UPPSALA
UNIVERSITET

Serienummer
Examensarbete 30 hp
Juni 2026

Super Resolution for Power Peak Estimation of an Industrial Application

Anna Ekstrand och Sofia Kvist

Civilingenjörsprogrammet i system i teknik och samhälle

UPPSALA
UNIVERSITET

Super Resolution for Power Peak Estimation of an Industrial Application

Anna Ekstrand och Sofia Kvist

Abstract

When dimensioning transformer stations, accurate estimation of peak power demand is crucial. However, the data provided by network operators is usually aggregated to 15-minute intervals, which risks overlooking short-duration power peaks that are critical for engineering dimensioning decisions. This master's thesis investigates whether machine learning-based super-resolution can be used to reconstruct high-resolution power demand data from low-resolution measurements, with a particular focus on preserving power peaks relevant for infrastructure dimensioning.

In this paper, machine learning-based super-resolution, a method that aims to reconstruct high-resolution data from low-resolution measurements, is applied to industrial power demand data. An SRP-CNN model was developed and trained on power measurements with a temporal resolution of 30 seconds collected at an industrial facility over a period of approximately nine days. Three loss functions were evaluated: Root mean square error (RMSE), Structural Similarity Index Measure (SSIM), and a weighted RMSE. The latter was specifically designed to assign greater emphasis on extreme power values. The hyperparameters of the model were optimized using Bayesian optimization, applied separately for each loss function.

All three models successfully reconstructed the overall structure of the high-resolution signal. At extreme power peaks, however, the reconstruction performance deteriorated, where all models systematically underestimated the peaks magnitudes. The weighted RMSE model demonstrated the strongest ability to reconstruct power peaks, although significant errors remained even for this model. The results indicate that SRP-CNN models can provide valuable support for general load analysis, but that the method should be used with caution when dimensioning as accurate assessment of extreme power peaks is critical.

Teknisk-naturvetenskapliga fakulteten

Uppsala universitet, Utgivningsort: Uppsala

Handledare: Erik Sterner Ämnesgranskare: Göran Ericsson

Examinator: Elísabet Andrésdóttir

UPPSALA
UNIVERSITET

Populärvetenskaplig sammanfattning

De elanvändare med störst effektbehov är ofta anslutna till elnätet via en egen transformatorstation. Transformatorstationen tar emot el från elnätet, omvandlar den till rätt spännings- och strömnivå och distribuerar elen vidare inom anläggningen. I princip alla elektriska komponenter behöver anpassas efter det system de ska ingå i. Det handlar främst om komponenternas kapacitet, där låg kapacitet innebär risk för överbelastning och därmed försämrad elsäkerhet, medan för hög kapacitet innebär onödigt höga investeringskostnader. Processen att bestämma lämplig kapacitet kallas dimensionering och utgör en central del av planeringen vid utbyggnad av transformatorstationer.

Problemställningen handlar i grunden om att säkerställa tillräckligt hög kapacitet för att garantera säker drift utan att bli onödigt dyr. Dimensioneringen baseras därför på hur stort effektbehovet är när det är som högst, eftersom alla komponenter måste kunna hantera denna nivå. Vidare är den temporala aspekten viktig. Elektrisk effekt beskriver den momentana energianvändningen, och det är det högsta värdena i denna, så kallade effekttoppar, som styr dimensioneringen. Olika elektriska komponenter är olika känsliga för effekttoppar, vissa klara kortvarig överbelastning medan andra är mer känsliga.

Att hänsyn behövs tas till både magnituden och varaktigheten på elanvändarens effekttoppar innebär att den som dimensionerar en transformatorstation behöver ha en god förståelse för hur det momentana elbehovet varierar över tid. Denna förståelse skapas med fördel baserat på data på historiskt effektbehov, något som ofta utgör en utmaning då sådan data sällan är tillgänglig. Den tillgängliga datan över elanvändning samlas in av elnätsägasaren och har en 15-minutersupplösning. I detta sammanhang, där momentant energibehov är av intresse, är denna data obrukbar på grund av sin låga tidsupplösning. Effekttopparna som förekommer slås ut över 15-minutersmedelvärden och raderar därmed de för dimensioneringen kritiska insikter.

I denna uppsats undersöks om Artificiell Intelligens (AI) kan användas för att återskapa detaljerade mätdata från glesare mätningarna och därmed fylla i den information som saknas. Detta är en metod som kallas super-resolution vilket är en teknik som ursprungligen utvecklades för att förbättra kvalitén på lågupplösta bilder men som i denna uppsats tillämpades på tidsseriedata. Metoden bygger på maskininlärning, där modellen tränas på par av detaljerade och glesa mätdata, vilket gör det möjligt för den att lära sig sambandet mellan lågupplösta och högupplösta effektsignaler.

Teknisk-naturvetenskapliga fakulteten

Uppsala universitet, Utgivningsort: Uppsala

Handledare: Erik Sterner Ämnesgranskare: Göran Ericsson

Examinator: Elísabet Andrésdóttir

UPPSALA
UNIVERSITET

För att få tillgång till högupplöst effektdata utfördes mätningar direkt vid en industrianläggning med hjälp av ett avancerat mätinstrument som registrerade elförbrukningen var trettionde sekund under 11 dagar. Denna upplösning ger en betydligt mer detaljerad bild av hur elanvändningen varierar under dygnet, inklusive snabba, kortvariga toppar.

Den insamlade datan användes för att träna de tre modellvarianterna som utvecklades inom ramen för denna uppsats. Skillnaden mellan de olika modellvarianterna låg i hur fel viktades under träningsprocessen. Denna viktning kallas för en förlustfunktion. Den första modellversionens förlustfunktion byggde på det genomsnittliga felet, den andra på strukturell likhet, och den tredje prioriterade fel vid höga effektnivåer. Modellerna utvärderades sedan i sin förmåga att återskapa den övergripande signalen samt förmågan att specifikt återskapa effekttoppar. Detta möjliggjorde en bedömning av modellernas praktiska användbarhet ur ett ingenjörsperspektiv.

Resultatet visade att alla tre modellvarianterna kunde återskapa den övergripande formen på effektförbrukningskurvan med relativt hög noggrannhet. Rekonstruktionsnoggrannheten försämrades dock avsevärt vid de högsta uppmätta effekttopparna. För de fyra största topparna varierade felen mellan 30 och 40 procent, vilket är för stora avvikelser för att resultaten ska kunna användas som ett tillförlitligt underlag vid dimensionering av transformatorer. Denna feluppskattning var konsekvent en underskattning av effekttopparnas magnitud. Samtliga modeller var alltså för konservativa i sina prediktioner. Den modellvariant som var särskilt utformad för att betona rekonstruktion av effekttoppar uppnådde bäst resultat vid höga effektnivåer, men på bekostnad av något lägre noggrannhet för resten av signalen. Detta återspeglar en inneboende avvägning mellan den övergripande rekonstruktionsförmågan och noggrannheten vid specifika effekttoppar.

Resultaten tyder på att AI-baserad super-resolution har potential som ett verktyg för att analysera generella mönster i industriell elförbrukning. Samtidigt bör metoden användas med försiktighet i sammanhang där en noggrann uppskattning av kortvariga extrema effekttoppar är avgörande, såsom vid dimensionering av transformatorer. Studien visar också behovet av att utveckla tydligare matematiska definitioner av effekttoppar som inkluderar både storlek och varaktighet. Detta kan bidra till bättre anpassade modeller och mer träffsäkra utvärderingsmetoder i framtida studier.

Teknisk-naturvetenskapliga fakulteten

Uppsala universitet, Utgivningsort: Uppsala

Handledare: Erik Sterner Ämnesgranskare: Göran Ericsson

Examinator: Elísabet Andrésdóttir

UPPSALA
UNIVERSITET

Acknowledgements

We would like to thank WSP for the opportunity to carry out this thesis in close collaboration. We also thank the client for their willingness to participate in the project, for welcoming us to their facility when arranging the conduction of measurements, and for making the data collection for this study possible. Access to high-resolution measurement data is often a major limitation in this type of research, and their openness and collaboration made this work possible.

We are especially grateful to the colleagues at WSP who generously shared their time and expertise through interviews, and who continuously contributed with valuable insights, advice, and encouragement throughout the process. A special thanks is directed to our supervisor at WSP, Erik Sterner, whose support made this work possible in many ways. Beyond practical guidance, the numerous discussions and reflections shared during our meetings have played an important role in shaping this thesis.

We would also like to thank Uppsala University and the STS program for providing a truly interdisciplinary environment, where questions at the intersection of energy systems and information technology can be explored. At the university, we are particularly grateful to our subject reviewer, Göran Ericsson, for his engagement, constructive feedback, and for strengthening the quality of this work.

Finally, we would like to thank our friends and family who have supported us along the way, and especially those who have taken the time to read and provide feedback on the report in its final stages.

Teknisk-naturvetenskapliga fakulteten

Uppsala universitet, Utgivningsort: Uppsala

Handledare: Erik Sterner Ämnesgranskare: Göran Ericsson

Examinator: Elísabet Andrésdóttir

Table of Contents

List of Abbreviations

Abbreviation	Meaning
AI	Artificial Intelligence
CNN	Convolutional Neural Network
cos(f)	Power factor
EO manager	Electrical Operations Manager
I	Current
GAN	Generative Adversarial Network
GCU	Growing Cosine Unit
MAE	Mean Absolute Error
MSE	Mean Squared Error
PreLU	Parametric Rectified Linear Unit
ProfileSR-GAN	Profile Super-Resolution Generative Adversarial Network
ReLU	Rectified Linear Unit
RMSE	Root Mean Squared Error
SDG	Stochastic Gradient Descent
SRCNN	Super-resolution Convolutional Neural Network
SRP-CNN	Super-resolution Perception Convolutional Neural Network
SRP-NSE	Super-resolution Perception Network for State Estimation
SSIM	Structural Similarity Index Measure
T	Number of iterations
Tailored-GAN	Tailored Generative Adversarial Network
V	Voltage

1. Introduction

Recent advances in Artificial Intelligence (AI) and digitization have contributed to the rapid development of data-driven technologies across a variety of application areas (Miraftabzadeh et al., 2019). One such area is the energy sector, which is simultaneously undergoing a major development characterized by the integration of renewable energy and the electrification of society (Xie et al., 2020). The increasing complexity of modern electricity systems imposes growing demands on system planning and design, thereby enabling the adoption of data-driven analytical methods within the energy sector (Miraftabzadeh et al., 2019).

The increased deployment of smart meters and sensors in the power system has enabled the collection of large amounts of heterogeneous, high-resolution energy data (Wang et al., 2019). This has not only contributed to increased opportunities for analysis of the energy system but also increased the need for methods capable of processing and extracting information from complex datasets. In this context, research interest in the integration of machine learning-based methods within the energy sector has received increasing attention, particularly in areas such as power flow forecasting, monitoring and modeling of smart grids (Miraftabzadeh et al., 2019).

For several of these application areas, the availability of high-resolution energy data is an important prerequisite for analyzing energy consumption patterns and performing more accurate load modeling and load forecasting (Iversen et al., 2023; de-Paz-Centeno et al., 2021). High-resolution data can provide a more detailed understanding of consumption patterns and enable the identification of short-term variations and peak loads in power demand. However, collecting high-resolution data poses some practical challenges, as the large amount of data places high demands on storage space and processing capacity. This is an important reason why energy data has traditionally been collected at lower temporal resolutions, for example in 15-minute or hourly intervals (Iversen et al., 2023).

To address the practical limitations associated with the acquisition and management of high-resolution energy data, super-resolution methods have gained increased attention in the research field as a promising method for reconstructing detailed load profiles. By leveraging machine learning-based super-resolution techniques, a more detailed representation of power demand can be generated (Iversen et al., 2023). More specifically, in the dimensioning of substations, it is particularly important to be able to identify and analyze load peaks in power consumption, as these greatly influence the choice of dimensioning capacity. It is therefore not sufficient to only reconstruct the overall structure of the data, since the system dimensioning largely depends on both the size and the temporal characteristics of power peaks (Interviewee 4, 2026).

Based on that background, this master’s thesis applies machine learning-based super-resolution techniques to reconstruct high-resolution power demand data, focusing particularly on the model's ability to preserve the magnitude and temporal structure of identified power peaks relevant for decision-making in system dimensioning.

1.1 Aim and Research Questions

The purpose of this thesis is to investigate how machine learning-based super-resolution techniques can be used to reconstruct high-resolution power demand data, with particular focus on accurately preserving peaks in the signal. The work presented in this thesis assesses how accurately the model reproduces peak magnitude and timing and thereby examines its engineering relevance in the context of system dimensioning and decision-making support for power demand analysis.

This aim is addressed through the following research questions:

1) How can a machine-learning model be developed to reconstruct high-resolution power data while accurately preserving peaks in the signal?

2) How well does the model preserve the magnitude and timing of power demand peaks?

2. Background

The following chapter presents WSP, the company with which this thesis was conducted, and an overview of their approach to substation dimensioning. Also, a brief background to the broader project within which this thesis was conducted is provided. Thereafter, fundamental concepts related to AI, relevant to the present study, are introduced. Lastly, background information on energy data analytics is presented.

2.1 Company Background

WSP is an international strategy and engineering consultancy, employing over 80,000 people. Within its general profile of built environment, the company encompasses more than 100 service areas and sectors (WSP, n.d.a). One of these is the Energy and Industry sector, which is further divided into sub-sectors addressing areas such as transmission, thermal power generation and distribution, and energy efficiency. A specific part of the organization is dedicated to projects related to industry and substations. These projects include, among others, collaborations with industrial clients that engage WSP to plan, design, and manage the construction or renovation of their grid connections, typically in the form of substations (WSP, n.d.b).

