medium can be a constraint.

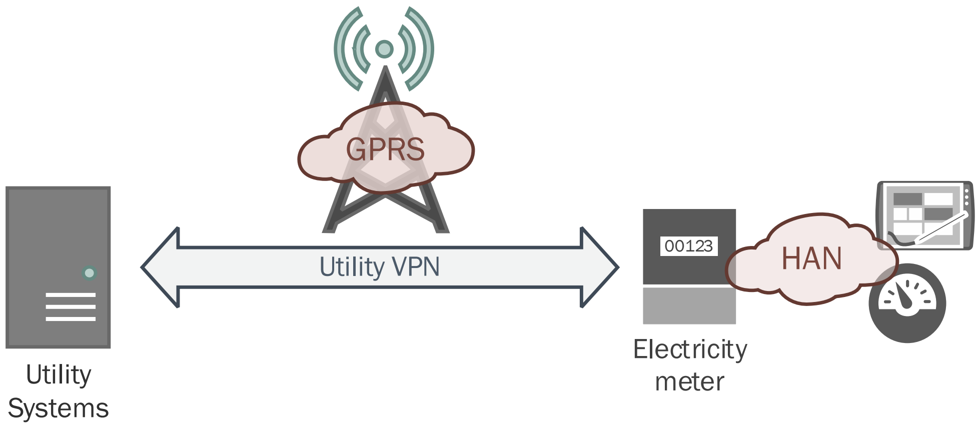


Figure 2.13: Virtual P2P network

- **Mesh:** compared with P2P, data is transferred to more than one central device directly or via other nodes (Figure 2.14), based on the best available route, allowing for a mesh network to provide higher reliability, availability, and feasibility; a disadvantage would be the communications' range though. This topology is a utilities' option for smart metering systems, using technologies that include Z-Wave and ZigBee.

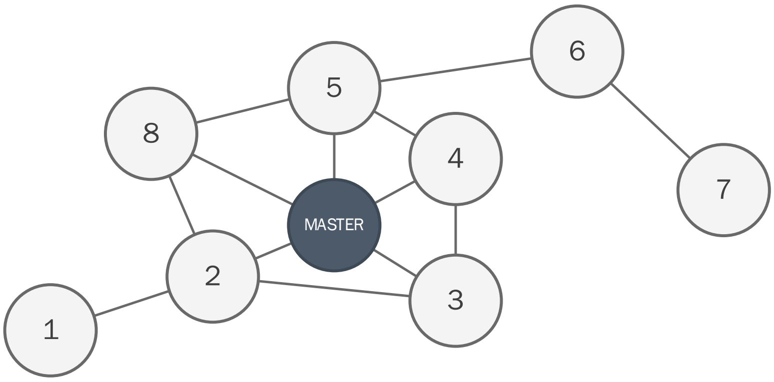


Figure 2.14: A mesh network

- **Broadcast:** while similar to concentrated P2P networks, it allows for a greater range of communication, up to tens of kilometres (Figure 2.15). Known implementations include RF and ripple-controlled systems in Europe, Africa and Australia. This topology has reduced installation and maintenance costs, but low signal bandwidth.

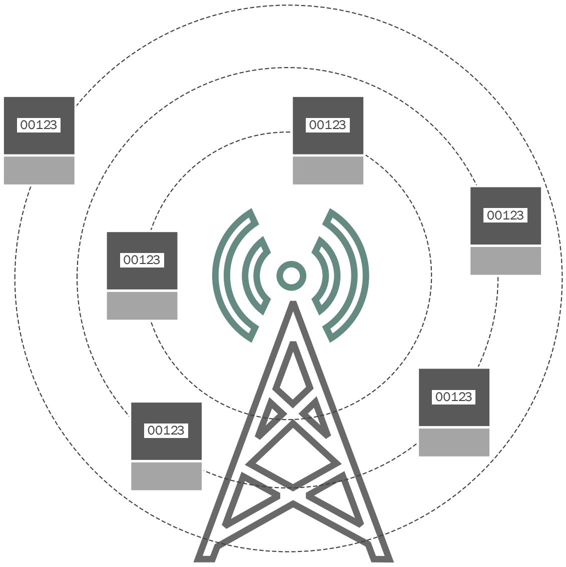


Figure 2.15: Broadcast network

2.2.3 Meter data management system

As an important element of an AMI, the Meter Data Management System (MDMS) collects data from smart metering devices (such as DCUs and meters) and stores it in a central database; this central repository, through the MDMS's analytical components, enables operations and management systems such as billing and demand response (DM), as illustrated in Figure 2.16 (Barai et al., 2016).

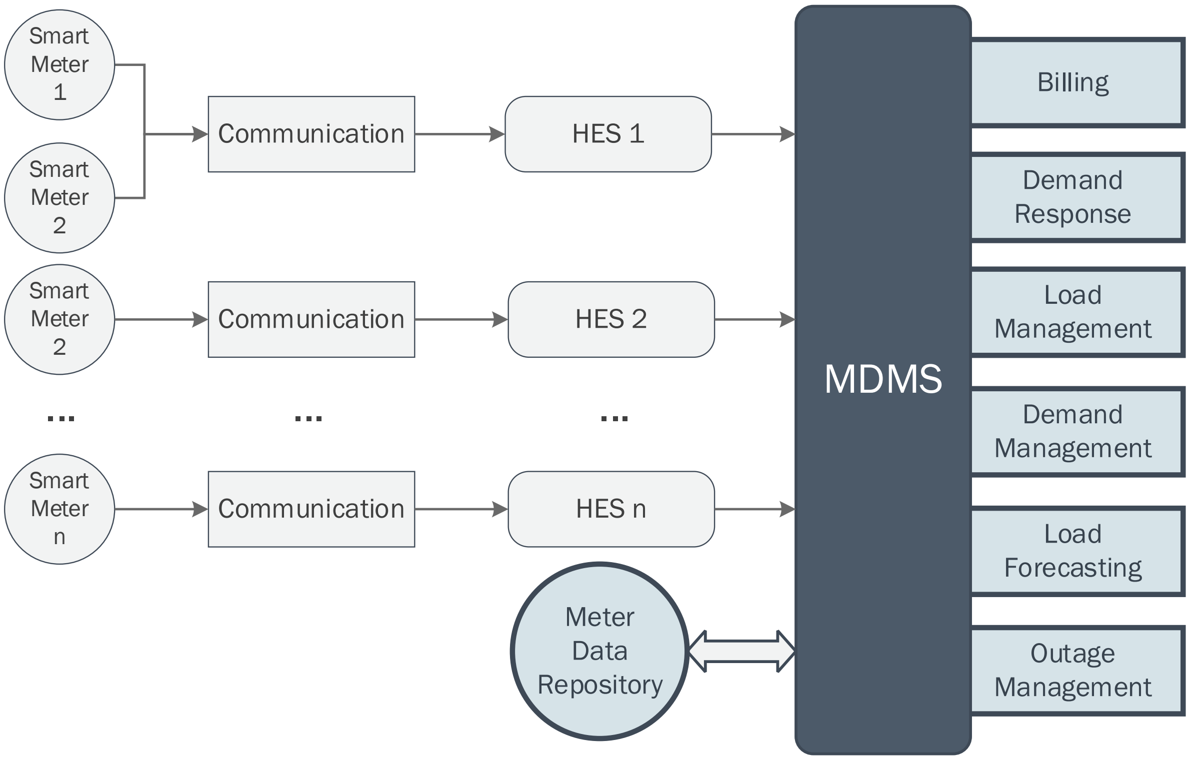


Figure 2.16: MDMS (based on Barai et al., 2016; ADB, 2018a)

According to ADB (2018a), related functions of the MDMS may include:

- Collect input data from different HESs, HHUs or manual readings.
- Data validation, estimation and editing (VEE).
- Billing calculation and determinants.
- Perform trend analysis, log exceptions, and generate reports.
- Integration with interfaces and systems.
- Consumer access to interval data, current and historical.

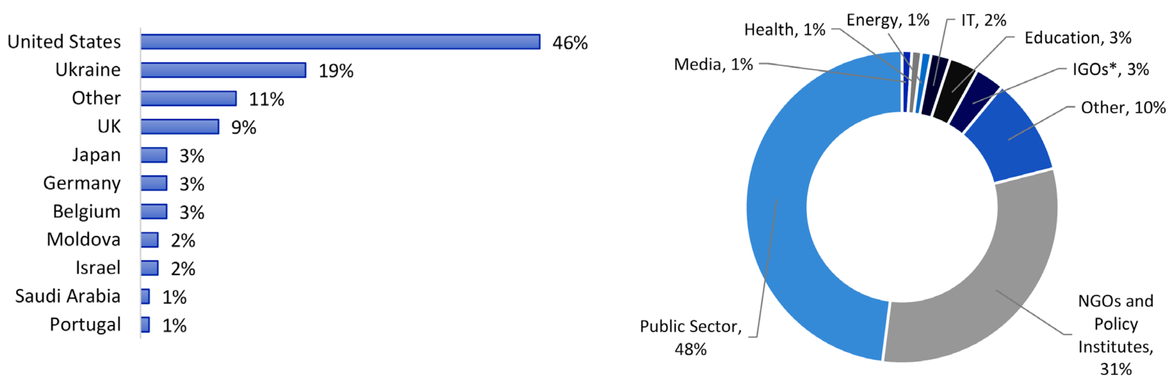
2.2.4 Smart metering and security

A communications network is an important element of a smart metering (SM) system; among other benefits, it mainly allows the smart meter to send energy data to the MDMS and utility systems, or to receive commands from them. However, these networks also bring disadvantages to security and privacy, such as Internet attacks and exposure to personal data (Sun et al., 2016). Regardless of where the data is applied, utility or energy access in this case, threats such as the ones described in Table 2.4 must always be considered.

Table 2.4: SM security threats (Živi et al., 2015; Hasan & Ibrahim, 2020)

Threat	Description	Impact
Man-in-the-middle attack	The attacker connects to both ends of the communication session, captures data from one end and sends it to the other, giving a false impression that both ends are communicating directly.	Being able to either eavesdrop or modify the data, false measurements could be provided.
Denial-of-service (DoS) attack	The attacker sends excessive commands to the gateways or utility servers, saturating the system to a point where it becomes unresponsive.	Disable the grid totally or partially, with severe impact on essential services.
Remote disconnect/connect attack	Turning off or on critical components of the grid, such as a meter or faulty shutdown equipment, respectively.	Disable the grid totally or partially, with severe impact on essential services.
Packet injection	The attacker injects false commands or packets into the network, to either compromise SM system components or the billing processes	Disable the grid totally or partially, with severe impact on essential services, and financial losses due to false bills.
Malware injection	Malware is injected by the attacker into the network, affecting device communications.	Compromises reporting and billing processes and may disrupt grid load.
Eavesdropping	The attacker "listens" to the data flowing between the system and the smart meter or SM gateway.	Compromises personal/customer data privacy
Firmware manipulation	A smart meter or SM gateway's firmware is manipulated, physically or via WAN (if supported); can be executed on a single or large scale.	Disturbs billing (e.g., prepayment manipulation) and meter measurement (e.g., false consumption).
Energy theft	Tampering with the meter's terminals, wiring, the mechanism (analogue meters), firmware (digital meters), bypassing the meter, etc.	Financial losses due to meter damage, false consumption data and resulting bills.

As illustrated in Figure 2.17, the country most targeted by cyber-threats is the United States, with the energy sector being one of the least targeted, and the public sector the most targeted. This indicates that while sub-Saharan Africa and SM systems might not be a favourite target for cyber-threats, attention should be given to the security of SM data stored by public utilities.



*Intergovernmental Organizations

Figure 2.17: Security threats globally (Lambert, 2021)

2.2.5 Global footprint and impact on COVID-19

Smart meter deployments have been growing across the world: countries like China are leading the trend (Figure 2.18) with further plans for over 300 million more of these devices, while projections for emerging market countries indicate 178 million smart meters projected deployment; within sub-Saharan Africa, some of those emerging countries such as Kenya, Ghana, Nigeria and South Africa plan on taking the same path to address AT&C losses, however with financing as the main challenge to overcome (Chakerian, 2021).

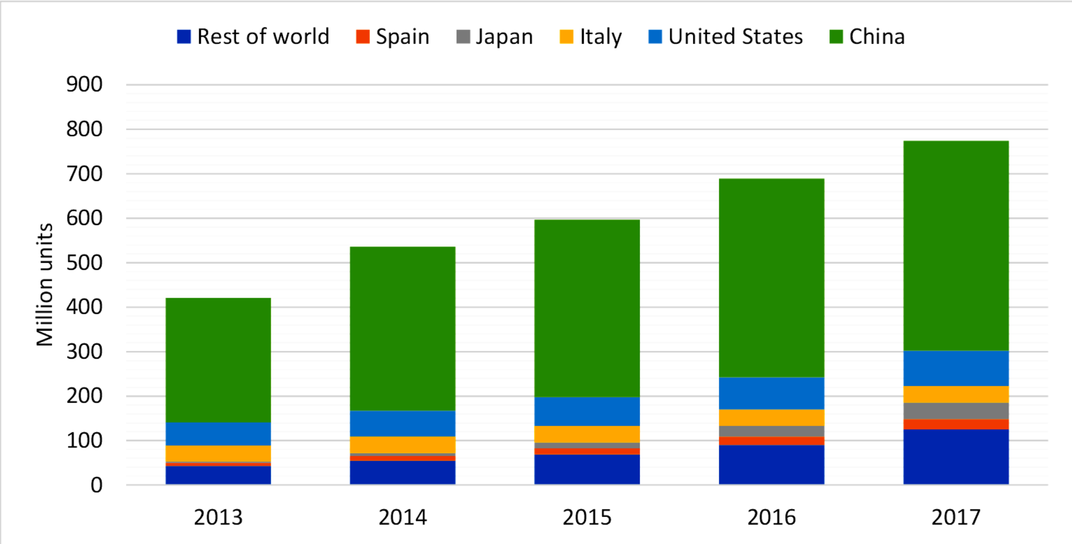


Figure 2.18: Smart meter deployment (IEA, 2019)

The emergence of COVID-19 highlighted the importance of SM systems, with two main benefits:

- Reduced physical presence to collect meter readings since these could be read remotely; therefore, besides minimizing personnel exposure to COVID, during lockdown periods utilities would still be able to timely and accurately bill consumers (ADD Grup, 2020).
- Energy usage data collected from the meters pre-, during and post- COVID would allow for further analysis into consumer behaviour during these 3 periods, as García et al. (2021) show in Figure 2.19, where it is noticed that during strict lockdown energy demand had decreased for non-residential consumers (businesses closing, employees working from home, etc.).

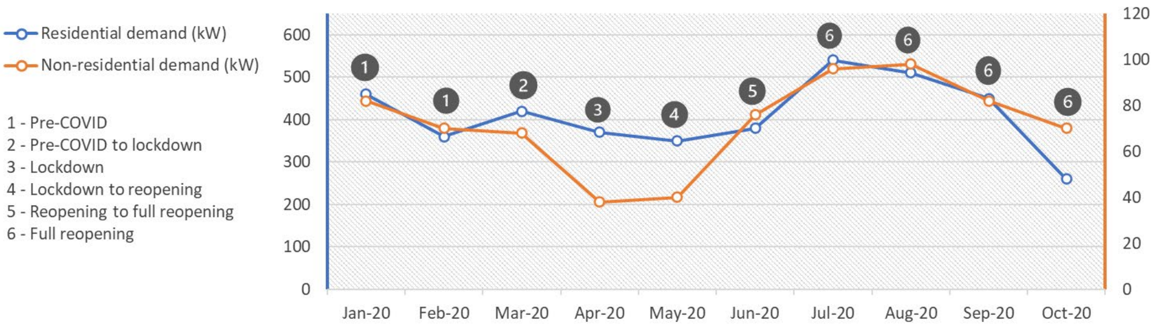


Figure 2.19: Demand profile during COVID-19 (based on García et al., 2021)

What stands out is the possibility of a smart metering system remotely accessing data (on-demand or at a schedule) reducing not only the constraints of health, safety and logistics for a physical site visit, but also the continuity of data and ability to study behaviours to address challenges.

2.3 Energy access through the SEAP framework

A key component in planning out the ideal solution for an EAP, the SEAP framework aims to identify technologies and resources which are sustainable and cost-effective, toward "providing universal access to basic energy services and to assess the affordability of cleaner-energy service options to energy-poor households" (Shrestha & Acharya, 2015); furthermore, it stands out from conventional energy and electricity planning frameworks through the following features, in particular for developing countries in the Middle East and sub-Saharan Africa regions:

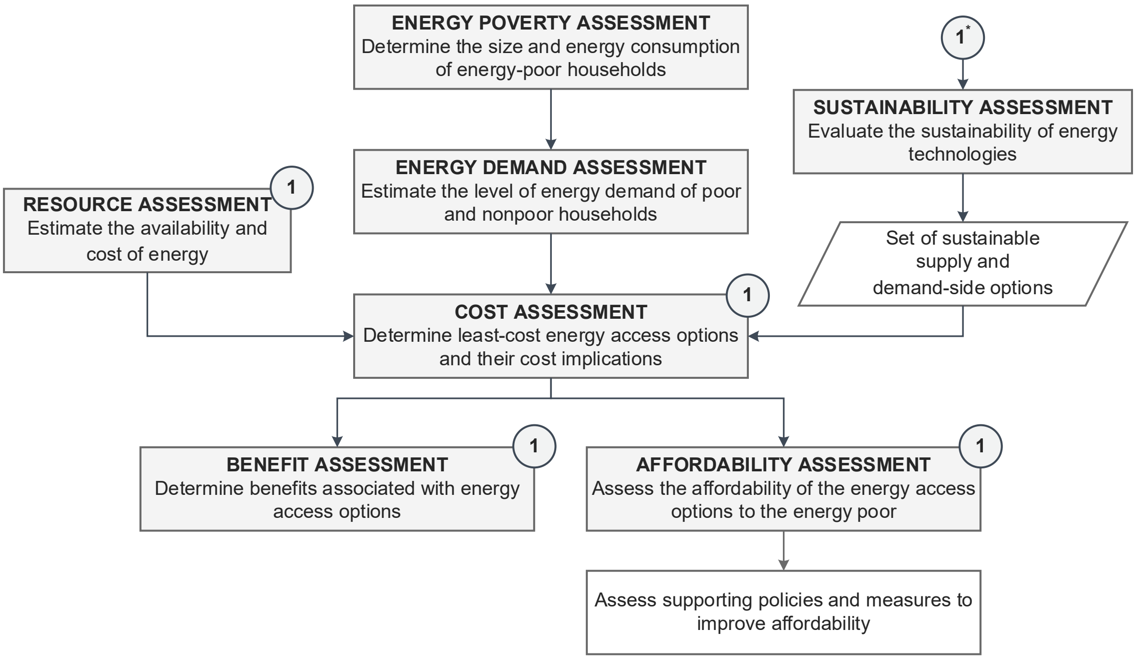
- Oriented to social inclusiveness, considering the poorest households in terms of access to electricity and other cleaner energies for necessities such as lighting, heating, and cooking.
- For the poor households, it contemplates the minimum acceptable level of essential energy services, while for the non-poor, it considers energy demand access through econometrics and other conventional approaches.
- Ensures affordability in the supply of electricity and other clean energies to the poor by evaluating the economic implications of even the more cost-effective options.
- Safeguards EAPs' quality and sustainability locally by analysing acceptable, reliable, and sustainable cleaner energy options.
- Generates key data on requirements for investment, along with EAP benefits concerning the reduction of energy inequality and greenhouse gases (GHGs), as well as improving environmental quality and social well-being.

- Ensures EAP continuity by analysing how acceptable, reliable, and sustainable the options of clean-energy services are.

