Universitetet
i Stavanger

DET TEKNISK-NATURVITENSKAPELIGE FAKULTET

BACHELOROPPGAVE

Studieprogram/spesialisering: Bachelor i ingeniørfag / Automatisering og elektronikkdesign	Vårsemesteret 2025 Åpen
Forfatter(e): Azat Kaplan, Philip Grønvold Andersen, Karl Kristian Holst Vigsø	
Fagansvarlig: Merkebu Zenebe Degefa Veileder(e): Merkebu Zenebe Degefa , Bjørn Grønning	
Tittel på bacheloroppgaven: Lastprognoser og fleksibilitetsberegninger for KI-datasentre Engelsk tittel: Load forecasting and flexibility estimations for AI data centers	
Studiepoeng: 3 x 20	
Emneord: AI, Flexibility, Forecasting	Sidetall: 73 + vedlegg/annet: 85 Stavanger 15. mai 2025

Contents

Abstract

Data centers are critical to the world's infrastructure while being energy intensive facilities. This study focuses on improving the understanding of the facility characteristics. These characteristics are used to further investigate data centers energy flexibility and efficiency capabilities.

A model was built in MATLAB Simulink using Ref Design 83 (1) to simulate power consumption. Recognizing the trends of the Simulink model can lead to better strategic decisions and actions in energy management to maximize operational efficiency. Implementing a smart strategy gives energy an opportunity to serve as a flexible resource for data centers. This allows the data centers to maintain stability and achieve sustainable growth. The evolution of technology will make effective energy management more vital in ensuring operational readiness of the centers in the future.

To better understand the energy usage of AI models, three different types is chosen to be compared. LSTM model produced the best predictive results with MAE 3.48, but it followed with increased power consumption as it exceeded 20W for most of the time and even reached 30W. XGBoost proved to be the most energy-efficient, operating under 10W, but this came at a loss of predictive accuracy though, and showed the lowest predictive accuracy (MAE: 4.79). NARX demonstrated an accuracy (MAE: 3.86) close to LSTM while maintaining energy usage in the 5-10W which was close to XGBoost.

The flexibility is explored through integrating demand-side management, battery energy storage systems (BESS), solar photovoltaic systems, and hybrid solutions. These models are used to show how these facilities can reduce peak loads, improve energy efficiency, and lower environmental impact. Simulations using real-world components, including reference battery(2) and reference solar panel(3), demonstrate that combining solar generation with battery storage provides the highest flexibility, enabling up to 200 kWh of energy shifting during peak hours. The results highlight that while individual systems offer specific advantages, hybrid models maximize self-consumption and grid independence, offering a promising path toward sustainable and cost-effective AI data center operation.

Acknowledgements

The bachelor's thesis concludes a three-year journey through the study program in Automation and Electronics Design at the University of Stavanger. Working on the project entitled Load Forecasting and Flexibility Estimations for AI data centers has been both rewarding and challenging. This thesis has proposed several challenges and has been demanding work, leading to some great learning experiences. One of the most valuable experiences we bring with us in the future, is the deep insight in how teamwork functions.

Firstly, We wish to express gratitude to our Main supervisor, Mr.Merkebu Z. Degefa. His assistance, guidance and advice has been very valuable to us. Secondly, We would like to thank Bjørn Grønning as well, for his guidance and support, especially in flexibility. Finally, We would like to extend a big thank you to our family and friends for encouragement and inspiration for this project.

Chapter 1

Introduction

The modern world is always advancing in technology. A key driver of these advancements are data centers. Data centers have become a fundamental need for every country in the world. This need comes from the expansion of global connectivity. Resulting from its connections to everything from cloud services to AI. Data centers have evolved from smaller server rooms into massive facilities capable of handling massive workloads in just a few decades. This growth comes from its connection with modern infrastructure. It is involved with business, schools and daily life. This growth has resulted in massive problems regarding their demand for electrical power.

The global energy that is being consumed by data centers is already accounting for around 1% (4)of the global energy usage. The forecasted loads as AI workloads increase are predicting big energy increases in data centers. The predictions include this: One ChatGPT query can consume up to ten times (5) more energy than a google search. By 2027 the projected energy consumption of AI servers is at least 85,4 terawatt hours yearly (6). This sums to more energy consumption than a few small nations. This energy consumption creates a big challenge especially for the power grid when it comes to how stable it can be. Other big challenges include emissions and operational costs. These problems show the background for the importance of energy efficiency in relation to the growth of AI

This bachelor will try to adress these problems by studying the major electrical components of a reference Tier 1 data center and study how its efficiency and flexibility can be improved. The study focuses on simulation of an electrical model created in Simulink. Simulations of different AI models showing their energy impact. Simulations of different flexibility strategies. By studying load forecasting and flexibility strategies the study shows insight into how data centers can reduce peak loads, increase efficiency and minimize environmental impact.

1.1 Data center Architecture

A data center refers to the building in which large amounts of data are stored, processed, and managed. This is one of several basic elements that make up the digital world, and we use it on a day to day basis. For example websites, apps, online banking, and more.

A data center consists of racks that hold servers, or computers that store or run the programs respectively. A stable power supply is essential for the operation of the data center. Minor power outages could lead to major problems. To avoid this scenario, they are equipped with advanced power systems that include backup and emergency power sources like UPS and generators. These are also designed to take over automatically when the main supply fails, thus giving continued operation for the data center without interruptions.

One element of data center operations is cooling because servers operate at high temperatures. This operation gives extra protection for the data center with cooler temperatures, protecting systems from overheating. Also, a cooler temperature creates a stable and efficient data center that can handle higher loads. Thus, controlling temperature means controlling operational efficiency for stable processing.

As the drive to make data centers more energy efficient increases, new opportunities arise to reduce energy usage and maintain performance needs with the flexibility of addressing future needs as well. Changing to renewables helps data center use less fossil fuel, reducing its carbon footprint. Some data centers are even located in colder climates where natural cooling can work without excessive energy use.

1.1.1 Tier Classifications and Specifications

The way to classify the design of a data center in a commonly accepted fashion is through a four tier system. (7)The system is made by ANSI boasting arround 400 member companies support a 4 tier classification based upon power distribution, uninterruptable power supply (UPS), cooling delivery and redundancy in the facility.

Tier I data center has one power and cooling distribution path with no redundancy (N).

Tier II data center adds redundant components to the configuration (N + 1) for better availability.

Tier III data center has one in use and one standby distribution path for power, cooling, and other utilities, with each path equipped with redundant components that can be repaired without shutting down. The distribution paths give redundancy, therefore avoiding downtime during routine checks .

Tier IV data center has two in use power and cooling distribution paths with redundant components in each path and is designed to sustain any single component failure without impacting the load .

The figure 1.1 below shows the different types of level tier classifications, explaining their meaning and function.

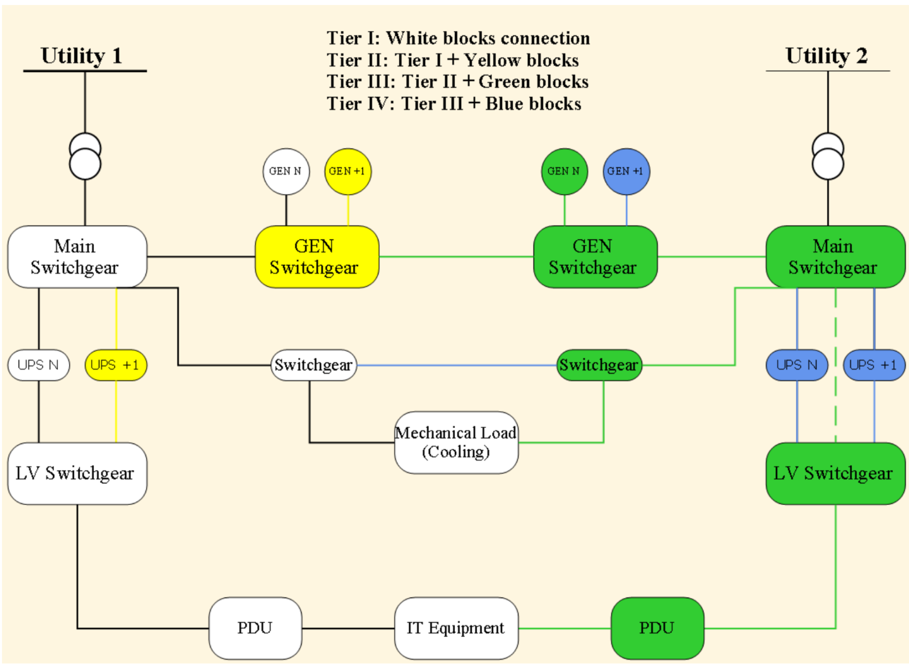


Figure 1.1: Reference Design 83(1) -Tier classification of power distribution

1.2 Major Components of Data Centers

The sizes of data centers vary widely, and many are categorized based on their footprint or critical power the amount of electric power that can be sustained to IT equipment on a continuous basis; for instance,(7; 8)About 66% of servers in the U.S. were located in data centers under 5.00 ft^2 (450 square meters), with a power limit below 1 MW.Big data centers are designed to keep many servers for different companies. They can deal with high energy demands, sometimes reaching dozens of megawatts. Cloud suppliers also perform their data centers similarly to support great computing needs . Below is a table 1.1 showing various data center types with their effective power capacities.

Data Center Size	Small	Medium	Large
Building size	5,000 - 20,000 sqft	20,000 - 100,000 sqft	100,000 sqft to millions of sqft
Server Count	500 - 2,000 servers	2,000 - 10,000 servers	10,000 to 100,000 servers
Power Capacity	1 - 5 MW	5 - 20 MW	20 - 100+ MW
Design / Efficiency	Energy control and cooling	Efficient power use balanced, performance	Optimized performance, green energy use

Table 1.1: The table classifies data centers as small, medium, and large. Four operating standards include square footage and number of servers, energy managament, and optimized perfomance. dgtlinfra(9) was used as the source.

1.2.1 Data Center Power systems

In this section, the major components of power systems in data centers are briefly introduced and explained:

- **Transformers:** Transformers adjust voltage levels to make sure the electricity from the main power source is right for the data center (10).
- **Power Distribution Units (PDUs):** PDU distribute power in a data center, taking electricity from the main source and delivering it to IT systems and other important devices.
- **Power Cables and Connectors:** These fundamental components connect power equipment to IT equipment.
- **Power Breakers:** Circuit breakers protect the electrical equipment from too much current or short circuits.
- **Uninterruptible Power Supplies (UPSes):** UPS devices offer power redundancy to IT equipment suffering power loss due to power outages, high voltage or low voltage, voltage dip, and voltage surge. A battery backup system and an automatic transfer switch automatically shift power to the battery when mains fails.UPS employs

AC-DC-AC double conversion. Incoming AC is converted to DC, which feeds a UPS-internal bus feeding battery strings(11). The output of this DC bus is inverted back to AC, which feeds the data center PDUs. When utility power fails, input AC is absent, yet input DC is present (from the batteries) so that AC output to the data center is uninterrupted. At this time, however, the generator kicks in and restores input AC power.

- **Backup Generators:** Generators create standby power for long term outages that UPS systems cannot accommodate.
- **Automatic Transfer Switches (ATSes):** ATS shifts power from the main source to a backup, such as a generator when there is an outage.

1.2.2 Essentials of IT Hardware

A server rack helps save space, allows for simple upgrades, improves cooling with better airflow and makes maintenance easier since professionals can easily access and manage the equipment.

Rack servers are designed to easily slide in and out of a server rack. This means that IT specialist, engineer, and operators can easily diagnose technical problems and instant module switching without having to power down and disassemble an entire system.

This feature is essential for critical systems, where long periods of down time can lead to serious consequences, from financial loss to safety risks or even life threatening situations.

A rack server is designed to optimize its performance, features and services based on the specific needs of a program or application. Figure 1.2 illustrates fundemental perspective and what is functionality of racks and server (12).

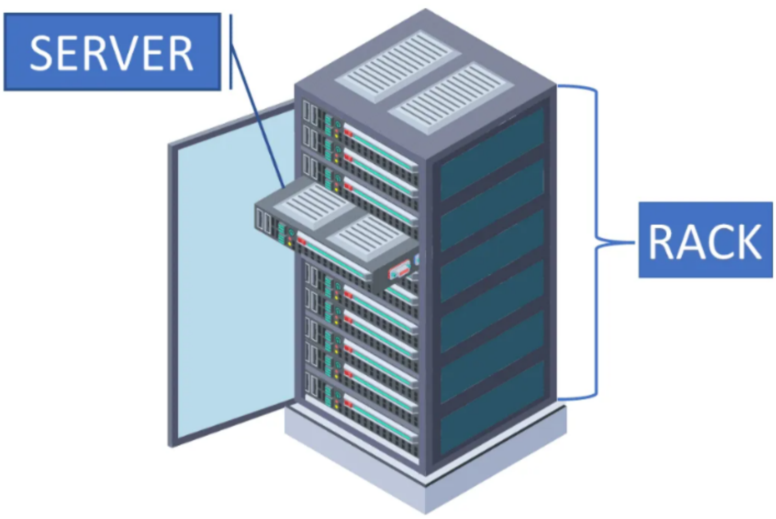


Figure 1.2: Block diagram of a data center, showing rack and server components and their functionality.Image sourced from nylyte(13).

1.2.3 Cooling System in The Data Centers

A Computer Room Air Conditioning system (CRAC) HVAC unit operates similarly to an air cooling system, using a cooling fluid and a pressure driven mechanism to regulate the room's temperature. These units made to cool data centers, push air over a cooling coil while a compressor keeps the refrigerant in the coil cool. The CRAC unit releases heat by using water or another cooling substance to carry it away. In a data center with a CRAC system, you'll find the CRAC unit across from the hot aisle. It takes in the hot air and can send it back out as cool air through holes in the raised floor cooling the servers from below. This also helps to build up pressure in the space under the floor. Most CRAC units have a simple on/off switch, but some newer ones can change the airflow.

Optimizing Cooling Systems with CRAC HVAC Units

There are several optimizations offered by a CRAC unit:

- Responsible for the circulation, dampnes, and temperature of the data center.
- Sucks heat out of servers and blows cold air at the servers to help cool them.
- Ideal for small data centers that do not require rapid scaling.
- Space-saving, especially when compared with CRAH units that provide full setback control.
- Works effectively in all conditions, whether in extreme cold or intense heat (14).

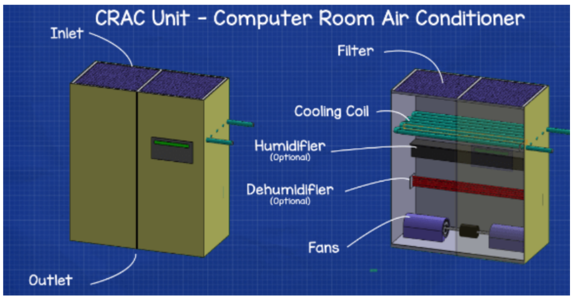


Figure 1.3: Components of a Computer Room Air Conditioning (CRAC) System.From the image above were sourced from theengineeringmindset (15).

Air Cooling Technology

Air cooling technology is one of the most traditional and commonly used solutions for large-scale data centers. Its key benefits include easy maintenance and a reasonable operating cost (16; 17). As shown in Figure 1.4 , the concept relies on an airflow cycle within the TCSS system, where the cooling is supplied by the computer room air conditioning system (CRAC).

The following sections focus on two specific air cooling methods: terminal cooling and mechanical air cooling. Both play an important role in enhancing the efficiency and performance of data center operations.

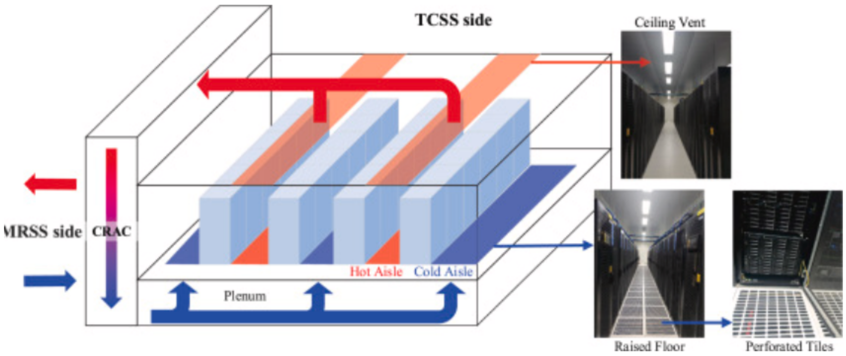


Figure 1.4: The fundamental process of air-cooling at the room level, designed for raised-floor setups.Image sourced from scrd (18).

Level	**Type**	**Advantages**	**Disadvantages**
Room Level	Raised floor design (19)	Modular tiles make it easy to reconfigure for setups like cold and hot aisles	Raised floors generally have a lifespan of around 20 years
Row Level	In-row cooling (20)	Energy efficiency is better compared to room-based systems	Costs can increase quickly
	Overhead cooling (20)	Quick installation reduces starting costs	Cooling air from above isn't always effective, especially if it causes areas to overheat.
Rack Level	Internal enclosure (21)	Simplifies the installation, management,and movement of equipment.	Expensive to install

Table 1.2: Cooling System Levels: Advantages and Disadvantages

Terminal Cooling

Data centers use air cooling to avoid their equipment from getting too hot. The way they do this depends on how close the machines are located together. There are different ways to save energy and help cooling works better. Chip cooling is the most advanced, but cooling on rack or server level is too expensive to configure and maintain. For this reason, many data centers do not use it.(16; 22).

Indirect Liquid Cooling

The mechanical refrigeration sub-system (MRSS) is important, while direct fluid cooling is mainly used in the terminal cooling sub-system(TCSS). Indirect fluid cooling is stable and reliable, while the liquid works very well because it removes the heat directly at the source. The use of special cooling fluids helps data centers to remain effective in energy and function without problems.(23).

1.3 Modeling of the Electrical System of Data Centers

The data center analyzed in this work is structured according to Reference Design 83(1). The reference data center gets power from the grid through a 9 kVA transformer and has a 500 kVA backup generator for stability . This goes to a 600-amp I-Line panelboard, which feeds the cooling and support equipment, as well as a 90 kW Symmetra PX Uninterruptible Power Supply (UPS). The UPS has only 6.7 minutes of battery runtime at full load, powering the IT racks. It has a total footprint of 1637.6 ft^2 and targets a PUE of 1.64. This has been a useful and clear design description to engineers and technicians.

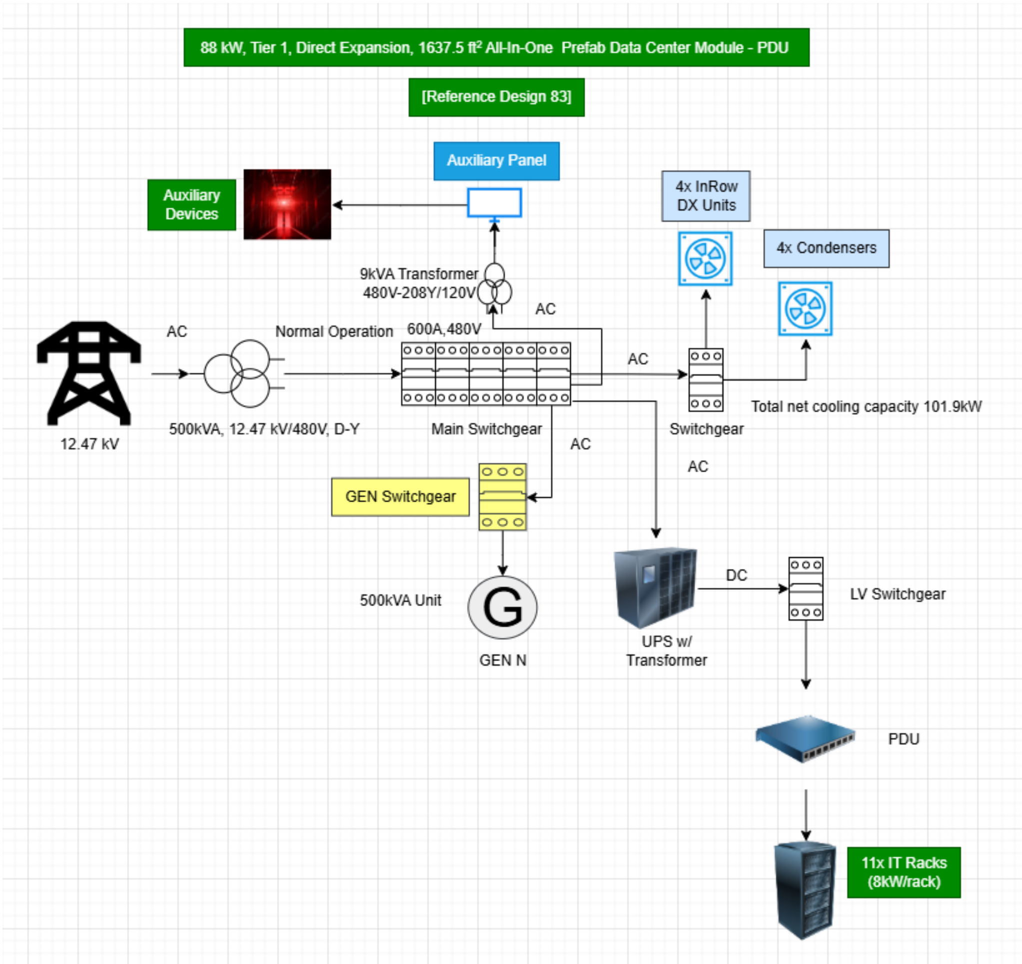


Figure 1.5: Modeling Tier 1 Data Center with 88 kW IT Capacity Reference Design 83.This block diagram summarizes the design and power architecture of a data center. It is an 88 kW IT capacity data center, which consists of 11 IT racks, with an average of 8 kW per rack. This is a Tier 1 system , with a power route and without redundancy. Use a UPS N (without transformer) that is consistent with Tier 1 specifications, instead of reference design 83 that made an UPS with transformer for Tier 2 facilities.