2.1.1 Substation Dimensioning WSP

At present, when a consultant at WSP is to design a substation, one key consideration must be addressed before any design decisions can be made: the estimation of power demand. All required components must be appropriately dimensioned to accommodate this demand (Glover et al., 2023). The following section outlines the framework typically used by WSP consultants to approach this question. The material is based on interviews with four WSP consultants with experience in projecting substations and estimations of dimensional loads:

- Interviewee 1, Senior Electrical Engineer
- Interviewee 2, Manager Electrical Engineering
- Interviewee 3, Senior Electrical Engineer
- Interviewee 4, Experienced Electrical Engineer

In substation design projects, load data is typically provided by the client's local network operator. However, this data cannot be used to any significant extent due to its low temporal resolution. It is recorded at 15-minute intervals since January 1, 2026, and hourly prior to that (EIFS 2025:1). This limitation also entails a risk of missing peak loads (Interviewee 1, 2026 & Interviewee 3, 2026). Transformers must be dimensioned with regard to the peaks in power demand, where both the magnitude and the temporal dimension are relevant (Interviewee 4, 2026). This is because transformers exhibit an inherent thermal inertia, meaning that short and infrequent periods of overload will not damage the transformer. Nevertheless, other components in the substation are more sensitive to currents exceeding the rated current (Interviewee 1, 2026, Interviewee 3, 2026 & Interviewee 4, 2026).

To determine the dimensioning demand, the pattern of the specific load being assessed must be understood. In general, all types of loads follow characteristic patterns in their power demand (Interviewee 2, 2026). Once an understanding of these patterns exists, and specifically knowledge of the peak magnitude and its temporal characteristics, an expert judgement can be made regarding what transformer size is reasonable (Interviewee 4, 2026). When selecting a reasonable transformer size, a loading margin of about 70 percent is typically included. This means that the estimated peak load magnitude should correspond to approximately 70 percent of the transformer's rated power capacity (Interviewee 1, 2026, Interviewee 2, 2026, Interviewee 3, 2026 & Interviewee 4, 2026). Additionally, considerations relating to this margin may arise if the load is expected to expand. In such cases, the estimated increase in peak magnitude is added to the current demand. The aggregated power demand should still constitute roughly 70 percent of the transformer's rated capacity (Interviewee 3, 2026). Even in estimating the magnitude of future loads, the temporal aspect is central. An expansion of electrical load does not necessarily imply an increase in peak power demand, such as if the new load operates when overall consumption is low. On the other hand, it may also lead to overlapping peaks and thus a significant increase in peak magnitude and/or duration (Interviewee 1, 2026).

In some cases, the present-day power demand can be used as an input for determining the dimensioning requirement. This presupposes that the load already exists, for example, that the facility is not a new industry or household being built from scratch. When this condition applies, measurements can be performed on the existing load, typically with a one or two minute resolution, collecting high frequency data on power demand over one to two weeks. The collected power data provides a detailed situational picture which in most cases is representative of the annual power demand, depending on how regular the load is and how much it is influenced by external factors and seasonal variations. Thus, it constitutes a valuable decision-making basis for establishing the dimensioning demand. For example, industrial loads can generally be assumed to display patterns during a one to two week measurement period that are representative of the rest of the year. How this data, when it exists at all, is used depends on the expert making the final assessment of the required capacity (Interviewee 1, 2026).

To develop an understanding of the power demand of an aggregated load, power consumption is broken down into its components and evaluated in relation to one another over time. A widely applied tool for this purpose is the diversity factor (Interviewee 1, 2026, Interviewee 2, 2026, Interviewee 3, 2026 & Interviewee 4, 2026). This approach assumes that all components of the load, for example machines in an industrial facility or appliances in a kitchen, will not operate simultaneously. Based on information on the power demand of all components, and an understanding of operational routines, an estimate is made of the degree to which different components will run concurrently. The result is a value between zero and one that represents the extent of simultaneous operation. This factor is multiplied by the sum of the rated power values of the components, yielding the dimensioning power demand. The diversity factor can be determined for an entire aggregated load or for sub-groups of components, depending on how the load in question behaves (Interviewee 1, 2026, Interviewee 3, 2026 & Interviewee 4, 2026).

2.2 Project Background

This master's thesis work has been conducted in collaboration with WSP and is embedded within an ongoing project involving one of the company's clients. Due to considerations of confidentiality and to protect the business interests of both WSP and its client, certain information in this thesis has been anonymized. Accordingly, the client's relevant industrial facilities and processes will be referred to as the Facility, the chief of operations at the Facility will be referred to as the Plant Manager, the electrical operations (EO) manager will be referred to as the EO manager, and the work carried out by WSP will be referred to as the Project.

The client is an internationally operating manufacturing firm. Several of the client's industrial processes are automated and electrically driven. In addition to production-related loads, the Facility's electricity demand is based on electric vehicle charging, on-site solar power generation (including associated cooling systems), and the offices' general electricity consumption (Plant Manager, March 17, 2026). The Facility's current electricity supply is provided by a substation that is approaching the end of its operational lifetime. At the same time, the client is planning to expand its operations and introduce new industrial processes at the Facility. Together, these factors give rise to the need for a new substation (Plant Manager, February 3, 2026).

WSP has developed a technical description for the Project in which several key parameters have been established. The external distribution grid operates at a voltage level of 12 kV (50 Hz), while the internal grid operates at 400 V (50 Hz). These values define the rated voltages for both the existing and the planned transformer. The current peak load is estimated at 800 A, based on information provided by the EO Manager. In addition, the planned new loads are expected to increase the peak load by a further 780 A, based on expert assessments using diversity factors. Finally, the new switchgear is designed to accommodate future expansion corresponding to approximately 30 percent of the installed capacity (WSP, 2025).

2.3 Energy Data Analytics

In recent years, a growing body of research has emerged around the application of data analytics to electrical load data. The widespread deployment of smart meters and energy loggers has enabled the collection of large volumes of energy data, creating new opportunities for advanced analytical methods (Wang et al., 2019). Such developments are particularly relevant in the context of modern energy systems, which are becoming increasingly complex and vulnerable due to the integration of small-scale and intermittent generation, as well as rising levels of electrification. Mitra et al. (2024) argues that effective data analytics holds significant potential for addressing key challenges related to secure grid operation, anomaly detection, as well as load and generation management.

One important approach within load analysis is load identification, or load disaggregation, which aims to decompose aggregated load profiles into their constituent components. This enables a more detailed understanding of the underlying drivers of power demand and may, similarly to the use of diversity factors, provide valuable insights for dimensioning purposes. In addition, load forecasting methods aim to predict future load behavior over short-, medium-, or long-term horizons, offering a data-driven complement to conventional safety margins, such as the 30 percent capacity margin applied at WSP. However, both approaches are strongly dependent on the quality and resolution of the available data. In particular, research on load disaggregation indicates a strong relationship between measurement frequency and achievable accuracy, as well as the feasibility of disaggregation itself (Tolnai et al., 2023). Similarly, load forecasting methods generally require historical data with a temporal resolution equal to or higher than that of the desired forecasts (Hamad et al., 2020; Liu et al., 2026). Consequently, accurate analysis and prediction of peak loads necessitate high-frequency historical data, on the order of one minute, to capture significant peaks.

2.4 Orientation on Artificial Intelligence

Kufel et al. (2023) define machine learning as a subcategory of AI that focuses on using data to train computer models to detect patterns and make decisions based on the data provided. By analyzing data, such models can continuously improve their performance and accuracy over time. Machine learning is commonly divided into two categories, supervised and unsupervised learning (Kufel et al., 2023).

Supervised machine learning involves training models on a dataset where both the input and the corresponding output are known. The goal is for the model to learn the relationship between input and output so that it can make predictions and decisions regarding new data (Kufel et al., 2023). Unsupervised machine learning, in contrast, relies only on input data without predefined outputs to discover underlying patterns and structures within the dataset. In this way, the model independently learns to identify components and relationships in the data (Naeem et al., 2023).

Deep learning is a subfield of machine learning built upon multiple computational layers that together learn different representations and patterns within data. Artificial neural networks are among the most widely used architectures in deep learning. Inspired by the structure of the human brain, these networks consist of large numbers of interconnected neurons that collectively form a predictive model (Gupta et al., 2022). In recent years, the use of Convolutional Neural Networks (CNNs) has increased significantly (Gupta et al., 2022), and they are currently among the most applied algorithms in deep learning (Yu et al., 2022). Common applications of CNNs include image classification, image segmentation, and object detection (Ayeni, 2022).

2.5 Super-resolution

Thanks to advances in deep learning techniques, super-resolution has emerged as a prominent application area within computer vision (Zhang et al., 2025). It aims to reconstruct high-resolution data from low resolution inputs. Although super-resolution is most commonly applied to images, it is also used in other areas of application, for example surveillance video. In some applications, hardware limitations can prevent the direct capture of high-resolution input data, creating a demand for super-resolution techniques. Super-resolution methods address this constraint by reconstructing fine details and textures from low-resolution input data (Hassan et al., 2025).

However, reconstructing high-resolution input data is an inherently challenging process, as multiple versions of high-resolution input data can correspond to the same low-resolution input data, which makes the problem ill-posed (Zhang et al., 2025). Several methods have been explored to enhance the resolution of low-quality input data. However, deep learning has proven to be particularly effective in reconstructing high-resolution input data. By training neural networks on large datasets consisting of paired low- and high-resolution input data, these models can learn complex patterns and relationships between them, enabling the generation of more detailed and accurate reconstructions (Hassan et al., 2025).

These challenges collectively provide a strong incentive to explore super-resolution techniques for energy data. This remains an emerging research field and, to the best of our knowledge, only one comprehensive review has been conducted to date. This review, presented by Iversen et al. (2023), surveys machine learning-based approaches for the super-resolution of smart meter data. The authors identify four particularly relevant and effective methods within the domain, comprising two models based on CNNs and two based on Generative Adversarial Networks (GANs).

The CNN-based approaches include Super-resolution Perception Convolutional Neural Networks (SRP-CNN), developed by Liu et al. (2020), and Super-resolution Perception Net for State Estimation (SRP-NSE), proposed by Liang et al. (2020). Methods based on GAN consist of Profile Super-resolution Generative Adversarial Network (ProfileSR-GAN), developed by Song et al. (2022), and Tailored Generative Adversarial Network (Tailored-GAN), introduced by Li et al. (2020). While all four methods demonstrate the ability to reconstruct high-frequency load data from low-frequency inputs, they exhibit different strengths and limitations in terms of accuracy, capacity to capture complex relationships, and computational efficiency. Moreover, important differences arise with respect to their applicability across datasets. In particular, both SRP-NSE and ProfileSR-GAN require large volumes of training data, while Tailored-GAN has limited generalizability to data originating from different geographical contexts (Iversen et al., 2023). These characteristics are of particular importance when assessing the suitability of such methods for real-world applications. Several previous studies have examined CNN-based models, such as the Super-resolution Convolutional Neural Network (SRCNN), primarily for image super-resolution (Zhang et al., 2025; Hassan et al., 2025). However, Liu et al. (2020) extend the CNN-based model to time series data by proposing a modified architecture capable of reconstructing high-frequency data from low-frequency power data, the so called SRP-CNN model. The core of the SRP problem is to enable the model to reconstruct high frequency power data from observed low-frequency measurements in power grids (Liu et al., 2020).

The proposed SRP-CNN model is inspired by traditional CNN architectures used for image super-resolution. However, adaptations are required to handle one-dimensional time series data instead of the two-dimensional structure of image inputs. The model proposed by Liu et al. (2020) consists of a network with three main components: feature extraction, information supplementation, and reconstruction. The first two components are built using one dimensional, instead of the typical two-dimensional, conventional layers, where feature extraction uses a layer to extract features from the original low-frequency input. The information supplementation uses seven convolutional layers to perform a nonlinear mapping between the low-frequency input and the high-frequency representations. The last component, reconstruction, uses a transposed convolution layer to aggregate the generated vectors and reconstruct the high-frequency output signal (Liu et al., 2020).

3. Theory

This chapter presents the theoretical foundation underlying the development of the super-resolution model implemented in this thesis. Firstly, the architecture of CNNs is described in detail. Then, the concept of loss functions and their central role in the learning process of machine-learning models are presented. Lastly, the chapter addresses hyperparameters and the process of hyperparameter tuning.

3.1 The Architecture of a Convolutional Neural Network

A CNN model is typically composed of an input layer, an output layer, and between these are several hidden layers. The CNN network consists of neurons that receive input and perform a calculation by computing a dot product with their weights. Based on this calculation, each neuron produces an output. When the network receives an input, it is propagated through the network's hidden layers. Each hidden layer consists of a set of neurons, where every neuron is fully connected to all neurons in the previous layer. Within a given layer, the neurons are independent of each other and do not share any connections (Sewak et al., 2018).

The hidden layers can include convolutional layers and fully connected layers (Gupta et al., 2022). In the convolutional layers a kernel moves across the data to extract its fundamental features. The movement of the kernel across the input data is determined by a parameter referred to as the stride length, which specifies the number of positions the kernel shifts between each operation. A smaller stride length allows the convolutional filter to examine overlapping regions of the input data, meaning that the same data points are processed multiple times from slightly different contexts. Literature on deep learning highlight that this can improve the model's ability to capture detailed patterns and local structures, particularly in cases where the available training data is limited, although at the cost of increased computational complexity (Sewak et al., 2018; Goodfellow et al., 2016). The fully connected layers connect all the neurons from previous layers. By using multiple layers, the model can identify patterns in input data, with the layers first detecting simpler structures and patterns and later revealing more complex structures (Ayeni, 2022).

During the training of a CNN model, the dataset is repeatedly propagated through the neural network in iterations referred to as epochs. One epoch corresponds to the model processing the entire training dataset once. Increasing the number of epochs generally enables the network to learn more complex relationships within the data and improve performance. However, excessively long training may cause the model to overfit the training data, reducing its ability to generalize to unseen data (Goodfellow et al., 2016). To mitigate this problem, a commonly applied strategy is early stopping, where the training process is terminated before all predefined epochs have been completed if the model's performance on validation data ceases to improve. Previous studies show that early stopping can improve generalization performance and reduce the risk of overfitting, particularly in deep learning applications with limited datasets (Prechelt, 1998).

3.2 The Role of the Loss Function

To evaluate how well a network represents data, a loss function is used. The loss function is defined as the difference between the predicted output and the actual value, the so-called ground truth, and serves as a mathematical measure of the model's performance (Raiaan et al., 2024; Elharrouss et al., 2025). It constitutes a vital part of the learning process, as it decides how the network's weights are updated through each iteration. The objective during training is to minimize the loss function (Raiaan et al., 2024; Elharrouss et al., 2025), since high loss values indicate poor accuracy, whereas low loss values suggest that the model has successfully produced predictions close to the true values (Raiaan et al., 2024).