2.3.1 SEAP framework assessments

Seven assessments establish the SEAP framework. As illustrated in Figure 2.20, the links between each assessment can be vertical, sequential, and horizontal as, for instance, the resource, cost, benefit, and affordability assessments provide inputs required for the sustainability assessment. Shrestha & Acharya (2015) define each assessment as:

- **Energy poverty** - the households considered as energy-poor and their energy consumption, providing the input to estimate the energy demand required by an energy supply system within an EAP.
- **Energy demand** - energy-poor household's present and future demand, for energy services such as water heating, cooking, use of other electrical appliances, lighting, and others, meeting the minimum acceptable level of basic energy services, such as the one indicated in the Global Tracking Framework (IEA&WB, 2014).
- **Energy resources** – indicates if sufficient energy resources are accessible to address present and future demand for the required volume of energy services, in the short, medium, and longer-term, and sustainably and reliably.
- **Cost** – cost impact for options and programs for cleaner energy, getting information around the investment total, as well as additional costs implicated in the development and implementation of a lower-cost EAP, and the poor households' burden for energy (affordability) in the program.
- **Benefits** – an EAP can have the following benefits: (i) improve the quality of the environment (especially the quality of air indoors); (ii) energy, health and social security; (iii) lessen energy inequality amongst countries, etc.
- **Sustainability** – the resource and technology options for providing energy access, measured in an EAP, are assessed for economic, institutional, technical, environmental, and social sustainability throughout their lifetime, from installation, operation, and maintenance next to pertinent indicators.
- **Affordability** – defines the amount to pay for basic energy services, those energy-poor households can afford (e.g., lighting, cooking, heating) and the population size that cannot.



*The resource, cost, benefit, and affordability assessments also provide inputs to the sustainability assessment.

Figure 2.20: SEAP framework's flow diagram (based on Shrestha & Acharya, 2015)

2.3.2 Relationship with energy data

Each assessment uses different approaches, with its applicable data requirements. Examining the SEAP framework and fifty-one diverse types of data required, it becomes apparent how each of the 7 SEAP assessments could, directly or indirectly, benefit from energy data collected from a smart metering system (Table 2.5).

Table 2.5: Matrix on SEAP assessments and energy data requirements (Shrestha & Acharya, 2015)

Data requirement	EP*	ED*	RSC*	CST*	SUS*	AFF*	BEN*
Electrified/Unelectrified Households	✓	✓		✓	✓	✓	✓
Basic Minimum Energy Requirement	✓	✓					
Energy Consumption by Fuel Type	✓	✓		✓	✓	✓	✓
Specific Electricity Consumption per Activity Level		✓					
Daily Load Profile				✓	✓		
Share of Renewable Energy in Electricity Supply					✓		
Hours of Electricity Supply					✓		
Potential for Renewable Energy			✓	✓	✓		
End-use Device Efficiency		✓		✓	✓		
Device power rating		✓					
Time of Use		✓		✓			
Losses in Transmission and Distribution				✓	✓		

*EP = energy poverty assessment, ED = energy demand assessment, RSC = resource assessment, CST = cost assessment, SUS = sustainability assessment, AFF = affordability assessment, BEN = benefit assessment.

2.4 Considerations

Smart metering and energy access have been addressed separately in the literature reviewed; moreover, a particular study by Casals et al. (2020) analysed a smart meter's capability to address energy poverty within the context of a developed world HH, with data provided by the utility.

This study takes instead an approach to energy poverty for the least developed and developing countries in SSA by directly using the smart metering system (meter, communications, software) to record and collect low-tier energy consumption data and then explore its applicability to an energy access framework that is relevant to those countries.

Therefore, smart metering and energy access become relevant if the earlier is applied to the latter, which is an approach that has not been explored, especially from the perspective of smart metering high-resolution data and detailed capabilities, applied to a framework like SEAP towards energy access implementation and sustainability. Shrestha & Acharya (2015: 21) noted: "In electricity access planning, besides the total electricity demand, the typical daily demand profile (load profile) … also contains valuable information, since electricity consumption varies during the day and the year … combined with information about household use patterns to arrive at the demand profile with an electricity access program. In an area without electricity supply, the demand profile of electrified areas that are similar enough to the area for which an EAP is being planned would indicate the likely demand profile of the area."; furthermore, Casals et al. (2020) have mentioned: "… big data mining from Smart Meter readings could provide valuable information to utilities about the consumption patterns of households allowing the identification of people at risk.". These statements highlight the value of using technologies such as smart metering for energy access planning.

CHAPTER THREE
METHODOLOGY

3.1 Introduction
3.2 Research design
3.3 Sampling method
3.4 Data collection
3.5 Data analysis
3.6 Ethical considerations
3.7 Summary

3.1 Introduction

This section explains the methods and processes to address the research questions. The literature review on the elements of smart metering and EAPs defined the context of the analysis. But instead of studying these elements separately, as has been done to date, the research aims at examining the relationship between them.

The following segments aim to describe the methodology proposed for the research, with an explanation of the research design, sampling method, data collection, data analysis, and ethical considerations.

3.2 Research design

This study involves primary and secondary data, from the smart meter and existing SSA country databases respectively, implemented EAPs, among others, therefore a quantitative research method will be used. The data will be primarily on energy consumption over time and other related variables. Saunders et al. (2007) indicate that any data collection or data analysis procedure that produces or utilizes numerical data defines as quantitative; furthermore states primary as new, purpose-specific data, and secondary as previously collected data for other purposes, but both significant towards answering the research questions.

3.3 Sampling method

Decreasing the amount of data by collecting it from a smaller group, instead of all probable cases, is achieved by implementing sampling techniques; moreover, sampling as allowing for improved accuracy when compared to a census (Saunders et al., 2007).

Simple random probability sampling was used in this research, based on the criteria described by Saunders et al. (2007: 216–218), with the resulting random numbering removing the bias factor from the data. The data was sampled as shown in Table 3.1.

Table 3.1: Selecting samples

Sampling frame	Country or country area in sub-Saharan Africa
Selection criteria for country	Availability and detail of secondary data, energy poverty severity
Case or element	Region/area within the country
Sample size	278 – 384 or 906 – 9595, within a 5% or 1% margin of error, respectively, based on a population range of 1 000 – 10 000 000 (Saunders et al., 2007: 212). The specific size will depend on the availability and detail of EAP and other relevant data necessary for the energy profile.

3.4 Data collection

Collected secondary data was derived from multiple sources, based on area and time series, such as government publications, books, journals, and industry statistics (Saunders et al., 2007). Some sources included organizations such as the country's local website and statistics, The World Bank, International Energy Agency, International Renewable Energy Agency, and others. Primary data was acquired from an experiment, based on the examination of energy profile data collected from the secondary data over a period of 1 to 3 months; the latter includes variables such as, but not limited to: (i) energy data – interval consumption profile per household, supplied electricity, quality of supply, renewable energy data; (ii) demographic data – population, number of households, electricity access, energy poverty.

Saunders et al. (2007) further state the experiment as a research method related to the natural sciences; and that it is often carried out in laboratories instead of in the field, allowing greater control over elements of the research process such as the selection of samples and the experiment occurrence context. Figure 3.1 shows the components of the experiment, as secondary data provides the energy profile to render detailed energy data scenarios on the smart meter using a variable load, resulting in new primary data for further analysis (which still includes secondary data).

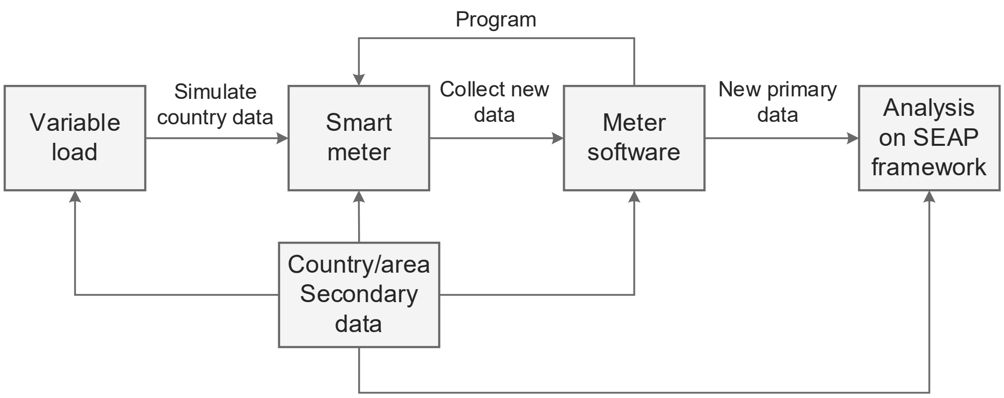


Figure 3.1: Steps of the proposed method to collect and apply the data to SEAP

3.5 Data analysis

Quantitative analysis was applied for both primary and secondary data in this study, using statistical methods and tools such as, but not limited to, Excel, Python, Jupyter, Azure ML and HOMER Pro. The analysis included the country/area data, the simulated data from the smart meter experiment, and the mapping/relationship with the SEAP framework components. Details on the approach are shown in

Table 3.2.

Table 3.2: Quantitative analysis approach, (Saunders et al., 2007)

Analysis component	Sub-component	Approach
Data preparation, input, and verification	Data type	▪ Quantifiable and categorical, sub-categories are applicable based on the country data and metering data available.
	Data layout	▪ Mostly tabular (data matrix), in software-compatible formats such as CSV, XLS and XLSX. Other less common formats may include PBI (Power BI) and IPYNB (Python).
Exploring and presenting data	Exploration	▪ Variable statistics, such as maximum, minimum, median, average, standard deviation, and distributions.
	Presentation	▪ Tabular, histogram, box plot, bar, line or pie charts, based on the data type.
Describe data applying statistics	Relationships	▪ Scatter chart, correlation, related to possible regression analysis on energy data
Examining differences, trends and relationships by applying statistics	Predict values	▪ Regression analysis on energy data provided by the smart meter (benefit highlight).

3.6 Ethical considerations

No ethical considerations need to be taken, as quantitative data will be extracted from publicly available databases (secondary data) and generated from an experiment using technology (primary data). Data bias, from a data ethics perspective, is also not a concern as random numbering will be used for sampling as part of the simple random probability sampling technique.

3.7 Summary

In this chapter, the methodologies and processes to carry out the research work were presented. From an SSA country/region selected based on available information, secondary data was acquired and applied as input for the smart metering experiment, with the generated primary household energy data analysed against the SEAP framework. The quantitative data were presented in different formats (tabular, CSV, XLSX, PBI, IPYNB), explored using statistical methods (distributions, average, standard deviation, among others), and presented using different visuals and charts based on the data type (line, pie, histogram, etc.). Data analysis included correlation, regression, and prediction of energy data. Tools used included Excel, Python, Azure ML and HOMER Pro, with no ethical concerns for the quantitative data.

CHAPTER FOUR
EXPERIMENTAL DESIGN

4.1 Introduction

Toncich mentions in his book that a clear mark of good research is "the ability to establish a series of experiments or studies that can push a concept to its operational limits and, thereby, expose those limits" (1999:264). This chapter attempts to explore boundaries by running experiments on energy data through its analysis, creation, prediction, and simulation, using different means and tools within the context of the energy access framework.

4.2 Secondary data: analysis

Located in Western sub-Saharan Africa, Niger has the third-lowest electric energy consumption per capita in the world (WB, 2022), with a population of nearly 24 million (CIA, 2022) of which most are rural as shown in the secondary data from table 4.1.

The number of households was calculated from the average household (HH) size. All secondary data was acquired for 2014, based on the latest available consumption data per HH in Niger, except for the required HH size (2012 latest); for this case, the HH size was normalized to 2014 by linear regression, based on the available historical data as shown in Figure 4.1.

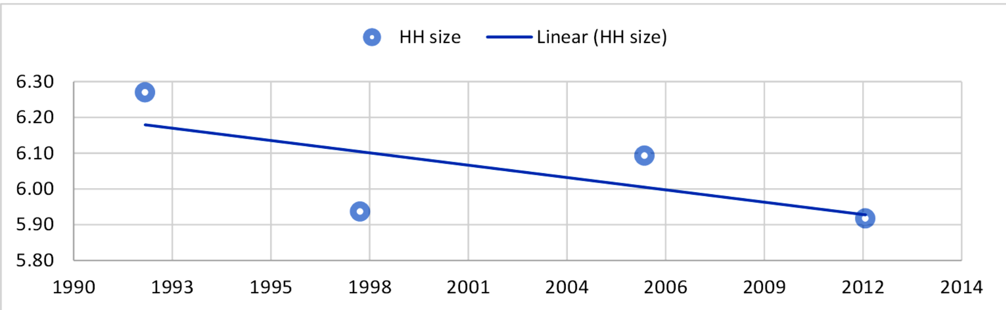


Figure 4.1: Historical HH size data and regression line

$$y = -3 \times 10^{5}x + 7.3421$$

Equation 4.1

With:

- Coefficient of determination, R^2=0.45
- Coefficient of correlation, ρ=-0.67

From linear regression Equation 4.1, the calculated HH size for 2014 is 6.08. Despite a coefficient of determination close to moderate effect (Karch, 2020), as all five data points were round to 6 and the number of people per household was used as an integer, the projected value was applied for the calculation.

Table 4.1: Niger energy and household data for 2014

Indicator	Value	Unit	Source
Access to electricity: total population	15.77	%	WB (2021)
Access to electricity: rural population	7.72	%	WB (2022a)
Rural population	84.00	%	WB (2022a)
Average HH size (projected)	6.00	Member/HH	UN (2019)
Number of HH (calculated)	4 034 439	Household	WB (2021); UN (2019)
Electricity consumption per HH	307.14	kWh/yr	UN (2019); CIA (2022)

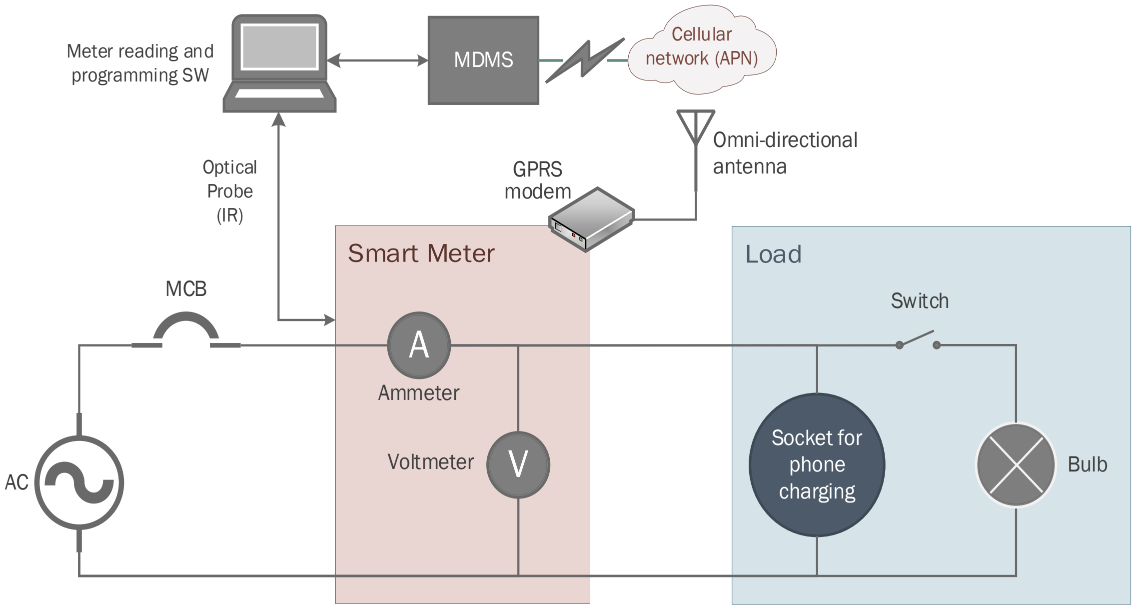
According to the multi-tier framework (MTF) for electricity access based on minimum thresholds for energy services (Table 4.2), the average household in Niger would fall under tier 2. However, seeing that majority of the population in the country, mostly rural, has no access to electricity (Table 4.1), consideration must be given to the lower tiers. This is corroborated by the fact that not only the rural areas lack access to modern energy services, but the urban areas with possible connectivity also face challenges with unreliable and poor access (WB, 2018); moreover, IEA & WB (2014) indicate that elements like the variety of appliances and energy efficiency would not accurately be reflected by a kilowatt-hour based indicator, with some household-level MTF as a key measure to get data on electricity access progress. This is where technology such as smart metering would aid with data collection and a more accurate energy profile.

Table 4.2: Multi-tier matrix of access to electricity services and consumption for a household (WB, 2015)

Tier	Electricity services	Annual consumption (kWh)
0	-	< 4.5
1	Task lighting + Phone charging	≥ 4.5
2	General lighting + Phone charging + TV + Fan	≥ 73
3	Tier 2 + any medium-power appliances	≥ 365
4	Tier 3 + any high-power appliances	≥ 1 250
5	Tier 4 + any very high-power appliances	≥ 3 000

4.3 Primary data: simulation

Based on the country profile collected, the system illustrated in Figure 4.2 was built to simulate and build primary data for a typical MTF tier-1 household in Niger.



*APN = Access Point Name, MCB = Miniature Circuit Breaker.

Figure 4.2: Designed system to simulate MTF tier-1 data for a household

System assembly and device parametrization was performed between Dec 2021 – Jan 2022, with the following aspects:

- Simulation start date: 16/01/2022 15:00
- Simulation end date: 26/04/2022 24:00
- Analysed data set period: 17/01/2022 00:00 - 26/04/2022 24:00
- Energy data to be collected:
  - Load profile (LP), also known as interval usage.
  - EOB cycles for the billing register (BR) data that include, among others, the monthly consumption based on current minus previous register calculation, and the maximum demand (MD) registers.
  - The following main events:
    - 29/01/2022: Mobile charging added to the load
    - 03/02/2022: Change to a low consumption bulb

The sections that follow provide detail on the main system components.

4.3.1 Smart meter

A complex and robust meter designed for commercial and industrial applications, the Itron SL7000 meter (Figure 4.5, component 1) includes the capabilities of a modern residential smart meter and more, as shown in Table 4.3.

Table 4.3: Technical specification of the Itron SL7000 meter (Itron, 2012)

Rating	3 x 57.7/100V to 3 x 277/480V, auto-ranging voltage In = 5A; Imax = 120A
Network Type	Direct Connection
Accuracy	Active energy: Class 1 Reactive energy: Class 2
Frequency	50 Hz
Meter firmware version	6.11
Standards	Full compliance with IEC 62052, IEC 62053, MID standard EN50470-1 and EN50470-3 and CE marking standards (electrical, mechanical, climatic, metrological, electromechanical), IEC 62056 (DLMS-Cosem Protocol).
Communication	IR port (IEC 61107) RS232 and RS485 ports (IEC 62056)
Input/Output (I/O)	4 pulse inputs 6 pulse outputs 2 control inputs 4 control outputs
Data	Bi-directional active, reactive, and apparent power Per phase measurement Load profile: 2 sets of 8 channels each Energy: 10 channels, 32 rates Demand: 10 channels, 24 rates Network monitoring

To log and collect the required energy data, the meter was programmed with the required parameters; Table 4.4 shows the energy data and corresponding parameters configured in the meter. Other related configurations included clock synchronization and register resets; every other parameter was left unchanged.