	Single-phase Power	Three-phase Power
Voltage	Commonly used in small systems	Commonly used in large systems
Power Capacity	Decreased power capacity	Increased power capacity
Power Distribution	Support small devices	Power servers and larges maschines
Efficiency	Unbalanced flow reduces efficiency	Efficient with balanced flow
Complexity and cost	Easier, cheaper to install and service	More complex and costly
Scability	Limited power growth	Scales easily for data centers
Industyr applicability	Used in small data centers and factories	Found in large data centers

Table 1.3: A comparison between a single-phase and a three-phase power. This table highlights the main differences and applications of both power systems, including voltage, capacity, power distribution, efficiency, complexity, cost, scalability and industry applicable. The table on the site is based on information from device42(10)was used as the source.

Table1.3 Compares single phase and three phase power systems by their main features and uses

1.3.1 Steady-State Operations in Data Center Systems

Usually the data center's primary power comes from the electrical grid. During extreme weather events or other disruptions, data centers might be able to tap into battery reserves or backup generators to ride things out. (24)The utility delivers power to the data center at moderate voltage, like 13.8 kV or 34.5 kV, and the data center uses step-down transformers to convert that to lower voltages, like 480V or 208V. After it's converted, the lower voltage power is delivered to the data center facility through uninterruptible power supplies (UPS), power distribution units (PDUs), and remote power panels (RPPs).

The electric power for most large data centers comes from the local utility company connected to the main electrical grid. In the data center, power is delivered using three-phase electricity, which is typically found in industrial settings to power motors and other machinery. Data Centers primarily use alternating current (AC). The electricity supply is 50 cycles per second, or 50 Hz, in the UK. In most US-standard regions, it's 60 Hz(25)

1.3.2 Grid Instability Impact On Data Centers

Data Centers are critical infrastructure, where even a brief interruption of electricity can lead to serious consequences. These impacts include hardware damage, disturbed data and operations that can lead to financial failures and services. To relieve these risks, backup systems, including uninterrupted energy sources (UPS) and support generators, play an essential role in maintaining soft operations during energy interruptions and transition guarantee without problems when the grid is restored. The UPS system acts as a reliable intermediary between the electricity network and the data center and provides immediate support during energy disturbances. It consists of components such as batteries, rectifiers, investors and control mechanisms, UPS moves without problems to support the battery when primary energy is interrupted .

This instantaneous change ensures that the critical load remains fed until the grid is restored or that the backup generator does not become functional. For data centers, this capacity is necessary to maintain operations during unpredictable interruptions. In addition to its main backup performance function, UPS systems are an integral part of energy conditioning. They manage frequency and voltage fluctuations that could otherwise damage sensitive electronic devices. By providing a clean and stable power supply, the operations of the UPS data center provide and extend the useful life of expensive hardware. This dual function underlines its value both to maintain the reliability of the system and to optimize the long -term performance of the critical team(26).

In the scenarios where the grid is restored, the role of backup generators becomes significant. Backup generators ensure the continuous energy supply during extended interruptions and act as a bridge until the electricity grid stabilizes. Its presence allows data centers to move gently between different energy sources, increasing operational flexibility and resistance. Together with the UPS systems, they form a complex safety network that protects the vital systems of the consequences of the interruption of performance and instability associated with the recovery of the network. It cannot overestimate the importance of UPS generators and backup in data centers. These systems allow data centers to stand out in difficult circumstances, which guarantees flexibility and continuous functions. By providing reliable energy during interruptions and stabilizing the transition back to network power, strengthens the compatibility and efficiency of data center operations.

Because energy reliability remains main for modern infrastructure, the integration of solid backup solutions is necessary to guarantee the provision of critical service data centers. In this way, data centers remain safe even during moments of uncertainty and maintain their role as a basis for technological continuity .

1.4 Energy Consumption per Unit

To ensure consistency, an assumption of the standardized power consumption based of Reference Design 83 (1) were made. The following estimates were applied across the 24-hour simulation period.

Table 1.4: Assumed Power Consumption for Reference Design 83 (1) Data Center

Component	**Description**	**Power Assumption**
Rack Power (IT Load)	11 racks, 8 kW per rack, varies with workload (50% idle, 100% peak)	Up to 88 kW
Cooling Power	Four InRow DX units and four outdoor condensers. Scales with IT load; 45% at peak (PUE $\approx$ 1.64)	Up to 45% of IT load
Auxiliary Power	Lighting, monitoring systems, PDU inefficiencies	$\sim$4 kW
Power Delivery Losses	Transformers, UPS, switchgear, and cabling losses	$\sim$6 kW (6.8% of IT load)
Backup Systems	90 kW Symmetra PX UPS, battery charging losses	$\sim$2.5 kW
Backup Generator	500 kVA diesel generator on standby	$\sim$1.5 kW
Miscellaneous	Security and other minor loads	$\sim$1 kW

Whole energy consumption table in the Appendix4

1.4.1 Data center Electric Model

An electrical model for a data center is a subject of this thesis, i.e. developing a clear view for how power is supplied, distributed, and managed. The model plays a crucial role in the design and operation of a data center because it smooths the operation, even in the event of a power outage. This ensures that the power system for the data center is realistic, functions well, and is able to be modified to meet new needs.

All the major electrical components connect together to form the electrical model. Power sources, distribution networks, batteries, servers, and racks are the major components covered by the model. Normal operation is demonstrated, and provisions for backup power systems, such as generators and UPSs (uninterruptible power supply), are included to assure that the operation could still continue even when an outage from the grid power supply occurs. Keeping the data center always up and running is the most important aim of this model since downtime causes expensive damages and losses. The electrical system has thus been designed robust and reliable.

Developing a Data Center Electrical model based on the above statement includes Tier 1 data center with an IT capacity of 88 kW as described in the reference here(1). Furthermore, the model will be developed and simulated in MATLAB Simulink so that it can analyze the power consumption with the plot figures showing the energy flow within the data center.

Chapter 2

Methodology

2.1 Simulink Model of a Data center

In this section, a MATLAB Simulink model of the power system is developed based on a reference data center(1). This is to study the load profiles in normal operation and sequence of actions in grid interruption scenario.

2.1.1 Complete Simulink Data Center Model for 24 Hours

Figure 2.1 represents a complete Simulink data center model of 24 hours. It shows the final design of data center Reference Design 83(1). The model uses Simulink blocks to simulate power grid and energy systems for a whole day. It includes important components such as generators, transformers, cooling systems and backup power. This model helps to analyze and understand the flow of power and performance to a data center.

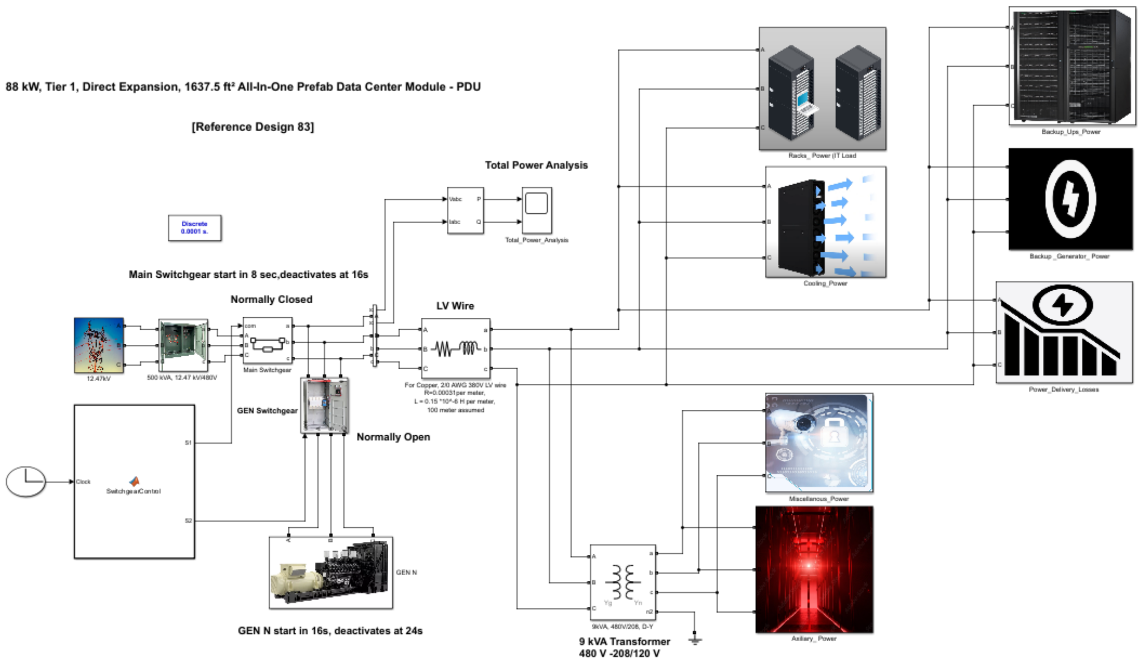


Figure 2.1: Simulink Model of Data Center Operations Over 24 Hours

2.1.2 24-Hour Transformer Simulation in MATLAB Simulink

Figure 2.2 illustrates a three-phase transformer that takes 12.47kV from the power grid and step it down to 480 V. The A, B, C connections bring high voltage electricity, while a, b, c connections send the lower voltage to the main switchboard, which then distributes power throughout the facility. The work of the transformer is to adjust the voltage from 12.47 kV to 480 V, before distributing it on the switch panel.

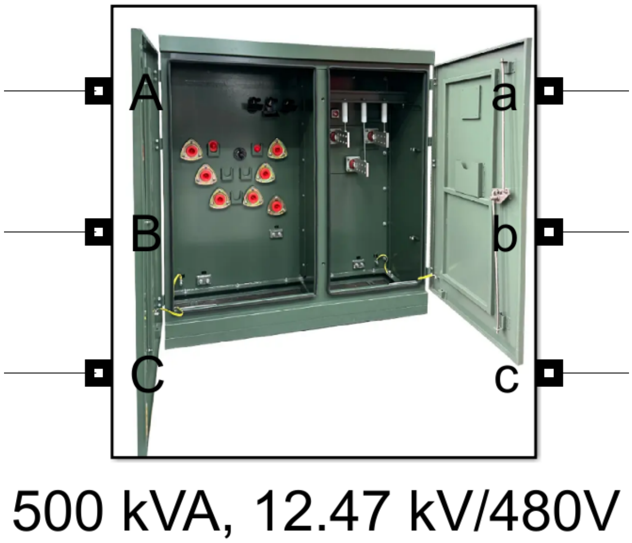


Figure 2.2: 500 kVA Transformer (12.47 kV/480 V)

2.1.3 Racks Power 24hr Simulation

Figure 2.3 illustrates a power simulation model of the Ref 83 data center (1), created with Simulink. It focuses on key aspects such as active power (P), reactive power (Q), power factor (PF) and total power (kW) to ensure a complete analysis. Throughout the day, the simulation traces the energy demand of the data center, breaking it step by step to examine the function of each component and how they interact. When analyzing each part, the model provides a clear representation of how energy flows, is regulated and optimized to ensure operation without problems

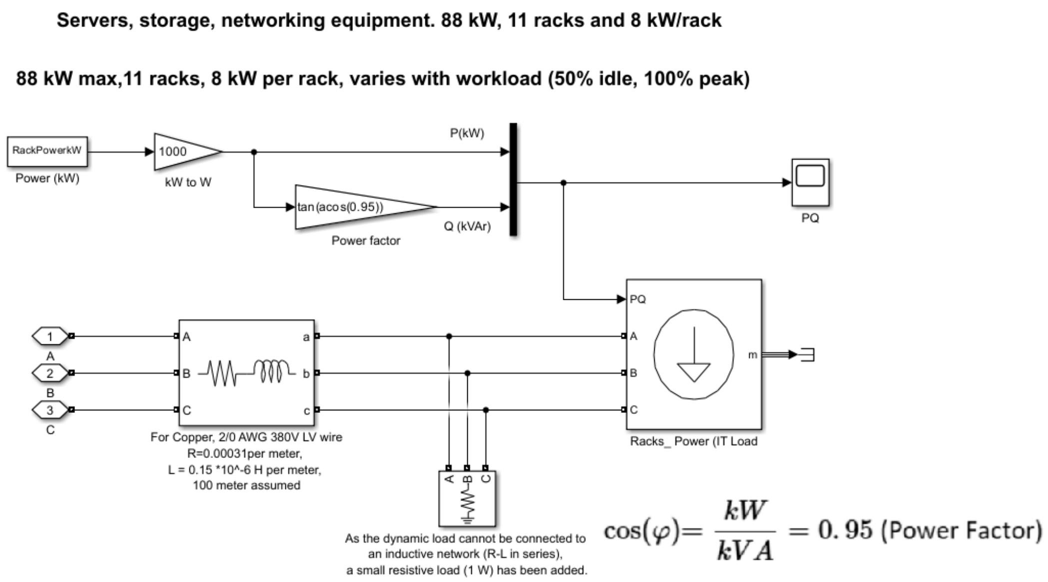


Figure 2.3: IT Racks Power (IT Load) Block Diagram Over 24 Hours

Active power values (P) come from an Excel file, where P is recorded in kilowatts (kW). To calculate the reactive power (Q), the formula given by:

$$Q = P \times \tan(\phi) \tag{2.1}$$

Here, ϕ (phi) represents the Power Factor. The Power Factor is 0.99 at full load, but it is approximately 0.95 under typical conditions. This information is referenced from 3dfs(27)

At this date center, the auxiliary energy is represented by LED lighting, which is used for lighting and monitoring. The move to LED technology has greatly reduced energy consumption, making the system more efficient. By decreasing energy use, LED lighting improves energy management and helps reduce costs. Auxiliary power in a power system, as well as any other system that does not require reactive power, just uses active power, making them efficient and energy conserving. An example of this can be a lamp or a light fitting with lower Power Factor. It generates additional reactive power which does not contribute to the light energy and space in the electrical system. This increases the amount of power needed by the system, making it difficult to deliver useful energy and causing an increase in electricity consumption (28; 29).

Model 2.3includes low voltage energy distribution as the main source to feed the data center racks. The model also incorporates an RLC load, which represents the combined resistive, inductive and capacitive components. Since a dynamic load cannot be connected directly to an inductive network (R-L in series), a small resistance load of 1 W for system stability is added. Three-phase dynamic load in the model ensures that active power change (P) and reactive power (Q) is changed depending on the positive sequence voltage (V).

This configuration follows the electrical standards of North America, specifically in Canada, where systems operate at a frequency of 60 Hz. The model accurately represents how power is managed in a real-world data center, following regional guidelines and guaranteeing efficiency in distribution and load balancing.

2.1.4 Data Center Rack Power Overview

After energy conversion, the system calculates reactive power (Q in kVAr) using the formula:

$$Q = P \times \tan(\arccos(\text{power factor})) \tag{2.2}$$

Each component on the data center uses a different power factor, depending on its efficiency and function. This helps to understand how energy is being used efficiently and allows adjustments to better power management. To simplify things, a mux block is used to combine P (kW) and Q (kVAr) into a single signal. This make the process and analysis easier. The final signal is sent to the scope, which shows a visual energy consumption graph over time. This shows how much electricity is used, find patterns and optimize energy use.

This method is used for all data center components, with each with its own power factor, making sure everything is measured in a way that reflects its energy efficiency. It helps track the flow of energy to racks, cooling systems, and other components. By using Excel for data, MATLAB for calculations and the scope of monitoring, this system helps data centers better manage energy, and improve efficiency.

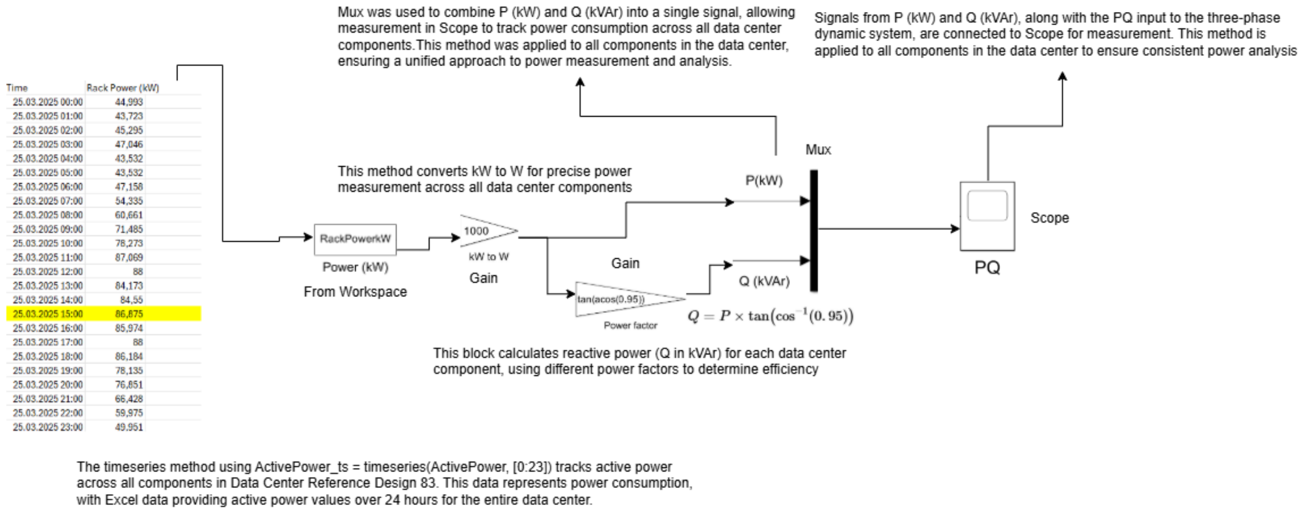


Figure 2.4: 24-Hour Rack Power Overview

2.1.5 IT Racks Energy Usage Throughout the Day

The energy used by IT racks on the data center follows a clear pattern throughout the day. It does not remain the same all the time, because it changes depending on activity levels. In this study, to generate synthetic load profiles, the following data center activity is assumed:

- **Midnight at early morning (00:00-05:00)**: IT racks are performed with a capacity of 50% as there is low activity in the data center.
- **Morning (06:00-11:00)**: To the extent that the day begins, the use of energy gradually increases from 50% to 100%, reaching total capacity.
- **Midday in the afternoon (12:00-17:00)**: The system operates at full power 100%, handling peak loads.
- **Night at night (18:00-23:00)**: The use of energy gradually decreases from 100% to 60% as the activity decreases.

To make the simulation feel more real, a small ± 2 kW variation is added each hour, since energy use in real life is never completely steady. This helps show how power is used, improve efficiency, and keep everything working well throughout the day.

Key Components of the Data Center

The main components of the data center and their roles in power distribution are explained as:

- **Rack Power (IT Load):** Servers, storage, and networking use up to 88 kW for 11 racks. Each rack uses 8 kW. Power changes based on usage—50% when idle and 100% at peak. The power factor (PF) is 0.95, as referenced from 3dfs (27)
- **Cooling Power:** Four InRow DX cooling units and four outdoor condensers handle up to 101.9 kW. Cooling uses about $\sim 45\%$ of IT load at peak (PUE ~ 1.64). The power factor (PF) is approximately 0.85, as referenced from mircevski(30)
- **Auxiliary Power:** Lighting, monitoring, and PDU losses are around ~ 4 kW. The power factor (PF) is 1.
- **Power Delivery Losses:** Transformers, UPS, switchgear, and cables lose about ~ 6 kW, which is around $\sim 6.8\%$ of IT load.
- **Backup Systems:** A 90 kW Symmetra PX IT space UPS uses ~ 2.5 kW for battery charging. A 500 kVA generator uses ~ 1.5 kW when on standby load. The power factor (PF) is 1.
- **Miscellaneous:** Security and small systems use about ~ 1 kW. The power factor (PF) is 1.

2.1.6 Data Center Timeseries Load Profile

Figure 2.5 shows the synthetic load profiles generated in this study based on the assumptions outlined in section 2.1.5 , focusing on the total active energy (kW) displayed on the graph. A list of how much energy each component uses, including cooling systems, rack and backup systems.

During working hours, servers use more energy and cooling systems adjust to maintain stable temperatures. Backup systems, like UPS and generators, remain ready for power outage. Other lower energy needs, such as lights and monitoring systems, consume very little energy and appear zero on the graph.

The graph and table work together to show how energy is distributed by components and effectively used on the data center, providing a clear and detailed overview of energy management over time.