The choice of loss function therefore affects the model's convergence speed, its ability to generalize, and the stability of the training process. According to Elharrouss et al. (2025), the choice of loss function must be guided by the specific characteristics of the problem, as no single loss function is universally suitable for all tasks. A well-designed loss function steers the optimization process so that the error is minimized (Elharrouss et al., 2025).

Elharrouss et al. (2025) categorize different loss functions into various categories based on the nature of the learning task. For regression problems, commonly used loss functions include Mean Squared Error (MSE) and Mean Absolute Error (MAE), where the purpose is to calculate the difference between the model's predictions and the actual values. In addition, both MSE and MAE are also commonly used as evaluation metrics for assessing model performance (Arabboev et al., 2024). For classification tasks, where the model aims to distinguish between different classes, loss functions are used that examine how well the model's probability distribution corresponds to the actual outcome, such as different types of cross-entropy (Elharrouss et al., 2025). MSE aims to measure the average squared difference between predicted and actual values. By squaring the errors, greater weight is given to larger deviations (Elharrouss et al., 2025). MAE, on the other hand, calculates the mean absolute difference between the predicted values and the actual values. In contrast to MSE, MAE is generally less influenced by extreme deviations. A closely related metric derived from MSE is Root Mean Squared Error (RMSE), which is defined as the square root of MSE and is used to express the error in the same unit as the original data (Hodson, 2022).

Furthermore, several previous studies have introduced a weighted variant of MSE, commonly referred to as Weighted Mean Squared Error (WMSE) (Sree and Vijaya, 2019; Jeong, Kim and Lin, 2010; Zhang et al., 2025). This weighted variant aims to assign different levels of importance to various components of the prediction error. Since the traditional MSE handles errors from all data points equally, it is effective on balanced datasets. However, for imbalanced datasets, MSE may overlook certain minority categories and instead prioritize reducing errors for the dominant categories of the dataset (Sree and Vijaya, 2019). By employing a weighted MSE, the loss function can therefore be adapted to the specific problem, allowing control over how different types of errors contribute to the overall loss (Sree and Vijaya, 2019; Jeong et al., 2010; Zhang et al., 2025).

Another loss function frequently used in image reconstruction and super-resolution tasks is the Structural Similarity Index Measure (SSIM) (Elharrouss et al., 2025). SSIM focuses on the structural and perceptual quality of images in super-resolution applications (Elharrouss et al., 2025; Anagun, Isik and Seke, 2019), in contrast to methods such as MSE, which focus primarily on pixel-wise differences (Elharrouss et al., 2025).

3.3 Hyperparameters of the Network

A CNN network consists of several hyperparameters that are set before training the network and are significant for the model's learning process and performance. Raiaan et al. (2024) divide these parameters into different categories based on the network's architecture, ranging from hyperparameters associated with the convolutional layers to those of the fully connected layers, as well as general hyperparameters. The general hyperparameters include, among others, optimizer, batch size and learning rate. The hyperparameters associated with the convolutional layers, include the number of layers, the number of kernels, kernel size and activation function (Raiaan et al., 2024).

The learning rate determines the magnitude of updates made to the model's weights during the training process and thus affects how quickly the model learns and reaches convergence. The choice of learning rate is critical, as an incorrect setting can degrade the model's performance (Raiaan et al., 2024). Learning rate is strongly related to the optimizer (Sewak et al., 2018), as the optimizer adapts the learning rate to reduce the error (Raiaan et al., 2024).

There are several different optimization algorithms for CNN networks, and their suitability depends on the characteristics of the dataset and the specific task (Sewak et al., 2018). Sewak et al. (2018) and Raiaan et al. (2024) highlight Stochastic Gradient Descent (SGD) as a commonly used optimizer. According to Raiaan et al. (2024), SGD is a variant of Gradient Descent (GD), an optimization method that updates the network's weights at each iteration to minimize the error. A limitation of SGD is that the learning rate must decrease proportionally to 1/T to ensure that the model converges, where T corresponds to the number of iterations. Sewak et al. (2018) argue that other optimization algorithms, such as Adam, instead adjust their step size to the order of magnitude of the gradient and can thus overcome this limitation. Adam is a commonly used optimizer that is particularly well-suited for problems involving large datasets or many parameters (Raiaan et al., 2024). Since Adam combines features from several other optimizers, Raiaan et al. (2024) argues that it is time-efficient during training, robust against noise, and converges quickly.

Batch size specifies how many data points are processed simultaneously when calculating gradients during the training process. The choice of batch size therefore affects learning, where a batch that is too small yields poorer results and a batch that is too large can cause the model to converge more slowly. Raiaan et al. (2024) note that batch sizes of 16, 32, and 64 are commonly used for neural networks but argue that smaller batch sizes often yield better model performance (Raiaan et al., 2024).

The hyperparameters associated with the convolutional layers determine the architecture of the model. The number of layers indicates the depth of the network (Ooi and Ibrahim, 2021) and affects the model's performance (Raiaan et al., 2024). Sewak et al. (2018) argue that one to two hidden layers may be sufficient for simpler problems, while an increased number of layers can lead to a more efficient use of the model's parameters. This is because greater network depth with fewer neurons per layer can represent more complex relationships. However, the authors emphasize that an excessively deep network can instead lead to overfitting, and suggest that in more complex cases, the number of layers can be increased gradually until overfitting is observed (Sewak et al., 2018). Dong et al. (2015) also argue that an increased number of layers does not necessarily lead to better results. The authors point out that as the number of layers increases, certain complications arise in the network. Deeper SRCNN networks are more difficult to train, require longer training times, and at the same time risk yielding marginal results (Dong et al., 2015).

The number of kernels determines how many different patterns the model can identify in the data. In CNN models, it is common to use multiple kernels, such as 32 or 64, to capture important patterns and thereby improve model performance (Raiaan et al., 2024). According to Dong et al. (2015), increasing the number of kernels generally improves the reconstruction quality of the network but also raises the computational cost per signal.

The choice of kernel size affects model performance, and different datasets may benefit from different kernel sizes. Smaller kernels are often advantageous because they are more computationally efficient, unlike larger kernel sizes, which require a higher computational cost (Raiaan et al., 2024). Similarly, Dong et al. (2015) argue that larger kernels lead to significantly higher computational cost while the model's performance decreases rapidly.

Activation functions help the neural network make decisions and identify complex patterns in the data, thereby improving the efficiency of the learning process. The most suitable activation function depends on both the task at hand and the network's architecture. Common activation functions include Rectified Linear Unit (ReLU), sigmoid, and tanh, as well as various variants of further-developed ReLU such as leaky ReLU and Parametric Rectified Linear Unit (PreLU). According to Raiaan et al. (2024), sigmoid and tanh are commonly used in classification problems, while ReLU is often used in hidden layers. Leaky ReLU helps prevent neurons from becoming inactive, and PReLU can be adjusted during training (Raiaan et al., 2024). Noel et al. (2025) propose an oscillating activation function, known as the Growing Cosine Unit (GCU), which, thanks to its multiple zero points, can create more complex decision boundaries. This in itself enables individual neurons to model more advanced relationships. According to Noel et al. (2025), the GCU can provide higher accuracy and faster convergence compared to more traditional activation functions.

3.4 Hyperparameter Tuning

While the choice of hyperparameters greatly impacts model performance, determining the set of hyperparameters is not a straightforward task. The amount of hyperparameters in combination with the various alternative values of these hyperparameters mean that there is a large number of possible combinations, all with different performance and computational cost. In the literature, different approaches to tackle this and determine the optimal hyperparameters are suggested. Generally, an experimental approach is recommended. For example, in Practical Convolutional Neural Networks, Sewak et. al. (2018) argues that although it is theoretically possible to determine all combinations of hyperparameters through exhaustive techniques such as grid search combined with cross-validation, this approach is neither practical nor necessary in most cases. The large number of hyperparameters involved would render such a process prohibitively time-consuming. Instead, the authors recommend selecting an initial set of hyperparameters based on general guidelines and insights reported in the literature for similar models and tasks. These initial choices are then refined iteratively by evaluating how adjustments affect model performance (Sewak et al., 2018).

In their comprehensive study on hyperparameter tuning, Raiaan et al. (2024) agrees that manual tuning of hyperparameters is a too time-consuming task to be recommended. Instead, more efficient optimization techniques in the form of automated algorithms are preferred. In their review, Raiaan et al. (2024) examines a wide range of hyperparameter optimization techniques, including metaheuristic, numerical, sequential, and statistical approaches. Among these, Bayesian optimization is identified as one of the most promising methods, particularly in applications where model evaluations are computationally expensive. The authors highlight that Bayesian methods have been successfully applied across several domains, including industrial applications. A key advantage of Bayesian optimization, as emphasized in the review, is that it does not require evaluating all possible combinations of hyperparameters. Instead, it builds a probabilistic model of the objective function and uses this model to guide the search toward promising regions of the parameter space. This significantly reduces computational cost compared to exhaustive approaches such as grid search, while still achieving strong performance. Additionally, the method incorporates information from previous evaluations, allowing for a more informed and efficient search process over time (Raiaan et al., 2024).

In a study conducted by Kolar et al. (2021), Bayesian optimization is applied to automatically tune hyperparameters of a CNN for fault diagnosis in rotary machinery. The model uses raw industrial sensor data in the form of one-dimensional vibration signals. Kolar et al. (2021) emphasize several strengths of Bayesian optimization. First, the method reduces the number of required model evaluations by leveraging prior observations, making it more computationally efficient than traditional search methods. Second, it is well-suited for optimizing black-box functions, where the relationship between hyperparameters and model performance is not explicitly known. Finally, it balances exploration and exploitation, enabling the algorithm to both investigate new regions of the search space and refine known promising configurations (Kolar et al., 2021).

Bayesian optimization operates by first defining a set of hyperparameters to be tuned. In Kolar et al. (2021), these include learning rate, momentum, activation function, kernel size, and number of kernels. For each hyperparameter, a predefined search space must be specified, determining the range of values the algorithm is to explore. In addition, a loss function must be selected, which serves as the objective that the tuning process seeks to minimize. The selected loss function is central to the execution, as it provides the feedback signal guiding the iterative search toward improved model performance. The algorithm then constructs a probabilistic surrogate model, commonly based on a Gaussian process, to approximate the relationship between hyperparameters and the loss function. Based on this model, an acquisition function is used to determine the next set of hyperparameters to evaluate. Through iterative updates, the algorithm efficiently converges toward an optimal or near-optimal configuration without requiring exhaustive search of the entire parameter space (Kolar et al., 2021).

4. Data

This section presents the data that constitutes the empirical basis for the development and evaluation of the model in this thesis. First, the data collection process is described, followed by a presentation of relevant metadata.

4.1 Data Collection

Data collection in this thesis follows the standardized methodology used at WSP, that is conducting a site visit and performing measurements of power demand, at a temporal resolution sufficient to capture peak loads that influence design decisions.

The power demand is derived from Equation 1:

$$P = VIcos(\varphi) \tag{1}$$

where P represents the active power, V represents the voltage, I represents the representative current value and cos (φ) represents the power factor. The voltage and power factor are assumed to be constant (Glover et al., 2023).

According to Interviewee 1, a sampling interval of one to two minutes is typically adequate for capturing significant peaks (Interviewee 1, 2026). In contrast, previous research on applied to electrical load data has employed significantly higher sampling frequencies, such as 100 Hz (Liang et al., 2020), 1/2 Hz (Iversen et al., 2025), 1/30 Hz (Liu et al., 2020), and 1/60 Hz (Li et al., 2020 & Song et al., 2022). These frequencies are at, or exceed, the upper end of the range suggested by Interviewee 1. Furthermore, Interviewee 1 recommends that high-frequency measurements of industrial loads should span one to two weeks, as such loads often exhibit pronounced weekly patterns but are less influenced by seasonal variations. Consequently, a measurement period of this duration is generally sufficient (Interviewee 1, 2026). A notable gap can therefore be identified between research and practice, as studies on super-resolution for load data typically rely on datasets covering at least one month (Iversen et al., 2023).

In this thesis, the primary limiting factor is the available measurement equipment, particularly its data storage capacity. Previous super-resolution studies have been based on datasets containing between 100 million and 500 million data points, corresponding to measurement periods spanning several months or even years, depending on the sampling frequency (Iversen et al., 2023). In contrast, the energy logger used by WSP, a VEGA 78, has an internal storage capacity of slightly above eight million data points (TEquipment, n.d.). In addition, the effective storage capacity available for the relevant dataset is further reduced by the measurement structure of the instrument itself. Several measurements are automatically recorded by the logger and cannot be disabled by the user. For example, when current measurements are enabled, the instrument automatically records values for all three phases, as well as minimum, maximum, and average values for each measurement interval. Furthermore, harmonic measurements are continuously recorded alongside the current measurements (TEquipment, n.d.). As a result, a substantial portion of the available storage capacity is occupied by auxiliary measurements that are not directly relevant to the present thesis, thereby limiting the amount of usable high-frequency data that can be collected. Consequently, the energy logger was configured to record current, including associated parameters such as harmonic components.

This constraint results in a trade-off between sampling frequency and measurement duration. A higher sampling frequency would align with academic practice but limit the duration of the dataset, increasing the risk of failing to capture weekly patterns and deviating from standard industry practice at WSP. Conversely, a lower sampling frequency would allow for a longer measurement period but would diverge from the methodologies commonly used in super-resolution research. However, the study by Liu et al. (2020), which shares several similarities with the present thesis, including the use of an SRP-CNN model, employed a sampling interval of 30 seconds. This frequency lies at the lower end of those used in super-resolution research, while remaining reasonably close to Interviewee 1's recommended range of one to two minutes. At this sampling rate, the VEGA 78 is capable of storing approximately 10 days of measurements, which falls within the range of Interviewee 1's suggested duration of one to two weeks. Accordingly, high-frequency measurements in this thesis were conducted at 30-second intervals over a period of slightly more than nine days. Data collection commenced on Monday, March 23, 2026, and concluded on Wednesday, April 1, 2026.