Table 4.4: Energy data parameters programmed into the meter

Energy data required	Parameter set	Unit set	No. of rates	Remarks
Energy registers	Import active aggregate	Wh	1	▪ Import - from the grid/supply ▪ Aggregate – total of all 3 phases ▪ Measured with no scaler (Wh instead of kWh) given the small size of the load.
	Export active aggregate	Wh	1	▪ Export - to the grid/supply
	Import reactive aggregate	VArh	1	
	Export reactive aggregate	VArh	1	
Demand registers	Import active aggregate	Wh	1	▪ Maximum demand register (top 5 for the configured cycle)
	Import apparent aggregate	VAh	1	
Billing register (energy register and	End of Billing (EOB): ▪ Trigger: internal clock	-	-	▪ EOB was set based on a calendar month billing cycle. At every reset date, the meter

Energy data required	Parameter set	Unit set	No. of rates	Remarks
demand register) sets	▪ Reset date: every 1st day of the month at midnight.			stored both energy and demand registers in a historical set, keeping the count for the earlier (cumulative) and resetting the latter.
Load profile (interval data) – set 1	Import active aggregate	Wh	-	▪ Recording interval to all channels: 30 minutes
	Export active aggregate	Wh	-	
	Import reactive aggregate	VArh	-	
	Export reactive aggregate	VArh	-	
	Import apparent aggregate	VAh	-	
	Export apparent aggregate	VAh	-	
	Power factor aggregate	-	-	
Load profile (interval data) – set 2	Voltage RMS value Phase 1	V	-	▪ Recording interval to all channels: 30 minutes
	Voltage RMS value Phase 2	V	-	
	Voltage RMS value Phase 3	V	-	
	Current RMS value Phase 1	A	-	
	Current RMS value Phase 2	A	-	
	Current RMS value Phase 3	A	-	
	Frequency	-	-	

4.3.2 Communications

Two different methods were used to communicate with the meter, with the respective devices as per Table 4.5:

- **Remote:** via GPRS modem (Figure 4.5, component 2), connected to the meter's RS232 port. This method was mostly used to read data using the MDMS and meter SW.
- **Local:** via IR optical probe, connected to the meter's IR port (Figure 4.5, component 3). This method was used to program the meter, or to read data if the remote method failed (e.g., due to cellular network issues).

Table 4.5: Technical specification of the communication devices (Itron, 2019; Tespro, 2022)

Modem: **Itron Sparklet GPRS**	**Rating**	▪ 10 VDC ▪ 100 mA to 500 mA ▪ Powered externally or via meter (data cable's RJ45 connector)
	Meter interface	▪ RS-232, via isolated RJ45 connector
	Indicator Light Emitting Diodes (LEDs)	▪ Power ▪ GSM signal strength ▪ Connection status
	GSM/GPRS	▪ Quad band 850/900/1800/1900 MHz2-bands 900/2100MHz
	Antenna	▪ External cellular antenna via FME connector
Optical probe: **Tespro TP-USB-IEC**	**Standard**	▪ IEC 62056-21
	Connector	▪ USB 2.0 connector
	Fastening	▪ Strong magnetic adhesion (IEC62056-21)
	Interface	▪ IR

4.3.3 Load

Based on MTF's tier 1 for a household, the load included:

- Bulbs (Figure 4.5, component 4), used in 2 stages:
    - Stage 1: 60 W, 548 lumen (lm) incandescent
    - Stage 2: low-cost 9W, 650 lm LED, replacing the incandescent bulb
- A socket (Figure 4.5, component 5):
    - 3-pin female socket with a 3-to-2 pin adaptor for mobile phone charging.

These elements were used randomly during the day, based on the need for lighting or charging. No specific schedule was applied to avoid bias in the data.

4.3.4 Software

To program, read and manage the devices, the following applications were used:

- **Itron MP Modem** SW: to check/program parameters (such as the APN details) in the GPRS modem, as well as monitor GSM signal strength (**Error! Reference source not found.**); the latter measured at a Received Signal Strength Indicator (RSSI) of -67 dBm, considered a good level within typical RSSI range (Azini et al., 2015).

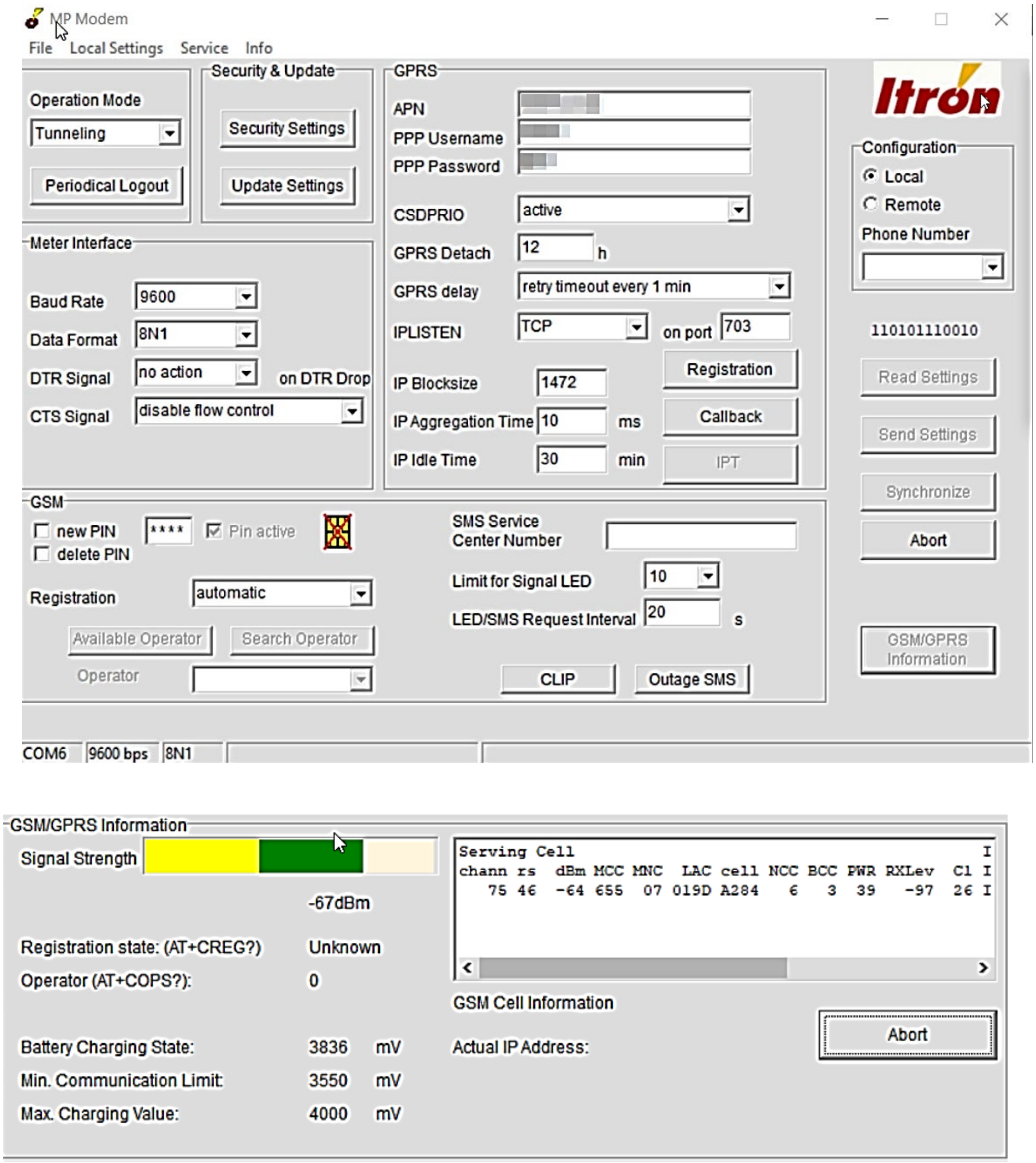


Figure 4.3: Parameters set for the GPRS modem, and signal strength measurement

- **ACE Pilot** meter SW: to check/program parameters in the meter (Table 4.4 and Figure 4.4), and to collect data either locally (via optical probe interface) or remotely (via GPRS modem). Besides the required energy data, specific data not accessible by the MDMS, such as events, fraud data, meter status and network history were collected using this specific meter application.

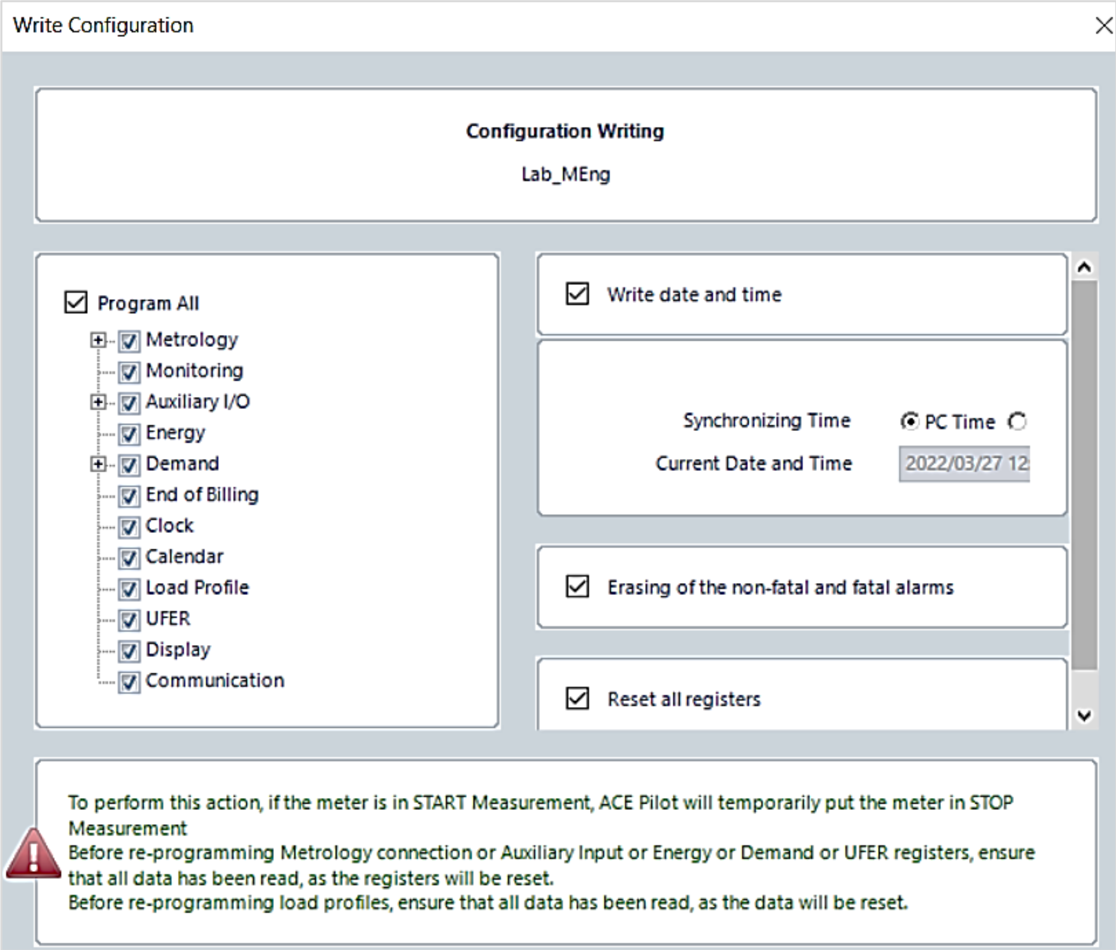


Figure 4.4: Programming the meter

- **PNPscada** MDMS: to remotely collect and analyse data from the meter, on an automated schedule, via GPRS modem; the following parameters were set on the MDMS:
  - Remote meter reading schedule: daily at 00:01
  - Data to read:
    - Load Profile
    - Real-time register totals
    - Real-time phasor data (RMS voltages, currents, and angles per phase)
    - Events (network monitoring, QoS).

The picture of the system, assembled with the components specified, is shown in Figure 4.5.

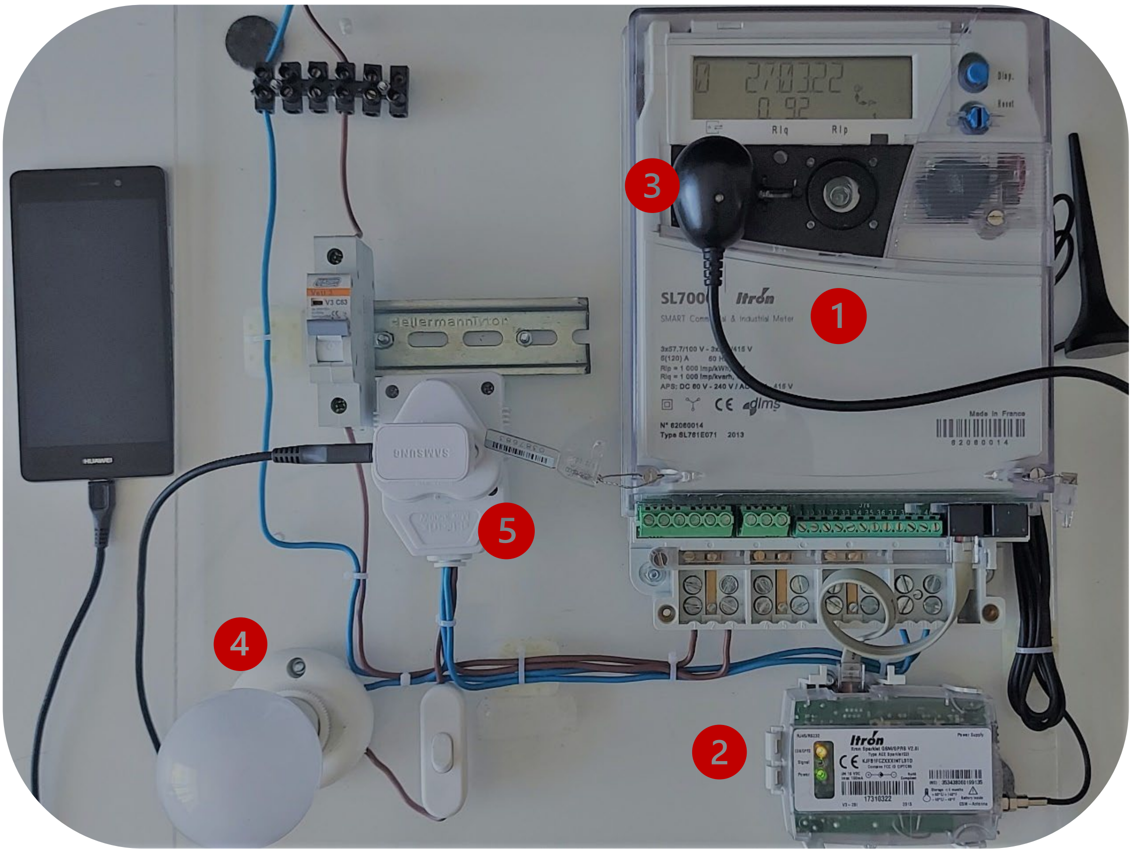


Figure 4.5: Designed simulation system

4.4 Energy data and the SEAP framework

The collected dataset from the simulation could then be applied as household data to the SEAP assessments, for both metered households in an area with access to electricity and for areas with no access to electricity but that have a profile similar enough to the metered, electrified area. For this purpose, the breakdown of energy data was defined as per Figure 4.6, mapping to the applicable assessments and use of the dataset to generate the required outputs, further described from a perspective of calculations or simulations; though each of the sections alone is fertile ground for further research, the current analysis was limited to the exploration of the different approaches to achieve such outputs (the journey), rather than just the outputs (the result). For simplicity, the acronyms defined for the SEAP assessments in Table 2.5 were used. The sections that follow cover different experiments, to analyse the different applications of energy data collected from a HH via smart meter.

Data requirements and SEAP framework		Applicable energy data per household	Benefit	Reference (where applicable)
Electrified and Unelectrified HHs	EP ED CST SUS AFF BEN	kWh interval consumption (LP), BR and MD	The database containing the energy data collected from the meter, used to ascertain which areas are supplied with electricity.	
Basic Minimum Energy Requirement	EP ED	kWh interval consumption (LP), BR and MD	A minimum threshold value is calculated based on the average kWh usage or maximum demand for the HHs of given area/region.	Equations 4.4, 4.6 - 4.8, and section 4.3.3
Energy Consumption by Fuel Type	EP ED CST SUS AFF BEN	kWh interval consumption (LP)	Usage periods for fuel-generated electricity assist with this categorization, as well as GHG emissions	Equations 4.13 - 4.15
Specific Electricity Usage per Activity Level	ED	kWh interval consumption (LP), BR and MD	Usage patterns for certain activities (e.g., small businesses), assists with this categorization.	Sections 4.3.2 and 4.3.3
Daily LP	CST SUS	kWh interval consumption (LP)	Daily LP data is used to characterize electricity demand	Section 4.3.3
Share of Renewable Energy in Electricity Supply	SUS	kWh interval consumption (LP), BR and MD	LP and BR on average daily and monthly load applied to a grid analysis tool	Section 4.3.3
Hours of Electricity Supply	SUS	kWh interval consumption (LP) and/or events	LP data accurately indicates the hours of electricity supply for a HH, with outage logs per event.	Equations 4.8 to 4.11
Potential Renewable Energy	RSC CST SUS	kWh interval consumption (LP), BR and MD	LP and BR on average daily and monthly load applied to a grid analysis tool	Equation 4.15, and section 4.3.3
End-use Device Efficiency	ED CST SUS	kWh interval consumption (LP), BR and MD	Track effectiveness of energy- and cost-saving initiatives, by comparing consumption before and after implementation.	Equations 4.13 and 4.14
Device power rating	ED	kWh interval consumption (LP)	With additional inputs from surveys, distinguish high and low power rating devices through usage patterns.	
Time of Use	ED CST	kWh interval consumption (LP)	Data is used to determine peak and off-peak consumption times, to help define tariff structures or implement load-shifting initiatives.	Section 4.3.3
Losses in Transmission and Distribution (T&D)	CST SUS	kWh interval consumption (LP), BR and MD	Determine energy losses by extending the smart metering system to T&D levels	

Figure 4.6: Energy data to SEAP mapping

4.4.1 Applicability: data to calculate

Deriving from figure 2.4’s matrix of energy data versus SEAP assessments, figure 4.8 linked specific energy data, the benefits, and which indicators could be calculated; while to some assessments the calculations based on energy data are applicable, to others it is a direct/indirect benefit. Hence applicable to the relevant assessments, the equations for those calculations were then defined.

4.4.1.1 Basic minimum energy requirement

The load profile (LP) acquired from the meter, per HH, comprises import active energy (kWh) consumption data in regular recording intervals, typically 30 minutes (min); hence for a recording interval I, in minutes, a HH’s total number of load profile intervals N recorded for the intended period can be calculated as:

$$N = \frac{1440 \times D}{I}$$

Equation 4.2

Therefore, a HH’s total electricity consumption for a period (such as a year), E_H in kWh, is given by the total of each interval consumption recorded by the meter E_i in kWh, determined as:

$$E_H = \sum_{i=1}^{N} E_i$$

Equation 4.3

Complemented with contextual data from the area/region (e.g., from HH surveys), the basic minimum electricity threshold can be taken as the average annual total electricity consumption per HH in the given area/region, i.e.:

$$\overline{E_{HE}} = \frac{\sum_{i=1}^{h} E_{H_i}}{h}$$

Equation 4.4

Where:

- $\overline{E_{HE}}$ = average electricity consumption per HH for the intended area/region and a given period, or if applicable the HHs minimum electricity threshold specific to the same area, in kWh.
- E_{H_i} = electricity consumption for each household, in kWh.
- h = total number of households.