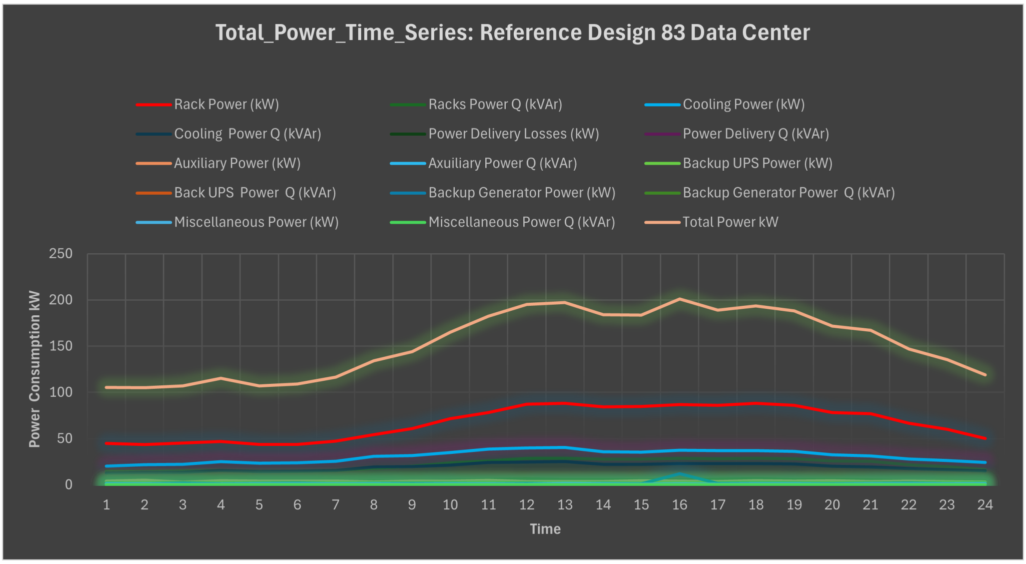


Figure 2.5: Excel Graph Showing Total Power Consumption Over a 24-Hour Period

2.2 AI Models and Energy Consumption

AI has made tremendous progress but its energy consumption poses significant environmental and supply problems. Google performs one search using 0.0003 kilowatt hours (31). A ChatGPT search uses energy between 0.3 watt hours and 3 watt hours(31). AI searches require power that exceeds standard web search requirements by up to ten times.

The dramatic energy consumption contrast demonstrates why AI systems need improved energy efficiency capabilities. AI technologies continues to expand into multiple application domains thus making energy efficient management of their power usage essential for developing sustainable solutions. The following research focuses then on examining various AI forecasting models to identify methods for minimizing power usage through choosing the best forecasting model.

Data Centers use a lot of power to operate, which means that the costs associated with power consumption are high. This makes accurate power forecasting crucial for effective energy management and for the flexibility potential.

Because of limited access to real data center, this project was conducted locally with an available PC. Utilizing available hardware parameters such as CPU utilization and temperature data from a standard personal computer. This will act as a proxy for a data center because a personal computer have similar basic construction as a data center.

The decision to use AI-driven models was due to the increased usage of AI in data centers. The models chosen is specifically long short-term memory (LSTM), non-linear auto-regressive with eXogenous input (NARX) and extreme gradient boost (XGBoost). The choice was motivated by the proven effectiveness of these models in time-series forecasting. Implementing these models on personal hardware provides a practical and cost effective solution to run and generate forecasts for flexibility.

2.2.1 Data collection

For accurate simulation and forecasting of power usage some parameters are needed. This project utilizes Intel Power Gadget to collect these parameters. It is a software tool designed by Intel specifically for real-time monitoring of various components of the PC. Those parameters include Processor power usage and more. In this protect three of those parameters is more important.

These are:

- Processor Power Consumption (Watt): Provides values of power usage of the CPU. This is the biggest power measurement in the pc.
- CPU Utilization: This tracks the percentage of CPU actively engaged at any given time.
- Temperature (°C): Measures the temperature of the CPU package internally.

These were the most stable and applicable measurements this computer configuration had to offer. While it would have been interesting to monitor GPU power usage, the GPU on this computer is always in use at 80-90% due to the high-definition display. GPU readings were therefore not utilized.

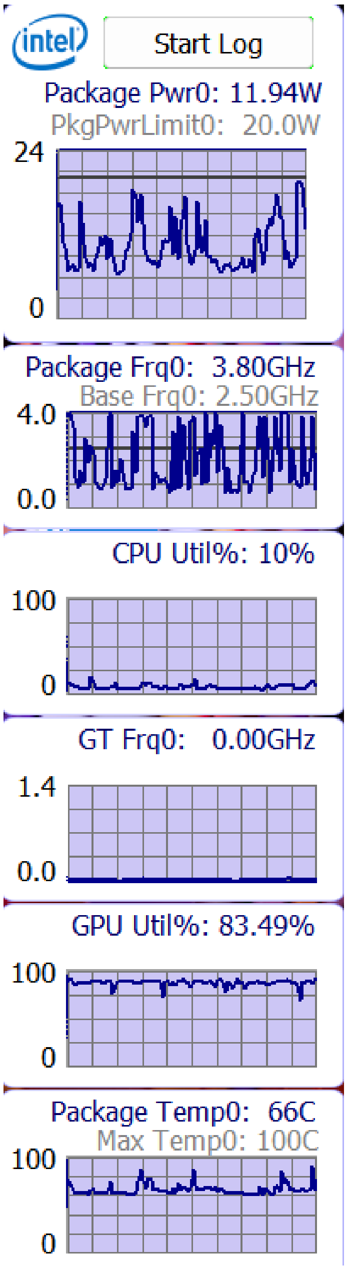


Figure 2.6: This is the interface of Intel Power Gadget.

The captured data is logged with a time interval of ≈ 0.1 seconds. Thus it is able to log rapid fluctuations. This will not be the case when forecasting for flexibility purposes, because the agreed upon time interval there is one hour. By using the datasets from the PC the models have a lot of values as some of the data logs are over 100 seconds. As a result more data to train and forecast. Which leads to easier detection of power increases when the models are running.

Data Preparation

After the data collection phase using Intel Power Gadget, the raw dataset undergoes some preparation steps. The preparation steps includes:

- **Removal of entries with missing data points.**
- **Selection of parameters** That is CPU utilization and package temperature. These two variables were identified as having the most correlation with the power-usage. Both of the parameters had a correlation of around 0.8. With some variance for dataset to dataset.
- **The correlation was determined by this formula:**

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2.3}$$

- **Normalization:** Utilizing the StandardScaler method from scikit-learn.This ensures that all variables scales properly. Meaning that even if the CPU value ranges from 10-80% the power value still ranges from 3-40W
- **Graphs** Some graphs are used to verify visually that the datasets are usable.

After the preparation the datasets are segmented into different sets for training and testing. The training dataset have all their parameters so the model can learn the correlation between the three. While the test dataset only have the first two (CPU Utilization and Package temperature). That means it needs to generate the possessor power based on the other two parameters. This serves as the unbiased performance evaluation of the forecasting models as the real values for the possessor power is still stored elsewhere and is ready to compare.

2.2.2 Case Study: Forecasting AI Models

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) designed to model sequential data. It is developed to overcome the shortcomings of traditional RNNs. LSTMs is good at learning and storing patterns without suffering from vanishing or exploding gradients. This makes them suitable for tasks such as forecasting, and time series analysis.

An LSTM network is made up of memory cells and gating mechanisms. These gating mechanisms are input gates, output gates, and forget gates. The gates regulate information flow, which means determining what information is retained, discarded, or updated with each timestep.

Mathematically, each LSTM cell's operation involves:

- Forget gate: Decides what information from previous timesteps should be discarded.
- Input gate: Determines what new information should be stored in the cell.
- Output gate: Controls the information flow from the cell to the subsequent cells and the output layer.

Application of LSTM

In this project, LSTM was implemented using PyTorch. PyTorch is a popular deep-learning framework in Python. The specific LSTM network architecture employed here consists of an input layer receiving two parameters: CPU utilization and package temperature. After the input layer a LSTM layer with 64 neurons is applied. Finally the output layer producing the forecasted processor power is applied.

The LSTM model was trained using sequences consisting of 20 consecutive data points to effectively capture temporal dependencies. Model parameters were optimized using the Adam optimizer as the data is pretty noisy. The loss function used during training was therefore Mean Squared Error (MSE).

The training was conducted over 30 epochs. This was determined as it had satisfactory predictive accuracy. Following training, the model predictions were inverse-transformed from normalized values back to real power measurements.

NARX models are specialized recurrent neural network structures used for time series forecasting. NARX models leverage past values of both the processor power and external inputs (CPU, Temp) to predict. The architecture of NARX consists of feedback loops where the model's past outputs are repeatedly fed back in as inputs, allowing the model to recognize and predict time-dependent patterns.

Mathematically, a NARX model can be represented by the function:

$$y(t) = f\big(y(t-1), y(t-2), \ldots, y(t-n_y),\ u(t-1), u(t-2), \ldots, u(t-n_u)\big) + \epsilon(t) \quad (2.4)$$

Where:

- $y(t)$ is the output at time t
- $f(\cdot)$ is a nonlinear function, typically approximated by a neural network
- $y(t-i)$ are previous outputs (autoregressive terms), for $i = 1$ to n_y
- $u(t-j)$ are previous inputs (exogenous variables), for $j = 1$ to n_u
- $\epsilon(t)$ is the noise term at time t

Application of NARX

In this project the NARX model was implemented using Python's scikit-learn library. More specifically using a linear regression method encapsulated within a Regressorchain structure. The implementation details include:

- Lag: which means that the prediction was also influenced by the previews prediction it had made. In this case the lag was 2.
- The NARX model was trained in open-loop mode using real past output values, and used in closed-loop mode during prediction by feeding back its own predictions.
- The input features and output values were standardized using the same approach as the LSTM implementation.

Extreme Gradient Boosting (XGBoost) is an type of gradient boosting machine, which means it combines many small neuron trees. The way it combines them is by first running one prediction through the first tree. After that the prediction from the first run, runs through the second and so forth. It is designed for efficiency, flexibility and scalability.

Key advantages of XGBoost include:

- Robustness: Effectively handles missing data and varied data distributions.
- Efficiency: Optimized for speed and performance through parallelization.
- Flexibility: Supports customized loss functions and hyperparameter tuning, enabling highly tailored predictive models.

XGBoost's predictions is achieved through gradient boosting. Where each new small tree of neurons is trying to correct the residual errors left by previous new tree. Resulting in improved accuracy over the number of trees used.

Application of XGBoost

XGBoost was implemented with Pythons API. Each model of XGBoost utilized regression trees configured with 50 estimators and a maximum depth of 3. The objective function used was MSE, to minimize the mean squared prediction error.As with the other models it was scaled using StandardScaler to standardize input.After training the output was in the original measurement unit (Watts), facilitating easy comparison with actual observed values.

Model Evaluation Metrics

Evaluating each AI model's performance involved several standard regression metrics to compare their predictive accuracy.

- Mean Squared Error (MSE): Measures the average squared differences between predicted and actual values giving a bias to larger errors.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2.5}$$

- Mean Absolute Error (MAE): Is average absolute differences between predictions and true values.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2.6}$$

- Coefficient of Determination ($\mathbf{R^2}$): Indicates how well it follows the data variability. Values closer to 1 indicating higher predictive accuracy.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2.7}$$

Where in equation 2.5, 2.6 and 2.7

- y_i is the true value
- $\hat{y}_i$ is the predicted value
- $\bar{y}$ is the mean of the true values
- n is the number of samples

Comparison

Plots to compare with are:

- Real vs XGBoost
- Real vs LSTM Predictions:
- Real vs NARX Predictions:
- Combined Comparison (Real vs. NARX vs. LSTM vs. XGBoost):

Real vs. XGBoost Predictions: The predictions shown in Figure 2.7 were reasonably accurate, tracking the general trend of processor power usage. The model had big errors, particularly with individual spikes. The previous errors indicate that although the model has a strong generalization capacity, it has problems modeling sudden shifts/spikes in the processor power usage.

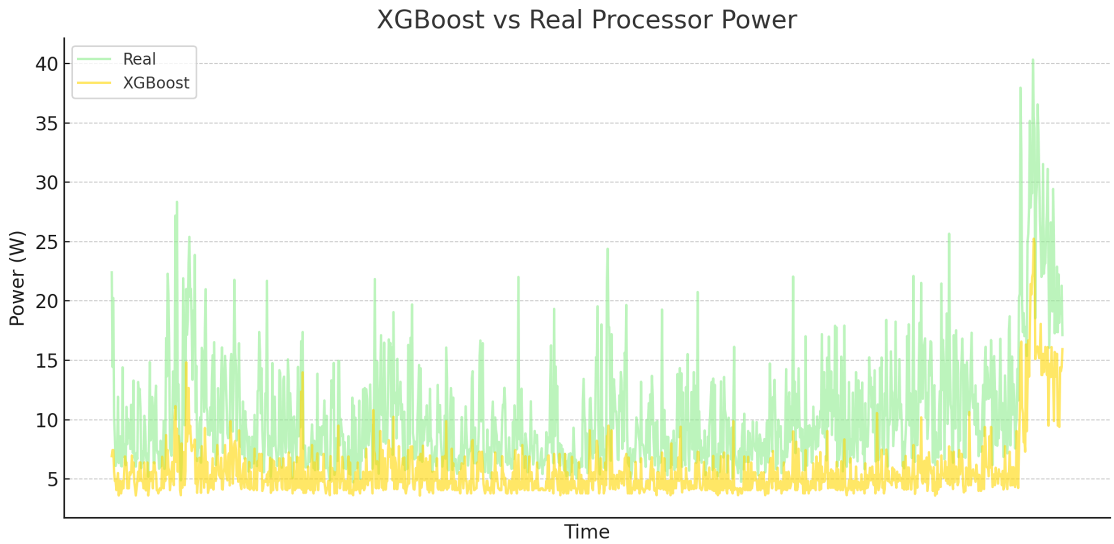


Figure 2.7: Energy load variation over time.

Real vs. LSTM Predictions: The LSTM model predictions shown in Figure 2.8 was smoother and it closely follow average trends but it does underperform with drastic fluctuations. Its predictive outputs indicate limitations in rapidly adjusting to sudden changes typical of high-intensity computational scenarios. This can be due to the use of 20 consecutive data points or 30 epochs when learning.

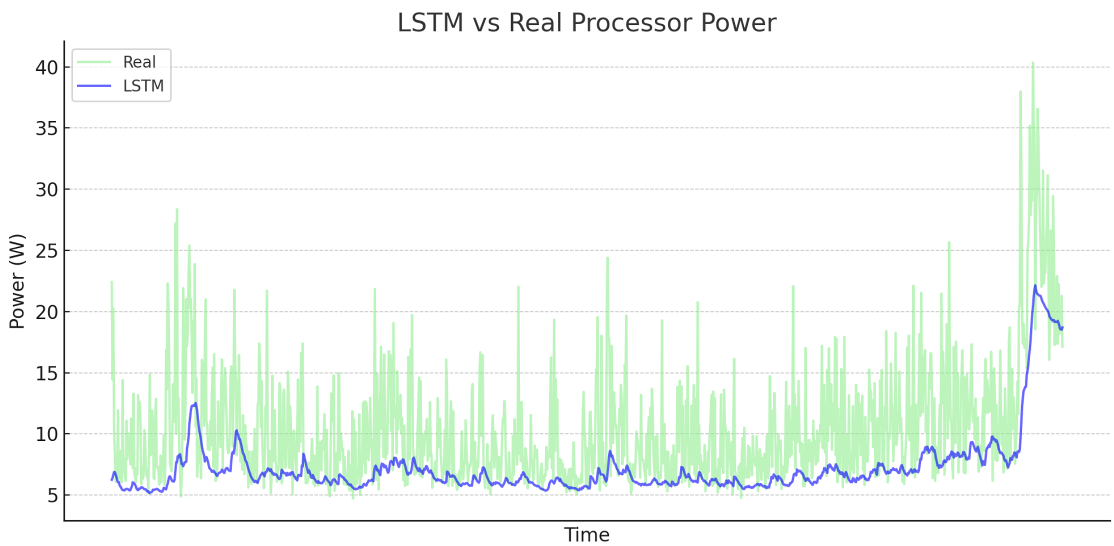


Figure 2.8: Energy load variation over time.

Real vs. NARX Predictions: NARX demonstrated in Figure 2.9 is a lot below the average at all times. This might suggest there is a bias between the training and test dataset. It does overall follow the trends well, yet not quite to the level of LSTM.

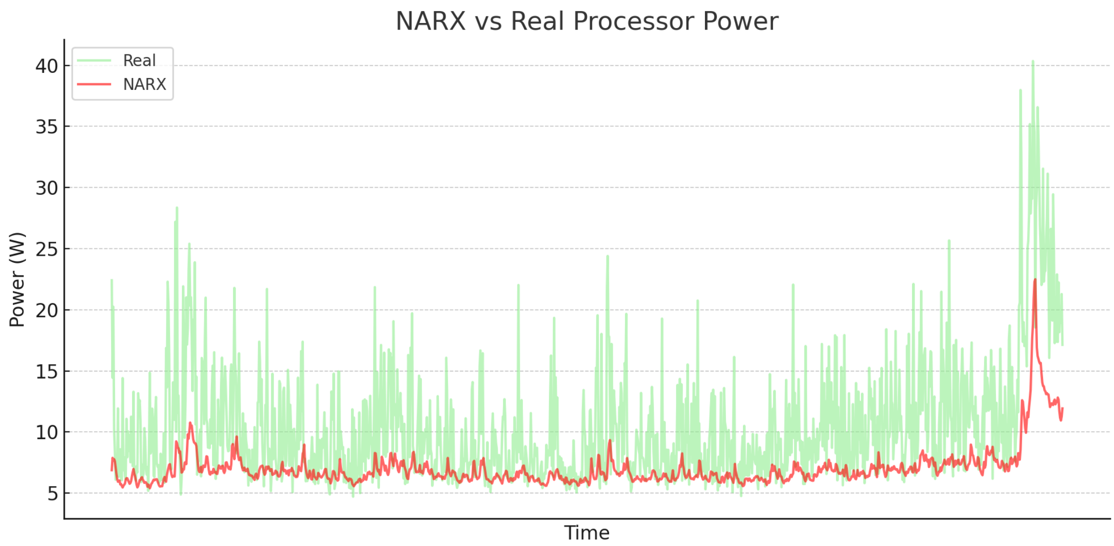


Figure 2.9: Energy load variation over time.

Combined Comparison (Real vs. NARX vs. LSTM vs. XGBoost): In Figure 2.10 you can easily see the differences. XGboost is by far the worst. As it is trying to follow the spikes. That means it also misses a lot of the spikes completely. Its easily shown the NARX at the bottom does not follow the real values as close as LSTM. All the predictions are shifted under the real value, this might suggest as previously stated that the correlation in the different datasets for training and testing is a little different. This difference makes a bias dragging the average of the predicted values down.

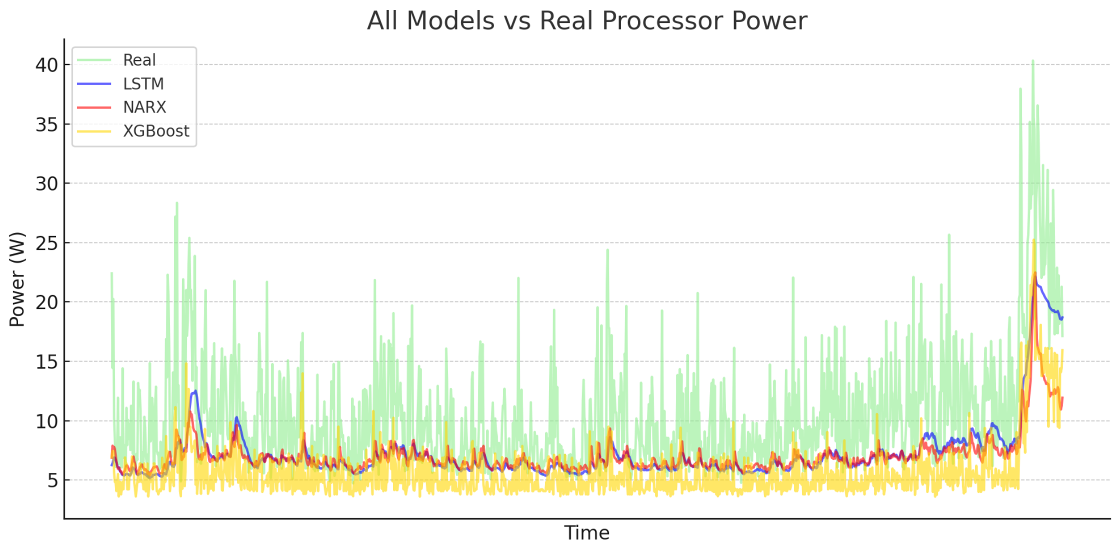


Figure 2.10: Energy load variation over time.

Quantitative Insights

Model	**MSE**	**MAE**	**R^2 Score**
LSTM	25.19	3.52	-0.063
NARX	26.94	3.58	-0.058
XGBoost	35.25	4.84	-0.379

Table 2.1: Performance metrics comparison of LSTM, NARX, and XGBoost models.

The Table2.1 show the LSTM performed the best with lowest mean squared error and mean absolute error with the highest R^2.The XGBoost performed the worst with the highest errors. NARX produced intermediate results, providing quite good but less consistently precise results in terms of error magnitude in comparison to LSTM. These findings give real-world insights into the selection of appropriate forecasting methods with regard to expected data volatility and desired accuracy.

Bias	**NARX**			**LSTM**			**XGBoost**		
	MAE	MSE	R^2	MAE	MSE	R^2	MAE	MSE	R^2
1	2.949	20.002	0.156	3.062	20.089	0.152	3.836	25.882	-0.092
2	2.740	16.262	0.314	2.899	16.989	0.283	3.037	19.328	0.185
3	2.797	14.522	0.387	2.985	15.889	0.330	2.612	14.775	0.377
4	3.056	14.782	0.376	3.266	16.788	0.292	2.513	12.221	0.484
5	3.470	17.043	0.281	3.704	19.688	0.169	2.665	11.667	0.508
6	3.999	21.303	0.101	4.256	24.588	-0.037	2.996	13.113	0.447
Best MAE	2.799	16.264	0.361	2.898	16.786	0.292	2.549	12.771	0.500
Best MSE	2.940	14.970	0.412	2.995	15.886	0.330	2.652	11.930	0.533

Table 2.2: Performance metrics for NARX, LSTM, and XGBoost models at various bias levels.