4.2 Metadata

The Facility operates on a two-shift schedule on weekdays, which means that industrial processes are carried out during two scheduled shifts per day, except on Fridays, when the factory operates on a half-day schedule. During these hours, the main operation of the industrial loads takes place, along with usage related to electric vehicle charging and office operations (Plant Manager, 2026). According to the EO manager, the aggregated load consists of three main loads that draw significantly more power than other loads. Out of these, two operate within the shift hours. The last one starts one hour before the first shift begins and shuts down a few hours after the second shift ends. However, on Fridays it shuts down when the shift ends (EO manager, 2026). According to the Plant Manager (3 February, 2026), the facility is currently operating at 50 percent of full capacity.

5. Methodology

The following chapter presents the methodology used in this thesis. It describes the approach used for model development and the key steps involved in this process. Section 5.1 describes the steps of the data preprocessing, including transforming raw data measurements and creating the datasets with different time resolutions. In section 5.2, the model development process is presented. This includes the model architecture, the selection of loss functions and the hyperparameter tuning process. Section 5.3 outlines the implementation of the model.

5.1 Data Preprocessing

The measurements of the raw data conducted with the VEGA78 instrument recorded current in three phases. To calculate the active power, the current values from the three phases were first combined by calculating their average, resulting in a single representative current value. The power demand was then calculated according to

Equation 1. Both the voltage level and the power factor were obtained from internal documents at WSP related to the project at the facility (WSP, 2025).

The processed data was then used to create two datasets with different temporal resolutions. The first dataset retained the original high-frequency resolution and was used without further aggregation. The second dataset was constructed by aggregating the data into 15-minute intervals. For each interval, the mean value of the active power was calculated. In relation to the high-frequency data, each aggregated 15-minute value was temporally aligned such that the corresponding low-frequency measurement was positioned at the midpoint of the 15 minute interval.

5.2 Model Development

Industrial power consumption data often exhibit complex temporal patterns spanning multiple time scales, including minutes, hours, days, and longer periods. Deep learning approaches have been successfully applied to perform super-resolution under such conditions. Based on this, a deep learning approach was selected as the foundation for the model developed in this thesis. The choice between supervised and unsupervised learning was guided by the fundamental nature of super-resolution, which relies on the relationship between corresponding high- and low-resolution data. In this thesis, the low-frequency dataset was generated directly from the high-frequency measurements by aggregating the original data over longer time intervals. As a result, paired high- and low-resolution datasets were naturally available, making the problem well-suited for a supervised super-resolution approach in which the model learns how low-frequency patterns relate to their high-frequency counterparts over time.

Previous studies on energy data super-resolution via supervised deep learning applications indicate that CNN-based models have been successfully implemented for this purpose, in particular in applications on industrial energy data. Specifically, SRP-CNN models have been successfully implemented in cases where training data is limited and generally demonstrate good generalizability. Consequently, this thesis implements an SRP-CNN model to generate high-frequency data from low-frequency measurements for the specific client and project context considered.

The model was implemented in Python (Python Software Foundation, 2023), utilizing the PyTorch library to enable the construction and training of the CNN architecture. The selection of model hyperparameters was determined through an algorithm-based tuning procedure. However, prior to this, a key design decision concerned the choice of loss function, as it significantly influences both model performance and training dynamics, particularly in relation to the evaluation of different hyperparameter configurations during the tuning process. In the literature on the choice of loss-function for CNN models, it is emphasized that the loss function should be selected based on the nature of the problem and the intended training objectives. To support this decision, a preliminary model was developed, allowing the selection of loss function to be informed by both experimental results and theoretical considerations. As the full dataset was not yet available at this stage, the preliminary model was instead trained on an artificially generated dataset.

5.2.1 Creation of Artificial Dataset

The artificial dataset used for the preliminary model was constructed based on the compiled metadata, that is, qualitative input from the Plant Manager and information from technical description derived by WSP. The construction of the model was also based on a set of simplifying assumptions.

First, a proportional relationship between maximum current demand and production level was assumed. This implies that when the facility operates at approximately half of its full production capacity, as would be during the measurement period, the maximum current demand is also assumed to be halved. That is, given the maximum current level of 800 A, the peak current demand measured during this period would be 400 A. In addition, the average current demand was assumed to be approximately half of the temporal maximum demand, hence 200 A. It was further assumed that power demand outside of active production hours is slightly above zero. As a result, the simulated power demand fluctuates between zero and approximately half of the global peak value, with an average around one quarter of the peak value.

The temporal variation in the dataset was designed to reflect the daily operational cycle of the facility. This includes an increase in demand during the morning as the first work shift begins, a temporary reduction around midday during the transition between shifts, and a gradual decrease toward the end of the day. To introduce realistic short-term fluctuations, variations in power demand during active production periods were modeled using a combination of three sinusoidal functions with different periods (15, 5, and one minute) and amplitudes. Additionally, one of the loads identified by the EO manager as among the largest was modeled separately, to reflect that its operation is not directly correlated with the daily shift schedule and remains active beyond standard working hours. Finally, random noise was added to the dataset to better approximate real-world measurement variability.

This dataset was then used to train an SRP-CNN algorithm identical to the one presented by Liu et al. (2020), that is, the same architecture and hyperparameters. This laid the foundation for the experimental determination of loss function.

5.2.2 Choice of Loss Function

The choice of loss function reflects the goal of capturing power peaks in the reconstructed high-resolution power data. According to previous studies, RMSE is an established loss function that is frequently used for both regression and image-related problems. Since it is critical in this thesis to capture large errors at power peaks, RMSE was considered relevant to apply in the model as it penalizes larger deviations more heavily than smaller ones. Furthermore, to facilitate the interpretation of the results, RMSE was chosen over MSE to express the error in the same unit as the original data.

In order to better adapt the loss function to the specific problem, where the greatest weight should be placed on extreme values such as power peaks, a weighted variant of RMSE was also selected. Previous studies have shown that weighted RMSE allows different parts of the data to be assigned different weights in the loss calculation. In this way, deviations at power peaks can be given an even higher priority than other deviations in the signal.

In this thesis, the weighted RMSE function is implemented using a peak-weighted MSE, where the weights are calculated based on the levels of the current signal. Since the objective in this thesis is to accurately reconstruct power peaks in the high-resolution data, higher weights are assigned to data points near power peaks. To define what should be classified as a power peak, a threshold value is computed using an upper quantile. This upper quantile was chosen to correspond to the top five percent of power values in the data in order to ensure that the weighting focuses on the most critical power peaks in the signal. Data points near or above this level are assigned higher weights in the loss function, with increasing weights for higher values.

The weighting is applied continuously via a sigmoid function. This means that data points approaching or exceeding the threshold value are gradually assigned higher weights. This approach was chosen since not only the peak values themselves, but also the surrounding areas of the power peaks, are important to reconstruct accurately as well. In this way, errors near power peaks can be emphasized while still considering the remaining parts of the signal. Based on the weighting of each data point, the total loss is then calculated as the mean of the sum of the squared errors. The weighted RMSE allows the model to focus on minimizing errors in critical areas with high power levels.

Furthermore, SSIM was selected as a suitable loss function for the developed model, as it takes into account the overall structural characteristics of the data. Given the objective of ensuring that the reconstruction of high-resolution data can accurately capture power peaks, it is important that the structure of the signal is preserved as accurately as possible, particularly in areas with such peaks. By using SSIM, the model is able to better preserve these structural properties in the high-resolution reconstruction. This enables a more accurate representation of the shape of the power peaks, which is central to the thesis' objective.

5.2.3 Hyperparameter Tuning

Literature on hyperparameter tuning underlines the importance of considering the individual task and goal of an artificial neural network for super-resolution. Although the study by Liu et al. (2020) is closely related to the present thesis and served as an important source of inspiration in the development of the preliminary model, the two cases are not identical. In particular, differences in the dataset characteristics, problem formulation, and choice of loss function imply that the optimal set of hyperparameters is not directly transferable. Consequently, the hyperparameters required for this specific case needed to be systematically evaluated and adapted. Based on the recommendations found in the literature, a set of candidate hyperparameters and associated search ranges was defined for the optimization algorithm to explore. The selected hyperparameters and their corresponding candidate values are presented in Table 1.

Bayesian optimization was selected as the tuning strategy due to the computational complexity associated with the chosen loss functions. Specifically, the weighted loss function requires identification of local maxima for each data point, resulting in approximately 8,300,000 extra computations per evaluation step. Given the limited computational resources available, exhaustive search methods were deemed infeasible. In addition to its computational efficiency, Bayesian optimization has been shown in previous studies to perform well in hyperparameter tuning for SRP-CNN models, particularly in industrial applications. This further supports its suitability for the present thesis. Lastly, Bayesian optimization has proven effective in cases where the relationship between hyperparameters and model performance is unknown, as is the case here.

Table 1: Hyperparameters and search ranges explored in the Bayesian optimization.

Hyperparameter	Search range
Optimizer	Adam, SGD
Learning rate	Between 10^{-5} and 10^{-2}
Batch size	8, 16 or 32
Activation function	prelu, relu, leaky relu or GCU
Number of layers	Between 2 and 10
Number of kernels	64, 128 or 256
Size of kernel	3, 5, 7, or 9

5.3 Implementation

The implementation of the model followed a sequential and iterative process. First, a preliminary model was developed using artificially generated data, which enabled an experimental evaluation of suitable loss functions. This empirical foundation, in combination with theoretical insights and findings from previous studies, resulted in the selection of three candidate loss functions. Subsequently, hyperparameter tuning was performed using Bayesian optimization, applied separately for each of the selected loss functions. During training, a stride length of one was used, meaning that the convolutional kernel moved one time step at a time across the input sequence, corresponding to the 15-minute resolution of the low-frequency data. This enabled the model to evaluate overlapping sections of the signal throughout the training process. The models were trained for a maximum of 100 epochs. However, to reduce the risk of overfitting and unnecessary computation, early stopping was implemented. Training was therefore terminated if the validation performance did not improve for 10 consecutive epochs. Once the optimal hyperparameter configurations had been identified, the final SRP-CNN models were constructed for each loss function using their corresponding tuned parameters. These models were then applied to the measured dataset and could then be evaluated.

5.4 Evaluation

The evaluation of the implemented models consists of two parts: an evaluation of the signal's overall reconstruction accuracy, and an evaluation of the model's ability to reconstruct critical power peaks. Since the primary objective of the model is to reconstruct high-resolution power data while preserving power peaks, the evaluation process focuses largely on assessing the model's performance in regions with high power levels.

To evaluate the model's performance in reconstructing the overall signal, MAE is used, which is calculated between the measured high-resolution data and the reconstructed high-resolution signal. MAE was chosen as the evaluation metric because it provides a representative measure of the average reconstruction error across the entire signal. This allows for an evaluation of the model's overall reconstruction capability.

Furthermore, to specifically evaluate how the model performs at high power levels, an analysis of the signal's highest measured points was conducted. The 15 highest measured values in the high-resolution measured data were identified and used as reference points to evaluate how well the model can reconstruct such power levels. Since the signal contains measurement noise, an exclusion window of 120 seconds was selected to avoid identifying multiple points within the same peak region. This means that the identified points represent only the distinct highest points distributed across the entire signal.

For all 15 identified values, the corresponding reconstructed value was calculated and compared with the measured value. In the comparison, the timestamp of each measured value, the power levels of both the measured and corresponding reconstructed value, as well as the absolute and percentage error between the two signals, were analyzed. This enabled a quantitative evaluation of the model's ability to reconstruct extreme power levels.

To further investigate the relationship between peak magnitude and reconstruction accuracy, a correlation analysis between measured peak magnitude and relative error was conducted. Pearson correlation coefficients and corresponding p-values were calculated both for the 15 identified peak values and for the complete dataset. This analysis was performed to examine whether reconstruction accuracy systematically deteriorates as the magnitude of the power demand increases, and to determine whether such a relationship is specific to extreme peak events or present throughout the entire signal.

Additionally, a visual evaluation of the model's reconstruction capability was conducted using both a plot of the entire signal and zoomed-in plots around the identified highest measured values. In the plot of the entire signal, the identified measured values were marked to facilitate an overall analysis of the model's performance in these areas. The zoomed-in plots, one for each identified measured value, were generated with a time window of 10 minutes before and after each selected point, to facilitate the visual assessment of how accurately the local structures around the highest values have been reconstructed.

6. Results

This chapter presents the results obtained in this thesis. First, the optimal hyperparameter configurations identified through the Bayesian optimization process are presented for each model. Thereafter, the overall performance of the models is evaluated from a broader perspective using general reconstruction metrics and visual comparisons between measured and reconstructed signals. Finally, particular attention is given to the reconstruction of peak power demand events, where the models are assessed based on their ability to preserve the magnitude and timing of the highest measured datapoints in the dataset.

6.1 Hyperparameters

The hyperparameter tuning conducted using Bayesian optimization identified optimal model parameters for the three investigated loss functions. The resulting hyperparameters obtained through the Bayesian optimization process are presented in Table 2.

For all loss functions, the optimizer Adam was selected as the optimizer together with a batch size of 32. More noticeable differences can, however, be identified among the remaining hyperparameters. The learning rate varies significantly among the three loss functions, with SSIM achieving the highest learning rate while the model using weighted RMSE resulted in the lowest. Furthermore, differences in the chosen activation function can be observed. For the models using RMSE and weighted RMSE, PReLU was selected as the activation function, whereas the SSIM-based model instead resulted in ReLU.

The hyperparameter tuning process also resulted in different model structures. The model using weighted RMSE obtained the highest number of kernels, 128, while the RMSE and SSIM models both resulted in 64 kernels. Since the number of kernels affects how many patterns the model can identify in the data, this indicates that the model with weighted RMSE can capture a greater variety of patterns in the signal compared to the other models.

Regarding the number of layers in the models, it can be observed that the SSIM-based model consists of only two layers, corresponding to a less complex network structure. The results also show that the RMSE and weighted RMSE models both resulted in a deeper network, with a total of five layers each. However, the models differ in their selected kernel sizes, where the SSIM-based model resulted in two larger kernels of size 7, while the RMSE and weighted RMSE models instead combined multiple different kernel sizes.

Table 2: Optimal hyperparameters for respective loss function.