Alternatively, the minimum electricity threshold can be determined from the monthly maximum demand, using the meter's billing registers recorded for the defined cycle (EOB), specifically the MD register. For this case, a typical calendar month cycle is considered, thus the maximum demand for a HH, for the considered period (e.g., a year), is determined as:

$$D_H = \max_{1 \le t \le M} D(t)$$

Equation 4.5

Where:

- D_H = maximum demand per HH for a period, in kW.
- $D(t)$ = monthly maximum demand per HH, in kW.
  - t = month of the year.
  - M= total number of months

Therefore, the minimum electricity threshold for the area $\overline{D_H}$, in kW, is calculated as:

$$\overline{D_H} = \frac{\sum_{i=1}^{h} D_H}{h}$$

Equation 4.6

Assuming 30min interval data, the threshold of Equation 4.6 can be converted to energy, through the following calculation:

$$\overline{E_{HD}} = \overline{D_H} \times 1/2$$

Equation 4.7

Where:

- $\overline{E_{HD}}$ = minimum electricity threshold for the area, in kWh.

4.4.1.2 Energy demand

According to (Shrestha & Acharya, 2015: 14–21), different methods in SEAP are used to determine energy demand. While the methods provide an estimate, smart meters installed at each of these endpoints would provide energy data (LP in particular), to determine:

- Historical consumption patterns
- Demand projection via forecasting techniques
- End-use of electricity

Comparatively, Equation 4.4 can be generally used to calculate a HH's annual electricity consumption, and with a previously defined minimum threshold (MTF-based, region-based, or calculated), determine the energy demand variance. Hence based on the steps by Shrestha & Acharya (2015:16), the additional electricity required ΔE_{DP} , in kWh, to raise that region's number of households h_P, from tier-n to tier-$(n+1)$, is given by a HH's average electricity consumption E_n and the minimum electricity threshold $E_{(n+1)}$, i.e.:

$$\Delta E_{DP} = h_P \times (E_{(n+1)} - E_n)$$

Equation 4.8

Where:

- $E_n = \overline{E_{He}}$
- $E_{(n+1)} > E_n$

4.4.1.3 Hours of Electricity Supply

The number of hours a HH has been supplied with electricity for a period (such as a day), t_S, is a SUS measure that indicates how reliable the energy supply or technology is (Shrestha & Acharya, 2015); now with the available LP data such variable is obtained with actual measured data based on the number of LP intervals, N_S, for which consumption was recorded by the meter during supply on the given day or period, and the recording interval I in minutes, i.e.:

$$t_S = \frac{N_S \times I}{60}$$

Equation 4.9

If required, the average daily HH hours of supply $\overline{t_S}$ for a given number of households h is determined by:

$$\overline{t_S} = \frac{\sum_{i=1}^{h} t_{S_i}}{h}$$

Equation 4.10

If instead calculating the number of hours a HH has been without electricity supply on a day or period, t_O, and the average daily HH hours for this scenario, $\overline{t_O}$, Equation 4.9 and Equation 4.10 can then be rewritten, respectively, as:

$$t_O = \frac{(N - N_S) \times I}{60}$$

Equation 4.11

$$\overline{t_O} = \frac{\sum_{i=1}^{h} t_{O_i}}{h}$$

Equation 4.12

4.4.1.4 End-use Device Efficiency

Assessing the efficiency of a device can be achieved in comparative, elemental actions, such as when in a HH an appliance or bulb is replaced. In section 4.3's experiment, an incandescent light bulb was replaced with a more energy-efficient, LED light bulb; in terms of energy consumption, the LP data set (30-min consumption recording) or BR data set (monthly consumption recording) was used to determine the difference before and after replacing the bulb, in three variants:

- Daily difference in electricity consumption ΔE_{de} , expressed as:

$$\Delta E_{de} = \sum_{i=1}^{N} E_{1_i} - \sum_{i=1}^{N} E_{2_i}$$

Equation 4.13

Where:

- N = number of recording intervals. For this case $N = 48$, as the LP recording interval is 30 min
- E_{1_i} = daily LP consumption at stage 1 (incandescent bulb)
- E_{2_i} = daily LP consumption at stage 2 (LED bulb)

- Period-based difference in electricity consumption ΔE_{pe} , expressed as:

$$\Delta E_{pe} = (R_{12} - R_{11}) - (R_{22} - R_{21})$$

Equation 4.14

Where:

- R_{12} = end register reading at stage 1 (incandescent bulb)
- R_{11} = start register reading at stage 1 (incandescent bulb)
- R_{22} = end register reading at stage 2 (LED bulb)
- R_{21} = start register reading at stage 2 (LED bulb)

- Cost difference, ΔC in South African Rand (R), determined as:
  $\Delta C = C_1 - C_2$

Equation 4.15

Where:

- C_1 = total monthly bill for stage 1 (incandescent bulb), obtained via MDMS provisional bill feature
- C_2 = total monthly bill for stage 2 (LED bulb), obtained via MDMS provisional bill feature)

4.4.1.5 GHG emissions

Shrestha & Acharya (2015) mention that the use of renewable energy sources and energy-efficient devices contributes to the avoidance of GHG emissions, and a comparison of the implementation can quantify such contribution, determined by:

$\Delta G = G_1 - G_2$

Equation 4.16

Where:

- ΔG = GHG emission reduction
- G_1 = GHG emission at stage 1 (before implementation)
- G_2 = GHG emission at stage 2 (after implementation)

4.4.2 Load forecasting: data to predict

The simulation for section 4.3.3 required LP data for a year, however, the current primary data simulated is for nearly 3 months. The current experiment attempted to fill that gap by applying electrical load forecasting (LF) through a machine learning (ML) model, built using Python (on Jupyter Notebook web interface) and Microsoft (MS) Azure ML Studio. The aim of this simulation was to, within the context of EP, ED, CST, and SUS, highlight smart metering energy data integration into machine learning and load prediction for future demand planning, preventive actions/maintenance, financial and tariff planning (Malik & Iqbal, 2021), especially when, according to Nti et al. (2020), close to 90% of the nine models mostly used in LF are based on artificial intelligence (AI).

4.4.2.1 LF components

The data-driven LF or electrical energy demand forecasting (EEDF) model in this experiment considers the following components, based on the inputs and related figures as described by Nti et al. (2020) and Malik & Iqbal (2021):

- Forecasting interval (lead time): medium-term forecasting (MTF), viz. 1-week to 1-year LF
- Error metric: Root-mean-square-error (RMSE), most used metric in LF (38%).
- Input parameters: weather and historical energy consumption; respectively, 50% and 38% of EEDF models are based on these variables.

4.4.2.2 Dataset definition

The small village of Gorou, in Niger, was selected for the exercise. Despite the limited information available for this community, the following key facts for the selection could be obtained from Plan International España (2017) and ECREEE (2018), in the context of a solar mini-grid project in Gorou:

- The village has about 4412 people.
- There was no access to electricity, i.e., on MTF tier-0 (Table 4.2).
- The community is far from the national electric grid, reinforcing the need for micro-grid renewable energy options.

To build the data set, as shown in Figure 4.7, the smart meter LP data was exported from the MDMS within the range 17/01/2022 18:00 - 17/04/2022 17:00; assuming it as the average demand per interval per HH, the total demand for the village was calculated and allotted to Gorou's historical 2021 weather data, from 01/01/2021 to 31/03/2021 (Freemeteo, 2022), as indicated in the sample from Table 4.6.

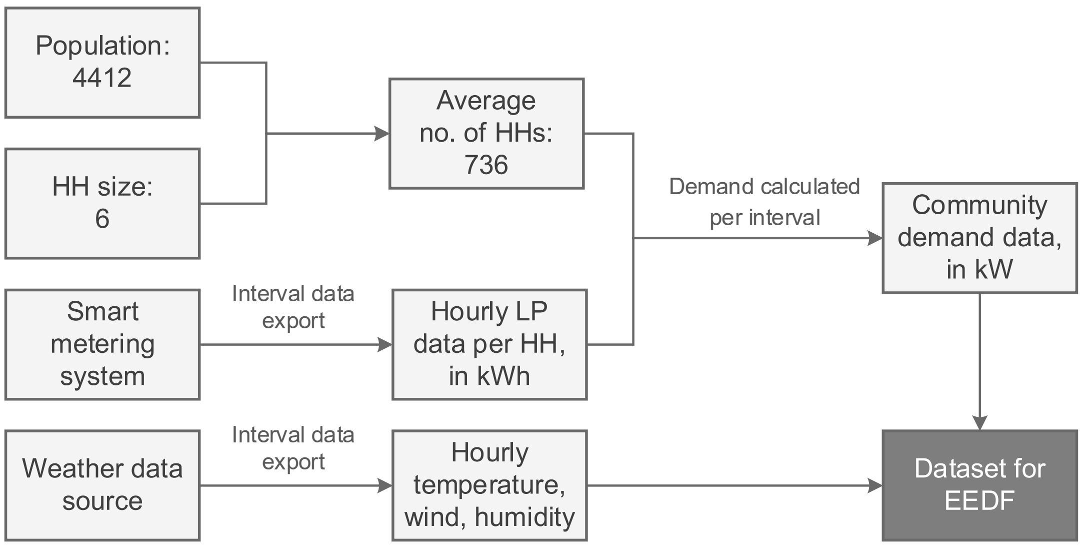


Figure 4.7: Dataset flow

Table 4.6: A sample from the 12-month dataset

Date	Time	Temperature	Wind	Humidity	Power_cut	Demand
2021/01/01	00:00	24.00	6.00	19.00	NO	0.00
2021/01/01	01:00	23.00	2.00	19.00	NO	0.00
2021/01/01	02:00	23.00	6.00	20.00	NO	44.16
2021/01/01	03:00	21.00	0.00	21.00	NO	44.16
2021/01/01	04:00	21.00	4.00	23.00	NO	44.16
2021/01/01	05:00	20.00	4.00	23.00	NO	0.00
2021/01/01	06:00	19.00	11.00	28.00	NO	0.00

4.4.2.3 Description and statistics

The analysis, performed using Python (Appendix B), was based on 2160 observations with three numeric features and one categorical feature (Table 4.7), using a regression model to predict the numeric label (demand data).

Table 4.7: Variable description

Variable/feature	Type	Description	Unit
Temperature	Numeric	Temperature	°C
Wind	Numeric	Wind speed	km/h
Humidity	Numeric	Relative humidity	%
Power_cut	Categorical	Power on (NO) or off (YES)	-
Demand	Numeric	Gorou hourly demand data	kW

Furthermore, numeric feature statistics that include the minimum, maximum, mean and standard deviation, indicated in

Table 4.8, assisted in the analysis of the dataset; the data showed a count lower than 2160 observations, indicating missing values, a key aspect to consider as it impacted the analysis and modelling, which was further addressed during the ML steps.

Table 4.8: Descriptive Statistics for numeric columns

Feature	Minimum	Maximum	Mean	Median	Range	Standard deviation	Count
Temperature	17.00	43.00	29.36	29.00	26.00	5.71	2113
Wind	0.00	39.00	12.96	11.00	39.00	7.85	2113
Humidity	2.00	50.00	11.75	11.00	48.00	6.02	2113
Demand	0.00	51.52	1.45	0.00	51.52	5.95	2128

The histogram in Figure 4.8 shows Demand values that are right-skewed, with a mean greater than the median value and a standard deviation larger than the mean; therefore, while electricity demand was mostly close to the value of a unit per hour, some few were distributed across a much broader range.

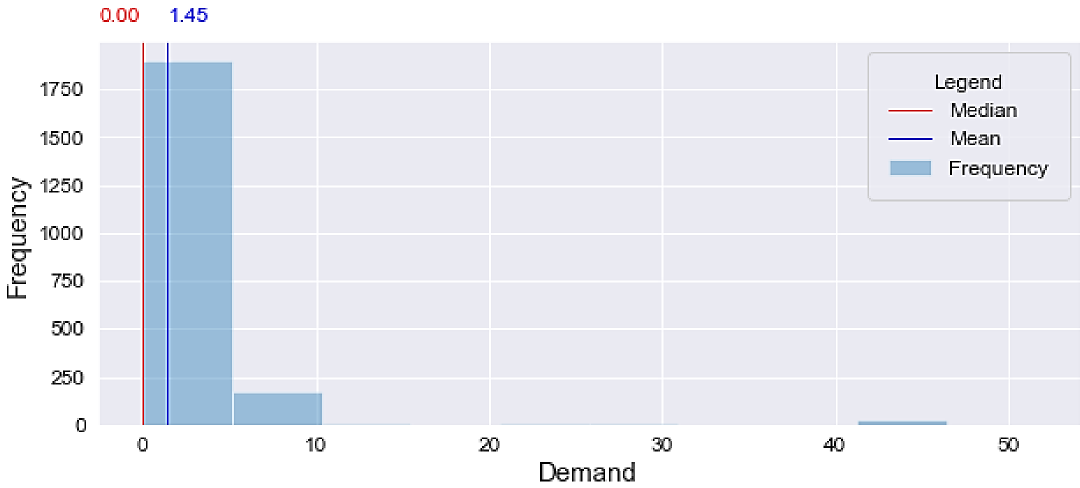


Figure 4.8: Histogram of Demand

The categorical feature Power_cut specified power supply to the households, indicating if the power supply was on (NO) or off (YES); with the meter also powered off, this resulted in missing demand data on the dataset.

4.4.2.4 Correlation and relationships

The relationships between the features in the data, and between the features and Demand, were analysed.

Numeric Relationships

The scatter-plot matrix in Figure 4.9 shows a comparison between the numeric features, including their distributions; when compared with Demand (the label), each of the features had a weak relationship with it, but from the relationship between the features a linear relationship could be identified between Temperature and Humidity, Wind and Humidity, and Temperature and Wind to some extent. Furthermore, Temperature is the only feature with a normal distribution, while the others are right-skewed.

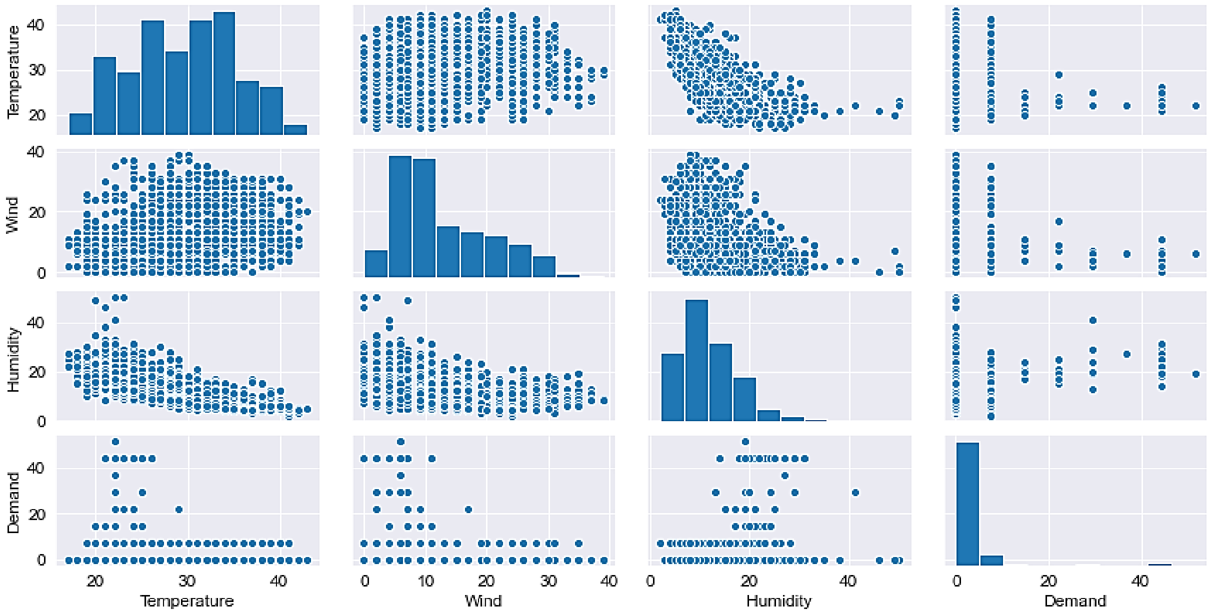


Figure 4.9: Scatter-plot matrix of numeric features

Figure 4.10 further shows these relationships, based on Pearson's correlation coefficient between the numeric features (Nettleton, 2014), of which magnitude is coded in shades of blue for positive and red for negative, darker as stronger the correlation is; here the correlation between the 3 features became apparent: (i) medium positive for Temperature and Wind; (ii) large negative for Temperature and Humidity; (iii) medium-large negative for Wind and Humidity.

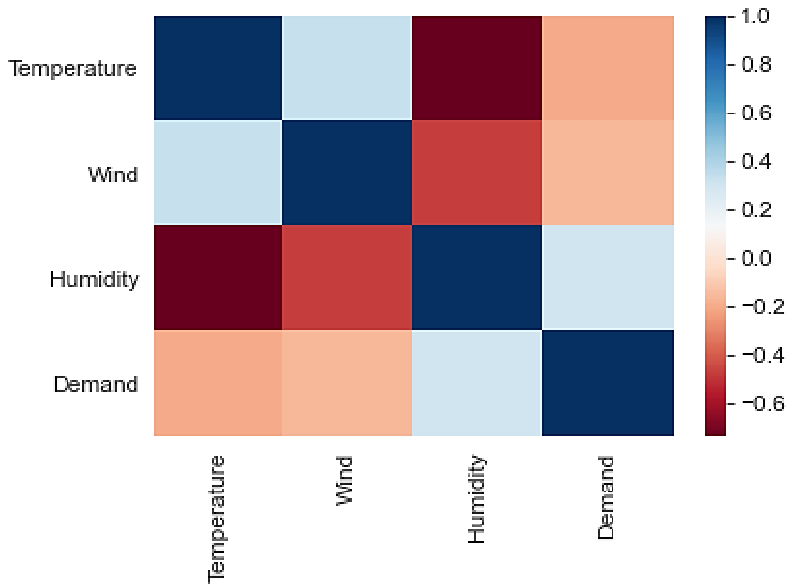


Figure 4.10: Correlation heatmap of numeric features

– *Categorical Relationships*

Missing Demand data could be caused by several factors, which would define the approach to handle it, either by removing or completing the missing values. The Power_cut feature clarifies this aspect, where the category YES confirms that the missing data is due to no electricity supply, with no data collected; furthermore, it provides information on the proportion of power cuts (missing data) versus no power cuts, as indicated in Figure 4.11. Missing values were therefore removed and not estimated, and while this feature had a key purpose for data preparation, it was not used as a predictive feature for Demand.

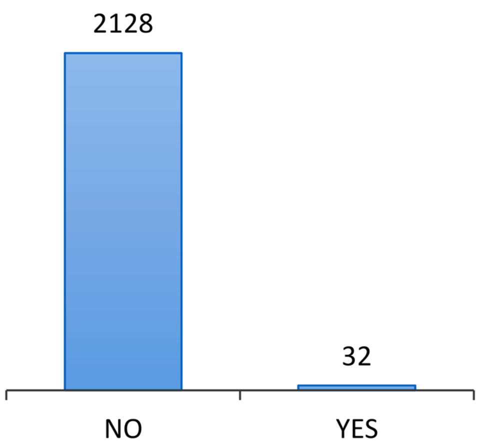


Figure 4.11: Count by Power_cut categories

– *Multi-Dimensional Relationships*

Besides the relationships between each of the features and Demand, a multi-faceted relationship was also analysed for a wider view of how the features define the label to be predicted, using the features with the highest correlation coefficient. As shown in Figure 4.12, while higher Demand values were clustered within lower temperatures (20 to 25°C) and medium humidity (23 to 33%), lower values were frequent and spread across the Temperature-Humidity relationship, i.e., the increase in Temperature or decrease in Humidity did not appear to have a strong impact over Demand.