Metric	**Before Bias**	**After Bias**
Bias Applied	0.00	+3.51
Mean Absolute Error (MAE)	3.58	2.56
Mean Squared Error (MSE)	23.83	11.50
Coefficient of Determination (R^2)	0.067	0.550

Table 2.3: Performance of the hybrid model before and after applying optimal bias correction.

Without applying biases LSTM performs just slightly better than NARX, however XGBoost performs noticeably worse. This changes drastically when a bias is applied, XGBoost error is significantly reduced making it seem XGBoost is the best predictor. It was not just XGBoost that needed a bias, the other two models needed bias too. Even the combined model's accuracy was greatly increased. This might suggest that there is a discrepancy between the training data and the testing data. Some factors that could contribute to those discrepancies are unforeseen loads in one of the data sets and temperature/CPU utilization not correlating 100%. LSTM and NARX models naturally capture patterns over time, giving them an edge in variable contexts. Which XGBoost does not have. Consequently, XGBoost relies on having highly accurate features. XGBoost without memory data would be inadequate compared to NARX and LSTM in a real-world setting or in this case since the date sets have some discrepancies.

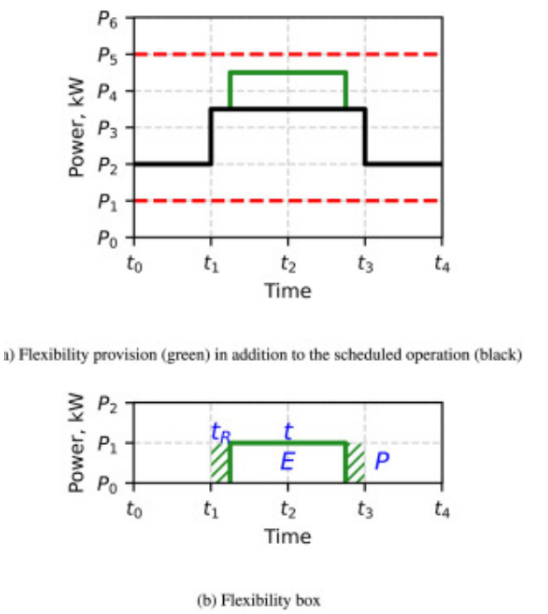
2.3 Energy Flexibility

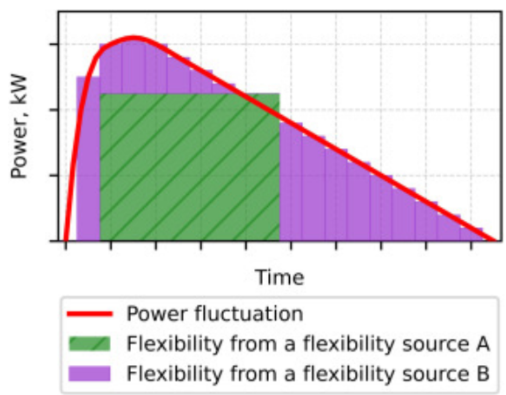
Energy systems in the future will be built on connected subsystems. These subsystems can be multiple things, for example buildings, industry or data centers. The main characteristic of a subsystem is the ability to decide their internal system and loads within their boundary.

The task of flexibility is to recognize fluctuations in energy, for both consumption and generation. After the fluctuations are recognized the use of utilities for compensating are crucial. The ability to equalize the fluctuations is called short-term flexibility. If the fluctuations can be forecasted, then a plan for the energy is created. The flexibility that remains can be used as a marketing chip, for example you can do business with other markets.

Figure 2.11a demonstrates a scenario where additional flexibility (green line) is provided beyond the scheduled operation (black line), within a defined upper and lower power boundary (red dashed lines). This highlights how an energy cell can increase or decrease its power consumption in response to fluctuations. Image description inspired by (32).

Figure 2.11b, on the other hand, shows how different sources can contribute to balancing fluctuating power demands. The green and purple areas represent flexibility provided by two different sources, working together to smooth the power fluctuation curve (red line). These visualizations form the basis for evaluating flexibility potential in this study. Image description inspired by (32).





Flexibility provision (green) in addition to the scheduled operation (black).Image taken from (32)

Flexibility from multiple sources balancing power fluctuations.Image taken from (32)

Figure 2.11: Illustrations of power flexibility modeling and provision

The existing requirement include variability and also uncertainty of demand. On the other side we have the additional flexibility requirement which include the variability and uncertainty of renewable energy output. Calculating the amount of power needed for the power grid using values from the forecast of demand and energy predictions, as the forecast error caused unexpected power ramps that have to be equalized. A result of this is using the applied power ramps in net load to measure the required flexibility in power systems.

2.3.1 Flexibility potential of Data Centers

Data Centers need stable and efficient energy supply to maintain operational. Modern energy systems are more variable due to fluctuations in demand and because of the renewable energy supplies. To face these challenges, flexibility methods like batteries, solar panels and diesel generators are implemented to better control the power consumption.

These components give different types of flexibility. Batteries offer temporal flexibility by storing energy. Solar systems use the supply-side flexibility. Diesel generators serve as backup power but can also contribute to load balancing. These systems are often integrated with each other, which makes the information on how they can work important.

Diesel generators are normally used as a emergency backup system. This is for the running of data centers if a grid failure happens. But in a flexibility view, the backup generators can be used to offset the peak grid demand. It can also be used to provide a stable energy source during high prices. While they are not environmentally good, they create a easy way to provide supply-side flexibility. (33)

BESS enables temporal flexibility by storing the excess energy from the grid or other renewable energy sources and discharge when electricity prices are high or the peak demand is high. The BESS creates an opportunity for uses of different methods like peak shaving, load balancing and contribution to grid stability. (34)

Solar photovoltaic systems create different types of flexibility, both temporal flexibility and supply side flexibility are potential types of uses. By supplying energy to the facilities during sunlight hours, they can reduce the grid dependence and also decrease emissions. By using it during sunlight hours it uses the supply side flexibility. However the solar variability requires either a real time control or a storage support for max effectiveness. (35)

Different hybrids are also possible to maximize the different flexibility methods. One hybrid is a solar panel with battery storage. By using this hybrid you maximize the effectiveness of the solar panels which increases the flexibility. By using a renewable energy supply with a storage system we can store the excess solar energy and use it when needed. This increases the self consumption potential but also the load shifting potential while reducing the grid reliance.(36)

Another hybrid is based on the same concepts and theory but instead of using a solar panel, we could use a diesel generator. The battery in this hybrid handles short term fluctuation and peak shaving. The generator provides a backup or extended coverage during long outages or high loads. This setup creates a balanced way of flexibility which is much easier to control than a solar-battery hybrid, but this hybrid will not be used in this study.

The usage of different flexibility technology shows the pathway to create a more sustainable data center. Each system has its own positive and negative sides in terms of cost, environmental impact and complexity.

2.3.2 Selection of Solar Panel and Battery

Battery Sizing Justification

The BESS capacity is an important value for the theories to be tested. The capacity must meet the expected energy demand during 4 peak load hours. It is also important to account for the efficiency losses and the depth of discharge(amount of usable energy). A common rule of thumb is to determine the battery size based on peak load duration and then account for the efficiency and depth of discharge to ensure longevity and real world credibility.

Sizing Formula

The battery capacity is calculated as:

$$\text{Battery Capacity} = \frac{\text{Required Energy}}{\eta \cdot \text{DoD}} \tag{2.8}$$

Where:

- Required Energy: Energy needed during peak demand (kWh)
- η: Battery efficiency (assumed 90%)
- DoD: Depth of discharge (assumed 80%)

Application Example

Assuming the system must meet a peak load of 150 kW over a 4-hour period:

$$\text{Required Energy} = 150\,\text{kW} \times 4\,\text{h} = 600\,\text{kWh}$$

$$\text{Battery Capacity} = \frac{600}{0.9 \cdot 0.8} \approx 833\,\text{kWh}$$

To account for sizing margins and real world credibility, this value is rounded up to **900 kWh**.

Chosen Battery Configuration

Figure 2.12: Battery energy storage system (BESS). Image taken from (2)
.

A 900 kWh battery energy storage system (BESS) was selected to meet the operational flexibility needs of the data center. The system is a modular and scalable solution designed for commercial and industrial applications. The reference (2) system offers integrated power electronics, energy management software, and safety controls, making it particularly suitable for environments with fluctuating loads and integrated renewable energy.
Key advantages:

- **Scalability**: The system allows future capacity expansion without major redesigns because it scales from 250 to 1000 kwh. this scalability gives room for growth in either load demand or a possible implementation of solar generation.
- **High Efficiency**: With peak efficiency near 98.8%, it aligns well with the assumptions used in sizing of 90% efficiency.
- **Commercial Readiness**: The system is proven in larger scale applications and supported by a global manufacturer with extensive experience in electrical infrastructure.
- **Safety and Longevity**: The system is also equipped with advanced battery management, thermal control, and fire protection systems, making it safe for continuous operation in critical infrastructure such as a data center. Even though this part of the battery information will be irrelevant it is still important qualities to note for real world application of the BESS.

Justification for Solar Panel Selection

Figure 2.13: Ground-mounted commercial solar panel. Image sources from (37)

The reference design (3) for the solar panel was selected based on efficiency, modularity, and commercial relevance. Solar panels generally have fewer adjustable operational parameters than battery systems. Still the choice of module still plays a significant role in estimating energy output and system scalability.

This model offers a competitive efficiency of 20.4%, which is suitable for maximizing output given the limited rooftop area of the data center. Its compact structure allows for a dense placement while maintaining enough space for thermal management and maintenance . The 430W power rating ensures a high energy per panel, which supports the data center's energy offset goals. This module is also used in commercial applications, which improves the credibility of the simulation results.

Still there are more advanced models with higher ,but this panel was selected for its balance between performance, availability, and relevance to real-world usage. The panel's performance profile also aligns well with the selected BESS (2)supporting efficient hybrid system simulations.

2.4 Different Types of Flexibility

2.4.1 Demand Side Flexibility

Theoretical Background:

Demand side flexibility is the ability of energy consumers to adjust their energy consumption in a fluctuating response to external factors. External factors like electricity prices, grid constraints and renewable energy availability are important focus areas. Demand side flexibility focuses on regulating energy demand to be able to optimize the efficiency in a system and reduce peak loads.

Key Strategies within Demand Side Flexibility:

- Load shifting: Moving energy consumption from peak periods to off-peak periods.
- Storage flexibility: Using batteries or energy
- storage systems to offset grid consumption.

Data Centers are an incredibly high energy consuming facility with a relatively predictable workload. this makes data centers well suited for load shifting, which makes it one of the most effective strategies. By dynamically scheduling the workloads or using other methods like optimizing cooling and storage solutions. Data Centers can decrease the cost of the electricity used and also reduce the strain on the grid. Resulting in a more energy efficient facility.

Application of Demand-Side Flexibility in Data Centers:

In a data center environment load shifting can be applied in the following ways:

1. IT Workload Scheduling:
    - Schedule more intensive and computational tasks(for example AI training) during low demand times.
    - Lesser critical worloads can be pre- or post delayed to off peak hours to optimize the electricity use.
2. Cooling System Optimization:
    - Direct Expansion (DX) cooling cycles can be adjusted based on external temperature conditions.
    - Pre-cooling strategies can shift cooling loads to times when electricity prices is cheaper than higher demand times.
3. Energy Storage Integration:
    - UPS systems and batteries can be used to store leftover energy or discharge power during peak grid demand periods.

Equations for Load Shifting Potential:

The effectiveness of load shifting can be shown by using the following equation:

$$BaselineLoad - FlexibleLoad = LoadShiftPotential \tag{2.9}$$

where:

- Baseline Load: The expected power consumption before flexibility adjustments.
- Flexible Load: The part of the load that can be rescheduled or reduced.

Peak demand reduction is calculated as:

$$PeakDemandReduction = BaselineLoad - \frac{FlexibleLoad}{BaselineLoad} \tag{2.10}$$

where:

- Baseline Load: The normal operating power of the data center.
- Flexible Load: The shiftable part of the power demand.

These equations provide a value for assessing the results of demand-side flexibility amounts on energy efficiency and cost savings.

2.4.2 Case study: Apartment Simulation as a Prototype for Data Center Flexibility

To test the functionality of demand side flexibility before applying it to a data center, an apartment model prototype will be used. The apartment simulation is focused on load shifting where household appliances were scheduled based on electricity spot prices during a day. This was primarily used to minimize cost and optimize energy usage. This smaller scale prototype functions as a simple version of the data center model used for a better understanding and realistic view on load shifting. Providing an easy way to analyze the effects of load flexibility before applying it to a data center workload.

Apartment Simulation Results:

The simulation adjusted the operating hours of major household appliances based on electricity pricing fluctuations. The primary findings include:

- Original vs. Flexible Energy Usage:
  - Energy demand was shifted from high-cost hours to lower cost hours.
- Cost Savings by Appliance:
  - High energy appliances such as washing machines, dishwashers, and ovens contributed the most to cost reduction.

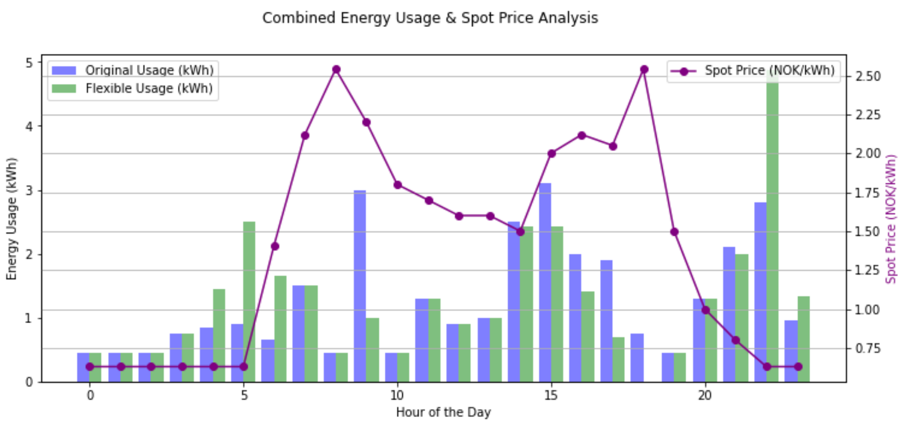


Figure 2.14: Energy Usage and Spot Price Analysis

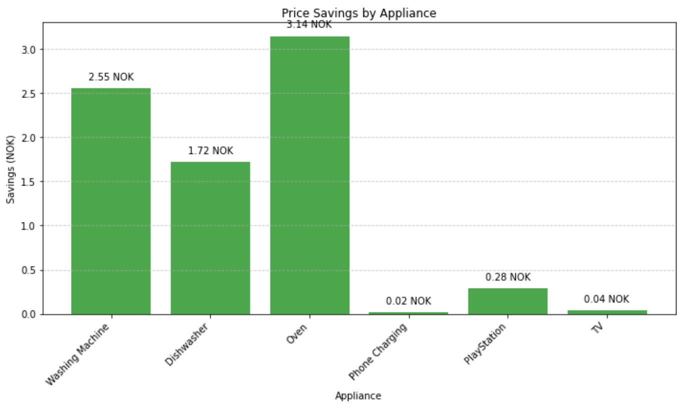


Figure 2.15: Savings by Appliance

Key Savings and Flexibility Outcomes:

- Total Energy Consumption:
  - No reduction in overall consumption. It was only a shift in energy demand.
- Cost Savings:
  - Significant reduction in electricity costs by shifting energy demanding tasks to off peak hours.
- Impact on Peak Load:
  - Peak electricity demand was successfully reduced leading to lower grid strain.

Metric	Value
Total Energy (Original)	30.95 kWh
Total Energy (Flexible)	31.20 kWh
Total Cost (Original)	46.33 NOK
Total Cost (Flexible)	39.09 NOK
Total Savings	7.24 NOK

Table 2.4: Summary of Total Energy and Cost

Appliance	**Original Energy (kWh)**	**Flexible Energy (kWh)**	**Flexible Cost (NOK)**	**Savings (NOK)**
Washing Machine	2.00	2.00	2.04	2.55
Dishwasher	1.20	1.20	0.76	1.72
Oven	2.00	2.00	1.26	3.14
Phone Charging	0.10	0.10	0.06	0.02
PlayStation	0.20	0.20	0.13	0.28
TV	0.15	0.15	0.09	0.04

Table 2.5: Appliance-Level Energy and Cost Breakdown

Comparison: Apartment Model and Data Center Model

While the apartment model effectively demonstrates how flexibility can reduce costs and shift load, there are key differences when applying the same approach to data centers:

Feature	**Apartment Simulation**	**Data Center Application**
Control Mechanism	User-driven (manual scheduling of appliances)	Automated (AI-based workload shifting, cooling control)
Flexibility Type	Primarily demand-side flexibility (load shifting)	Combination of demand-side, temporal, and spatial flexibility
Storage Capacity	Limited (no large-scale batteries)	Can integrate UPS and battery storage for temporal flexibility
Scaling Potential	Small-scale, individual household impact	Large-scale, real-time energy optimization across multiple systems

Table 2.6: Comparison of Flexibility Between Apartment Simulation and Data Center Application

The apartment model simulation shows the effectiveness of load shifting and therefore demand side flexibility. This prototype offers an understanding of energy consumption in a more change affected environment. However data centers require a more advanced automated approach to be able to implement flexibility on a higher scale.

2.4.3 Case study: Load Shifting using Forecasting Models

The strategy for the load shifting was designed based on the trends observed in the forecasted load profile. Load is shifted away during time periods where the forecast indicates an upward trend during peak demand. By trying to anticipate these trends, non critical loads will be preemptively shifted to avoid peak consumption. When it indicates a downward trend the load shifting will not commence, because of the assumption that it won't reach peak consumption. This strategy will prevent unnecessary shifting when the trends are upward or downward, because the shifting won't happen unless it is nearing peak demand. The shift will always be fixed in size because of our lack of testability and our lack of determining the non critical load size.

By using the theory discussed before, the use of forecasting models is fitting for load shifting. As later concluded in this study, NARX is the most fitting model for this purpose. By using this model we will se trends that make the baseline load shiftable.

2.4.4 Temporal Flexibility

Theoretical Background

Temporal flexibility refers to the ability to store and shift energy use over time to optimize efficiency and cost. Temporal flexibility relies on an energy storage system to create the ability to shift consumption dynamically over hours. The two primary mechanism of temporal flexibility are:

- Energy storage: Storing electricity when there is an energy surplus.For example a renewable energy supply or low cost hours. This surplus can be used at a later time when demand or cost is higher.
- Delayed consumption: Based on load forecasting it is possible to adjust power usage based on the forecasted energy availability.

For data centers the most effective temporal flexibility strategy revolves around battery systems like UPS batteries and BESS or grid connected energy storage.

Application of Temporal Flexibility in data centers

Data Centers can use temporal flexibility like this:

1. Battery storage for load shifting:
   - Data Centers can charge the battery when electricity costs is low
   - The batteries can discharge stored energy during peak usage hours.
2. Energy business and Cost optimization:
   - By using the pricing market, data centers can store energy at low prices which can be used during the high cost periods
3. Peak decrease and grid stability consumption:
   - UPS and battery storage can reduce peak loads by supplying power needs during peak times.

By using these methods data centers contribute to grid stability and energy efficiency by responding to dynamic fluctuations in supply.

Equations for Battery Storage

The usable energy in a battery over time is defined by:

$$EnergyStored = P_{charge} \cdot \eta \cdot t_{charge} \tag{2.11}$$

- $EnergyStored$ = Total energy stored (Wh)
- P_{charge} = Charging power (W)
- t_{charge}= Time the battery is charged (h)
- η = Battery efficiency

The energy available for discharge is:

$$E_{discharge} = P_{discharge} \cdot t_{discharge} \cdot \eta \tag{2.12}$$

Where:

- $E_{discharge}$=Total energy available for use (Wh)
- $P_{discharge}$=Discharging power (W)
- $t_{discharge}$=Time the battery is used (h)
- η = Battery efficiency

The battery flexibility potential can be measured by:

$$BatteryFlexibility = \frac{E_{discharge}}{BaselineLoad} \tag{2.13}$$

Equation 2.13 shows an estimate of the amount that can be shifted of the data center power demand.
Temporal Flexibility made by the battery storage systems allows data centers to strategically manage energy consumption. By charging and discharging ,based on needs like grid conditions and load forecasts, data center can reduce strain on the grid.

2.4.5 Case study: Battery Model

To be able to implement temporal flexibility in the data center model a battery energy storage system (BESS) will be used as a core component of load management. The battery functions as a buffer storing surplus energy in low demand or low price and discharges during peak hours or prices. This ability allows the system to smooth load profiles and decrease dependence on the grid.

For this model a reference battery(batterykilde) will be used. The system uses lithium iron phosphate and achieves a peak round trip efficiency of 98.8% making it well suited for high cycle environments like data centers.

2.4.6 Solar Energy

Theoretical Background

Solar energy plays an important role in increasing supply side and demand side flexibility. Solar panels differ from other power sources because of its dependency on weather, which in turn creates the need for storage to fully leverage its potential.
Flexibility in solar energy integration is categorized by:

- Self-consumption Optimization: Maximizing the use of solar energy instead of taking from grid.
- Grid interaction: Giving surplus solar energy into the grid or adjusting usage based on solar availability.
- Storage based: Using batteries to store excess solar energy for use during peaks.