Hyperparameter	RMSE	SSIM	Weighted RMSE
Optimizer	Adam	Adam	Adam
Learning rate	$118.97 * 10^{-5}$	$998.44 * 10^{-5}$	$13.863 * 10^{-5}$
Batch size	32	32	32
Activation function	prelu	relu	prelu
Number of layers	5	2	5
Number of kernels	64	64	128
Size of kernels	5, 3, 7, 3, 5	7, 7	3, 5, 3, 9, 5

6.2 Model Performance

The MAE values obtained for the three investigated loss functions when evaluating the reconstruction accuracy across the entire signal are presented in Table 3. The results show relatively small differences between the models, indicating that all models demonstrated a comparable reconstruction capability for the overall signal.

Table 3: MAE [kW] for the models with the respective loss functions

RMSE	SSIM	Weighted RMSE
7.97	8.35	8.68

The model resulting in the lowest MAE value was the RMSE-based model, with an MAE of 7.967, followed by the SSIM-based model with a value of 8.351. The weighted RMSE model resulted in the highest MAE value, 8.684, indicating that this model achieved the lowest overall reconstruction accuracy in relation to the measured high-resolution data.

Figures 1-3 present the reconstructed high-resolution signals generated by the RMSE-, SSIM-, and weighted RMSE-based models, together with the measured high-resolution signal and the original 15-minute input signal. The figures also highlight the 15 highest measured data points. As illustrated, all three models are capable of reconstructing the overall structure of the measured signal, including its temporal variations, with relatively high accuracy.

However, differences can be observed in how well the models reconstruct local variations and higher measured values in the signal. In Figure 2, it can be observed that the SSIM model reconstructs a slightly smoother signal with smaller local fluctuations, in contrast to the weighted RMSE-based model in Figure 3, which results in a more variable reconstruction with larger variations. Visually, the weighted RMSE model appears to follow the variations in the measured high-resolution signal with slightly higher accuracy than the other models. The RMSE-based model, shown in Figure 1, visually exhibits an intermediate behavior between the two other models, with some fluctuations while the overall signal structure is reconstructed with relatively high consistency with the measured signal.

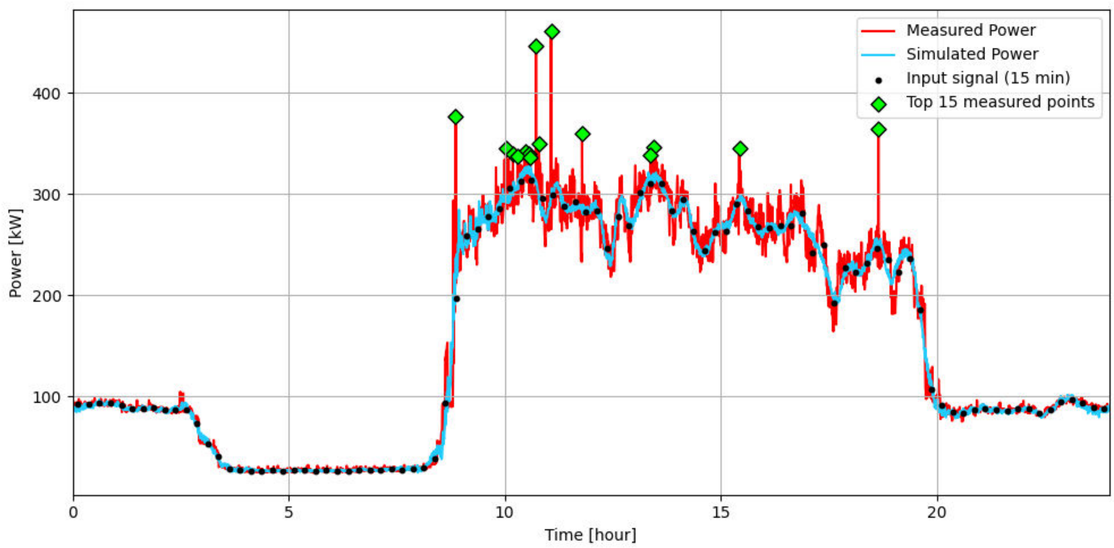

Figure 1: Full signal of measured high-resolution power and the reconstructed signal generated by the RMSE-based model, including the original 15-minute input signal and the identified top 15 measured power values.

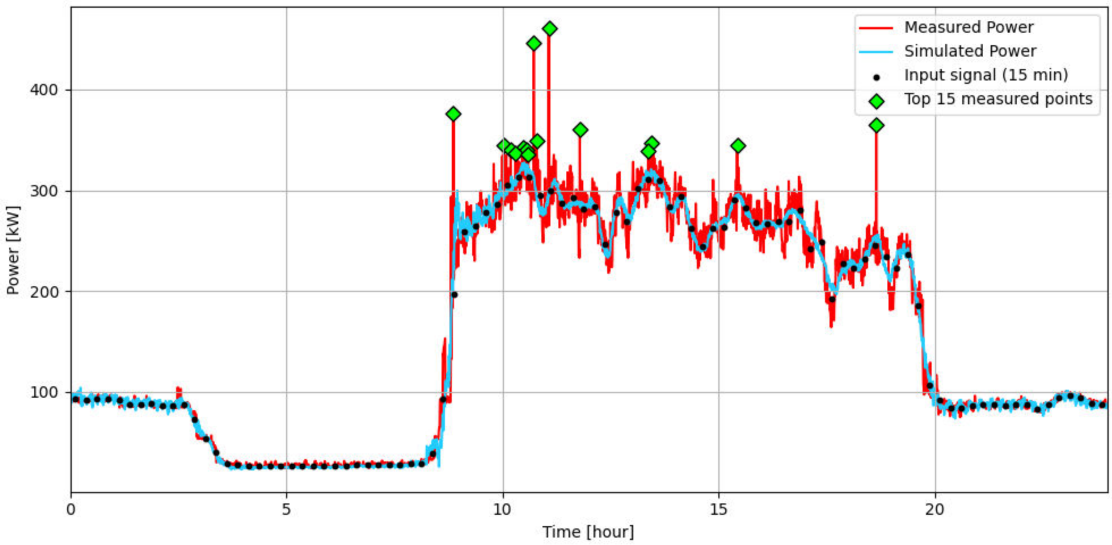

Figure 2: Full signal of measured high-resolution power and the reconstructed signal generated by the SSIM-based model, including the original 15-minute input signal and the identified top 15 measured power values.

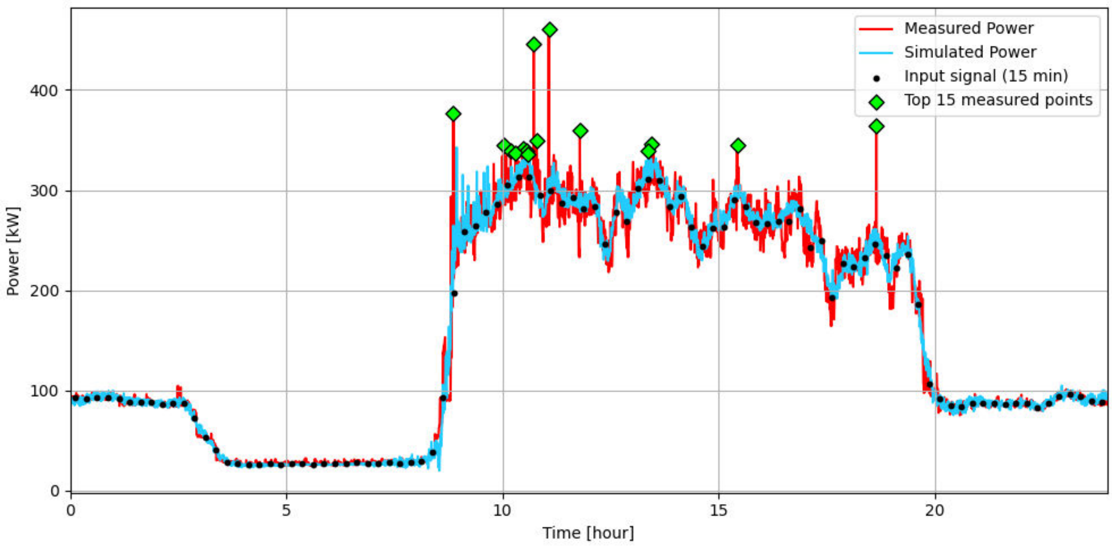

Figure 3: Full signal of measured high-resolution power and the reconstructed signal generated by the weighted RMSE-based model, including the original 15-minute input signal and the identified top 15 measured power values.

6.3 Peak Estimation

In accordance with the stated aim and research questions, particular emphasis is placed on the models' ability to capture peak values. As shown in Table 3, the MAE ranges between 7.967 and 8.684. However, Figures 1-3 indicate that the prediction error is substantially larger at the highest measured peaks. Tables 4-6 therefore present the model performance specifically at these points, expressed in terms of both absolute and relative error. The results show that, for the four largest peaks (rank 1-4), the relative error ranges approximately between 30 percent and 40 percent, whereas the remaining peaks exhibit lower errors, ranging from around 20 percent to values close to zero. Another notable observation is that all model variants display a clear tendency to underestimate peak values, as none of the predicted peak magnitudes exceed the corresponding measured values.

Table 4: Comparison of the 15 largest measured power demand peaks and their corresponding predicted values generated by the SRP-CNN model with the RMSE loss function.

Rank	Time [h]	Measured [kW]	Predicted [kW]	Error [kW]	Error [%]
1	11.07	460.3	298.8	-161.5	35.09
2	10.72	445.8	313.3	-132.5	29.72
3	8.87	376.7	236.1	-140.6	37.32
4	18.65	364.5	252.5	-112.0	30.73
5	11.79	360.2	288.4	-71.8	19.92
6	10.79	349.3	294.4	-54.9	15.72
7	13.46	346.6	315.9	-30.6	8.83
8	15.43	345.0	293.2	-51.8	15.02
9	10.03	344.8	299.4	-45.4	13.17
10	10.48	344.0	320.9	-21.1	6.17
11	10.19	339.6	302.1	-37.5	11.05
12	10.55	339.6	326.6	-13.0	3.834
13	13.38	338.8	312.3	-26.5	7.817
14	10.30	337.1	311.0	-26.1	7.738
15	10.59	335.9	324.9	-11.0	3.283

Table 5: Comparison of the 15 largest measured power demand peaks and their corresponding predicted values generated by the SRP-CNN model with the SSIM loss function.

Rank	Time [h]	Measured [kW]	Predicted [kW]	Error [kW]	Error [%]
1	11.07	460.3	297.7	-162.6	35.32
2	10.72	445.8	314.2	-131.6	29.51
3	8.87	376.7	227.8	-148.9	39.52
4	18.65	364.5	253.7	-110.7	30.38
5	11.79	360.2	286.1	-74.1	20.56
6	10.79	349.3	291.3	-58.0	16.61
7	13.46	346.6	318.6	-27.9	8.065
8	15.43	345.0	290.5	-54.4	15.78
9	10.03	344.8	291.1	-53.7	15.56
10	10.48	342.0	322.5	-19.5	5.700
11	10.19	339.6	302.8	-36.9	10.85
12	10.55	339.6	325.2	-14.4	4.229
13	13.38	338.8	314.7	-24.1	7.117
14	10.30	337.1	307.3	-29.7	8.824
15	10.59	335.9	322.1	-13.8	4.108

Table 6: Comparison of the 15 largest measured power demand peaks and their corresponding predicted values generated by the SRP-CNN model with the weighted RMSE loss function.

Rank	Time [h]	Measured [kW]	Predicted [kW]	Error [kW]	Error [%]
1	11.07	460.3	310.6	-149.7	32.53
2	10.72	445.8	313.0	-132.7	29.77
3	8.87	376.7	264.6	-112.1	29.76
4	18.65	364.5	260.8	-103.6	28.43
5	11.79	360.2	292.9	-67.3	18.70
6	10.79	349.3	302.0	-47.3	13.54
7	13.46	346.6	321.2	-25.3	7.303
8	15.43	345.0	292.5	-52.5	15.22
9	10.03	344.8	301.8	-43.0	12.48
10	10.48	342.0	321.8	-20.1	5.881
11	10.19	339.6	303.4	-36.2	10.67
12	10.55	339.6	339.6	-0.01	0.002945
13	13.38	338.8	319.9	-18.8	5.558
14	10.30	337.1	313.0	-24.0	7.127
15	10.59	335.9	334.4	-1.6	0.4703

All three model variants performed poorly in predicting the peak ranked third in magnitude. At this point, the largest prediction errors observed in the analysis were obtained, with the SSIM-based model exhibiting the highest relative error (39.5%), closely followed by the RMSE-based model (37.3%). In comparison, the weighted RMSE-based model achieved a noticeably lower relative error of 29.8%. Figures 4-6 present a zoomed-in view of the model performance around this peak.

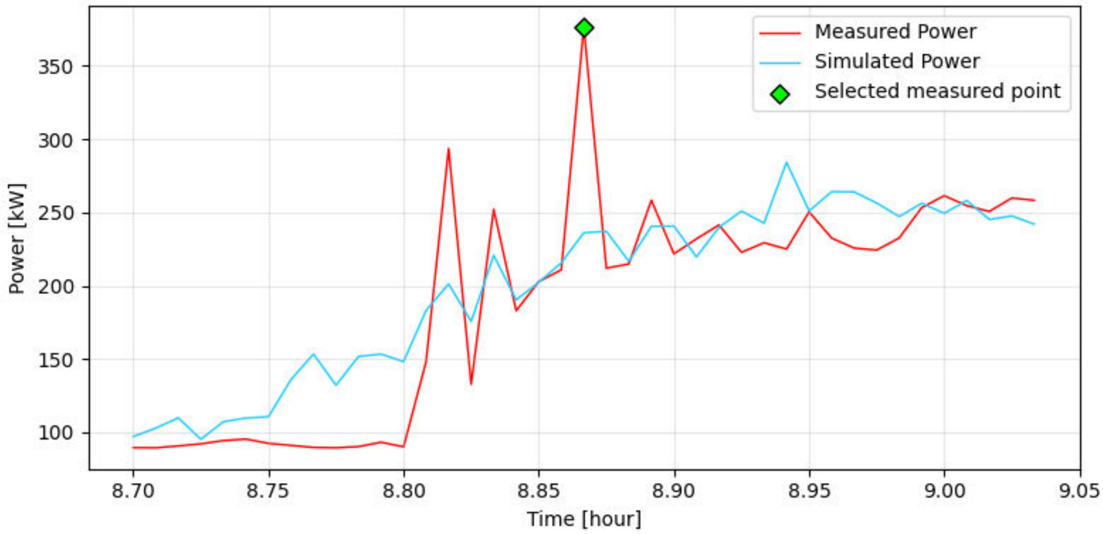

Figure 4: Zoomed view of predictions made by the RMSE-based model at the third-largest peak.