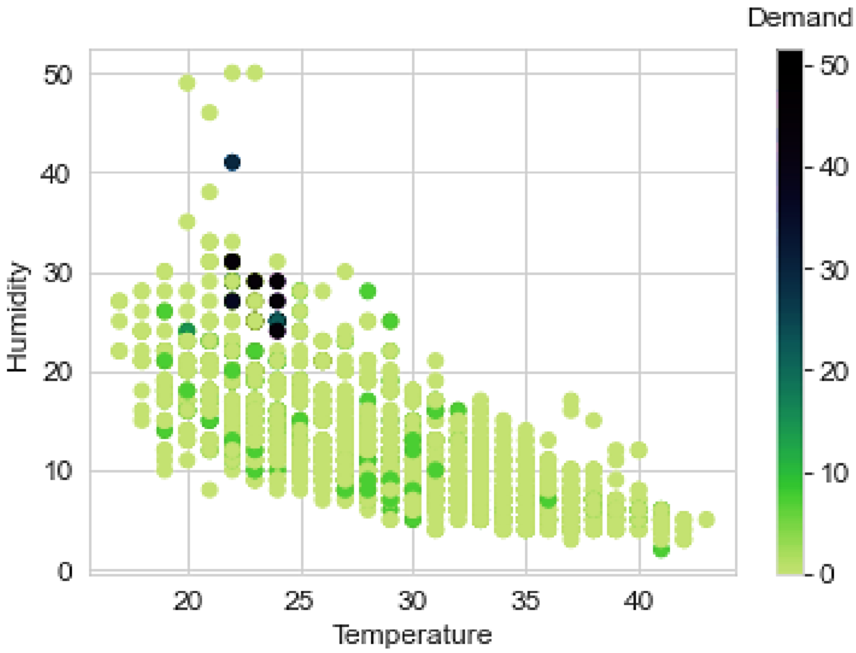


Figure 4.12: Influence of Temperature and Humidity interactions on Demand

4.4.2.5 Prediction of Demand

From the analysis of the features and the target label Demand, a model was defined to predict Demand where the value was not known, using MS Azure ML Studio.

- *Data preparation*

The following steps were applied to the data before creating the model:

- Treat missing values on affected features:
  - Demand: rows removed.
  - Temperature: mean imputation with data calculated from the column, applicable to normal distributions (Zest AI, 2018).
  - Wind and Humidity: median imputation with data calculated from the column, applicable to imbalanced distributions (Zest AI, 2018).
- Transformations to improve distribution properties and predictive ability:
  - Wind: square root.
  - Humidity: natural logarithm plus one.
- Normalization for modelling:
  - Temperature: Zscore method, for normal distributions (Microsoft, n.d.).
  - Wind and Humidity: MinMax method, for imbalanced distributions (Microsoft, n.d.).

- *Regression model*

Following the prepared data, a regression model (Figure 4.13) was created to predict Demand, with the data split into training and test sets at 70% and 30% of the dataset, respectively.

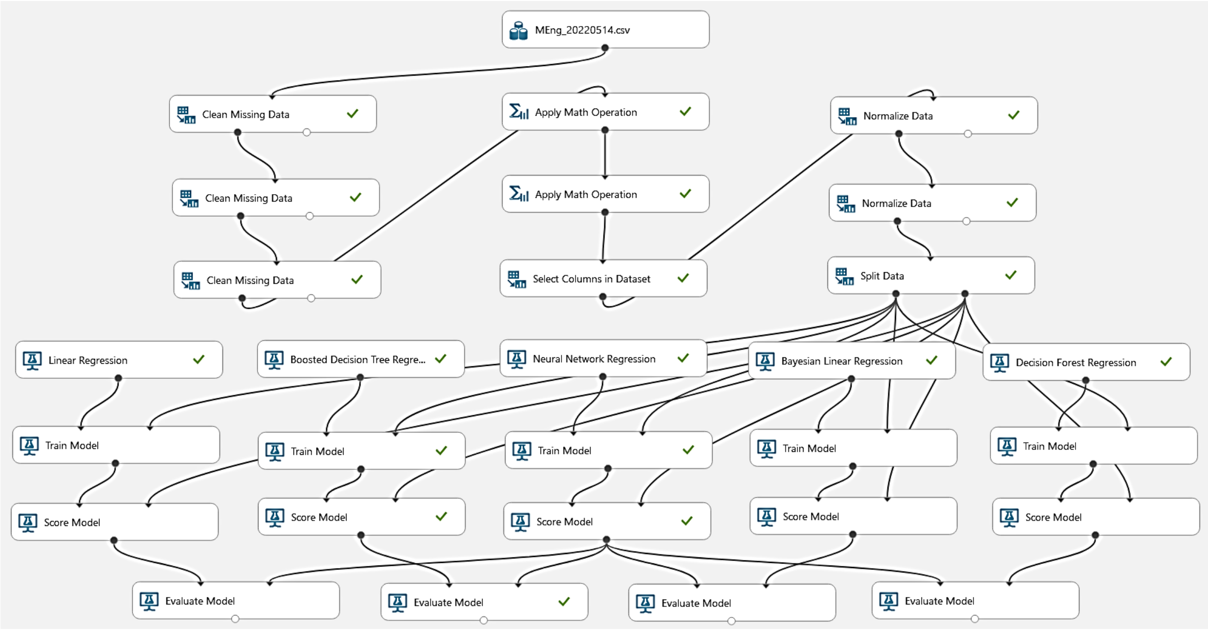


Figure 4.13: Machine learning model

A total of five models were run, with the Neural Network Regression (NNR) model being selected based on, among other metrics, the lowest RMSE and highest R-squared, as shown in Table 4.9.

Table 4.9: Regression models and metrics applied

Metric	Linear	Boosted Decision Tree	Neural Network	Bayesian Linear	Decision Forest
Mean Absolute Error	2.69	2.25	2.27	2.68	2.23
Root Mean Squared Error	6.04	6.03	5.74	6.04	6.07
Relative Absolute Error	1.03	0.86	0.87	1.03	0.86
Relative Squared Error	0.93	0.93	0.84	0.93	0.94
Coefficient of Determination (R^2)	0.07	0.07	0.16	0.07	0.06

The plot in Figure 4.14 indicates a slight linear relationship between predicted and actual values, with a dispersion tendency for larger values. A low R^2 (0.16 on a scale of 0 to 1) could be normal (Microsoft, 2019); further experiments, not within the boundaries of the current study, would contribute to improving the model's performance.

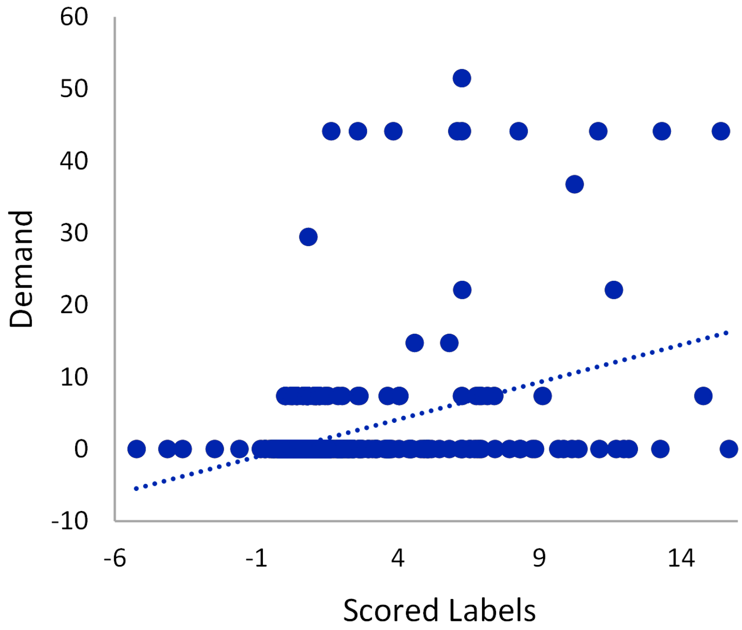


Figure 4.14: Scored results compared to actual values for the Neural Network model.

– *Prediction of Demand for the full year*

Using the ML model by Neural Network Regression, the Demand values were predicted for the remainder of the year, i.e., 01/04/2021 00:00 - 01/01/2022 00:00.

4.4.3 Energy potential and renewables: data to optimize

For the analysis of RSC (energy resources), CST (cost-effective), BEN (beneficial), AFF (affordable) and SUS (sustainable), rather than manual calculations an equivalent techno-economic simulation tool was used with the LP dataset as input. HOMER Pro runs simulations on energy systems for one year, evaluating and optimizing design, costs, load profiles, components, and environmental variables (HOMER Energy, 2017); while it covered all the elements of the assessments as indicated in Table 4.10, it also allowed the use of raw energy data collected from a HH, by using LP as direct input for the load setup, instead of the pre-defined synthetic load. The latter benefit was one of the main advantages to be shown with this approach.

Table 4.10: Assessments versus software features

Assessment	Description	HOMER Pro
RSC	▪ For off-grid power supply, identify renewable energy resources within the area where the EAP will be implemented; the grid-connected power supply is also considered. ▪ The economic potential of energy resources. ▪ Input for CST and SUS.	▪ Identify technology options based on location ▪ Renewable penetration ▪ Sensitivity analysis
CST	▪ Cost-effective options of access to electricity.	▪ Sensitivity analysis
BEN	▪ The benefit of access to cleaner energies.	▪ Emissions
AFF	▪ HH affordability to use electricity services	▪ Sensitivity analysis
SUS	▪ Energy resource options that are environment-friendly, reliable, affordable, cost-effective, and socially acceptable. ▪ Gets its input from RSC, CST, BEN and AFF.	▪ Emissions ▪ Sensitivity analysis ▪ Identify technology options based on location

4.4.3.1 Inputs

The interval-based (LP) demand dataset for the year 2021 results from the following outputs:

- Smart meter data (section 4.3), for the range 01/01/2021 00:00 - 31/03/2021 23:00, with missing demand values replaced with zeros.
- Predicted demand (section 4.4.2), for the range 01/04/2021 00:00 - 01/01/2022 00:00, with negative demand values replaced with zeros.

Besides the LP dataset, other inputs, as indicated in Table 4.11, included the solar and wind resources acquired directly from the software based on the location input, as well as economic variables. General system default values (costs and specifications) were used for this simulation.

Table 4.11: Defined inputs for the simulation

Variables	Settings/inputs
Location	▪ Community of Gorou, Niger. ▪ GPS coordinates: 14°3.0'N, 1°47.0'E.
Electric load	▪ Imported from time series (LP dataset for 2021).
Grid	▪ Set as connected, to include a grid-connected power supply in the analysis.
Fuel-based generator	▪ Considered for the base case for the analysis. ▪ Default values considered.
Renewables	▪ Solar photovoltaic (PV) ▪ Wind turbine
Sensitivity inputs	▪ Discount rates (%): 3, 6 and 12 ▪ Inflation rates (%): 1, 2 and 4 ▪ Diesel fuel price (R/L): 7.83, 15.70 and 31.30 ▪ Wind scaled average (m/s): 3.00. 5.68 and 8.00 ▪ Project lifetime (years): 25
Currency	▪ Default costs used, in rands ▪ $ to R conversion based on an exchange rate of 15.66 on 25/04/2022 (XE, 2022)

4.4.3.1 Importing the load and building the design

Using synthetic (predefined) loads (Figure 4.15), is a quick approach to generate a relatively realistic load for the simulation (HOMER Energy, 2017), but actual data from the source (the household) would provide a more accurate reflection of the load, thus a better approach.

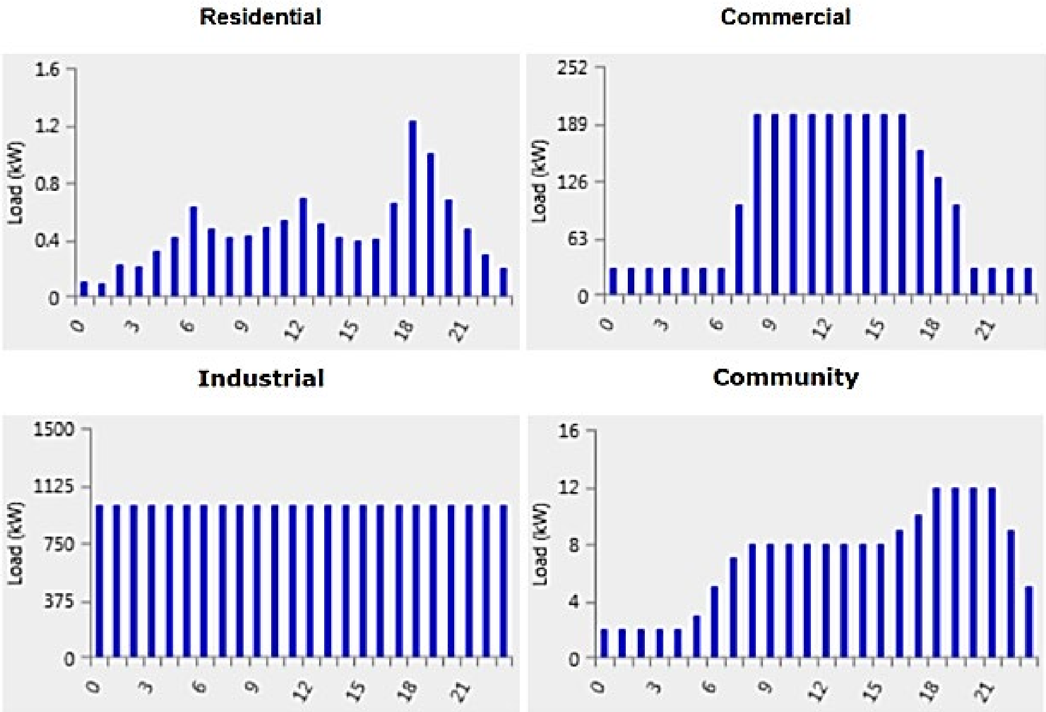


Figure 4.15: HOMER Pro synthetic loads (HOMER Energy, 2017)

The interval-based energy data was pre-processed and imported using HOMER's time-series import feature, and added to the design as daily profile, illustrated in Figure 4.16; the data was shifted to start on the first day of the year before the average daily profile is determined, as per criteria described by HOMER Energy (2017).

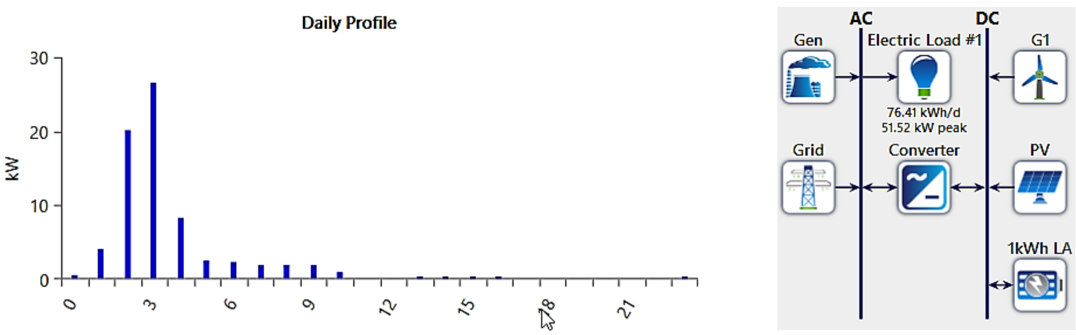


Figure 4.16: Generated daily profile (left) and the designed system (right)

The system had six electrical supply elements that included the renewable solar and wind options, the related storage (battery), and current conversion components described in Table 4.12, with the respective input values of both capacity and costs. Once the design was set, the simulation was run to perform the optimization and sensitivity calculation, and analyses.

Table 4.12: System components

Component	Description	Capital (R)	Replacement (R)	O&M* (R/year)	Lifetime (years)
Electric load	Average load demand based on the imported data, 51.52 kW peak.	-	-	-	-
Gen	Electrical supply: auto-sizing diesel generator	7 830.00	7 830.00	4 115.45	1.71
PV	Electrical supply: generic 1 kW flat-plate PV system	39 150.00	39 150.00	156.60	25.00
G1	Electrical supply: generic 1 kW wind turbine	109 620.00	109 620.00	1 096.20	20.00
1kWh LA	Electrical supply: 1 kWh Lead Acid (LA) battery.	4 698.00	4 698.00	156.60	10.00
Converter	Electrical supply: Rectifier for AC-to-DC conversion, and inverter for DC-to-AC conversion.	4 698.00	4 698.00	-	15.00
Grid	Electrical supply: a grid-connected system for grid extension analysis, power priced at 1.57 R/kWh	125 280 (per km)	-	2 505.6 (per km)	-

*O&M = Operation and maintenance

4.5 Summary

The experiment was defined in this chapter. With one of the lowest electricity consumption levels in Africa and the World, Niger was the selected SSA country, from which household and electricity access data were collected and analysed. Some linear regression was applied to normalize data to a particular year. The multi-tier framework was an important aspect analysed as it measured the level of access to electricity in the country, more severe in rural areas. The tier-2 level identified was then applied to the smart metering experiment designed to measure energy data from a household, collected remotely by an MDMS.

Following assembly of the smart metering experiment, an Itron SL000 smart meter was configured with the required measurement parameters using Itron ACE Pilot meter configuration software over an IR probe (local access); the parameters included energy and demand registers, profile interval channels and clock synchronization. For remote access, the Itron Sparklet modem was used for remote communication over the cellular network (GPRS), on which parameters such as the APN and signal were checked (at -67 dBm, good level signal strength) and configured and using Itron MP Modem software. The household energy, register and event data logged by the smart meter was then collected daily at midnight, via automated schedule set on the PNPscada MDMS. The load to be measured was set up as per identified tier-2 electricity access level: a bulb and an energy socket for cell phone charging at two stages, the first with an incandescent bulb and the second with an LED bulb replacement, defined to evaluate the impact of efficient appliances/devices in energy demand, cost, benefit, and sustainability; moreover, the load followed a random household usage to avoid data bias.

The applicability of the data to the SEAP framework was then analysed by mapping the data requirements to the measured household energy data, within the range 17/01/2022 00:00 - 26/04/2022 24:00. Where then applicable, energy access indicators were calculated based on the framework's criteria: the basic minimum energy requirement, energy demand, hours of electricity supply, end-use device efficiency (the bulb load stages of incandescent VS LED), and GHG emissions.

The household energy data measured was further explored by medium-term electrical load forecasting (LF) through a machine learning (ML) model, taking the existing 3-month dataset (17/01/2022 18:00 - 17/04/2022 17:00) allocated to the year 2021's weather data, to then train and predict demand data to cover that full year; extending the data set was with the intent of not only exploring the potential of smart metering energy data but also as input for the other aspects of the study. Inputs were local weather data, population, household size and demand data; the small rural village of Gorou, in Niger, was selected for this exercise, based on the criterion of village size, no access to electricity (tier-0) and distance from the grid. Statistics, correlation, and relationships were analysed in Python, and regression models were trained and built based on the 3-month dataset on MS Azure ML Studio. Five regression models were simulated, with the neural network model being selected based on the RMSE and R-squared metrics.

The complete year dataset, comprising 3 months of smart metering data and 9 months of predicted data, was then applied to run optimization and cost sensitivity analysis of hybrid renewable energy systems using HOMER Pro software. The designed energy system to be optimized comprised six electrical supply components that included a diesel Genset, renewable solar and wind based on Gorou's resources, storage (batteries), converter (DC-to-AC inverter and AC-to-DC rectifier), and grid supply; furthermore, sensitivity inputs were based on interest, inflation, diesel costs, and wind speed, over a 25-year project lifetime.

CHAPTER FIVE
EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Introduction

The designed experiment aimed to explore the energy data through its different applications relevant to the energy access framework. This chapter provides and discusses the outputs at the boundaries explored in terms of the SEAP framework assessments applicable through the measured household energy data from the smart metering system, the calculation results for the applicable SEAP indicators, the predicted demand data, and the optimized energy system.