For data centers to use solar energy integration, it needs to balance variable solar power production with stable IT load demands. this balance is achieved through real time energy management, battery storage and load forecasting.

Application of Solar Flexibility in Data Centers

Data Centers can integrate solar energy through:

1. Direct solar utilization for IT and cooling loads:
   - Power IT equipment and CRAC systems directly through solar power.
2. Solar driven load shifting.
   - Scheduling high energy computing tasks when solar production is high.
3. Solar storage systems:
   - Surplus solar energy can be stored in batteries to be used during peak hours.
4. Grid export:
   - If there is energy surplus, data centers can sell the energy back to the grid.

By using these methods data centers can reduce grid strain and carbon emissions while maintaining operational.

Equations for Solar Energy Utilization in Data Centers

The total solar power generated is:

$$P_{solar} = A_{panel} \cdot I_{solar} \cdot \eta \tag{2.14}$$

Where:

- P_{solar} = Solar power output (W)
- A_{panel} = Total panel area (m^2)
- I_{solar} = Solar irradiance ($\mathrm{W/m}^2$)
- η = Panel efficiency (

The effective solar energy contribution is:

$$E_{solar} = P_{solar} \cdot t_{sunlight} \tag{2.15}$$

where:

- E_{solar} = Total usable solar energy (Wh)
- $t_{sunlight}$ = Duration of effective sunlight (h)

The SCR(self consumption ratio(how much is directly used)) is:

$$SCR = \frac{Esolarused}{Esolarproduced} \tag{2.16}$$

where:

- Esolarused = Solar energy consumed by the data center.
- Esplarproduced=Total solar energy generated.

This ratio is important because it gives a value on how efficiently the data center uses the solar energy.

Solar energy when combined with battery storage and load shifting increase the data centers flexibility by giving it a renewable energy source, that can be managed dynamically. By optimizing the solar self consumption and the grid interaction, data centers can reduce its strain on the grid.

2.4.7 Case Study: Solar Panel

For testing the solar energy and its impact on flexibility a prototype solar simulation is developed. this prototype is used as a simplified representation of how solar panel energy could be implemented in a larger data center. Instead of modeling a facility scale photovoltaic system, the simulation uses a real world solar panel as a reference to estimate flexibility contribution and energy production.

The chosen reference panel (3) is a high-efficiency monocrystalline module commonly used in both residential and commercial settings.

This simulation uses irradiance values from a irradiance measurement tool (38) over a 24 hour period to model the panels energy production. This is used to:

- Direct usage during solar hours
- Potential to reduce peak demand or interact with the grid

Time (Hour)	**Global Irradiance $G(i)$ [W/m^2]**
07:45	14
08:45	104
09:45	222
10:45	315
11:45	381
12:45	426
13:45	404
14:45	384
15:45	324
16:45	239
17:45	131
18:45	23

Table 2.7: Hourly Global Irradiance on a Fixed Plane March Day (PVGIS Data, Latitude 58.897, Longitude 5.652) (38)

2.4.8 Hybrid Model

Theoretical Background

Solar panels offer a valuable source of clean and renewable energy. The solar panel production is variable/unstable and is limited by irradiance conditions and daylight hours. This variability leads to a mismatch between energy supply and the constant energy demand of a data center. During periods of high solar output the data center might not require all the energy produced. On the other hand during peak IT load hours solar generation could not be providing a sufficient amount of energy.

To fix this issue between production and consumption, battery energy storage systems (BESS) play a significant role. Using the battery we create an opportunity to store surplus energy generated during hours where instant usage is inefficient. The storage can later be used during peak hours or when solar generation is unavailable. This ability makes variable solar energy into a dispatchable source, which improves the self consumption rate and also the overall flexibility.

Application of Hybrid Model in Data Center

The integration of solar panels and battery storage creates a hybrid system capable of delivering both supply-side and temporal flexibility:

- Supply-Side Flexibility: Solar panels reduce grid dependency by supplying energy directly during solar hours.
- Temporal Flexibility: Batteries enable load shifting by redistributing solar energy to match high demand hours.

This combination is very important to a tier 1 data center like the one used in this project because the maximum IT load capacity is low. The low capacity creates an opportunity for peak IT load reduction, during non-sun hours. The data centers low capacity is the main reason an implementation of a battery-solar hybrid is important.

Equation for hybrid model

$$SCR_{\text{with storage}} = \frac{E_{\text{solar-used}} + E_{\text{solar-stored}\rightarrow\text{used}}}{E_{\text{solar-generated}}} \tag{2.17}$$

Where:

- $E_{\text{solar-used}}$ is the solar energy directly consumed by the data center.
- $E_{\text{solar-stored}\rightarrow\text{used}}$ is the portion of solar energy stored in the battery and later used.
- $E_{\text{solar-generated}}$ is the total solar energy produced by the system.

Without the battery storage, much of the solar energy that is not instantly used by the data center is lost or exported with low financial return. With battery support the energy can be used later. This support increases the flexibility and reduces the grid dependence. Also with an economical view the battery storage helps maximize return on investment.

2.4.9 Case study: Hybrid model

The Hybrid model's battery uses exclusively the solar surplus energy to be charged. No grid energy is being consumed as in the battery model. This is to focus on the hybrid interaction between solar and BESS. This strategy therefore has a focus on maximizing self consumption ratio and reducing dependency on the grid.

To evaluate the effectiveness of integrating solar energy and battery storage in a data center environment, a simulation was created using a realistic combination of available solar panels and a high capacity battery system.

Table 2.8: Summary of Solar Panel and Battery Storage Specifications

Number of Panels	40 units
Total Installed Capacity	17.2 kWh
Panel Efficiency	20.4%
Usable Battery Capacity	600 kWh
Round-Trip Efficiency	98.8%
Battery Chemistry	Lithium Iron Phosphate ($LiFePO_4$)

The simulation tracked solar production and battery usage over a 24-hour period to:

- **Demonstrate flexibility through energy storage:** The battery enables temporal flexibility by extending the usefulness of solar energy into non-sunlight hours.
- **Support demand-side management:** Stored solar energy is used to reduce grid power consumption during peak load periods.
- **Increase self-consumption:** The hybrid system enhances the self-consumption ratio of generated solar power, making the data center more energy autonomous.
- **Simulate realistic hybrid scenarios:** Using real components allows the simulation to reflect practical implementation conditions.

By combining renewable generation with storage, this case study allows the methodology to evaluate the true potential of solar energy as a controllable and strategic flexibility source in modern data center design.

Chapter 3

Results

3.1 Power Consumption Analysis of the Data Center in Simulink MATLAB

Figure 3.1 illustrates the total power used by the data center. In the first graph, the data calculated in Excel show that the total power peaked at 149.11 kW at 15:00 and then fell to 87,235 kW for 23:00. By comparing this with Matlab's graph, measurements show slightly different values, with power reaching 154 kW at 15:00 and falling to 89 kW at 23:00. This is due to the variations implemented.

Usually the data center receives energy from the grid. If there is a grid failure, the generators switching panel is automatically activated to provide backup energy, making sure the data center continue to run without interruption. When the grid supply returns, the system smoothly alternates to the grid energy and continues normal operation.

This configuration is designed to make the data center smart and efficient, quickly adapting to energy changes. It helps stable operations, even by switching between backup energy and grid energy. The simulation demonstrates how the system deals with these transitions without problems, without causing interruptions.
Next to the Figure3.1, 3.2 compares the total power measured in the transformer within 24 hours with the expected energy calculated in Excel. This configuration does not use three-phase dynamic loads, making the operation simpler as it focuses on stability.

Table 3.1 clarifies how the system works during grid supply and backup supply. The table provides details about energy use and how the system behaves at different times. The focus of this configuration is to make the data center as efficient as possible, using energy wisely, keeping the operations stable.

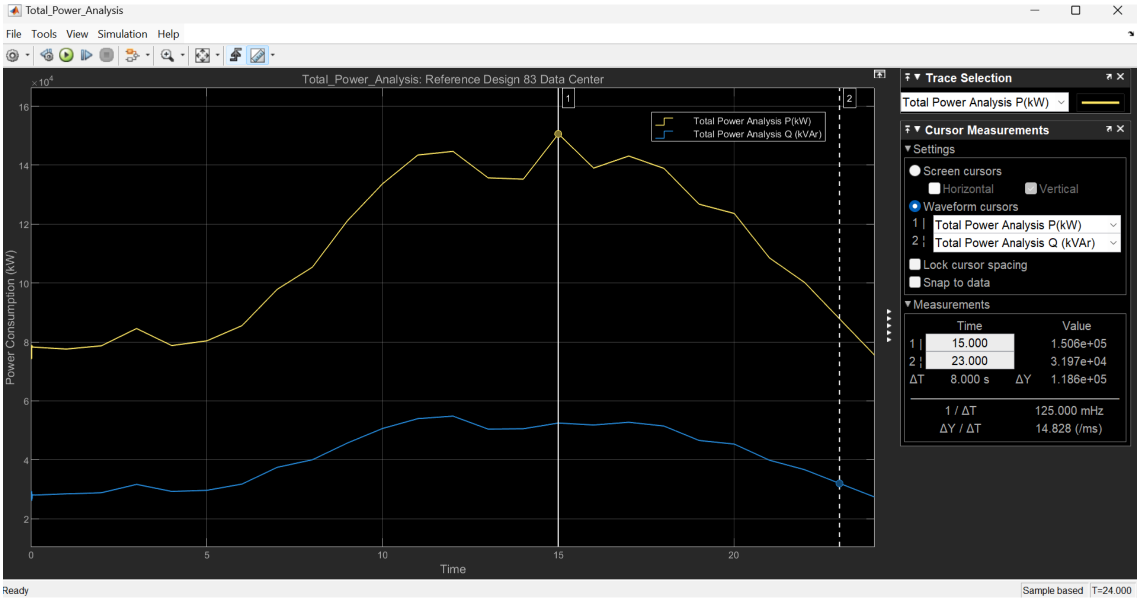


Figure 3.1: Total Power Analysis

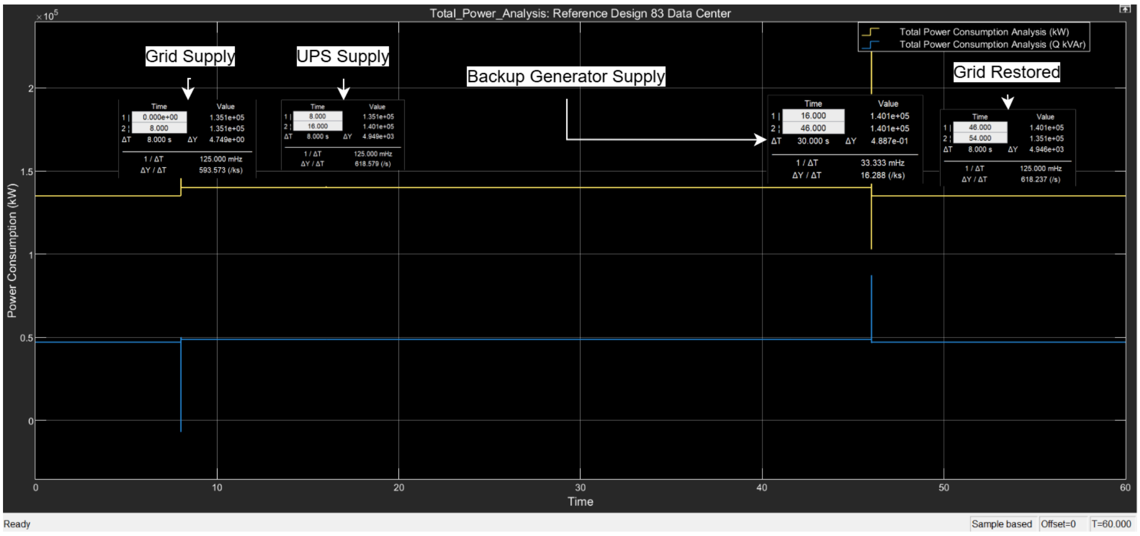


Figure 3.2: Energy Supply switching

Time (s)	Grid Switch	Gen Switch	UPS Switch	Operation
0-8	Closed	Open	Open	Normal Operation
8	Open	Closed	Open	Grid Interrupted; Gen Connect
8-16	Open	Open	Closed	UPS Supply; Gen Disconnect
16	Open	Closed	Open	UPS Disconnect; Gen Connect
16-46	Open	Closed	Open	Generator Supply
46	Closed	Open	Closed	Grid Restored
46-54	Closed	Open	Open	Normal Operation

Table 3.1: Compact Table of Switch States

The table 3.1 explains how an energy system works in different situations, showing the status of three switches, Grid Switch, Generator Switch and UPS Switch and how much energy is provided. Time intervals show when each event happens. The switch ”states” say whether they are connected (closed) or disconnected (open).

- **Normal Operation (0 to 8 and 46–54s):** The grid switch is closed, and the energy comes from the grid. Generator and UPS switches remain open because they are not needed.
- **Grid Interruption (8s):** The grid switch is opened, and the generator switch closes. The backup generator begins to provide energy.
- **UPS Supply (8–16s):** The UPS switch closes briefly to provide energy while the generator prepares to take control.
- **Generator Supply (16–46s):** The generator becomes the main source of energy. The generator switch is closed, and the UPS switch is opened.
- **Grid Restored (46s):** The grid switch closes, and the system goes back to normal operation.

The total energy column (kW) shows how much electricity is provided, with minor changes during transitions, but stable in general. The table 3.1 makes it easy to see how the system adjusts to keep the energy running.

3.2 Comparison of Energy Consumption of AI models

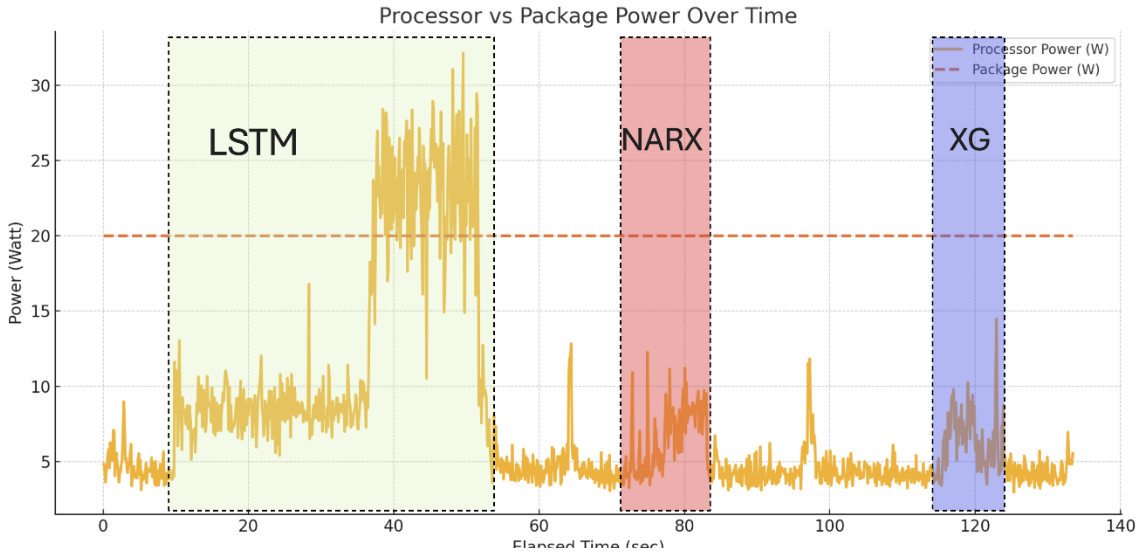


Figure 3.3: Energy load variation over time.

3.2.1 Analysis

The graph 3.3 shows that LSTM consumes much more power than both NARX and XGBoost. Especially during the interval between 10 and 55 seconds where power usage consistently exceeds 20W with some spikes peaking above 30W. The results aligns well with the theory that LSTM networks are more complex and computationally intensive. Therefore with nearly 10000 time-steps in the training, the dataset alone imposes a substantial computational load. Furthermore, LSTM's can not use parallelism because it has sequential processing. That means it cannot take advantage of multi-core execution, which XGBoost can. The number of trainable parameters are also greater, which also contributes to longer training time and more memory usage. It also requires more epochs to converge. In other studies, LSTM has been shown to draw around 100-125W on CPUs, and up to 200W when trained on GPUs (39).

In contrast, NARX is lightweight and efficient. The power usage of NARX remains in the 5W-10W range, between 70 and 85 seconds on the graph 3.3. The results show it needs less computational power and therefore also is a simpler model. This makes it fast to train, with a small memory footprint, and it is also able to predict within a few epochs. In other studies, it typically consumes 5-15W on average, with total energy use under 0.05 kWh so this is within the expectation of previews research (40).

XGBoost is the most power efficient of the three. This is shown in the graph during the 115-127 second window. Here it maintains power usage under 10W with only one spike over, which can be caused by other variables. XGBoost benefits from being parallelizable and making use of more CPU cores. It also trains faster than LSTM due to its tree based structure and uses built-in optimizations to minimize computation. This is because it completes training tasks quickly and does not rely on a GPU. The total energy usage is often under 0.02 kWh per run in other studies (41).

In the simulations of the models a clear constraint was the setup. The constraint cam from not having access to a real data center. However, there are some strengths and similarities. This setup allowed for full control over software and it was easy to measure with Intel Power Gadget. Also relative power trends will likely be the same regardless of scale, especially when comparing CPU-intensive workloads because this was only CPU testing.

In Laptops cooling usually accounts for minor (42) contribution of the power used. That is in stark contrast to data centers as cooling might be around 45% of the IT Load. Thus, the data in Figure 3.3 would have a lag increase when the cooling is on to compensate for the heat created. There is also other components that is not accounted for when using a laptop. Laptops are in addition more prone to throttle when under a high load. As a result it limits the power to the CPU if the heat becomes to high. However, the controlled environment that a PC provided was sufficient for comparing model energy consumption trends and performance in a constrained environment.

The research examined energy usage and predictive outcomes of LSTM, NARX and XGBoost AI models during power prediction applications. The evaluation showed LSTM generated the most precise predictions (MSE: 24.01, MAE: 3.48) while requiring more than 20W and exceeding 30W during operation. The model's sequential processing scheme and extensive trainable parameters and memory requirements lead to high computational expenses which extend training periods.

XGBoost emerged as the most power-efficient model in this study since its power consumption stayed under 10W and training finished under 0.02 kWh. The model displayed the lowest accuracy among the three (MAE: 4.79) thus potentially affecting applications needing exact forecasts.

The NARX model offered an optimal equilibrium between performance and power usage. The model achieved satisfactory accuracy through its MSE value of 28.12 and MAE value of 3.86 while using much less power than LSTM and staying within the 5-10W testing range with total power consumption below 0.05 kWh. The model combines easy training capabilities with low resource requirements alongside satisfactory performance in prediction tasks.

The results from this study indicate NARX should be chosen for data center deployments that require energy efficiency and reasonable forecasting results. NARX provides an excellent middle ground between performance and power efficiency, which makes it suitable for deployment in AI-powered forecasting systems that require minimal energy usage without accuracy loss.

While this project could use more tuning, its core target was to measure and compare power consumption. Which it succeeded at. The results proved to closely follow the litterateurs predictions and it demonstrated that power analysis is possible even on a Local PC.

In later applications, NARX is the most reasonable choice Especially with better parameter choices and more diversified input features, but in many applications the best predictive model of the tree may be a combination of them as that is the most common way to use them.

3.3 Energy Flexibility

This section shows the results of the case study simulations. The results are meant to indicate how different strategies can help manage energy consumption. By reducing peak loads and improving grid interaction. The results include demand-side load shifting, battery based temporal flexibility, solar energy and a hybrid model between solar and battery. These simulations are based on a reference Tier 1 data center and aim to reflect semi-realistic operation.

3.3.1 Load Shifting

The load shifting simulation is used to show how demand side flexibility could be used in a data center environment by shifting the energy consumption to avoid peak load periods. This was done by shifting the load from the high demand hours (12:00–16:00) to non peak hours (03:00, 04:00, 06:00 and 23:00). The shifts are computed by using forecasted and data center loads.

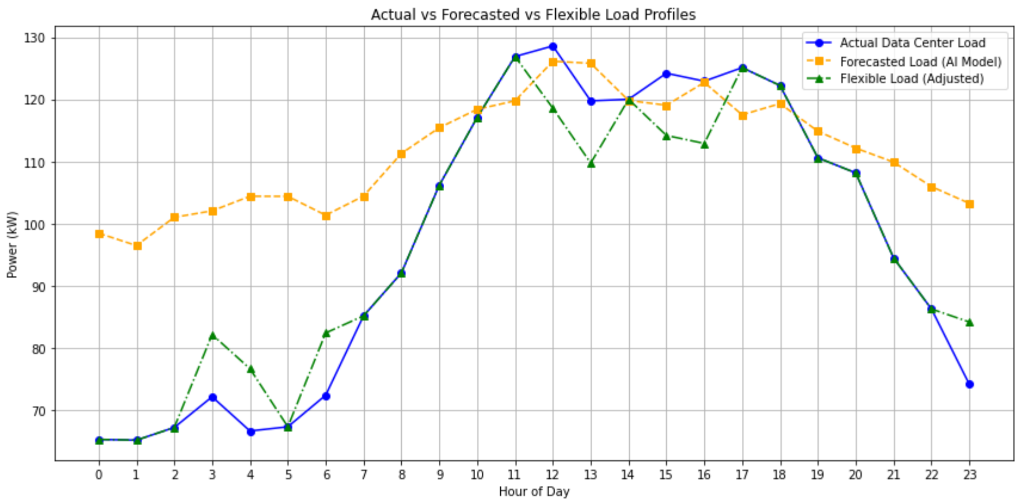


Figure 3.4: Actual, forecasted, and flexible load profiles for a 24-hour period. The flexible profile demonstrates how load can be adjusted in response to predicted peaks.