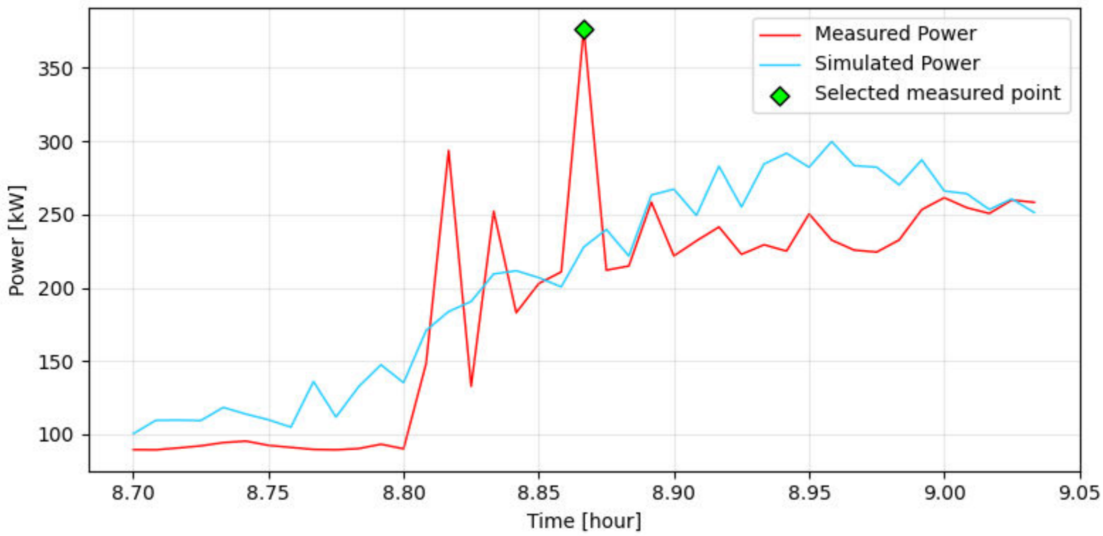

Figure 5: Zoomed view of predictions made by the SSIM-based model at the third-largest peak.

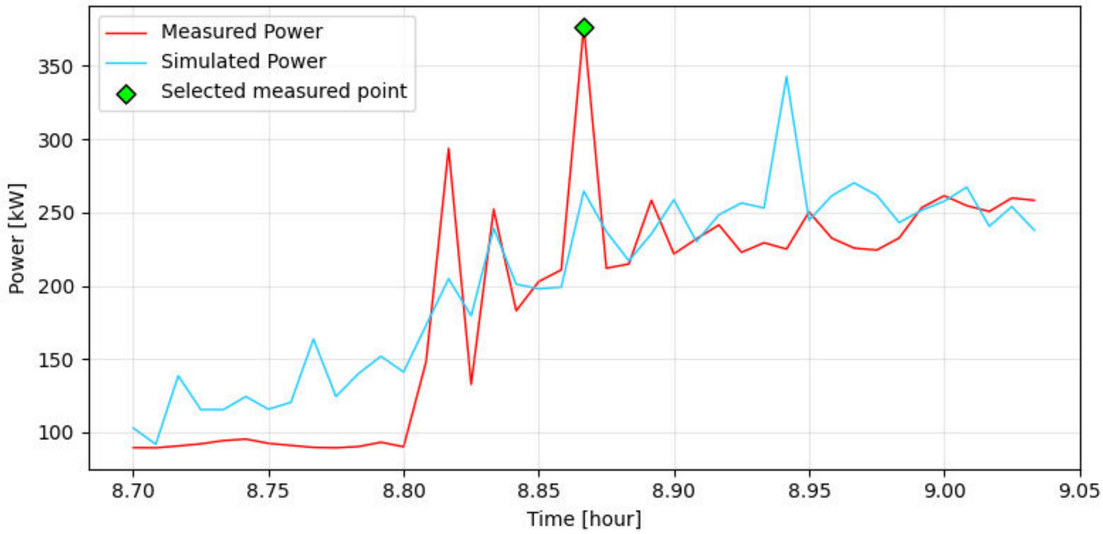

Figure 6: Zoomed view of predictions made by the weighted RMSE-based model at the third-largest peak.

In contrast, all three models performed considerably better for the peak ranked fifteenth in magnitude. The RMSE-based model reconstructed this peak with a relative error of 3.3%, while the SSIM-based model achieved an error of 4.1%. The weighted RMSE-based model reproduced the peak almost perfectly, with a relative error of only 0.5%. Figures 7-9 provide a close-up visualization of the model performance around this point.

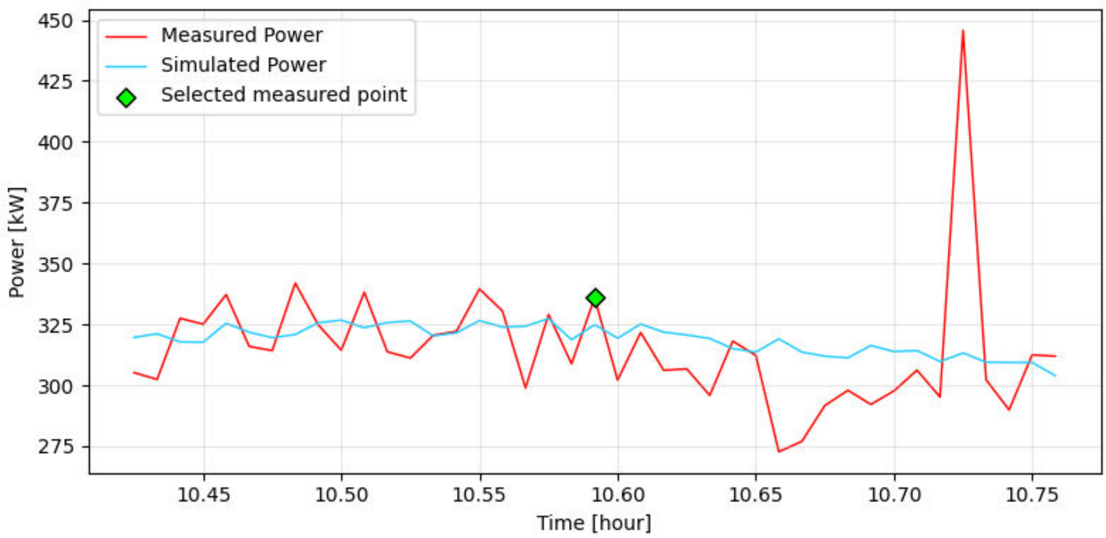

Figure 7: Zoomed view of predictions made by the RMSE-based model at the peak ranked 15.

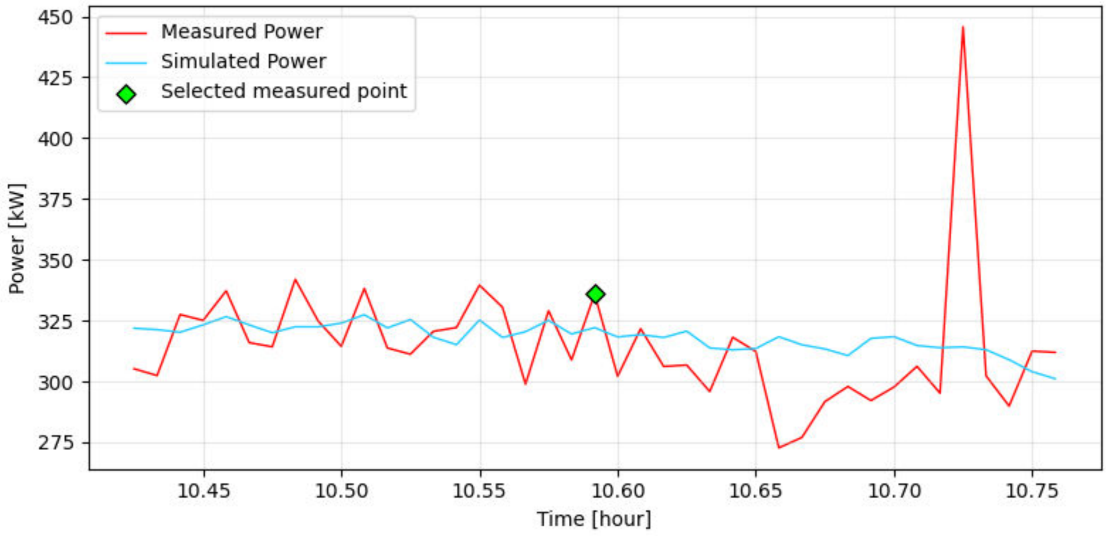

Figure 8: Zoomed view of predictions made by the SSIM-based model at the peak ranked 15.

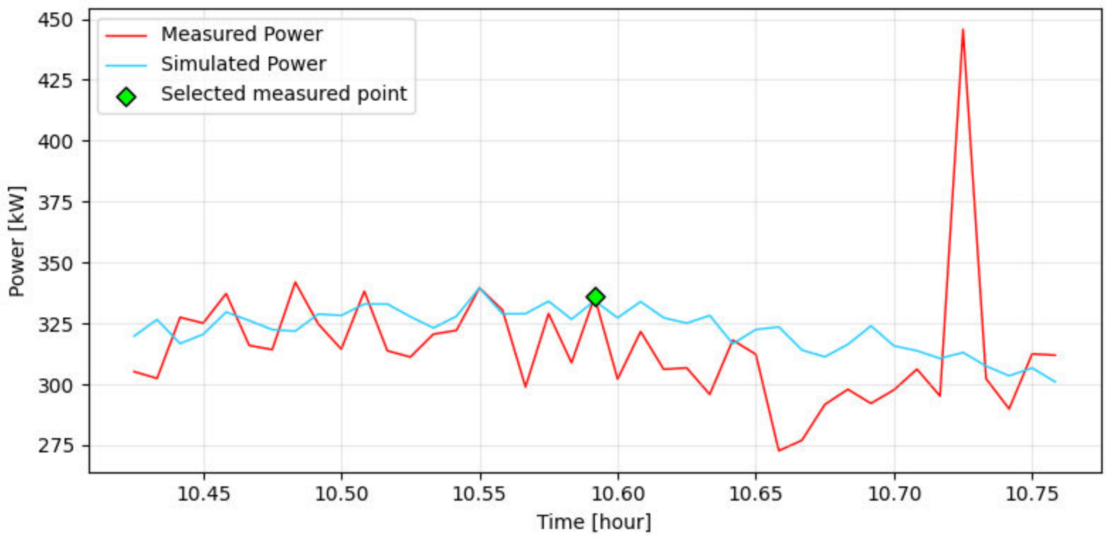

Figure 9: Zoomed view of predictions made by the weighted RMSE-based model at the peak ranked 15.

Figures 4-9 confirmed the pattern earlier identified in Figures 1-3, that the RMSE and SSIM based models have a greater tendency to smooth the signal whilst the weighted RMSE better captures the signals volatile pattern.

The observed decrease in percentage error with increasing rank suggests a potential relationship between peak magnitude and relative error. Figure 10 presents scatter plots illustrating the relationship between relative error and peak magnitude for the different model variants. Statistically, the correlation coefficient was found to be -0.774 for the RMSE-based model, -0.753 for the SSIM-based model, and -0.801 for the weighted RMSE-based model. These results indicate a strong negative correlation between peak rank and relative error, implying that larger peak magnitudes are generally associated with greater relative prediction errors. Furthermore, the corresponding p-values were 0.0007, 0.0012, and 0.0003, respectively, all well below the 0.01 significance level, indicating that the observed correlations are statistically significant.

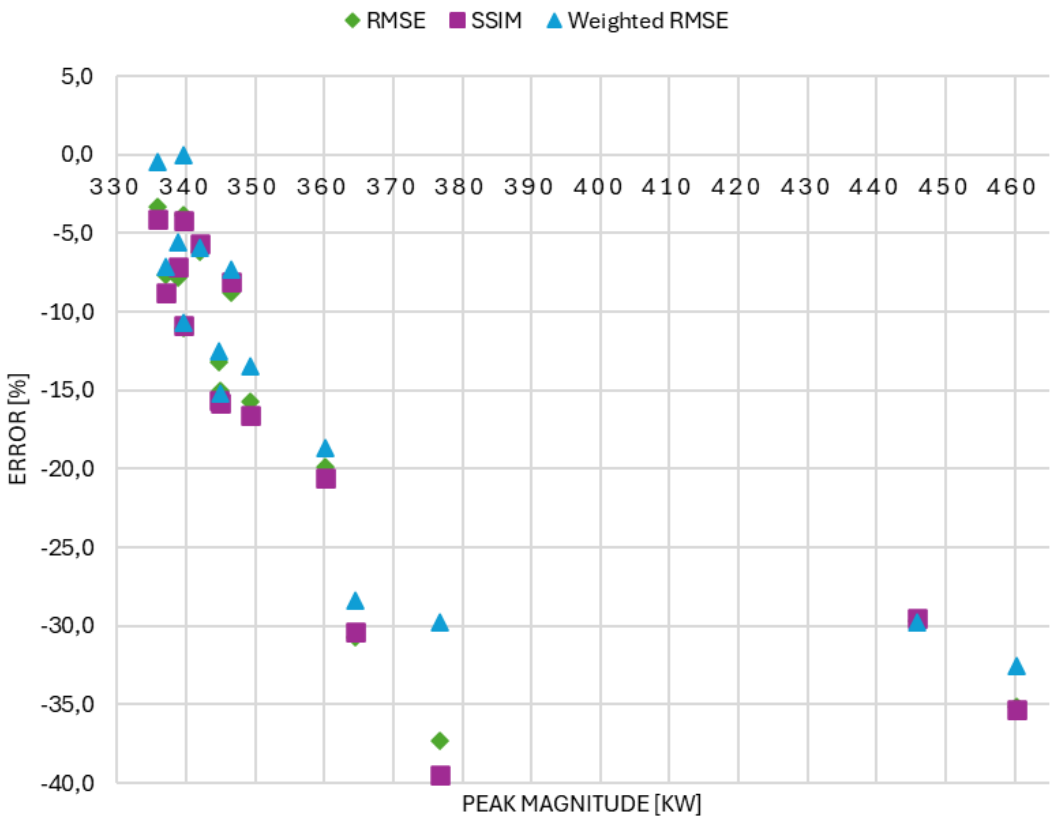

Figure 10: Scatter diagram of the peak magnitude and percentage error

Furthermore, when the analysis of the relationship between peak magnitude and relative error is extended to the entire dataset (2,880 data points), rather than being limited to the 15 highest peaks, a markedly different result is obtained. The correlation coefficients for the respective loss-function-based models are −0.0264, 0.0730, and 0.0469, indicating that no substantial correlation can be observed across the dataset as a whole. The correlation coefficients and corresponding p-values are presented in Table 7.