5.2 Meter data collection

In the built experiment a smart meter collected energy usage data from a household load based on MTF tier-2. Using the different features available from smart metering technologies, diverse data was acquired using local and remote access methods to the smart meter, each providing different levels of access to the meter, the detailed data in it, and the analysis and reporting from the MDMS system.

5.2.1 Local access by meter SW over IR probe

Even though in an SM system most of the access to a smart meter is accomplished remotely, it is occasionally required that the meter is locally accessed; such cases include:

- Initial configuration of the meter.
- Collect data from the meter that the MDMS might not be able to collect, such as detailed event and alarm data, and firmware version.
- Collect data from the meter when remote communication is not possible.

With the IR optical probe attached to the meter's optical sensor, and using the configuration SW, energy data on the load, device and installation was collected from the meter at 14/02/2022 20:42; such data included real-time power (at 14/02/2022 20:40), energy, and event logs, with the observations as indicated in Table **5.1**. For a measurement-transformer-connected meter, parameters such as the current transformer (CT) and voltage transformer (VT) ratios would've also been included.

Table 5.1: Energy data, status, and observations

Type	Parameter	Value/message	Observation
Real-time power usage	Import active, P+	8 W	Only the 9W LED bulb was on (no phone charging).
	Export active, P-	0 W	No exported quantities, only imported from the grid.
	Import reactive, Q+	0 var	No reactive quantities, despite LED bulb.
	Export reactive, Q-	2 var	No exported quantities, only imported from the grid.
	Power factor	1.0	No reactive quantities, despite LED bulb.
Energy registers	Import active, A+	2 899 Wh	Energy register count since the start of measurement
	Export active, A-	0 Wh	No exported quantities, only imported from the grid. No reactive quantities, despite LED bulb.
	Import reactive, R+	0 VArh	No reactive quantities, despite LED bulb.
	Export reactive, R-	0 VArh	No exported quantities, only imported from the grid. No reactive quantities, despite LED bulb.
Logs	**Date and time:** meter	2022/02/14 20:40:00	Synchronized, otherwise, it would result in inaccurate data.
	Date and time: PC	2022/02/14 20:40:11	The PC's time was used to synchronize the meter time.
	Alarms: Fatal		Not present. A fatal alarm indicates a severe problem with the meter, requiring removal and testing.
	Alarms: Non-fatal	• Network voltage unbalance detected. • Voltage cut detected in phase 2. • Voltage cut detected in phase 3.	Only 1 phase is being used, hence the absence of voltage in the remaining phases.

Event data is key to monitoring the network for factors such as QoS. One detailed case in Table 5.2 showed a power cut event, accurately logged from the time when the alarm was raised for a phase-cut, then a power cut, to restoration; the power cut lasted for exactly 6 minutes and 58 seconds.

Table 5.2: A sample meter configuration tool export of event data (network monitoring)

Type	Setting	ID	Date	Time
Register data saved in Flash memory	SAVE POWER UP	762	Sunday, 13 February 2022	19:22:49
AC fail signal disappearance	PRESENCE PHASE	761	Sunday, 13 February 2022	19:22:46
Non-fatal alarm disappearance	Alarm Type: VOLTAGE CUT PHASE 1	760	Sunday, 13 February 2022	19:22:46
Power up		759	Sunday, 13 February 2022	19:22:46
Power fail signal raised		758	Sunday, 13 February 2022	19:15:48
AC Fail signal raised	PHASE CUT	757	Sunday, 13 February 2022	19:15:48
Non-fatal alarm raised	Alarm Type: VOLTAGE CUT PHASE 1	756	Sunday, 13 February 2022	19:15:48
Non-fatal alarm raised	Alarm Type: VOLTAGE SAG PHASE 1	755	Sunday, 13 February 2022	19:15:48

While a smart meter in a household is an important element in an SM system, it is as equally important to make sure that such device is deployed with the correct parameters to measure and log the required energy data for the required purpose, and that further to the regular collection of the data, the device, installation, load, and supply aspects are monitored. Thus, the aspects of meter configuration, real-time data and event logging are important measures to the SEAP framework as they quantify and track the aspects of reliability, convenience, and adequacy of supply to a household. Now despite the advantages of monitoring through data collection, i.e., proactively identifying issues and addressing them, the maintenance actions to be taken on-site to fix those issues would require local know-how to use the required technologies and tools, and availability of resources.

5.2.2 Remote access by MDMS over GPRS

Once the meter was configured for the measurement and the modem configured for communications, the MDMS was scheduled to collect energy data regularly and provide different types of reports and views for the data analysis.

5.2.2.1 Daily load profile

The graph in Figure 5.1 showed some of the insights based on collected load profile data, with three load scenarios and related events:

- The initial load from 17/01/2022: a 60 W, 548 lumens (lm) incandescent bulb applied.
- An increase in usage from 29/01/2022: introduction of phone charging.
- A decrease in usage from 03/02/2022: incandescent bulb replaced by a 9W, 650 lumens Light Emitting Diode (LED).

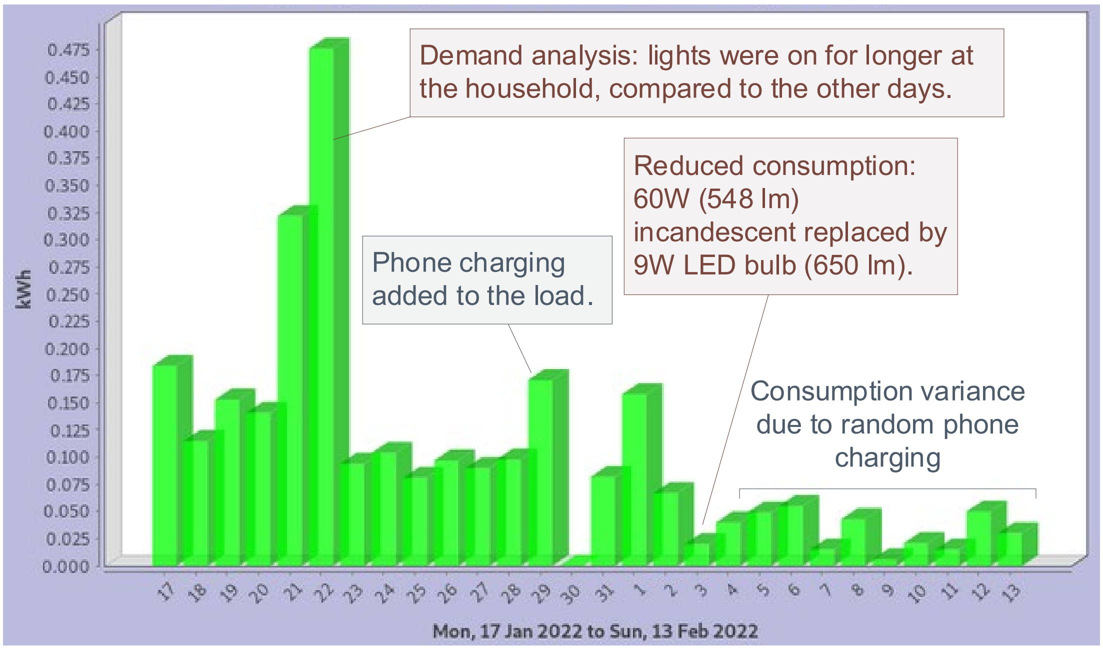


Figure 5.1: Daily consumption analysis (kWh per day).

The data from the last scenario measures and highlights the use of efficient options for demand-side technology in the SEAP framework (ADB, 2018b: 86): at a consumption 6 times lower (9 W) while providing more light, the LED lamp is more efficient and cost-effective (Desai, 2015), addressing important factors of affordability (for both lamp and electricity bill) and adequacy (low power rating), where aspects such as power generation capacity requirement and cost of providing electricity access would be reduced.

5.2.2.2 Hourly load profile

Observing the load profile further to a higher data resolution (every few hours to an hour, to half an hour), consumption patterns were identified, as illustrated in Figure 5.2:

- Accessibility and affordability through demand and ToU: a demand higher for both charging and lighting occurring in the evening, after 18:00 hours.
- Reliability through QoS monitoring: a stable voltage for phase 1.

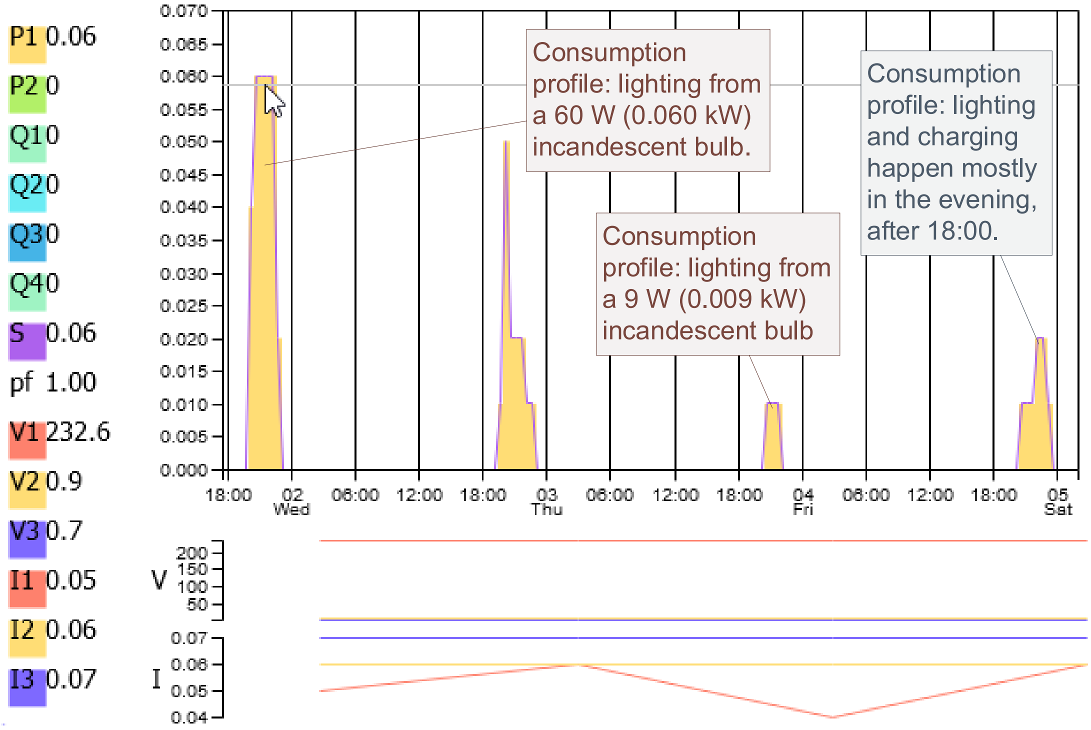


Figure 5.2: Half-hourly usage analysis (kW per day).

This level of granularity would assist in addressing affordability and affordability challenges by implementing improved electricity tariffs (ADB, 2018b) that benefit the energy-poor, optimize tariffs for the non-energy poor, and QoS (reliability). Furthermore, the interval-based, load profile data opens several possibilities to explore usage data, as indicated in section 4.4 and further determined in section 5.3 of this chapter.

5.2.2.3 Environmental impact

The United Nations refers to the importance of providing, where viable, energy access from low GHG emission sources (Shrestha & Acharya, 2015). With the energy data collected, environmental impact analysis data was generated from the MDMS, constants set out by the local utility and the specified consumption period, and the emission and utilization values for the different compounds, as indicated in Table **5.3**; considering aspects of carbon emissions and water use as an example, for the nearly 27 days of household consumption, close to 3 kg of CO_2 was released and 4 litres of water were used to generate the required electricity.

Table 5.3: Environmental impact (based on South Africa)

Implication	Unit	2022-01-17 16:00 to 2022-02-13 20:30	Daily Average	Monthly Average
Ash produced	kg	0.4280	0.0157	0.4791
CO2 emissions	kg	2.7334	0.1005	3.0601
Coal use	kg	1.4633	0.0538	1.6383
NOx emissions	kg	0.0115	4.2450	0.0129
Particulate emissions	kg	9.1113	3.3513	0.0010
SO2 emissions	kg	0.0214	7.8704	0.0240
Water use	kl	0.0039	1.4218	0.4327

Besides just energy, the data acquired from the smart metering system can also be used as input to generate environmental data and any other data that can be derived from HH usage. Applicable to the benefit and sustainability elements of the SEAP framework, and as further determined in section 5.3 (based on section 4.4), the environmental data can be used to determine the reduction in GHG emissions based on different energy access implementations, as defined by Shrestha & Acharya (2015).

5.3 Assessment calculations

For the applicable indicators, the values were calculated based on the equations directly or indirectly originating from the SEAP framework. Moreover, the input energy data measured from the experiment: (i) was about a quarter (99 days) of the annual range typically used for these indicators (section 4.3), which according to Saunders et al. (2007: 212) falls within an acceptable 1% margin of error; (ii) was collected from only one HH, thus used as an average of HHs.

From the results, shown in Table 5.4, the following relevant aspects are discussed:

- The average result from Equation 4.4 was applied for a year, equal to a value of approximately 16 kWh, which was within MTF tier-1 for task lighting and phone charging as per load used for the experiment.
- Niger's number of rural HHs with access to electricity (Table 4.1) was considered as the sample for the energy demand calculation; moving these HHs to MTF tier-2 required additional energy of nearly 15 GWh (Equation 4.8).
- Based on the result from Equation 4.10, electricity was available for 97% of the considered supply period.
- The consumption difference, based on the change from an incandescent to a more efficient LED bulb (Figure 5.1), indicated:

- Based on the periods before and after the change (February and March), a daily and monthly reduction by 47 Wh (Equation 4.13) and 246 Wh (Equation 4.14) respectively.
- From an MDMS-simulated tariff and bill, applied to the energy consumption for the months before and after the change (Figure 5.3), a monthly saving of 8.37 R on the HH bill (Equation 4.15).

- Considering South Africa’s grid emission factor of 0.928 kg CO2 per kWh (Carbon Footprint, 2020), the monthly reduction in consumption resulting from the use of a more efficient end-device saw a monthly decrease in carbon emissions of 228.29 kg CO_2 (Equation 4.16); furthermore, a daily comparative breakdown was extracted from the MDMS based on the same factor and consumption values, as indicated in Table 5.5.

Table 5.4: Calculation results for the applicable assessments.

Indicator	Applicable assessment	HH energy data input (SM)	Equation	Main parameter and unit	Result
Basic minimum energy requirement	EP, ED	Load profile, maximum demand	4.4	$\overline{E_{HE}}$, kWh	4.08
			4.7	$\overline{E_{HD}}$, kWh	37.50
Energy demand	EP, ED	Load profile, maximum demand	4.8	ΔE_{DP}, GWh	14.83
Hours of electricity supply	SUS, AFF	Load profile	4.10	$\overline{t_S}$, h	2 308.00
			4.12	$\overline{t_O}$, h	68.00
End-use device efficiency	ED, CST, SUS, BEN	Load profile, billing registers	4.13	ΔE_{de}, Wh	47.00
			4.14	ΔE_{pe}, Wh	246.00
			4.15	ΔC, R	8.37
GHG emissions	RSC, CST, SUS, BEN	Load profile	4.16	ΔG, kg CO_2	0.044

Calculating the minimum electricity threshold ($\overline{E_{HE}}$) using either 30-min consumption data or the monthly maximum demand billing registers ensured higher granularity and accuracy, when with non-metered data it would not have been possible. Furthermore, it was also noticed that the calculation of the same indicator using demand data ($\overline{E_{HD}}$) was about 9 times higher than that based on consumption; now monthly MD values look at demand peaks, which the supply energy system should be able to cater for; thus if such criteria are considered then the calculation could be used as an indicator for supply requirement (adequacy).

Indicators such as $\overline{t_S}$, complemented by logged events such as those indicated in section 5.2.1, particularly indicated when supply was available or not, and for how long. Moreover, with the measured energy data, improvements in end-device efficiency were quantified in consumption, cost, and carbon emissions.

Provisional Bill

Customer: TEST_ACC_M.E. **Document date:** 2022-08-14 18:56

Meter Account: ACC62060014

Period: From 2022-03-01 00:00:00.000 to 2022-04-01 00:00:00.000

Tariff: CAM_ALL

Meter Totals: **62060014** (SL7000) Start reading: 3.193kWh End reading: 3.818kWh

34155 Electricity: Consumption total: 0kWh

Tariff	Description	Units	Rate[R]	Amount
Availability Charge [Amp/month]		1month	8.4100	R8.41
Energy charge [R/kWh]	TEST_MTR_M.E., 62060014	0.625kWh	1.4328	R0.89
Demand Charge [R/kVA]	1.000pf 2022-03-10 22:00	0.014kVA	138.7100	R1.94
Energy charge exported [R/kWh]	TEST_MTR_M.E., 62060014	0kWh	0.9543	R0.00
			Sub Total:	R11.24

Total before VAT: R11.24
VAT(15.0%): R1.69
Total: R12.93

Provisional Bill

Customer: TEST_ACC_M.E. **Document date:** 2022-08-14 19:03

Meter Account: ACC62060014

Period: From 2022-02-01 00:00:00.000 to 2022-03-01 00:00:00.000

Tariff: CAM_ALL

Meter Totals: **62060014** (SL7000) Start reading: 2.322kWh End reading: 3.193kWh

34155 Electricity: Consumption total: 0kWh

Tariff	Description	Units	Rate[R]	Amount
Availability Charge [Amp/month]		1month	8.4100	R8.41
Energy charge [R/kWh]	TEST_MTR_M.E., 62060014	0.871kWh	1.4328	R1.24
Demand Charge [R/kVA]	1.000pf 2022-02-01 21:30	0.064kVA	138.7100	R8.87
Energy charge exported [R/kWh]	TEST_MTR_M.E., 62060014	0kWh	0.9543	R0.00
			Sub Total:	R18.52

Total before VAT: R18.52
VAT(15.0%): R2.78
Total: R21.30

Figure 5.3: MDMS simulated household bills with and without an energy-efficient device.

Table 5.5: MDMS-generated daily savings report on usage, costs, and carbon emissions.

Parameter	Period 1 2022-02-01 00:00:00.000 to 2022-03-01 00:00:00.000	Period 2 2022-03-01 00:00:00.000 to 2022-04-01 00:00:00.000	Difference
Energy (kWh)	0.031	0.02	0.011
Max. Demand (kVA)	0.064	0.014	0.05
Money (R)	0.662	0.363	0.299
Carbon (kg)	0.029	0.019	0.01

5.4 Predicted demand

The ML exercise by NNR model (Figure 4.13) was applied to predict the demand values for the period of 01/04/2021 00:00 - 01/01/2022 00:00, based on temperature, wind, and humidity. The output model in Figure 5.4 and sample of predicted data in Table 5.6

shows the process and data, where it was noticed that not only transformations were applied to the input data (features) to improve predictive power, but that some of the predictive values were negative; as there can be no negative demand, they were replaced by zeros for the final dataset.