In Figure 3.4, the adjusted (green) and forecasted (orange) loads are shown with shaded areas representing Watt that have been shifted away and to different time periods. The visual clearly confirms that the peak load between hours 12 and 16 was successfully reduced. This is done while energy consumption at hours 3, 4, 6 and 23 increased slightly. These changes reflect a typical demand response strategy aimed at flattening the load curve.

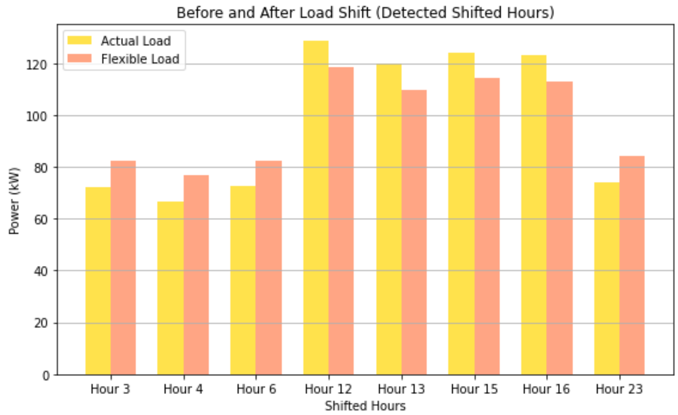


Figure 3.5: Focused comparison of selected hours where load shifting occurred.

Figure 3.5 focuses only on those key hours. It focuses on the drop in power demand between hours 12 and 16 and the resulting increase at hours 3, 4, 6 and 23. This is further highlighting the flexibility visually.

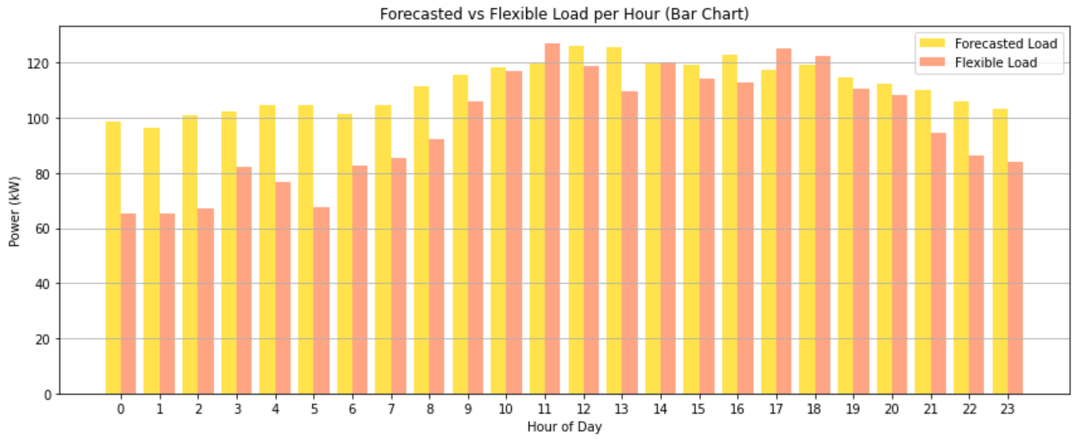


Figure 3.6: Bar chart comparison of forecasted and flexible load across all hours.

Figure 3.6 presents the actual vs. flexible loads for each hour in a bar chart format, making the differences more clear. Especially around hours 3, 4, 6, 16, and 23 the bar height differences highlight the effects of the load shift.

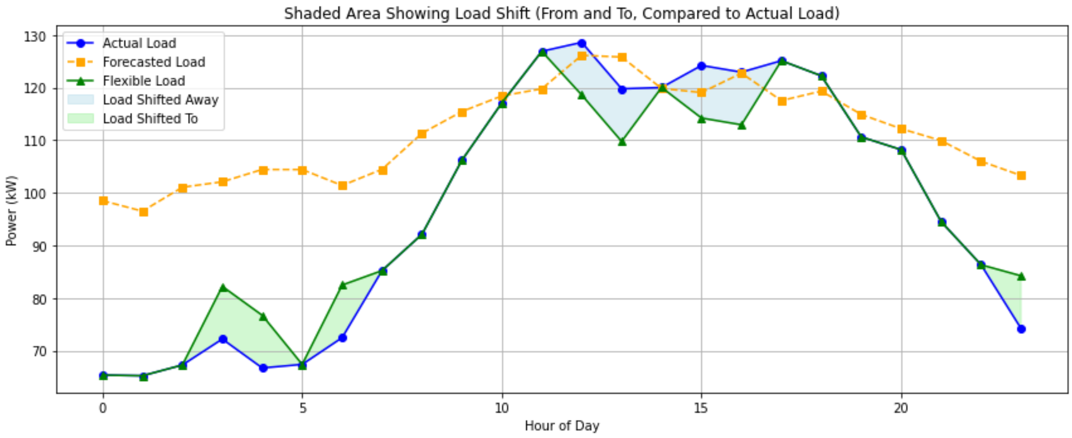


Figure 3.7: Shaded area plot highlighting the energy that was shifted. Blue areas indicate hours where load was reduced, and green areas show where load was added.

Figure 3.7 compares actual, forecasted and flexible loads. It shows clearly that the flexible profile closely follows the actual and forecasted values outside of the most important hours.This validates that load shifting is a controlled and deliberate adjustment rather than a random fluctuation or convenient shifts.

Metric	**Value / Insight**
Original Peak Load (Actual)	127.20 kW
Flexible Peak Load	121.00 kW
Peak Load Reduction	6.20 kW (4.87%)
Total Energy Shifted	38.40 kWh
Main Peak Hours Reduced	Hours 12-16
Load Shifted To (Off-Peak Hours)	Hours 3, 4, 6, 23
Mean Absolute Forecast Error (MAE)	3.75 kW
Mean Absolute Percentage Error (MAPE)	2.11%
Flexibility Strategy	Peak shaving via forecast-based load shifting

These results show that the data center has a clear potential for short term flexibility using simple load shifting strategies. Even though it is limited in absolute energy savings, the reduction in peak demand can lead to an important improvement in grid stability.

This type of flexibility is especially relevant as data centers grow in size and power demand. Particularly in future AI-dependent data centers. Efficient load shifting can help avoid the dependence on highly complex cooling infrastructures and can provide help to the grid.

The load shifting strategy simulated in this project demonstrates how flexible IT loads can be shifted from peak hours to non peak hours to reduce demand during grid stress periods. By shifting 40 kWh from hours 12–16 to hours 3, 4, 6 and 23, a noticeable peak reduction was achieved without sacrificing overall performance. The bar chart comparing selected hours further shows that the system can offload demand from high cost afternoons. The offload then gets put in the low cost hours when demand is typically lower based on the forecasting.

These results aligns with the demand side flexibility (DSF) theory that was used to compute the results. Where adjusting non critical operations during peak hours helps reduce strain on the grid and operational costs.

While the simulation assumes perfect control and predictability of load, which in reality would depend on more advanced AI forecasting and dynamic load scheduling. The impact of load shifting is limited in capacity, because only a portion of their load is shiftable. It still remains a highly cost effective and low difficulty flexibility strategy. Especially when integrated better predictive models that can accurately determine optimal shift windows and quantity.

3.3.2 Battery Simulation

Charging occurs between 00:00 and 04:00 and again between 20:00 and 23:00, when electricity demand and grid prices are typically lower. Discharging takes place from 10:00 to 18:00, which follows the typical afternoon peak in data center activity

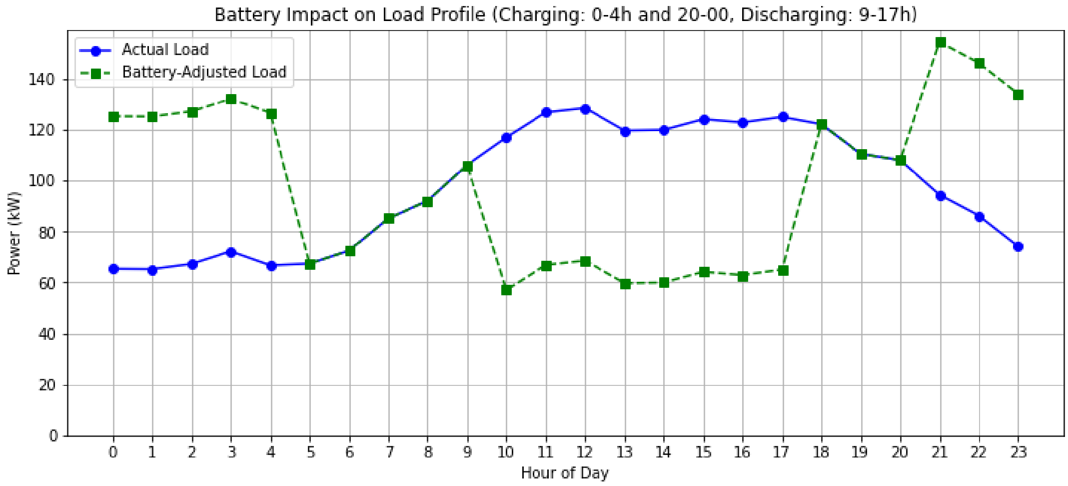


Figure 3.8: Battery impact on the data center load profile. The battery charges between 00:00 and 04:00 and 20:00-23:00 (green spike), slightly increasing load during low-demand hours. It discharges from 10:00 to 18:00, significantly reducing the peak load by up to 60 kW.

The baseline load (blue) in Figure 3.8 is significantly reduced during discharge hours. Where the adjusted load (green) dropps by up to 60 kW per hour. The charging phase slightly increases the load during early morning and late evening hours, which is an efficient strategy since those hours coincide with lower energy costs and lower demand.

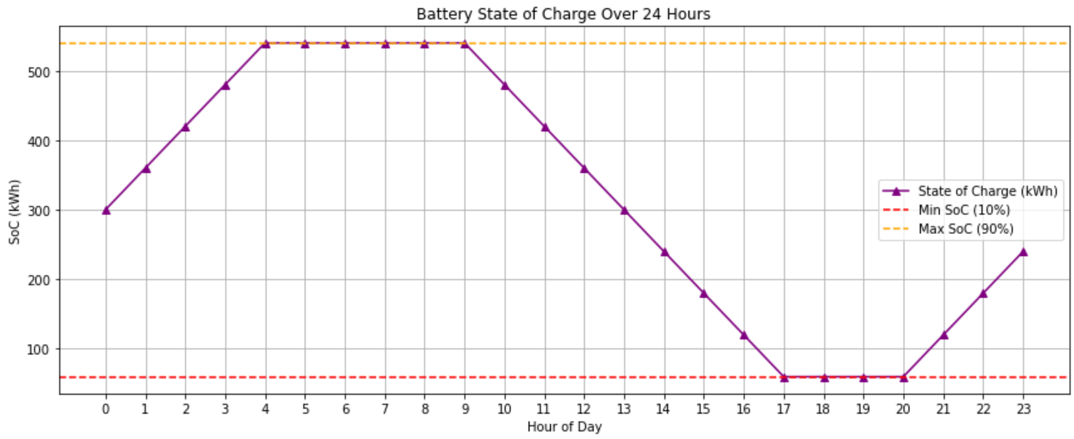


Figure 3.9: Battery State of Charge (SoC) throughout the 24-hour cycle. The battery charges to full during hours 0-4 and 20-23 and discharges during peak hours. The system respects the defined SoC limits of 10-90% for the 600 kWh battery.

The State of Charge (SoC) profile in Figure 3.9 shows the energy flow through the battery across the day. The battery begins at 240 kWh. It reaches full capacity at 04:00 and then discharges gradually during the afternoon peak. The SoC reaches the system's minimum limit of 60 kWh at 18:00 confirming that the battery is used optimally within its safe operational range of 10-90%.

Table 3.2: Battery-Based Temporal Flexibility Summary

Metric	**Value / Insight**
Initial State of Charge (SoC)	300 kWh
Charge Window	00:00-04:00, 20:00-23:00
Discharge Window	10:00-18:00
Max Charge/Discharge Rate	60 kW
Peak Load Reduction Achieved	Up to 60 kW/hour
Total Energy Discharged	480 kWh
SoC After Operation	240 kWh
Flexibility Type	Temporal flexibility

The simulation of the battery energy storage system (BESS) demonstrates how temporal flexibility can be used within the data center to reduce energy consumption during peak hours and optimize energy usage across the day. This simulation applies a fixed charging and discharging schedule using the reference battery (2) with a usable capacity of 600 kWh.

The temporal flexibility application confirms that batteries are well suited for peak shaving in data centers, offering a dynamic method to reduce load on the grid during critical times. It also provides operational benefits by allowing the facility to better align energy usage with grid availability and pricing. The battery's cycle (fully charging and discharging within one day) shows efficient energy usage.

The battery storage using a 600 kWh system with operational SoC constraints (10–90%), shows a huge potential for both load leveling and grid support. Charging was scheduled between hours 0-4 and 20-23 and discharging during peak demand hours (10–18). The charging results in a visible reduction in peak load on the grid.

The battery adjusted load profile proves how non peak energy can be used for peak reduction. The reduction contributed up to 480 kWh in flexibility. The state of charge plot confirms that the battery operated within its safety margins while still providing value to the load profile.

Effectiveness of the battery is tied to its initial SoC and charging window. The battery is therefore based on assumptions which could obstruct a real installment. There are also other factors like oversizing the battery relative to actual shiftable load that may reduce efficiency. Undersizing may fail to deliver desired peak shaving. The degradation and cycle limitations must be considered in real deployments. The results from the simulation are still a good indicator of the theory and potential the BESS has on flexibility.

3.3.3 Solar Simulation

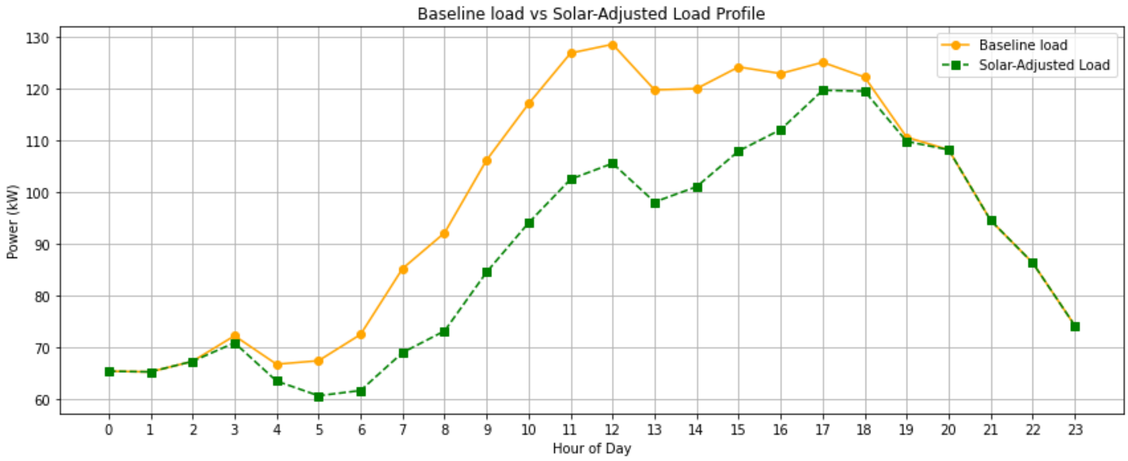


Figure 3.10: Comparison between baseline and solar-adjusted load profiles. During daylight hours, solar generation reduces the facility's grid demand, with a visible offset between 06:00 and 18:00.

Figure 3.10illustrates the difference between the baseline and solar-adjusted load profiles. During peak irradiance a part of the data center's energy demand is covered by on-site solar generation. The baseline load (orange) consistently sits above the adjusted profile (green) from 06:00 to 18:00, indicating a clear offset of grid power by solar energy. This appears in Figure 3.12where the shaded green area represents the total energy saved approximately 226.2 kWh.

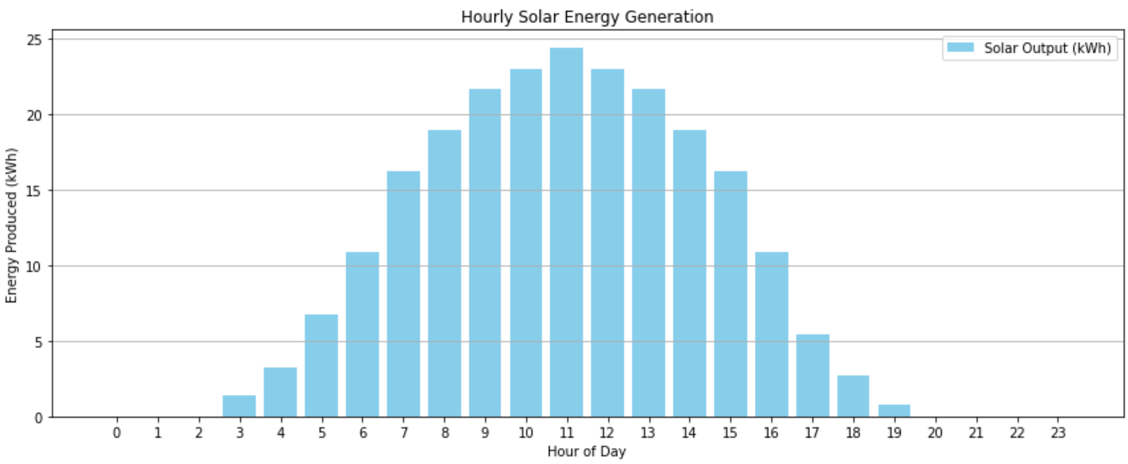


Figure 3.11: Hourly solar energy generation using 40 panels. Peak generation occurs between 10:00 and 13:00, closely matching peak irradiance conditions.

The solar output profile in Figure 3.11confirms this. Peak generation is achieved between 10:00 and 13:00, reaching nearly 25 kWh per hour. This trend follows with expected irradiance curves and supports the logic behind solar being a reliable mid-day flexibility resource.

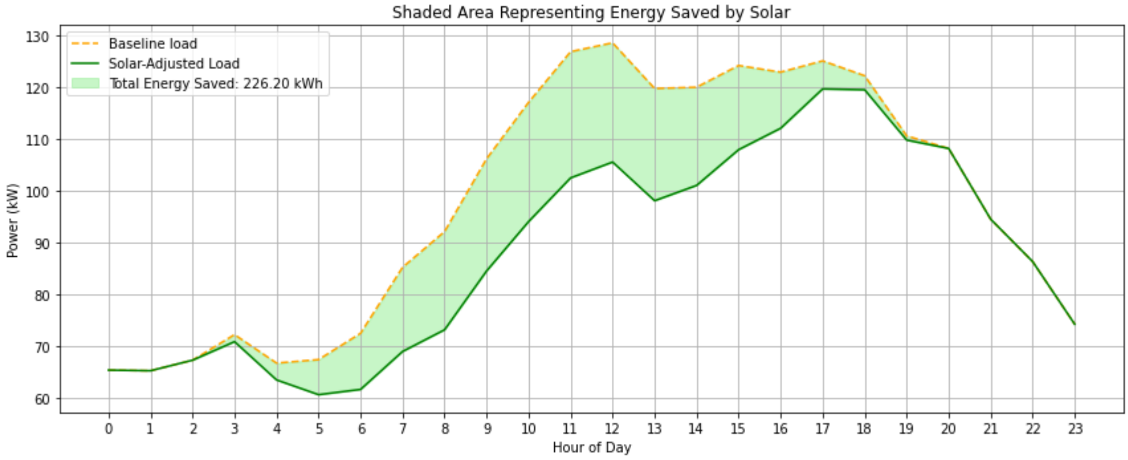


Figure 3.12: Shaded area representing energy saved by solar generation over 24 hours. A total of 226.2 kWh was offset from the grid, significantly lowering daytime demand.

Table 3.3: Summary of Solar Energy System Implementation

Metric	**Value / Insight**
Number of Panels	40
Total Installed Capacity	17.2 kWh
Panel Efficiency	20.4%
Estimated Panel Area	132.4 m^2
Peak Hourly Generation	∼24.5 kWh
Total Daily Solar Generation	226.2 kWh
Energy Offset Method	Direct usage
Load Reduction Window	06:00 – 18:00
Flexibility Type	Renewable / Generation-side flexibility

The self consumption rate in this case is high, because all generated solar energy is consumed directly inside the data center. The energy consumed is used for applications like the IT load or cooling loads. This represents an efficient use of generated power.

These results prove the possibility of solar integration in Tier 1 data centers. Solar energy provides a sustainable flexibility source during daytime hours. When the solar panels are paired with storage solutions, the flexibility can be further extended to critical peak loads or high prices. This in theory enables the data center to reduce its reliance on the grid. Solar energy helps its environmental sustainability especially during peak solar irradiance.

The solar energy model introduced 40 rooftop panels. Solar panels generate a maximum of 24.6 kWh/hour. The solar-adjusted load profile showed significant reduction in power drawn from the grid during the daytime hours. A total of 226.2 kWh was offset using solar generation over the day.

The shaded graph clearly displayed the energy saved due to solar. Especially during the middle of the day peaks. This strategy is environmentally sustainable and scalable which is a big bonus. It can also be used to directly reduce the carbon footprint of the data center. Main constraint lies in solar irradiance and weather conditions as all solar panels .