Table 7: Correlation between data magnitude and relative error, including corresponding p-values, for both the top 15 data points and the full dataset.

	15 largest measured power demand peaks		All data points	
	Correlation coefficient	p-value	Correlation coefficient	p-value
RMSE	-0,77422	0,00070	-0,02642	0,15621
SSIM	-0,75261	0,00121	0,07295	0,00089
Weighted RMSE	-0,80130	0,00033	0,04692	0,01177

7. Discussion

The following chapter discusses the obtained results in relation to the objectives of the thesis. The discussion is divided into three parts. First, the impact of different loss functions on hyperparameter selection and model architecture is analyzed. Secondly, the reconstruction performance of the models is discussed with a particular focus on the reconstruction of power peaks and its relevance for infrastructure dimensioning. Finally, the limitations of the thesis and directions for future research are presented.

7.1 Hyperparameter Selection and Model Architecture

A notable outcome of the Bayesian optimization process is that the different loss functions resulted in substantially different network architectures and training configurations. In particular, the weighted RMSE-based model converged toward the most complex architecture, consisting of the largest number of kernels, a deep network structure, and the lowest learning rate among the investigated models. Since weighted RMSE assigns greater importance to errors associated with high-magnitude peaks, this may indicate that peak-oriented reconstruction constitutes a more complex learning problem than general signal reconstruction. Capturing short-duration extreme events likely requires the model to identify more subtle and localized patterns within the signal, which in turn may necessitate greater representational capacity in the form of additional layers and kernels. The superior peak reconstruction performance observed for the weighted RMSE-based model further strengthens this interpretation.

In contrast, the SSIM-based model resulted in the simplest architecture, with only two convolutional layers, fewer kernels, and a substantially higher learning rate. This may suggest that reconstructing the overall structural characteristics of the signal is a less demanding task than accurately reproducing extreme values. Since SSIM primarily evaluates structural similarity rather than point-wise deviations, the optimization process may favor architectures that capture the general temporal shape and smooth variations of the signal rather than highly localized peaks. This interpretation is also supported by the overall reconstruction results, where all models reproduced the broader signal structure relatively well while simultaneously struggling to accurately reconstruct the largest peaks. The findings therefore indicate that global signal reconstruction and peak reconstruction may represent fundamentally different learning challenges.

The results further suggest that the choice of loss function may implicitly guide the optimization process toward different architectural solutions. The weighted RMSE and RMSE-based models both converged toward deeper architectures using PReLU activation functions, whereas the SSIM-based model instead selected ReLU and a considerably shallower structure. Since activation functions affect how nonlinear relationships are represented within the network, this may indicate that the reconstruction of peak behavior requires a more flexible representation of nonlinear patterns compared to the reconstruction of overall signal structure. Similarly, the larger number of kernels selected for the weighted RMSE-based model may reflect an increased need to identify a broader range of local signal characteristics associated with extreme events. In this sense, the loss function does not merely determine how model performance is evaluated, but may also indirectly shape the type of architecture that is most effective for the corresponding reconstruction objective.

Another important observation concerns the learning rates identified through Bayesian optimization. The weighted RMSE-based model resulted in the lowest learning rate, whereas the SSIM-based model obtained the highest. This outcome appears theoretically reasonable, as weighted RMSE amplifies the contribution of large errors during optimization. Consequently, excessively large parameter updates could destabilize the training process, making a lower learning rate advantageous. In contrast, SSIM-based optimization focuses more on broader structural similarity and may therefore permit more aggressive parameter updates without destabilizing convergence. The fact that Bayesian optimization consistently identified Adam as the preferred optimizer across the investigated loss functions also aligns with previous literature emphasizing Adam's robustness, fast convergence, and suitability for problems involving large parameter spaces and noisy gradients.

The results additionally illustrate how Bayesian optimization can function not only as an optimization tool, but also as an indirect analytical instrument for understanding the relationship between hyperparameters and model performance. One of the motivations for selecting Bayesian optimization in this thesis was that the relationship between hyperparameters and reconstruction performance was not known beforehand. By iteratively constructing a probabilistic approximation of this relationship, the method identifies promising regions of the search space without exhaustively evaluating all possible parameter combinations. Although the resulting hyperparameter configurations cannot be interpreted as globally optimal solutions, they may nevertheless provide indications regarding which architectural characteristics appear beneficial for different reconstruction objectives. For example, the results suggest that peak-oriented reconstruction may benefit from deeper architectures, larger representational capacity, and more conservative learning dynamics.

At the same time, important limitations associated with Bayesian optimization must be acknowledged. Since the method does not evaluate the entire search space, there is no guarantee that the identified configurations correspond to the true global optimum. Instead, the obtained hyperparameters should be interpreted as near-optimal solutions within the explored parameter space. Furthermore, the optimization process is inherently dependent on the search space defined by the researchers. The algorithm can only evaluate parameter combinations within the predefined intervals and categories included in the optimization setup. Consequently, different search ranges or additional architectural alternatives may potentially have produced different results. The identified configurations therefore reflect not only the characteristics of the reconstruction problem itself, but also the assumptions and constraints embedded within the experimental design.

7.2 Peak Reconstruction Performance and Engineering Relevance

The results regarding how well the various models were able to reconstruct the measured high-resolution data indicate a clear trade-off between the model's overall reconstruction capability and its ability to reconstruct specific areas with high power levels. Although the weighted RMSE model resulted in the highest MAE value, the results indicate that this model was better able to reconstruct local differences in areas with high power levels compared to the RMSE- and SSIM-based models. This suggests that the model prioritizes the reconstruction of extreme power levels over minimizing the average error across the total signal. This behavior is consistent with how the weighted RMSE function is designed, where data points near extreme measurements are penalized more heavily than other data points. Thus, the results can be interpreted as indicating that an increased focus on specific areas of the signal appears to occur at the expense of the model's overall reconstruction capability, as observed through the MAE evaluation.

However, the results suggest that the weighted RMSE-based model visually succeeds in preserving a greater degree of local variation in the reconstructed signal compared to the RMSE- and SSIM-based models. Since the measured signal exhibits oscillatory behavior, the weighted RMSE-based model appears to be able to mimic these fluctuations more accurately than the other models. The fact that the MAE value is nevertheless highest for the weighted RMSE model may be due to small, local deviations between the measured and reconstructed signals contributing to the total MAE value, even when the model reproduces the overall oscillatory behavior of the measured signal relatively accurately.

With regard to the evaluation of the highest measured values, several of these extreme values occurred during very short time intervals, isolated from surrounding power levels. Since these values deviate sharply from the rest of the signal, this may indicate that they are affected by measurement noise. From a dimensioning perspective, such isolated extreme values may be of lesser significance compared to longer periods characterized by elevated power levels. According to WSP dimensioning procedures, transformers are not dimensioned solely based on the highest measured values, but rather significantly considering the temporal aspect of load profiles. Since transformers have an inherent thermal inertia, they can handle shorter and infrequent periods of higher load. Consequently, the highest isolated measurement points in the signal are generally not as critical for dimensioning purposes as longer periods with high power values.

The results indicate that the reconstruction problem becomes fundamentally more difficult as the signal deviates further from its normal operating regime. While all three models demonstrated relatively strong overall reconstruction performance across the complete dataset, substantial errors were consistently observed at the highest peak values. One possible explanation for this behavior is the limited representation of extreme events within the training data. Since high-magnitude peaks occur relatively infrequently, the models are exposed to comparatively few examples of such events during training. Consequently, the learning process becomes dominated by normal operating conditions and moderate variations in the signal, while the rare extreme events contribute less to the overall parameter optimization.

This challenge is particularly relevant in this thesis, where the available dataset was relatively limited. In addition to the overall number of data points, the distribution and frequency of extreme peaks within the dataset likely influenced the results. During the 24-hour evaluation period, only a small number of pronounced peak events were observed, resulting in limited representation of such events during training. A different dataset, for example from another industrial process or operating environment, could exhibit substantially different load characteristics and peak distributions, thereby providing either a more or less favorable basis for model training and reconstruction performance.

To partly compensate for the limitation relating to rare occasions of extreme peaks, a stride length of one was used during training, allowing the convolutional kernels to process overlapping regions of the signal and thereby increasing the effective number of training observations. Nevertheless, although this approach likely improved the model's ability to learn local temporal structures, it does not fully address the scarcity of extreme peak events in the dataset.

Additionally, another attempt to mitigate this problem was made through the introduction of the weighted RMSE loss function, which assigns greater importance to errors occurring at higher values. As discussed previously, this approach improved the reconstruction of extreme peaks compared to the RMSE- and SSIM-based models. However, despite this improvement, the weighted RMSE model still exhibited substantial errors for the largest peaks. This suggests that changing the loss function alone may not fully compensate for the limited representation of extreme events in the training data.

Another possible explanation is that extreme peaks may partly originate from stochastic or processes that are inherently difficult to infer from low-resolution measurements. Such events may be driven by temporary industrial loads, random operational fluctuations, or human-related factors that do not follow stable and recurring patterns within the data. In such cases, there may simply be insufficient information available in the low-resolution signal for the model to reconstruct the corresponding high-resolution peak accurately, regardless of model complexity or optimization strategy. This interpretation also aligns with the fundamental limitations of super-resolution methods, where high-frequency information may be partially lost during temporal aggregation.

A further notable observation is the systematic underestimation of peak values across all model variants. None of the reconstructed peaks exceeded the corresponding measured values, indicating a clear smoothing tendency within the models. Such behavior is common in regression-based deep learning models, where the optimization process favors minimizing overall prediction error across the full dataset. As a result, the models tend to converge toward more conservative predictions and avoid highly uncertain extreme estimations. While this behavior improves overall reconstruction stability, it simultaneously reduces the ability to accurately reproduce short-duration high-magnitude events.

The correlation analysis further strengthens these observations. When the entire dataset was considered, the correlation between signal magnitude and relative error was negligible, indicating that the models generally maintained stable reconstruction performance across normal operating conditions and moderate load variations. In other words, no substantial general relationship between higher signal values and increasing reconstruction error could be identified across the dataset as a whole.

However, a markedly different pattern emerged when the analysis was restricted to the highest peaks. For the 15 largest measured values, all models demonstrated a strong and statistically significant negative correlation between peak magnitude and relative error, indicating that reconstruction accuracy deteriorated as the magnitude of the extreme events increased. This suggests that the challenge is not general reconstruction, or even reconstruction during periods of elevated power demand, but rather the reconstruction of rare extreme events. Consequently, the results indicate that super-resolution models may successfully reconstruct the overall temporal structure of industrial power signals while simultaneously struggling to preserve the most critical peaks relevant for infrastructure dimensioning applications.

7.3 Future Studies

An important limitation of the present thesis is the relatively short measurement period compared to previous studies on super-resolution for industrial load data. To establish more robust knowledge regarding peak-aware super-resolution methods for power demand reconstruction, future studies should therefore investigate longer time periods and include datasets from multiple industries with varying load characteristics and operational patterns. As of now, the generalizability of the results remains limited, since the observed performance may be strongly influenced by the specific load behavior and peak distribution present in the investigated dataset.

Furthermore, the question of how peaks should be defined emerged as a central methodological challenge throughout this thesis. Although there is a clear and well-established understanding of energy peaks among engineers in professional practice, this concept must be expressed in precise mathematical terms to enable its application in machine learning. Such a formal definition is necessary both for the development of a weighted loss function and for the evaluation. In this thesis, peak classification was primarily based on magnitude. The weighted loss function was defined with reference to the top 5% highest values relative to measurements within the surrounding 24-hour period. For evaluation purposes, the 15 highest observed values were selected for detailed analysis. However, these thresholds remain to some extent arbitrary, as the analysis could equally have been based on alternative cut-offs, such as the top 2.5% or 10%, as well as the 10 or 20 highest values. Moreover, some extreme observations may originate from measurement noise or temporary disturbances rather than representing physically meaningful load events

At the same time, very short-duration spikes may not necessarily be relevant from an engineering perspective, since electrical infrastructure such as transformers is influenced not only by instantaneous peak magnitude but also by load duration due to thermal inertia. Together, these considerations indicate that a purely magnitude-based definition of peaks may not be fully sufficient for evaluating engineering-relevant reconstruction performance.

Several alternative approaches for mathematically defining peaks were explored during the course of the thesis. One potential approach involved defining peaks based on the derivative of the signal, where a peak could be characterized as a sequence in which the derivative is initially positive, reaches zero at the local maximum, and subsequently becomes negative. Such an approach would shift the focus from isolated data points toward peak events as temporal sequences. However, the high level of noise present in the signal complicates derivative-based approaches, as meaningful peak detection would likely require some degree of signal smoothing. This introduces a methodological contradiction, since smoothing risks suppressing precisely the short-duration high-frequency variations that the super-resolution model aims to reconstruct and preserve. Consequently, the problem of defining engineering-relevant peaks remains an important area for future research within machine learning-based super-resolution of industrial power demand data.

8. Conclusion

The results in this thesis demonstrate that an SRP-CNN model can be developed to reconstruct high-resolution power demand data from 15-minute measurements using supervised learning. The choice of loss function and hyperparameters was shown to be central to the reconstruction performance. Moreover, Bayesian optimization proved to be effective for adapting the SRP-CNN architecture to the industrial dataset. The results indicate that the weighted RMSE loss function achieved the strongest ability to reconstruct local variations and power peaks.