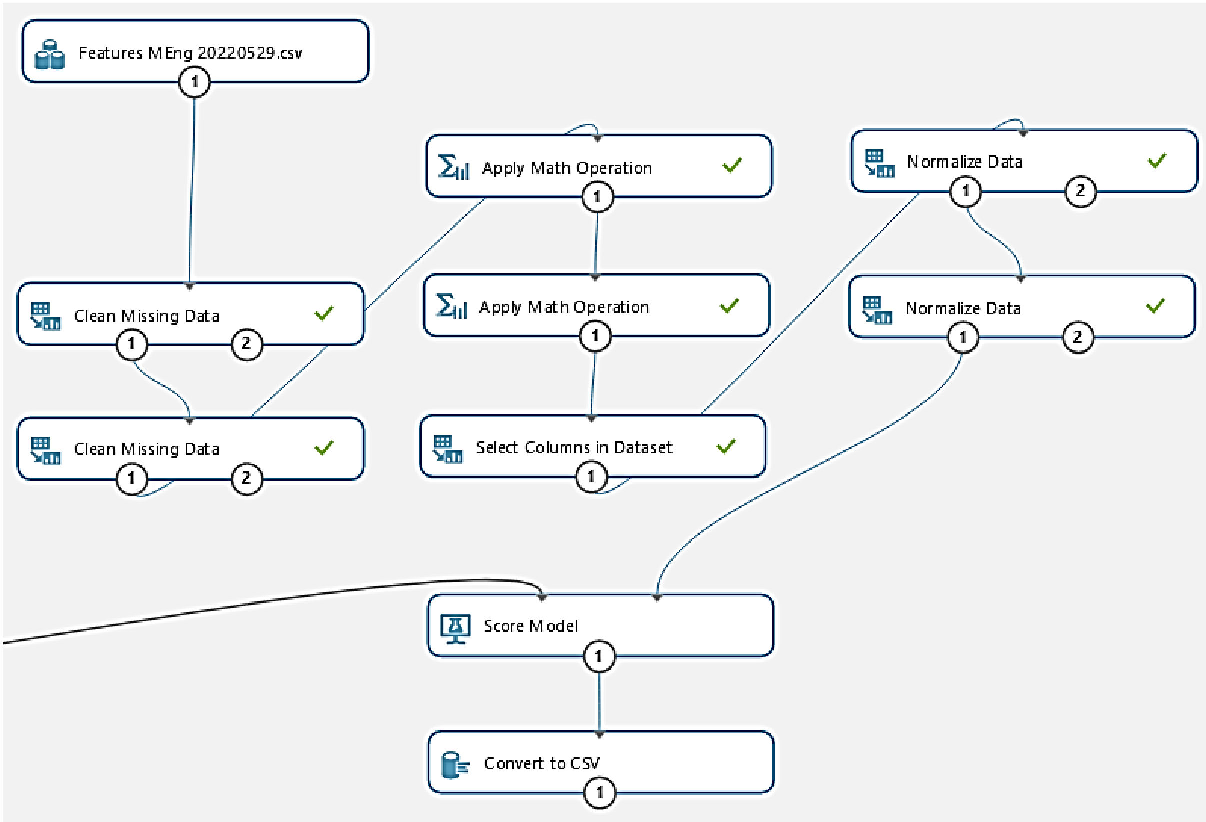


Figure 5.4: Predicting demand data with the built ML model.

Table 5.6: Sample of predicted data (scored labels).

Temperature	Sqrt(Wind)	LnPlus1(Humidity)	Scored Labels
-0.855093474	0.293972368	0.498921986	0.669818521
-1.049094918	0.293972368	0.498921986	1.927718997
-1.049094918	0.333333333	0.476092568	-0.176963925
-1.243096362	0.368513866	0.476092568	0.534485459
-1.243096362	0.368513866	0.476092568	0.534485459
-1.437097806	0.368513866	0.498921986	0.554687619
-1.631099251	0.368513866	0.498921986	-0.192573905
-1.631099251	0.368513866	0.498921986	-0.192573905
-1.243096362	0.368513866	0.476092568	0.534485459
-0.467090585	0.521157307	0.421637992	0.542003989
0.114913747	0.587944736	0.388236767	0.576710582

The predicted data was then combined with the input primary data (January to March 2021) for the final dataset covering the full year, which can be seen graphically in Figure 5.5. The initial spike in January related to the use of a higher load (incandescent bulb), then changed to a lower load in early February (LED bulb); these stages were part of the experiment on end-use device efficiency for affordability and sustainability, as per section 67. Also observed with the predicted data was the increase in consumption during the Winter and Autumn seasons (July to November), typical for annual usage patterns.

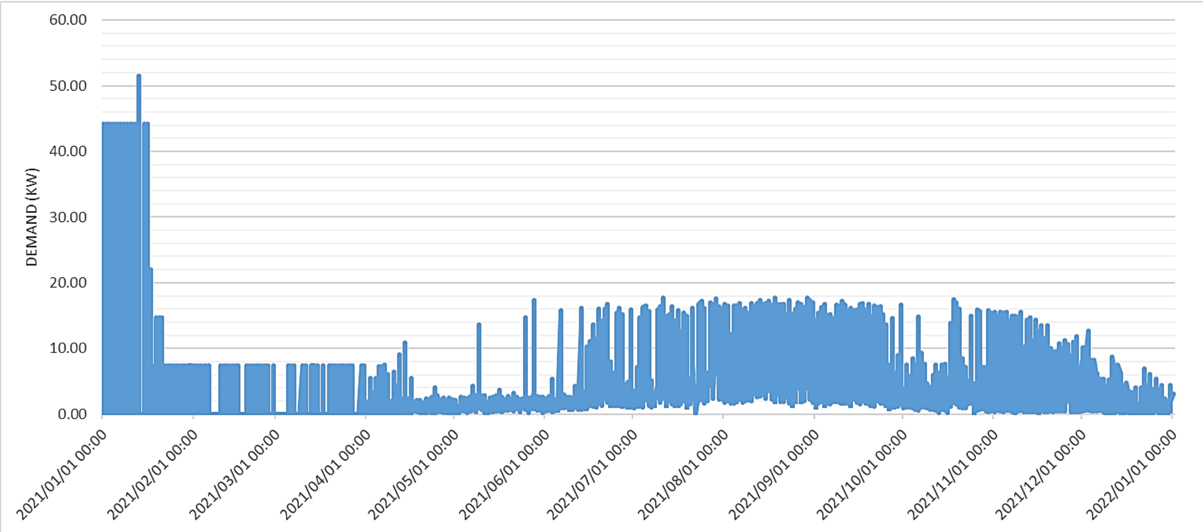


Figure 5.5: Full-year demand dataset (primary and predicted).

5.5 Energy system optimization and cost analysis

The full dataset resulting from the meter and prediction exercises was then used as load input to HOMER Pro to perform a resource and cost analysis of an energy system (ES) to implement for energy access, in line with SEAP frameworks' resource, cost, benefit, affordability and sustainability assessments. The results in Table 5.7 compare the base energy system supplying electricity via diesel fuel-based generator, with the optimized systems:

- Base versus best cases: the optimized energy systems reduce operating costs by 83% and 78%, and electricity costs by 35% and 67%, respectively. The hybrid renewable energy system (HRES) option provided an off-grid supply system composed of non-renewable (Gen) and renewable (PV), the latter responsible for nearly 31% of the supply
- Renewable-only versus HRES: the PV-only scenario was explored to assess the pros and cons of a fully renewable energy system; the environmental benefit is

evident, with no carbon emissions, no fuel consumption costs and 26% less in O&M costs, but all these come at a higher investment and an electricity cost about double that of the combined option. Thus, if considered, a carbon-free, renewable system is a clean but costly option, with a payback period of 4 to 5 years (4 times longer than the HRES).

Table 5.7: Results on base system against HRES and full PV systems

Metrics	Unit	Base system: Gen	Best-case system (optimized)	
			Renewable-only: PV + 1 kWh LA + Converter	HRES: PV + Gen + 1 kWh LA + Converter
Net Present Cost (NPC)	R	9.04M	5.69M	3.02M
Capital	R	446 310.00	4.28M	1.11M
O&M	R/year	1.01M	167 058	225 317.00
Levelized Cost of Electricity (LCOE)	R/kWh	38.18	24.78	12.77
Internal rate of return (IRR)	%	-	22.20	109
Discounted payback	Year	-	4.65	1.05
Simple payback	Year	-	4.28	1.01
Renewable fraction	%	-	100.00	31.10
CO_2 emission	kg/year	108 665.00	0.00	20 805.00
Fuel Consumption	L/year	41 513.00	0.00	7 948.00

Furthermore, on the HRES there were periods of the year with carbon emissions close to none; Figure 5.6 illustrates a cleaner supply of electricity during Summer/Autumn, with nearly no use of fossil-fuel-based power generation.

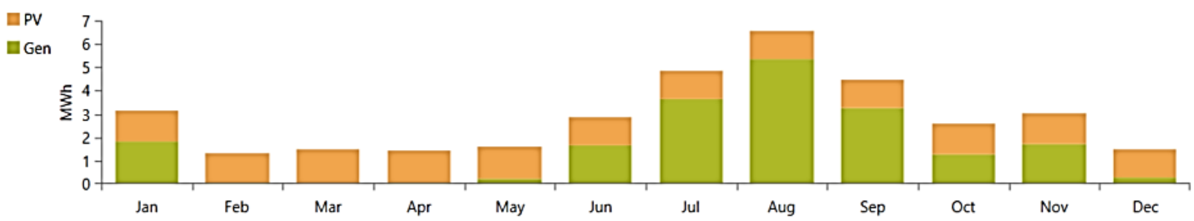


Figure 5.6: Monthly electric production

In the same scenario, fuel (resources) and storage are to be considered in terms of recurrent costs of electricity supply that are high in the mix, where the fuel costs over the EAP's lifetime were almost as high as the initial capital investment for all components combined, as indicated in Figure 5.7.

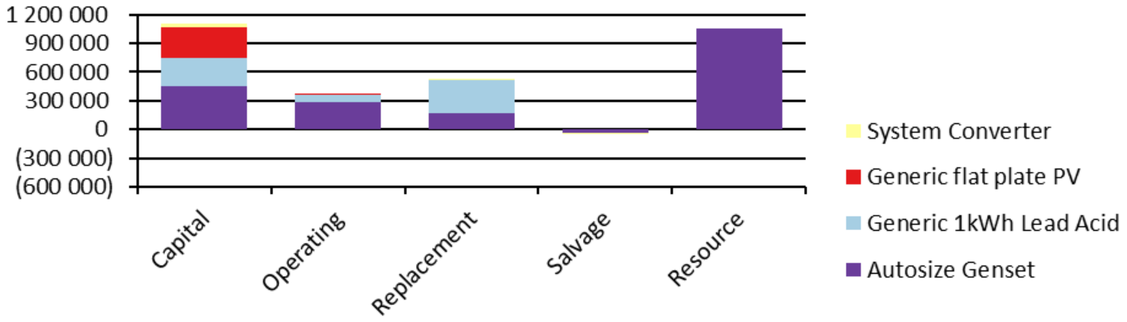


Figure 5.7: Cost breakdown, in USD.

5.6 Summary

This chapter presented and discussed the experiment results. Accessing the smart meter locally using an optical probe allowed for the initial configuration of the meter (parameterize the device to collect the intended household data) and extraction of data that would not be possible by the MDMS alone, either due to data size, the capability for reading it, or remote communication issues; such data included billing register, real-time active and reactive power, alarms, and date/&time synchronization. With the MDMS on the other hand, emphasis was on the scheduled, remote data acquisition as well as the MDMSs data processing, visualization and reporting. Analysis of daily and more granular half-hourly consumption provided usage patterns that included reduced consumption due to improved efficiency (incandescent to LED load replacement), consumption variance related to cell phone charging, higher demand during evenings, and environmental impact of the usage pattern (GHG emissions); all these aspects were relevant to the accessibility, affordability, and reliability components of the SEAP framework.

The calculated SEAP indicators, based on the household energy data collected and applied to the defined formulas, quantified aspects of the seven assessments of the framework: energy poverty, energy demand, sustainability, cost, benefit, resource, and affordability; the differentiator was using actual energy data measured from the household, in contrast to estimated or old data.

The neural network model trained from the measured household energy data was then applied to predict the remaining data based on weather input (temperature, wind, humidity). The result was demand data for a full year, combining the 3-month initial dataset and the 9-month predicted dataset; within the energy access planning context, weather data from a similar area could equally be applied to predict energy data applicable to that same area. Furthermore, the full-year dataset was required and applied to the simulation on energy system optimization and cost sensitivity analysis. The best-case scenario was an HRES of solar, genset, storage and converter, with the lowest NPC payback period and LCOE, however with a third of renewable energy penetration. Thus, a second best-case was considered for a fully renewable scenario (no genset), which had in contrast 100% penetration and zero emissions, but at almost double the NPC and LCOE, and a payback period nearly 4 times longer. These scenarios, based on smart metering energy data, present cases for analysis of costs versus sustainability and benefit, within the framework’s assessments.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusion

In this study, data-acquisition features of smart metering technologies were coupled with energy access (SEAP) methods through energy data. Using input data acquired from a selected country in SSA, a smart metering system was assembled, with a smart meter measuring electricity usage data and an MDMS wirelessly collecting it. The applicability of this data to the energy access framework was then analysed against its assessments and explored by applying it to different, but interdependent applications that included the calculation of indicators using the energy access framework as the baseline, data prediction (demand forecasting) via machine learning exercise to build a full year dataset from the available meter data, and energy systems optimization and cost sensitivity analysis by applying the full dataset (measured plus predicted) to simulation software.

The analysis and exploration of household energy data from smart metering through the SEAP framework highlighted the following applicability aspects: (i) usage patterns identified from interval-based consumption data that indicate how the household uses electricity, used to make electricity more affordable over improved tariffs or increased efficiency of household appliances/devices – demand (ED), affordability (AFF) and cost (CST); (ii) energy and local supply data that translates into environmental indicators such as carbon emissions, quantifying the impact of energy supply technologies to the environment – benefit (BEN) and sustainability (SUS); (iii) real-time power usage and alarms (e.g., voltage cuts, sags and swells, per phase) not only to monitor the electricity supply (QoS) but also to identify issues and take preventive action on the implemented energy system – resource (RSC) and sustainability (SUS); (iv) interval, register and/or alarm data used to directly (calculations) or indirectly (input for analysis) determine energy access indicators (e.g., basic minimum energy requirement, hours of electricity supply, GHG emissions, etc.), continuous and accurate as measured from the household – applicable to all assessments; (v) register and interval-based data used as proxy demand profile for completely unelectrified areas with similar demand profiles – demand (ED); (vi) demand prediction from interval-based and geographic area data (such as weather), of advantage for demand forecasting (additional data based on features from the same area), for new demand profile (non-electrified area based on features from that area), and for proactive action (demand planning, preventive maintenance, financial and tariff planning) – applicable to all assessments; (vii) daily load profile from interval-based and geographic area data (such as energy resources: solar, wind, etc.) to use for the optimization and cost analysis of energy systems, including microgrids and renewables – resource (RSC), cost (CST), benefit (BEN), affordability (AFF) and sustainability (SUS).

With the contribution provided in this study, energy planners not only can effectively extract, interpret, and use real energy data toward implementing and sustaining an optimal energy system for the energy poor, but also benefit from aspects such as data acquisition on-demand, building big data with household usage that is systematically collected and stored in a database, monitoring the implemented energy system and proactively acting on issues identified by events and alarms, tracking changes in behaviour and implemented measures (e.g. energy efficiency initiatives) through energy data, and reducing costs of physical presence for data collection. It is a ripple effect that will be valuable to the EAP, implementation, and tracking, but ultimately and most importantly, to those who need access to modern energy.

6.2 Further work and recommendations

The results of this thesis have the potential to be extended in several directions, namely: Limitations in wireless communications, building local know-how, Increasing the accuracy of forecasted demand data, applying demand profile data to non-electrified areas, and assessing the cost-benefit of fully renewable versus HRESs. These issues are briefly discussed below.

Limitations in wireless communications:

While continuous and accurate energy data collection will help build a database for reference and forecasting, the remote communications component of smart metering technologies plays a key role in access to the smart meter; the best technology, applicable to the reality and available services of the area where energy access is to be implemented, are aspects that must be explored.

Building local know-how:

The smart meter not only needs to be installed at the household or distribution point but also requires maintenance (preventive and corrective); to ensure the sustainability of both energy system and on-site smart metering technologies such work implies giving local personnel the required know-how; furthermore, adding knowledge-transfer contributes to the social, empowering aspect of a community. Consideration should thus be given to the incorporation of smart metering technologies and local empowerment into the energy access framework and planning.

Increasing accuracy of forecasted demand data:

Applying different sets of ML models and other methods not only to improve predictive power but also to cater for the uncertainty on the availability of renewable energy supply, such as the cases of solar and wind.

Applying demand profile data to non-electrified areas:

Analysing how existing household demand data obtained from electrified areas, using smart metering, could be applied to non-electrified areas that are similar.

Assessing the cost-benefit of fully renewable versus HRESs:

Ensuring electricity access to the energy-poor is looking at energy access programs that are not only cost-effective but also sustainable and beneficial, e.g., improved health, reduced air pollution, increased energy security, etc.

REFERENCES

ADD Grup. 2020. COVID-19 a trigger for smart metering rollouts. https://www.smart-energy.com/industry-sectors/smart-meters/covid-19-a-trigger-for-smart-metering-rollouts/ [9 April 2022].

Ali@gwc.org.uk. 2004. CurrentTransformers. https://commons.wikimedia.org/wiki/File:CurrentTransformers.jpg. CC BY-SA 2.5 [23 July 2022].

Ali@gwc.org.uk. 2005. ElectricityMeterMechanism. https://commons.wikimedia.org/wiki/File:ElectricityMeterMechanism.jpg. CC BY-SA 2.5 [24 July 2022].

Asian Development Bank. 2018a. *Smart Metering Roadmap For Nepal*. Manila: Asian Development Bank.

Asian Development Bank. 2018b. *Sustainable Energy Access Planning: A Case Study*. Manila: Asian Development Bank.

Azini, A.S., Kamarudin, M.R. & Jusoh, M. 2015. Transparent antenna for WiFi application: RSSI and throughput performances at ISM 2.4 GHz. https://www.researchgate.net/figure/Wireless-channel-over-RSSI-range-with-solar-source_tbl2_277581008.

Bajer, M. 2019. Building Advanced Metering Infrastructure using Elasticsearch database and IEC 62056-21 protocol. https://www.researchgate.net/publication/333248821_Building_Advanced_Metering_Infrastructure_using_Elasticsearch_database_and_IEC_62056-21_protocol.

Barai, G.R., Krishnan, S. & Venkatesh, B. 2016. Smart metering and functionalities of smart meters in smart grid - A review. *2015 IEEE Electrical Power and Energy Conference: Smarter Resilient Power Systems, EPEC 2015*: 138–145.

Benedictine University Library. 2021. Public Health Research Guide: Primary & Secondary Data Definitions. https://researchguides.ben.edu/c.php?g=282050&p=4036581 [14 April 2021].

Carbon Footprint. 2020. *Country Specific Electricity Grid Greenhouse Gas Emission Factors*. https://www.carbonfootprint.com/docs/2020_06_emissions_factors_sources_for_2020_electricity_v1_1.pdf.

Casals, L.C., Herrero, S.T., Barbero, M. & Corchero, C. 2020. Smart meters tackling energy poverty mitigation: Uses, risks and approaches. *2020 IEEE Electric Power and Energy Conference, EPEC 2020*, 3.

Central Intelligence Agency. 2022. The World Factbook. https://www.cia.gov/the-world-factbook/countries/niger/#geography [13 February 2022].

Chakerian, S. 2021. Emerging markets poised to achieve smart metering despite COVID-19. https://www.smart-energy.com/industry-sectors/smart-meters/emerging-markets-poised-to-achieve-smart-meter-deployment-despite-covid-19/ [9 April 2022].

Desai, S. 2015. *Handbook of Energy Audit*. New Delhi: McGraw Hill Education.

ECOWAS Regional Centre for Renewable Energy and Energy Efficiency. 2018. Inauguration of a Solar Mini-Grid Project in Niger. http://www.ecreee.org/news/inauguration-solar-mini-grid-project-niger [27 April 2022].