Limitations are based on the trends of peak loads and the solar energy produced isn't always the same. This creates a possibility of the sun being covered by the sky during peak hours or a situation where the irradiance is peaked and its loads are at a minimum.

These situations can create a huge efficiency problem. So for the solar panels to achieve maximum efficiency an energy storage system is crucial. Another solution could be a coordination between load profiles and irradiance, which would be more limiting and complex. This solution, even though it's complex, will be cheaper than the investment of a storage system.

3.3.4 Hybrid Model Simulation

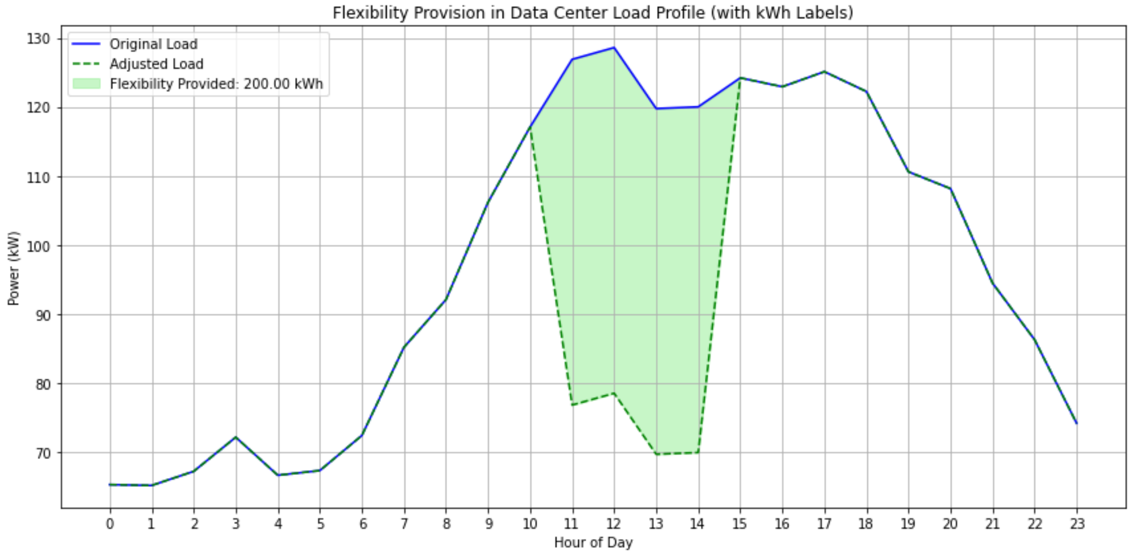


Figure 3.13: Flexibility provision through battery discharge during peak hours (11-14). The green shaded area shows 200 kWh of energy shifted from peak hours, reducing grid stress.

Figure 3.13 shows a clear example of targeted flexibility provision. The adjusted load profile reduces the peak load during hours 11–14 by approximately 200 kWh. This green-shaded area represents a direct reduction in grid demand. This area shows a controlled simulation of a flexibility event.

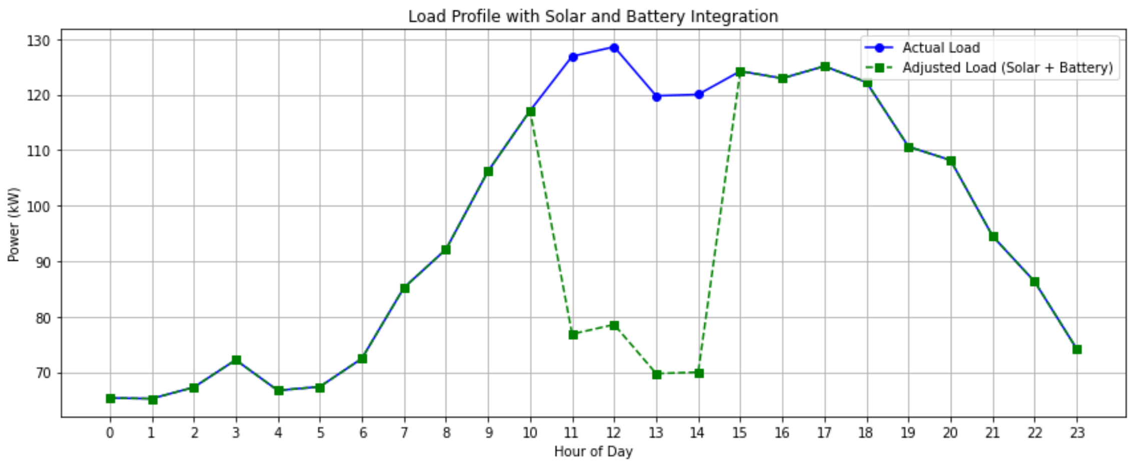


Figure 3.14: Load profile with solar and battery integration. Solar offsets the load during daylight hours, while the battery discharges to further reduce the afternoon peak.

Figure 3.14 shows the solar and battery integration. The adjusted load drops during the middle of the day hours compared to the actual data center load. This comparison confirms that the system is effectively shifting or offsetting load during high-demand periods. This supports the theory that solar energy and stored battery power work together to smooth out the data center's power demand curve.

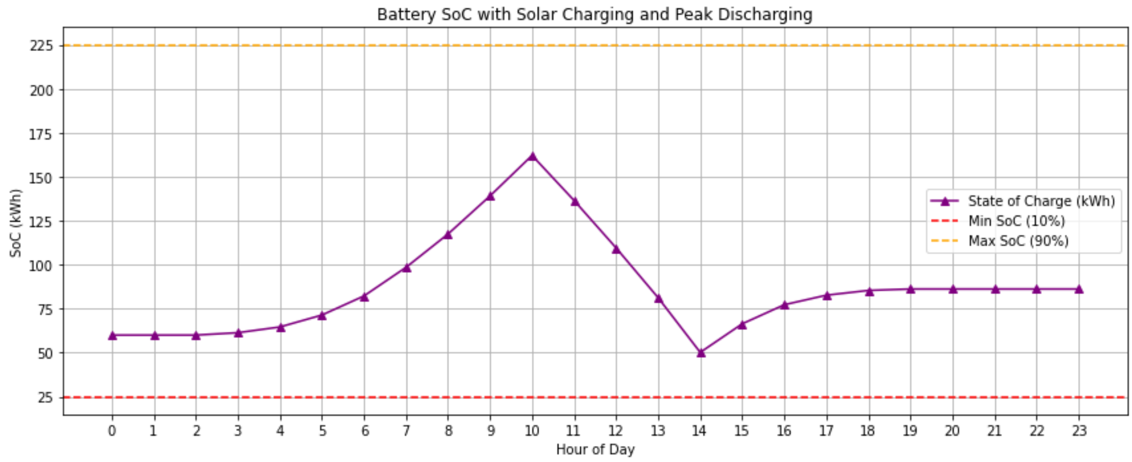


Figure 3.15: Battery State of Charge (SoC) profile. The battery charges during solar peak and discharges during the afternoon. The SoC remains within operational limits (10%-90%).

In Figure 3.15, the focus is specifically on the battery's State of Charge. The battery starts the day at 60 kWh and steadily charges until hour 10. The SoC peaks at approximately 165 kWh.This charging behavior correlates with the increased solar output and low demand conditions. Between the hours 11 to 14, the SoC drops, indicating full battery discharge during the peak load period.

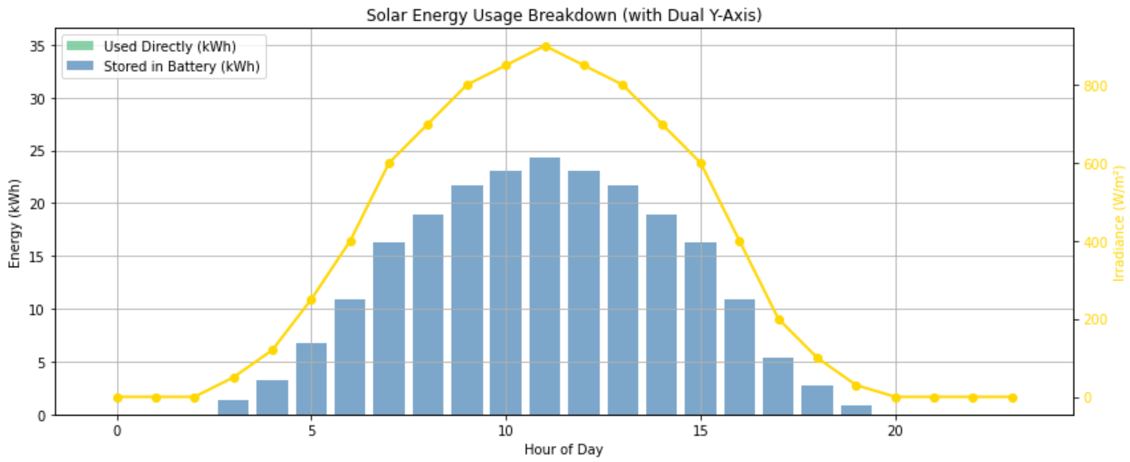


Figure 3.16: Breakdown of solar energy usage. Blue bars show how much solar is used directly vs. stored. The yellow line tracks solar irradiance throughout the day.

Figure 3.16 breaks down how solar energy is used throughout the day. The blue bars represent the energy being stored in the battery. The yellow irradiance curve confirms that solar production peaks between 10:00 and 13:00.

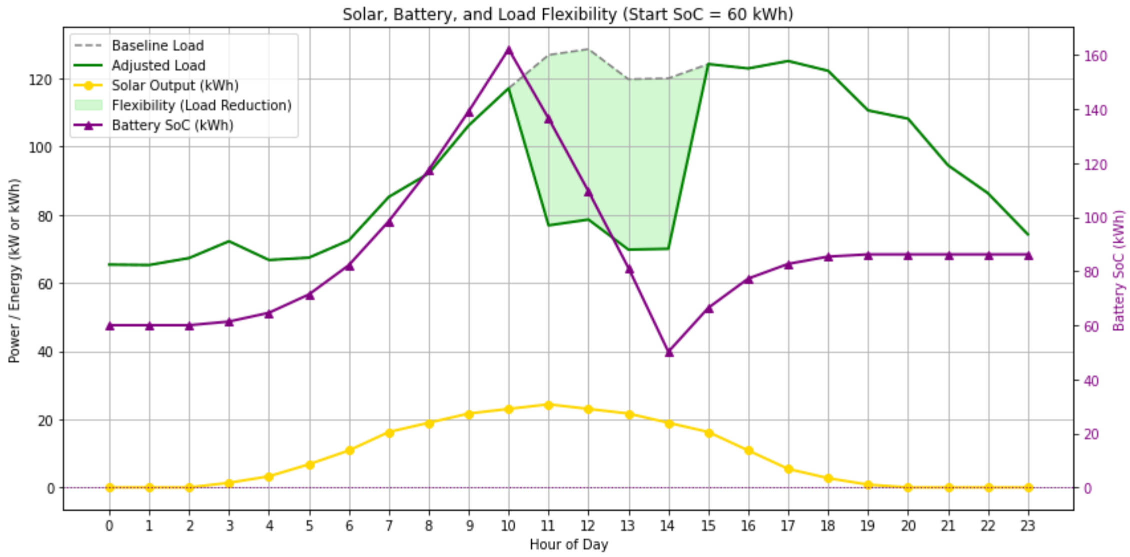


Figure 3.17: Combined plot of solar output, battery SoC, and adjusted load profile. this overview demonstrates the coordination between solar production, battery use, and load reduction.

The combined simulation output in Figure 3.17 shows a clear view of how solar energy and battery storage can work together to improve load flexibility in a Tier 1 data center. At the beginning of the day, the battery starts with a moderate SoC of 60 kWh. As solar irradiance increases between 06:00 and 10:00, the battery begins charging using the solar energy. The charge reaches a peak SoC of approximately 165 kWh by 10:00.

The discharge starts between 11:00 and 14:00. The discharge follows the data center's peak load period. The adjusted load in this period shows a peak load reduction from 128kW to 78kW. The shift ability comes from the combined discharge of solar energy that is stored and also the direct solar input. This usage provides approximately 200 kWh of flexibility. This is visualised as the green shaded area in the plot. This improves self consumption and reduces environmental impact.

The flexibility strategy ensures that the battery never drops below a safe operating threshold. This is to preserve the battery's long-term health and to keep reserve capacity. After the discharge window between 11:00 and 14:00, the battery recharges because the irradiance is still affecting the solar panels and increases to 88kWh at the end of the day.

Table 3.4: Detailed Insights from Hybrid Flexibility Simulation (Figure 3.17)

Metric	Value / Observation
Starting Battery SoC	60 kWh (at 00:00)
Maximum Battery SoC	∼165 kWh (around 10:00)
Minimum Battery SoC	∼40 kWh (at 14:00)
Final Battery SoC	∼88 kWh (at 23:00)
Battery Charging Window	05:00-10:00 (charging from solar excess)
Battery Discharging Window	11:00-14:00 (supporting peak load)
Total Flexibility Provided	∼200 kWh (shaded green area)
Peak Solar Output	∼26 kWh(at 11:00-13:00)
Load Reduction Achieved	From ∼128 kW to ∼78 kW (peak shaving)
Flexibility Type Enabled	Combined temporal (battery) + renewable (solar)
System Coordination	Solar used first; excess stored in battery and later used
Grid Dependence (11-14)	Significantly reduced due to solar + battery discharge
Energy Efficiency	Maintains SoC above 40 kWh while flattening load peaks

The simulation proves that the theory about combining solar energy with battery storage significantly increases the flexibility potential of a Tier 1 data center. By using solar energy and the battery for strategic peak shaving. The figures show that the hybrid model reduces the grid dependence during high demand hours.

The battery creates a controlled discharge during peak demand (11:00-14:00). this resulted in a contribution of 200 kWh of flexibility. Additional contribution comes from that solar energy is either consumed directly or stored in the battery. This improves self consumption and reduces environmental impact.

The system is made to avoid a full battery depletion, because of the battery health. The system also uses already defined State of Charge parameters. These battery charges are based on solar irradiance. The results show that the hybrid model using solar and battery systems is an effective tool for creating hybrid systems, but also is compatible with real-world data center requirements.

The big challenge with the implementation of the hybrid model is the module specifications. The solar panels and battery system used in the model are based on commercial specifications, but without physical testing or hardware setups their results are always going to show trends or be used as a prototype.

Even though the process had some challenges, the simulation still completes its primary task of demonstrating the hybrid strategy in a data center. The results clearly show how the integration between solar panels and BESS can support the theoretical ideas. Insights from the model revolve around mostly how this model can be valuable for planning and it provides a realistic baseline for future models that are to be tested.

The energy amount available for discharge is limited by the solar generation. The BESS used has a capacity of 600kWh and the maximum of SoC is 165 kWh. This leaves untapped potential of 435 kWh. Solutions to this are increasing the number of solar panels, but that would assume the data center owns land which is a stretch for this study. Another solution would be to connect the grid, but as described before the intention was to show the interaction between the solar and battery. In a real world scenario this capacity would be leveraged by grid charging. However, in this study this interaction was intentionally excluded to focus on solar driven flexibility.

Chapter 4

Conclusion

This study has explored the major electrical components in a reference Tier 1 data center while trying to better the energy- efficiency and flexibility. The exploration of major electrical components looked at power use in the data center over 24 hours. The highest power observed was 150 kW at 15.00. Showing how the system handles energy changes during the day. The data center's power management system easily shifts between grid energy and backup energy while keeping stable. Reliability is maintained successfully within the system's designed capabilities. Uninterrupted functionality is achieved by automatic shifts between energy sources. Sources like grid, generator, and UPS are used. The system quickly adapts to interruptions within set parameters.

The energy usage of the three forecasting models with respect to the predictive performance when applied to the forecasting tasks was compared. The comparison between the models showed LSTM produced the best predictive results, but that came a increasing power consumption consequence. This was due to a higher complexity training and memory demands. XGBoost, on the other hand was the least power intensive, but that also was shown in the predicted results as it lacked compared to the other models. NARX executed the forecasting with good predictive results with respect to LSTM, but at the same time using a fraction of the power. This means NARX is the clear choice for energy-constrained environments based on this study A combination of the three might show better predictive results. The power usage of all there models fairly similar to previous studies with when taking in to account the energy constraints of a laptop. The NARX yielded the best predictive results pr watt, which is why the NARX was later used for flexibility purposes.

The flexibility results from the simulations show the different flexibility methods that leverage demand-side, temporal and solar energy. Load shifting using the forecasting model reduces the peak demand by relocation of the peak hours, enhancing grid stability. The battery simulation that a controlled charge and discharge can easily smooth out load profiles, reducing the peak loads. The solar simulation offset the daytime grid demand, providing sustainable power even though there are a lot of variability challenges. The hybrid model offers the highest flexibility potential. Enabling peak reductions, a higher SCR with stronger independence from the grid. This comes as no surprise as it is also the most expensive to facilitate. Together, the strategies clearly show a potential for energy flexibility in a Tier 1 data center.

To conclude this study shows the possibilities of improvements, while maintaining its core principles. The findings from the flexibility models creates a platform for where newer and better technology can prosper. With things like dynamic load management or multi-site energy sharing networks. The AI models used have huge potential for improvements, with reinforced learning the models can improve their forecasting accuracy and enable real-time optimization of the energy usage. In addition to this, huge data centers including solar panel parks can be a possible strategy.

Appendix

Inrow Module Spacing for Sizing

In solar panel system design, the spacing between solar panels is a key factor. This is because it directly affects system performance, which includes light reception, heat dissipation, and maintenance convenience. Proper panel spacing enhances energy efficiency but also extends the system's lifespan. The main reasons are as follows:

- Hot Spot Effect and Thermal Management
- Enhancing System Stability and Safety
- Maintenance and Cleaning:
- Preventing Shadows and Obstructions

For this study the most important factor is Preventing shadows.

Preventing Shadow

During sunrise and sunset, The angle of sunlight is lower and if the spacing between PV panels is insufficient, the front-row panels may cast shadows on the rear-row panels. This results in reducing their power generation efficiency. Properly designed spacing ensures that each panel receives the right solar radiation, while minimizing the negative impact of shadows on solar panels.

To calculate the space needed we will use Figure 1:

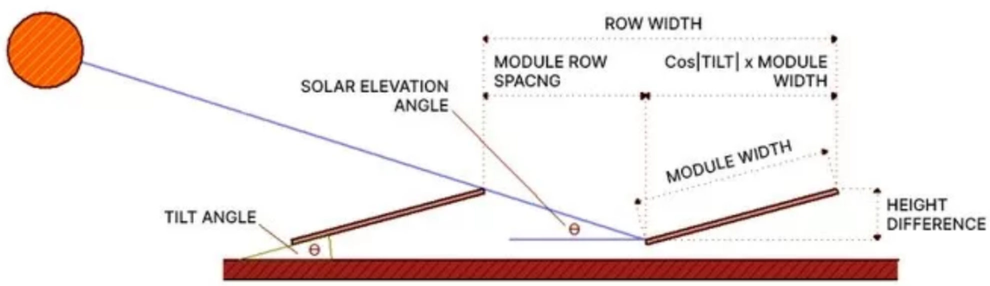


Figure 1: Model image for equations taken from (43)

Cooper Equation(Solar Declination)

$$\delta = 23.45^\circ \cdot \sin\left(\frac{360}{365} \cdot (284 + n)\right) \tag{1}$$

Where:

- n = Day of the year
- δ = Solar declination angle

Solar Hour Angle

$$\omega = 15^\circ \cdot (\text{Solar Time} - 12) \tag{2}$$

This gives the angle of the sun east or west of solar noon.

Optimal Tilt Angle

$$\tan(\beta) = \frac{-\sin(\delta)\cos(\phi)\cos(\gamma) + \cos(\delta)\sin(\phi)\cos(\gamma)\cos(\omega) + \cos(\delta)\sin(\gamma)\sin(\omega)}{\sin(\delta)\sin(\phi) + \cos(\delta)\cos(\phi)\cos(\omega)} \tag{3}$$

Where:

- β = Optimal tilt angle
- ϕ = Latitude
- γ = Surface azimuth angle
- ω = Solar hour angle
- δ = Solar declination angle

Height Difference Between Panel Edges

$$\text{Height Difference} = \sin(\beta) \cdot W \tag{4}$$

Where:

- β = Panel tilt angle
- W = Panel width

Minimum Row Spacing

$$\text{Minimum Module Row Spacing} = \text{Module Row Spacing} \cdot \cos(\text{Azimuth Correction Angle}) \tag{5}$$

Total Row Width

$$\text{Row Width} = \text{Minimum row spacing} + \cos(\beta) \cdot W \tag{6}$$

Where:

- β = Panel tilt angle
- W = Panel width

.1 Case study: Finding Row Space

Finding the tilt angle is the most important part. The winter solstice period is being used as the day because that is the worst case scenario, where the sun is at its lowest illustrated by Figure 2.