Freemeteo. 2022. Gorou - Daily weather history. https://ke.freemeteo.com/weather/ [30 April 2022].

García, S., Parejo, A., Personal, E., Guerrero, J.I., Biscarri, F. & Léon, C. 2021. A retrospective analysis of the impact of the COVID-19 restrictions on energy consumption at a disaggregated level. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7843085/.

Hasan, M. & Ibrahim, A. 2020. Design & Development of Secure Smart Metering Systems: Practical Guide and Detailed Study.

HOMER Energy. 2017. HOMER Pro 3.14. https://www.homerenergy.com/products/pro/docs/3.14/index.html [25 April 2022].

International Electrotechnical Commission. 2002. INTERNATIONAL STANDARD - IEC 62056-21 Electricity metering – Data exchange for meter reading, tariff and load control – Part 21: Direct local data exchange.

International Energy Agency. 2019. Smart meter deployment in key markets, 2013-2017. https://www.iea.org/data-and-statistics/charts/smart-meter-deployment-in-key-markets-2013-2017 [9 April 2022].

International Energy Agency & The World Bank. 2014. *Sustainable Energy for All 2013-2014: Global Tracking Framework Report*. Washington, DC.

Itron. 2012. ACE SL7000: Industrial & Substation Electricity Meter Range. https://www.itron.com/-/media/feature/products/documents/brochure/ace-sl7000el-0019-1en0212.pdf.

Itron. 2019. Sparklet 3G Modem: Commercial and Industrial. https://www.itron.com/-/media/feature/products/documents/brochure/brochure-sparklet-2-rel3-3g_dec2019-v2-eng.pdf.

Karch, J. 2020. Improving on Adjusted R-Squared. , 6: 1–11.

Kibelka, C. 2013. 13. Dezember (11352026703). https://commons.wikimedia.org/wiki/File:13._Dezember_(11352026703).jpg. CC BY-SA 3.0 [23 July 2022].

Lambert, J. 2021. Microsoft Digital Defense Report shares new insights on nation-state attacks. https://www.microsoft.com/security/blog/2021/10/25/microsoft-digital-defense-report-shares-new-insights-on-nation-state-attacks/ [9 April 2022].

Malik, H. & Iqbal, A. 2021. Intelligent Data Analytics for Time-Se- ries Load Forecasting Using Fuzzy Rein- forcement Learning (FRL).

Mapondera, F.A. 2015. Landis+Gyr E450 G3 1-Phase Smart Electricity Meter. https://commons.wikimedia.org/wiki/File:Landis%2BGyr_E450_G3_1-Phase_Smart_Electricity_Meter.jpg. CC BY-SA 3.0 [23 July 2022].

Microsoft. Big Data with Azure Machine Learning. https://raw.githubusercontent.com/MicrosoftLearning/Developing-Big-Data-Solutions-with-Azure-Machine-Learning/master/Labs/DAT228x-Lab2-Training Predictive Models.pdf [14 May 2022].

Microsoft. 2019. Evaluate Model. https://docs.microsoft.com/en-us/previous-versions/azure/machine-learning/studio-module-reference/evaluate-model?redirectedfrom=MSDN [15 May 2022].

Nettleton, D. 2014. Selection of Variables and Factor Derivation. *Commercial Data Mining*: 79–104.

Nti, I.K., Teimeh, M., Boateng, O.N. & Adekoya, A.F. 2020. Electricity load forecasting : a systematic review. *Journal of Electrical Systems and Information Technology*, 8. https://doi.org/10.1186/s43067-020-00021-8.

Plan International España. 2017. Access to electricity in Gorou, Niger. https://www.youtube.com/watch?v=7bAPLyPIe-I&list=PL3YbsrlfhsraDjYargV0oTfZM9GYOVPEH&index=3 [25 April 2022].

RODALCO. 2012a. AEG hybrid electromechanical meter with electronic maximum demand readings. https://commons.wikimedia.org/wiki/File:AEG_hybrid_electromechanical_meter_with_electronic_maximum_demand_readings.JPG. CC BY-SA 3.0 [24 July 2022].

RODALCO. 2012b. Landis & Gyr 2 rate electromechanical meter. Type YL11d. https://commons.wikimedia.org/wiki/File:Landis_%26_Gyr_2_rate_electromechanical_meter._Type_YL11d.JPG. CC BY-SA 3.0 [24 July 2022].

Saunders, M., Lewis, P. & Thornhill, A. 2007. *Research Methods for Business Students*. 4th ed. Harlow: Pearson.

Shrestha, R.M. & Acharya, J.S. 2015. *Sustainable Energy Access Planning: A Framework*. Manila: Asian Development Bank.

Sun, H., Hatziargyriou, N., Poor, H.V., Carpanini, L. & Fornié, M.A.S. 2016. *Smarter Energy: From Smart Metering to the Smart Grid*. London: The Institution of Engineering and Technology.

Tespro. 2022. TP-USB Optical Probe for Meter Communication. http://www.tespro.hk/en/TP-USB.php [27 March 2022].

The United Nations. 2019. Household Size and Composition. https://population.un.org/Household/index.html#/countries/562 [13 February 2022].

The World Bank. 2022a. Access to electricity, rural (% of rural population) - Niger. https://data.worldbank.org/indicator/EG.ELC.ACCS.RU.ZS?locations=NE&type=shaded&year=2014 [14 February 2022].

The World Bank. 2021. Access to electricity (% of population). https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=ZG [24 January 2022].

The World Bank. 2015. Beyond Connections: Energy Access Redefined. : 1–244.

The World Bank. 2022b. Electric power consumption (kWh per capita) - Sub-Saharan Africa (excluding high income). https://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC?end=2014&locations=ZF&start=1971&type=shaded&view=map [13 February 2022].

The World Bank. 2018. *International Development Association Project Paper on a Proposed Additional Credit in The Amount of Eur 48.0 Million (US$56 Million Equivalent) and a Proposed Additional Grant in The Amount of SDR 9.9 Million (US$14 Million Equivalent) to the Republic of*.

Toledo, F. 2013. *Smart Metering Handbook*. Tulsa: PennWell.

Toncich, D.J. 1999. *Key Factors in Postgraduate Research - A Guide for Students*.

Weranga, K.S.K., Kumarawadu, S. & Chandima, D.P. 2014. *Smart Metering Design and Applications*. Singapore: Springer.

Wordtwist. 2013. Instrument Transformer. https://commons.wikimedia.org/wiki/File:Instrument_Transformer.jpg. CC BY-SA 3.0 [23 July 2022].

XE. 2022. 1 USD to ZAR - Convert US Dollars to South African Rand. https://www.xe.com/currencyconverter/convert/?Amount=1&From=USD&To=ZAR [25 April 2022].

Zest AI. 2018. 6 Methods for Dealing with Missing Data. https://www.zest.ai/insights/6-methods-for-dealing-with-missing-data [14 May 2022].

Živi, N.S., Ur-rehman, O., Ruland, C., Member, S., Meter, A.M. & Mmr, R. 2015. Evolution of Smart Metering Systems. : 635–638.

Zuzu. 2008. Elster Type R15 electricity meter. https://commons.wikimedia.org/wiki/File:Elster_Type_R15_electricity_meter.jpeg. CC BY-SA 3.0 [23 July 2022].

APPENDICES

Appendix A: Types of electronic meters and grid connections

**Figure A.1: Electronic meter types and related components, left-to-right: single module smart meter, plug-in modular prepayment meter, and socket modular meter (Zuzu, 2008; Kibelka, 2013; Mapondera, 2015). **

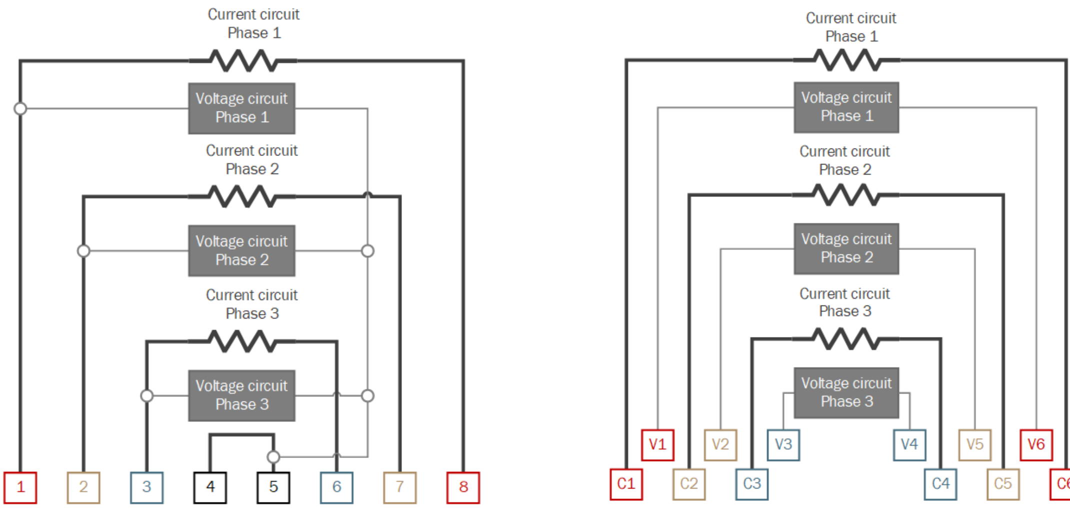


Figure A.2: Connection to the grid, left-to-right: Direct meter circuit, and indirect meter circuit, current transformers, voltage transformers (based on Toledo (2013)

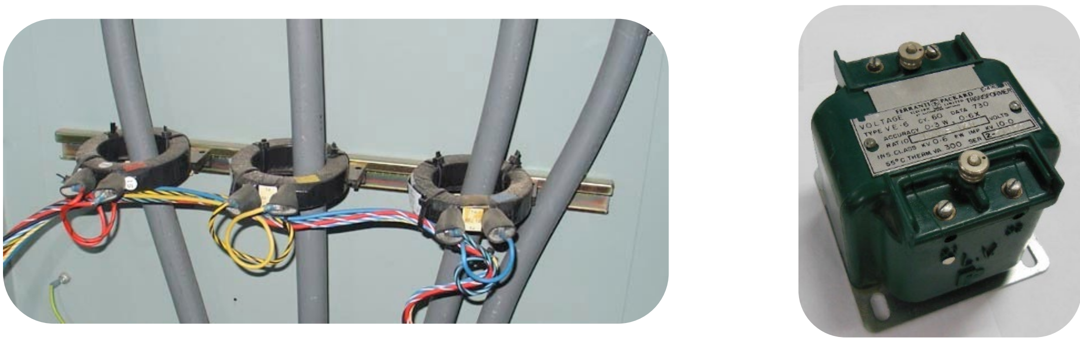


**Figure A.3: Instrument transformers, left-to-right: current transformer, and voltage transformer (Ali@gwc.org.uk, 2004; Wordtwist, 2013). **

Appendix B: Python commands used for data analysis

```
#!/usr/bin/env python
# coding: utf-8

# In[1]:


import pandas as pd
from sklearn import preprocessing
import sklearn.model_selection as ms
from sklearn import linear_model
import sklearn.metrics as sklm
import numpy as np
import numpy.random as nr
import matplotlib.pyplot as plt
import seaborn as sns
import scipy.stats as ss
import math
from tabulate import tabulate

get_ipython().run_line_magic('matplotlib',
'inline')


# In[2]:


#Import dataframe

df =
pd.read_csv(r"C:\Users\Bbacar\Desktop\Pyt
hon files\Dataset.csv", sep=',')


# In[3]:


#Show imported data

df.head(10)


# In[4]:


print(df.info(),'\n')
print('>Unique entries (for duplicates):
',df.UID.unique().shape, '\n')


# In[5]:


#Define features

num_col=['Temperature','Wind','Humidity','D
emand']

num_col_f=['Temperature','Wind','Humidity']

cat_col=['Power_cut']

# In[6]:


#Summary statistics (median = 50% quartile)
#Use transpose() or T to transpose.

df[num_col].describe().round(4).T


# In[7]:


# For Dissertation

median1 = df['Demand'].median()
mean1 = df['Demand'].mean()

font = {'family' : 'arial',
        'weight' : 'normal',
        'size'   : 12}

def plot_density_hist(df, cols):
    for col in cols:
        sns.set_style("darkgrid")
        fig = plt.figure(figsize=(10,4)) # define
plot area
        ax = fig.gca() # define axis
        sns.distplot(df[col], ax = ax, bins = 10,
rug=False, hist = True, kde = False,
norm_hist = False)
        #Vertical lines and labels
        plt.axvline(x=median1, linewidth=1,
color='r')
        plt.text(median1-
2.5,2100,'0.00',fontsize=12, color = 'r')
        plt.axvline(x=mean1, linewidth=1,
color='b')
        plt.text(mean1,2100,'1.45',fontsize=12,
color = 'b')
        #Axis labels
        #plt.title('Histogram of ' + col) # Give the
plot a main title
        plt.xlabel('Demand',fontsize=15) # Set
text for the x axis
        plt.ylabel('Frequency',fontsize=15)# Set
text for y axis
        #font size for axis labels and legend
        plt.rc('font', **font)
        #Axis legend
        plt.legend(('Median', 'Mean',
'Frequency'), title = 'Legend', frameon = True,
borderpad = True)#frameon = True,
framealpha =120
        #plb.plot(x=, label='Frequency')
        #pylab.plot(x, y2, '-r', label='cosine')
        #plb.legend(loc='upper right')
        #plt.yscale('log')
        plt.show()

plot_density_hist(df, ['Demand'])
```

In[8]:

```
def plot_density_hist(df, cols):
   for col in cols:
       sns.set_style("darkgrid")
       fig = plt.figure(figsize=(10,2)) # define
plot area
       ax = fig.gca() # define axis
       sns.distplot(df[col], ax = ax, bins = 25,
rug=False, hist = True, kde = True)
       plt.title('Histogram of ' + col) # Give the
plot a main title
       plt.xlabel(col) # Set text for the x axis
       plt.ylabel('Density')# Set text for y axis
       plt.rc('font', **font)
       #plt.xscale('log')
       plt.show()

plot_density_hist(df, num_col)
```

Notes from histograms: right-skewed. Needs transformation (log, ln, sqrt, etc)
#
[log(x) for natural (base-e) of x, log2(x) for base-2 of x, log10(x) for base-10 of x, log1p(x) for natural (base-e) of (x+1)]
#
After different trials, sqrt and log1p were the ones that improved some of the distributions (other logs didn't like the zeros):

In[9]:

```
#Temperature - no transformation required

def plot_density_hist(df, cols):
   for col in cols:
       sns.set_style("darkgrid")
       fig = plt.figure(figsize=(10,2)) # define
plot area
       ax = fig.gca() # define axis
       sns.distplot(df[col], ax = ax, bins = 25,
rug=False, hist = True, kde = True)
       plt.title('Histogram of ' + col) # Give the
plot a main title
       plt.xlabel(col) # Set text for the x axis
       plt.ylabel('Density')# Set text for y axis
       plt.rc('font', **font)
       #plt.xscale('log')
       plt.show()

plot_density_hist(df, ['Temperature',
'Demand'])
```

In[10]:

```
#Wind transformation - sqrt

df_TR_wind = 0

df_TR_wind =
df[num_col].applymap(math.sqrt)

def plot_density_hist(df_TR_wind,
cols):
   for col in cols:
       sns.set_style("darkgrid")
       fig = plt.figure(figsize=(10,2)) #
define plot area
       ax = fig.gca() # define axis
       sns.distplot(df_TR_wind[col], ax =
ax, bins = 25, rug=False, hist = True,
kde = True)
       plt.title('Histogram of ' + col) #
Give the plot a main title
       plt.xlabel(col) # Set text for the x
axis
       plt.ylabel('Density')# Set text for y
axis
       #plt.xscale('log')
       plt.show()

plot_density_hist(df_TR_wind, ['Wind',
'Demand'])
```

In[11]:

```
#Wind transformation - log1p

df_TR_hum = 0

df_TR_hum = np.log1p(df[num_col])

def plot_density_hist(df_TR_hum, cols):
   for col in cols:
       sns.set_style("darkgrid")
       fig = plt.figure(figsize=(10,2)) # define
plot area
       ax = fig.gca() # define axis
       sns.distplot(df_TR_hum[col], ax = ax,
bins = 25, rug=False, hist = True, kde = True)
       plt.title('Histogram of ' + col) # Give the
plot a main title
       plt.xlabel(col) # Set text for the x axis
       plt.ylabel('Density')# Set text for y axis
       #plt.xscale('log')
       plt.show()

plot_density_hist(df_TR_hum, ['Humidity',
'Demand'])
```

```
# In[12]:


#Build transformed dataframe for analysis

#Convert to array
col_temp = df.Temperature.to_numpy()
col_tr_wind = df_TR_wind.Wind.to_numpy()
col_tr_hum =
df_TR_hum.Humidity.to_numpy()
col_tr_demand = df.Demand.to_numpy()

#Confert to dataframe
df_TR =
pd.DataFrame({'Temperature':col_temp,
'TR_Wind':col_tr_wind,
'TR_Humidity':col_tr_hum,
'TR_Demand':col_tr_demand})
df_TR


# In[13]:


sns.pairplot(df[num_col], height = 1.5, aspect
= 2)
plt.show()


# In[14]:


#With transformed features

sns.pairplot(df_TR, height = 1.5, aspect = 2)
plt.show()


#                    Confirm using
Cohen guidelines (Pearson method):
#
# | # | Correlation | Meaning |
# |---|--------------|------------|
# | 1 | 0.0 - 0.1 | Negligible |
# | 2 | 0.1 - 0.3 | Small |
# | 3 | 0.3 - 0.5 | Medium |
# | 4 | 0.5 + | Large |

# In[15]:


corr = df[num_col].corr(method =
'pearson').round(2)
corr


# In[16]:


corr_tr = df_TR.corr(method = 'pearson').round(2)
corr_tr
# No significant correlation of any of the
features with the label.
#
# Temperature VS Wind: MEDIUM (positive)
#
# Temperature VS Humidity: LARGE
(negative)
#
# Wind VS Humidity: MEDIUM-LARGE
(negative)
#
# The above shows inter-feature correlation
that varies between 0.3 and 0.7, therefore
confirming a medium to large correlation
(>=0.3)

# #Correlation heatmap - transformed VS
untransformed:
#
# sns.heatmap(corr_tr, cmap='RdBu')
# plt.title('Correlation matrix (transformed
features)')
# plt.yticks(rotation='horizontal')
# plt.xticks(rotation='vertical')
# plt.show()
#
# sns.heatmap(corr, cmap='RdBu')
# #plt.title('Correlation matrix (untransformed
features)')
# plt.yticks(rotation='horizontal')
# plt.xticks(rotation='vertical')
# plt.show()

# In[17]:


x2 = np.array(df['Temperature'])
y2 = np.array(df['Humidity'])
z2 = np.array(df['Demand'])

#color = sns.cubehelix_palette(dark=-4,
as_cmap=True, gamma = 2)
color = sns.cubehelix_palette(8, start=.5,
rot=-.75, dark=0, as_cmap=True, gamma =
2)
#color = sns.cubehelix_palette(8, start=.5,
rot=-.75, as_cmap=True, reverse=True)
("")
sns.set_style("whitegrid")
f, ax = plt.subplots()
points = ax.scatter(x=x2, y=y2, c=z2,
cmap=color)
f.colorbar(points)
#plt.xscale('log')
#plt.yscale('log')
ax.set_xlabel('Temperature') # Set text for
the x axis
ax.set_ylabel('Humidity')# Set text for y axis
plt.legend((''), title = 'Demand', fontsize = 'x-
small', frameon = False,
        borderpad = False, loc='upper right',
bbox_to_anchor=(1.2, 1.1))
```