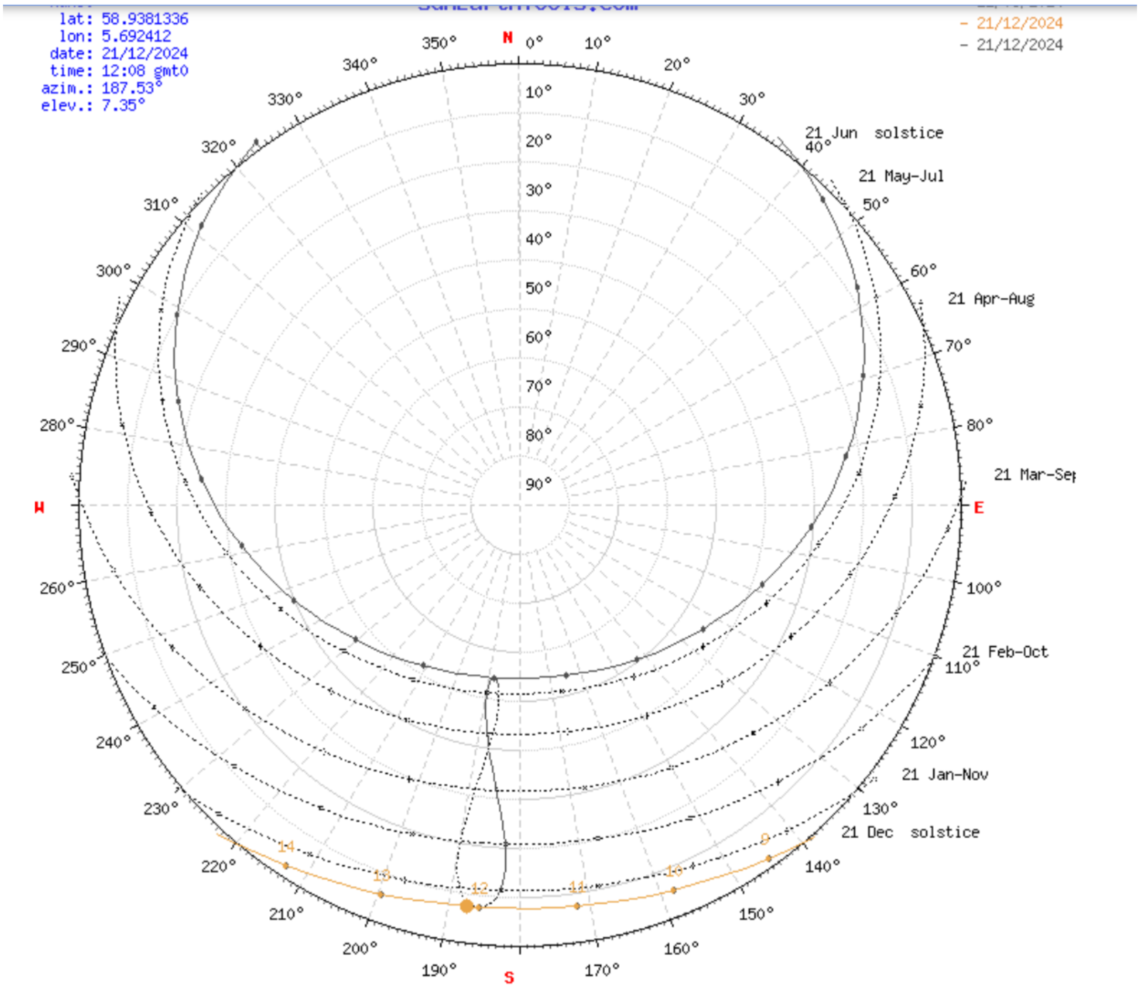


Figure 2: Sunplacement during a year

The solar movement is taken from a sunplacement tool (44). The latitude and longitude angles also need to be determined. Since we are students of UIS the university has been chosen as the place of the data center. The correlation between the sun and place is easily shown by Figure 3:

Figure 3: Clearly shown the latitude and how the sun moves around the university, with the yellow line showing winter solstice and the black line is showing the summer solstice

The solar elevation angle is very important data to have. This is to find the azimuth correction angle which is crucial. The solar elevation angle is found by the sun path chart program (44). Shown by Figure Figure 4:

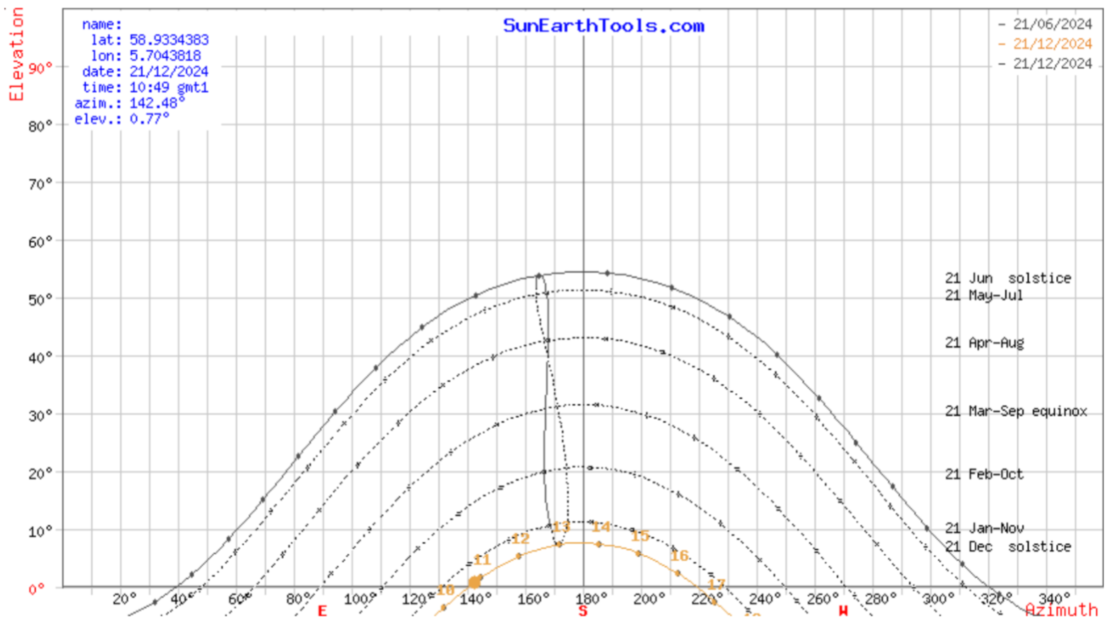


Figure 4: Image of sun path chart. Clearly shows the sun is up from 11 to 1630

By using Figure 4 we know the angle of the sunrise and the sunset. By using these values we can determine the azimuth correction angle. The correction angle is taken from what the azimuth angle is at 11 and at 16, this is found from Table 1:

Table 1: Solar Elevation and Azimuth at Stavanger (58.9334383°N, 5.7043818°E) on 21 December 2024

Hour	Elevation (°)	Azimuth (°)
10:29:27	-0.833	138.3
11:00:00	1.61	144.86
12:00:00	5.3	158.14
13:00:00	7.3	171.82
14:00:00	7.47	185.69
15:00:00	5.79	199.43
16:00:00	2.39	212.79
16:41:18	-0.833	221.7

To be confident of optimal performance and avoiding shading loss, the solar panel array is designed based on the worst case scenario. The worst case scenario is the winter solstice period, where the geometrical condinitions are the worst based on latitude 58.9°N(Stavanger). By using the real azimuth and elevation data, the azimuth correction angle and solar elevation between 11:00 and 16:00 were used to find the optimal tilt angle and minimum space.

Table 2: Solar Geometry and Row Spacing Summary

Parameter	**Value / Description**
Location	Stavanger, Norway (58.9°N, 5.7°E)
Date Used	Winter Solstice - 21 December
Min Solar Elevation (11:00)	~1.6°
Azimuth Range (11:00-16:00)	144.9° to 212.8°
Panel Tilt Angle (β)	81.9°
Panel Width (W)	1.01 m
Height Difference ($H = W \cdot \sin\beta$)	~1.01 m
Azimuth Correction Angle	~35°
Minimum Row Spacing	~30 m

The calculated minimum row spacing of 30 meters ensures complete shading avoidance during the worst case scenario(winter solstice). Since we are stationed in Norway this measurement is not practical and often dismissed in tight areas, especially for roof top mounts. Therefore, the focus is more on the core productive hours and seasons where solar energy is significant. The 2.26m spacing reflects a more efficient use of land, which is important when dealing with a flat 152 square meter roof. This spacing optimizes space usage and amount of panels that fit.

The diagram in Figure 5 shows the solar angle geometry with its values. It also shows how the tilt angle, solar elevation, module width, and calculated row spacing are used to avoid inter-row shading.

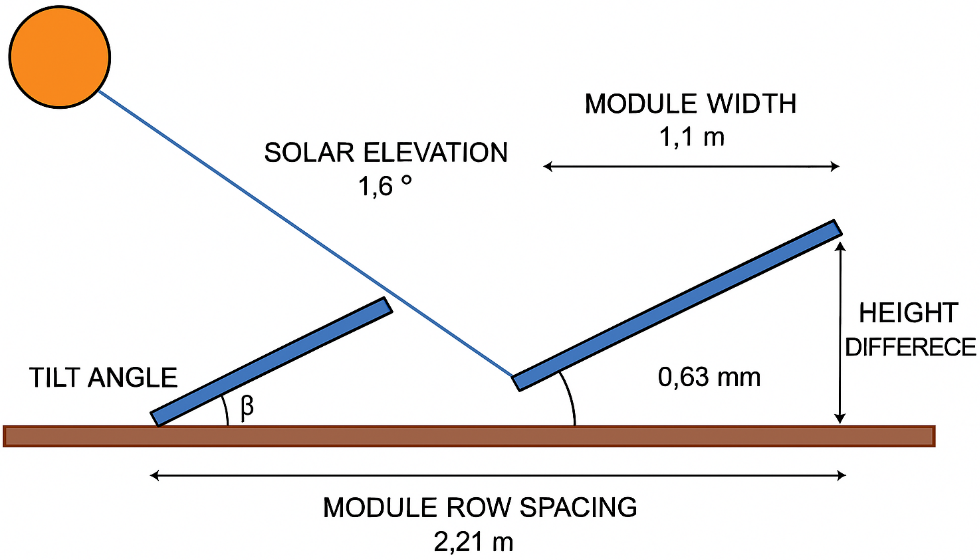


Figure 5: Solar panel layout based on optimal space usage.

Table 3: Roof Area and Solar Panel Layout Calculation

Parameter	**Value**
Roof Width	5.689 m
Panel Width	1.10 m
Panels per Row	5
Roof Length	26.74 m
Panel Length	1.60 m
Rows Available	8.11
Total Panels Fit	$5 \times 8 = 40$
Total Installed Capacity	40×430 W $= 17.2$ kW

Table 4: Data Center Component Loads per Hour

Hour	Rack	Cooling	Losses	Aux	UPS	Gen	Misc	Total	PUE
0	44.99	20.38	3.23	3.98	2.56	1.66	1.05	77.85	1.73
1	43.72	21.52	2.09	4.78	2.55	1.32	1.16	77.15	1.76
2	45.30	21.99	3.24	2.69	2.50	1.78	0.75	78.25	1.73
3	47.05	25.18	3.01	4.41	2.45	1.22	0.74	84.05	1.79
4	43.53	23.19	2.62	4.04	2.22	1.62	1.10	78.33	1.80
5	43.76	23.12	3.20	3.30	2.19	1.66	1.06	78.26	1.79
6	45.65	22.31	2.59	3.83	2.30	1.79	1.19	79.66	1.75
7	47.97	26.01	3.45	4.40	2.71	1.30	0.98	86.82	1.81
8	47.65	24.50	2.94	3.17	2.41	1.42	1.20	83.30	1.75
9	48.29	25.42	3.10	4.42	2.57	1.61	1.09	86.51	1.79
10	46.96	24.87	3.20	4.03	2.59	1.42	1.18	84.85	1.81
11	44.99	21.62	3.20	3.69	2.32	1.37	0.99	78.18	1.74
12	45.89	25.47	3.44	3.74	2.38	1.35	0.89	83.15	1.81
13	46.60	23.74	3.02	3.49	2.40	1.48	0.84	81.57	1.75
14	45.13	25.41	3.06	3.26	2.41	1.62	1.20	82.09	1.82
15	48.32	26.12	3.10	4.57	2.34	1.44	0.80	86.69	1.79
16	44.84	22.30	2.85	4.15	2.38	1.59	0.97	79.07	1.76
17	44.69	23.26	2.96	3.84	2.42	1.33	1.04	79.53	1.78
18	45.52	24.24	3.12	3.97	2.61	1.55	1.18	82.19	1.81
19	46.21	25.13	3.34	4.27	2.48	1.48	0.89	83.79	1.80
20	47.75	26.21	3.49	4.15	2.42	1.38	1.06	86.47	1.81
21	45.39	24.16	3.18	3.85	2.50	1.64	0.85	81.58	1.80
22	43.80	21.78	2.92	4.07	2.21	1.33	1.11	77.22	1.76
23	44.26	22.56	3.20	3.51	2.51	1.58	0.99	78.61	1.77

Code

The used code is available on GitHub (45).

Bibliography

[1] S. Electric, *Reference Design 83 - Technical Guide*, 2025, accessed: 11 May 2025. [Online]. Available: https://download.schneider-electric.com/files?p_Doc_Ref=RD83DSR0_EN&p_enDocType=Other+technical+guide&p_File_Name=RD83DSR1.pdf

[2] Eaton, "xstorage battery energy storage system datasheet," 2025, Eaton xStorage BESS Datasheet, Accessed: 2025-05-10.

[3] Q. CELLS, "Q.peak duo l-g8.3 data sheet," 2025, Q.PEAK DUO L-G8.3 Solar Panel Datasheet, Accessed: 2025-05-10.

[4] E. Masanet, A. Shehabi, N. Lei, S. Smith, and J. Koomey, "Recalibrating global data center energy-use estimates," 2020, Recalibrating Global Data Center Energy-Use Estimates, Accessed: 2025-05-10.

[5] A. Anthony, M. Hsu, P. Henderson, A. Wu, E. Nachmani, X. Sun, N. Thain, D. Kiela, J. Pineau, and K. Lo, "The carbon emissions of writing and illustrating: Are large language models green?" 2023, The Carbon Emissions of Writing and Illustrating, Accessed: 2025-05-10.

[6] L. Leffer, "The ai boom could use a shocking amount of electricity," *Scientific American*, 2023, The AI Boom Could Use a Shocking Amount of Electricity, Accessed: 2025-05-10.

[7] L. A. Barroso, U. Hölzle, and P. Ranganathan, *The Datacenter as a Computer: Designing Warehouse-Scale Machines.* Springer Nature, 2019, The Datacenter as a Computer: Designing Warehouse-Scale Machines.

[8] L. B. N. L. (LBL), "Small data centers — center of expertise for energy efficiency in data centers," 2025, Small Data Centers — Center of Expertise for Energy Efficiency in Data Centers, Accessed: 21 March 2025.

[9] Dgtl Infra, "Data center power," 2025, accessed: 12 May 2025. [Online]. Available: https://dgtlinfra.com/data-center-power/

[10] Device42, "Data center power," n.d., Data Center Power, Accessed: 2023-10-05.

[11] S. Electric, "How double conversion online ups operate?" n.d., How Double Conversion Online UPS Operate?, Accessed: 2023-10-05.

[12] T. Systems. (2021) What is a rack server? What is a Rack Server?, Accessed: November 19, 2023.

[13] Nlyte, "Data center basics," 2025, accessed: 12 May 2025. [Online]. Available: https://www.nlyte.com/blog/data-center-basics/

[14] C. T. Group, “Crac vs crah: Understanding the differences,” 2025, CRAC vs CRAH: Understanding the Differences, Accessed: 2025-04-06.

[15] The Engineering Mindset, “Crac unit - computer room air conditioner,” 2025, accessed: 12 May 2025. [Online]. Available: https://theengineeringmindset.com/data-center-hvac-cooling-systems/crac-unit-computer-room-air-conditioner/

[16] A. H. Khalaj and S. K. Halgamuge, “A review on efficient thermal management of air-and liquid-cooled data centers: From chip to the cooling system,” *Applied Energy*, vol. 205, pp. 1165–1188, 2017, A Review on efficient thermal management of air-and liquid-cooled data centers: From chip to the cooling system.

[17] A. C. Kheirabadi and D. Groulx, “Cooling of server electronics: A design review of existing technology,” *Applied Thermal Engineering*, vol. 105, pp. 622–638, 2016, Cooling of server electronics: A design review of existing technology.

[18] I. Virtosu and L. Chen, “Smart life and sustainable development: a comparative analysis on energy and water efficiency in china and the eu,” *Smart Cities and Regional Development (SCRD) Journal*, vol. 8, no. 3, pp. 19–40, 2024. [Online]. Available: https://www.scrd.eu/index.php/scrd/article/view/477/439

[19] Dataspan, “Reasons you may be experiencing issues with your data center floor,” 2025, Reasons You May Be Experiencing Issues With Your Data Center Floor, Accessed: 14-April-2025.

[20] R. P. UPS, “Comparing in-row cooling and overhead cooling solutions for data centers,” 2025, Comparing In-Row Cooling and Overhead Cooling Solutions for Data Centers, Accessed: 14-April-2025.

[21] OneChassis, “Rack mount enclosures: The ultimate guide to cabinets & chassis,” 2025, Rack Mount Enclosures: The Ultimate Guide to Cabinets & Chassis, Accessed: 14-April-2025.

[22] J. Wan, X. Gui, S. Kasahara, Y. Zhang, and R. Zhang, “Air flow measurement and management for improving cooling and energy efficiency in raised-floor data centers: A survey,” *IEEE Access*, vol. 6, pp. 48 867–48 901, 2018, Air flow measurement and management for improving cooling and energy efficiency in raised-floor data centers: A survey.

[23] W. Tech, “Assessing cooling technologies in datacenters,” 2024, Assessing Cooling Technologies in Datacenters, Accessed: 2025-04-14.

[24] TierPoint, “Data center power distribution,” Data Center Power Distribution, Accessed: 08 April 2025.

[25] M. N. Insights, “Data center power infrastructure,” Data Center Power Infrastructure, Accessed: 08 April 2025.

[26] N. Events, “Ups in data centers: How can they use it for grid services,” 2025, UPS in Data Centers: How Can They Use It for Grid Services, Accessed: 12 April 2025.

[27] 3DFS, “Vavr-9ej262 r0 en,” 2015, accessed: 12 May 2025. [Online]. Available: https://3dfs.com/wp-content/uploads/2015/11/VAVR-9EJ262_R0_EN.pdf

[28] A. Hermanto, "Data center lighting," *EYPMCF Inc.*, 2021, Data Center Lighting, Accessed: 2025-05-04.

[29] A. Lighting, "Power factor for led light sources," *Ansell Lighting News*, 2025, Power Factor for LED Light Sources, Accessed: 2025-05-04.

[30] S. Mircevski, "The impact of reactive power on energy efficiency in electric drives - plenary lecture 3," 2015, accessed: 12 May 2025. [Online]. Available: https://www.researchgate.net/profile/Slobodan-Mircevski/publication/280246904_The_Impact_of_Reactive_Power_on_Energy_Efficiency_in_Electric_Drives_Plenary_Lecture_3/links/5675535308ae125516e700db/The-Impact-of-Reactive-Power-on-Energy-Efficiency-in-Electric-Drives-Plenary-Lecture-3.pdf

[31] Kanoppi, "Search engines vs ai: energy consumption compared," *Kanoppi Blog*, February 2025. [Online]. Available: https://kanoppi.co/search-engines-vs-ai-energy-consumption-compared/

[32] N. Maitanova, S. Schlüters, B. Hanke, and K. von Maydell, "An analytical method for quantifying the flexibility potential of decentralised energy systems," Applied Energy, Volume 362, Article 122078, 2024, Accessed: 14 May 2025.

[33] DataCenterDynamics, "Powering tomorrow: What role do diesel generators play in the future of clean energy solutions?" DataCenterDynamics Article, Published: 2024, Accessed: 14 May 2025.

[34] Schneider Electric, "Specifying and deploying energy storage systems in data centers," Schneider Electric White Paper 185, Version 1.1, 2020, Accessed: 14 May 2025.

[35] LandApp, "Can data centers be powered by solar energy?" LandApp Blog Post, Published: 2024, Accessed: 14 May 2025.

[36] Nordic Solar, "Battery energy storage system (bess)," Nordic Solar Website, Accessed: 14 May 2025.

[37] Solar Electric Supply, "Q cells q.peak duo l-g8.2 430w solar panel," Solar Electric Supply Website, Accessed: 14 May 2025.

[38] E. C. JRC, "Photovoltaic geographical information system (pvgis)," 2025, PVGIS Tool, Accessed: 2025-05-10.

[39] Y. You, Z. Zhang, C. Hsieh, J. Demmel, and K. Keutzer, "Understanding and optimizing gpu energy consumption of dnn training," *USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2023, Understanding and Optimizing GPU Energy Consumption of DNN Training, Accessed: 2025-05-10.

[40] R. Kashef, S. Emshagin *et al.*, "Short-term prediction of household electricity consumption using customized lstm and gru models," *arXiv preprint arXiv:2212.08757*, 2022, Short-term Prediction of Household Electricity Consumption Using Customized LSTM and GRU Models, Accessed: 2025-05-10.

[41] Z. Dong, J. Liu, B. Liu, and X. Li, "Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification," *Energy and Buildings*, 2021, Energy Consumption of a Building Using the XGBoost Algorithm: A Forecasting Study, Accessed: 2025-05-10.

[42] D.-Y. Chou, Y.-S. Chang, C.-Y. Chen, and M.-Y. Cheng, “Improvement in laptop heat dissipation with taguchi method,” *Electronics*, vol. 13, no. 5, p. 882, 2024. [Online]. Available: https://www.mdpi.com/2079-9292/13/5/882

[43] Greentech Renewables, “Determining module inter-row spacing,” Greentech Renewables Website, Accessed: 14 May 2025.

[44] SunEarthTools, “Sun path chart tool,” 2025, Sun Path Chart Tool, Accessed: 2025-05-10.

[45] P. G. Andersen, “Bachelor thesis code repository,” https://github.com/phiand03/AI-models-and-flexibility, 2025, gitHub repository.