The results further show that all investigated SRP-CNN models were capable of reconstructing the overall temporal structure of the measured high-resolution signal with relatively high accuracy. However, reconstructing power peaks in the dataset proved to be significantly more challenging than reconstructing the overall structure of the signal. All models exhibited a strong negative correlation between increased peak magnitude and reconstruction accuracy for the fifteen largest measured values. In contrast, no substantial correlation was observed between signal magnitude and reconstruction error across the full dataset. These results demonstrate that the reconstruction errors were primarily associated with the extreme power peaks rather than with the general signal reconstruction.

This thesis demonstrates that machine learning-based super-resolution models provide valuable support for general power demand analysis and the reconstruction of broader load patterns. However, the investigated models reconstruct individual peaks with errors of such magnitude that they are not sufficiently reliable for decision-making purposes. The results also suggest that the models were more successful in reconstructing longer periods of elevated load compared to isolated spikes, indicating greater practical relevance in transformer dimensioning applications due to the thermal inertia of electrical equipment.

Beyond the empirical findings, this thesis contributes to the research field of machine learning-based super-resolution with a specific focus on the reconstruction of power demand peaks. In contrast to previous studies that have primarily examined overall reconstruction accuracy, this study investigates how well super-resolution models can preserve extreme power values relevant for engineering applications. The findings provide new insights into both the potential and limitations of deep learning-based super-resolution techniques for peak-sensitive power demand analysis.

References

Published Sources:

Anagun, Yildiray, Isik, Sahin and Seke, Erol (2019). SRLibrary: Comparing different loss functions for super-resolution over various convolutional architectures. *Journal of Visual Communication and Image Representation*. 61, pp. 178-187, doi: https://doi.org/10.1016/j.jvcir.2019.03.027

Arabboev, Mukhriddin, Saydiakbarov, Saidakmal, Begmatov, Shohruh, Nosirov, Khabibullo and Rikhsivoev, Mokhirjon (2024). A comprehensive review of image super-resolution metrics. *Acta IMEKO*. 13(1), p. 1-17, doi: https://doi.org/10.21014/actaimeko.v13i1.1679.

Ayeni, Joshua Ayobami (2022). Convolutional Neural Network (CNN): The architecture and applications. *Applied Journal of Physical Science.* 4(4), p. 42-50, doi: https://doi.org/10.31248/AJPS2022.085

de-Paz-Centeno, Iván, García-Ordás, María Teresa, García-Olalla, Oscar, Arenas, Javier and Alaiz-Moretón, Héctor (2021). M-SRPCNN: A Fully Convolutional Neural Network Approach for Handling Super Resolution Reconstruction on Monthly Energy Consumption Environments. *Energies*. 14 (16), 4765, doi: https://doi.org/10.3390/en14164765

EIFS 2025:1 *Energimarknadsinspektionens föreskrifter och allmänna råd om mätning och rapportering av överförd el*. Eskilstuna: Energimarknadsinspektionen. https://ei.se/download/18.5094606019684fe27e01895/1746156902909/EIFS-2025-1-om-m%C3%A4tning-och-rapportering-av-%C3%B6verf%C3%B6rd-el.pdf [Accessed: April 5, 2026]

Elharrouss, Omar, Mahmood, Yasir, Bechqito, Yassine, Serhani, Mohamed Adel, Badidi, Elarbi, Riffi, Jamal and Tairi, Hamid (2025). Loss Functions in Deep Learning: A Comprehensive Review. *arXiv*. https://arxiv.org/pdf/2504.04242v1

Glover, J. Duncan, Sarma, Mulukutla S., Overbye, Thomas and Birchfield, Adam (2023). *Power System Analysis and Design.* 7. ed. Cengage Learning. https://www.cengage.uk/c/power-system-analysis-and-design-si-edition-7e-glover-sarma-overbye-birchfield/9780357676196/?filterBy=Higher-Education [Accessed: May 11, 2026].

Goodfellow, Ian, Bengio, Yoshua, & Courville, Aaron (2016). *Deep Learning*. Cambridge: MIT Press. https://mitpress.mit.edu/9780262035613/deep-learning/ [Accessed: May 11, 2026].

Gupta, Jaya, Pathak, Sunil and Kumar, Gireesh (2022). Deep Learning (CNN) and Transfer Learning: A Review. *Journal of Physics*. 2273(1), doi: https://doi.org/10.1088/1742-6596/2273/1/012029

Hammad, Mahmoud A., Jereb, Borut, Rosi, Bojan & Dragan, Dejan (2020). Methods and models for electric load forecasting: A comprehensive review. *Logistics, Supply Chain, Sustainability and Global Challenges*. 11(1), p. 51-76, doi: https://doi.org/10.2478/jlst-2020-0004

Hassan, Esraa, El-Rashidy, Nora, Elbedwehy, Samar, Abd El-Hafeez, Tarek, Saber, Abeer and Shams, Mahmoud Y. (2025) Exploring the frontiers of image super-resolution: A review of modern techniques and emerging applications. *Neural Computing and Applications*. 37(22), p. 17913-17961, doi: https://doi.org/10.1007/s00521-025-11331-1

Hodson, Timothy O. (2022) Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development.* 15, pp. 5481 - 5487, doi: https://doi.org/10.5194/gmd-15-5481-2022

Iversen, Malin, Khan, Mehak, Miraki, Amir and Arghandeh, Raza (2023) Advancements in super-resolution methods for smart meter data. *Front. Energy Res.* 11, doi: https://doi.org/10.3389/fenrg.2023.1288683

Iversen, Malin, Khan, Mehak, Miraki, Amir and Arghandeh, Raza (2025). T2SR: Super-Resolution in Smart Meter Data Using a Transformer-Based Framework. *IET Smart Grid*. 8(1), doi: https://doi.org/10.1049/stg2.70010

Jeong, In-Jun, Kim, Kwang-Jae and Lin, Dennis K.J. (2010). Bayesian analysis for weighted mean-squared error in dual response surface optimization. *Quality and Reliability Engineering International*. 26(5), pp. 417-430, doi: https://doi.org/10.1002/qre.1058

Kolar, Davor, Lisjak, Dragutin, Pająk, Michał and Gudlin, Mihael (2021). Intelligent fault diagnosis of rotary machinery by convolutional neural network with automatic hyper-parameters tuning using Bayesian optimization. *Sensors*. 21(7), p. 2411, doi: https://doi.org/10.3390/s21072411

Kufel, Jakub, Bargieł-Łączek, Katarzyna, Kocot, Szymon, Koźlik, Maciej, Bartnikowska, Wiktoria, Janik, Michał, Czogalik, Łukasz, Dudek, Piotr, Magiera, Mikołaj, Lis, Anna, Paszkiewicz, Iga, Nawrat, Zbigniew, Cebula, Maciej, Gruszczyńska, Katarzyna (2023) What is machine learning, artificial neural networks and deep learning?-Examples of practical applications in medicine. *Diagnostics*. 13(15), 2582, doi: https://doi.org/10.3390/diagnostics13152582

Li, Fusheng, Lin, Dan & Yu, Tao (2020). Improved Generative Adversarial Network-Based Super Resolution Reconstruction for Low-Frequency Measurement of Smart Grid. *IEEE*. 8, p. 85257-85270, doi: https://doi.org/10.1109/ACCESS.2020.2992836

Liang, Gaoqi, Liu, Guolong, Zhao, Junhua, Liu, Yanli, Gu, Jinjin, Sun, Guangzhong & Dong, Zhaoyang (2020). Super Resolution Perception for Improving Data Completeness in Smart Grid State Estimation. *Engineering*. 6(7), p. 789-800, doi: https://doi.org/10.1016/j.eng.2020.06.006

Liu, Guolong, Gu, Jinjin, Zhao, Junhua, Wen, Fushuan & Liang, Gaoqi (2020). Super Resolution Perception for Smart Meter Data. *Information Sciences.* 526, p 263-273, doi: https://doi.org/10.1016/j.ins.2020.03.088

Liu, Tianhao, Li, Fangning, Zhang, Dongdong, Shan, Linke, Zhu, Hongyu, Du, Pengcheng, Jiang, Meihui, Goh, Hui Hwang, Kurniawan, Tonni Agustiono, Huang, Chao, Kong, Fannie (2026). Intelligent load forecasting technologies for diverse scenarios in the new power systems. *Renewable and Sustainable Energy Reviews*, 226, doi: https://doi.org/10.1016/j.rser.2025.116366

Miraftabzadeh, Seyed Mahdi, Foiadelli, Federica, Longo, Michela and Pasetti, Marco (2019). A Survey of Machine Learning Applications for Power System Analytics. *IEEE.* p. 1-5, doi: https://doi.org/10.1109/EEEIC.2019.8783340

Mitra, Somalee, Chakraborty, Basab & Mitra, Pabitra (2024). Smart meter data analytics applications for secure, reliable and robust grid system: Survey and future directions. *Energy*. 289, p. 129920, doi:https://doi.org/10.1016/j.energy.2023.129920

Naeem, Samreen, Ali, Aqib, Anam, Sania and Ahmed, Muhammad Munawar (2023) An unsupervised machine learning algorithms: Comprehensive review. *International Journal of Computing and Digital Systems*. 13(1), p. 911-921, doi: https://doi.org/10.12785/ijcds/130172

Noel, Mathew Mithra, Arunkumar, L., Trivedi, Advait and Dutta, Praneet (2025). Oscillating activation functions can improve the performance of convolutional neural networks. *Applied Soft Computing*. 175, doi: https://doi.org/10.1016/j.asoc.2025.113077

Ooi, Yoong Khang and Ibrahim, Haidi (2021) Deep learning algorithms for single image super-resolution: a systematic review. *Electronics*. 10(7), p. 867, doi: https://doi.org/10.3390/electronics10070867

Prechelt, Lutz (1998). *Early Stopping - But When?*. 1.ed. Salem: Springer. https://link-springer-com.ezproxy.its.uu.se/book/10.1007/3-540-49430-8 [Accessed: April 22, 2026]

Python Software Foundation (2023). *Python*, version 3.10.11. Wilmington, Delaware: Python Software Foundation. https://www.python.org/

Raiaan, Mohaimenul Azam Khan, Sakib, Sadman, Fahad, Nur Mohammad, Mamun, Abdullah Al, Rahman, Md. Anisur, Shatabda, Swakkhar and Mukta, Md. Saddam Hossain (2024). A systematic review of hyperparameter optimization techniques in convolutional neural networks. *Decision Analytics Journal*, 11, 100470, doi: https://doi.org/10.1016/j.dajour.2024.100470

Sewak, Mohit, Karim. Md. Rezaul & Pujari, Pradeep (2018). *Practical Convolutional Neural Networks*. Birmingham: Packt Publishing Ltd. https://www.oreilly.com/library/view/practical-convolutional-neural/9781788392303/ [Accessed: April 28, 2026]

Song, Lidong, Li, Yiyan & Lu, Ning (2020). ProfileSR-GAN: A GAN Based Super-Resolution Method for Generating High-Resolution Load Profiles. *IEEE*. 13(4), p. 3278-3289, doi: https://doi.org/10.1109/TSG.2022.3158235

Sree, Laxmi B.R. and Vijaya, M.S. (2019). A Weighted Mean Square Error Technique to Train Deep Belief Networks for Imbalanced Data. *International Journal of Simulation: Systems, Science and Technology*. 19(6), doi: https://doi.org/10.5013/IJSSST.a.19.06.14

TEquipment, (n. d.). *HT Instruments VEGA 78 4 CTs HTFLEX33D for CAT IV Power Quality Analyzer and Energy Logger*. TEquipment. https://www.tequipment.net/HT-Instruments/VEGA-78/Power-Quality-Analyzer/?srsltid=AfmBOopbDIftRHiBe1794FAPSwJ-a8SFX61PKzrZmzZGnx-L9TtwKVLQ#description [Accessed: April 6, 2026]

Tolnai, Balázs András, Ma, Zheng & Børregaard Jørgensen, Bo (2023). A Scoping Review of Energy Load Disaggregation. In: *Moniz, Nuno och Vale, Zita och Cascalho, João och Silva, Catarina and Sebastião, Rui (red.) Progress in Artificial Intelligence - EPIA 2023*, p. 209-221, doi: https://doi.org/10.1007/978-3-031-49011-8_17

Wang, Yi, Chen, Qixin,Hong, Tao & Kang, Chongqing (2019). Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. *IEEE*. 10(3), p. 3125-3148, doi: https://doi.org/10.1109/TSG.2018.2818167

WSP (2025). *E-63-TB-0001 Teknisk beskrivning*. Linköping: WSP Sverige AB.

WSP (n.d.a). *Vi är samlingsplatsen för ambitiösa och passionerade innovatörer.* WSP. https://www.wsp.com/sv-se/vilka-vi-ar [Accessed: April 2, 2026]

WSP (n.d.b). *Elkraft - Transmission och distribution*. WSP https://www.wsp.com/sv-se/tjanster/elkraft [Accessed: April 2, 2026]

Xie, Jian, Alvarez-Fernandez, Inalvis and Sun, Wei (2020). A Review of Machine Learning Applications in Power System Resilience. *IEEE Power & Energy Society General Meeting*. p. 1-5, doi: https://doi.org/10.1109/PESGM41954.2020.9282137

Yu, Jiyeon, de Antonio, Angelica and Villalba-Mora, Elena (2022). Deep Learning (CNN, RNN) Applications for Smart Homes: A Systematic Review. *Computers*, 11(2), p. 26, doi: https://doi.org/10.3390/computers11020026

Zhang, Le, Li, Ao, Hou, Qibin, Zhu, Ce and Eldar, Yonina C. (2025). Deep Learning Empowered Super-Resolution: A Comprehensive Survey and Future Prospects. *IEEE*. 113(6), p. 516-556, doi: https://doi.org/10.1109/JPROC.2025.3613233

Interviews and Spoken Communication:

EO manager, manager of electrical operations at the Facility. 2026. Interviewed on March 23, 2026.

Interviewee 1, Senior Electrical Engineer at WSP Sweden, Group 3 Sthlm El. 2026. Interviewed on the February 6, 2026.

Interviewee 2, Manager Electrical Engineering at WSP Sweden, Industry and Substation 1. 2026. Interviewed on the February 19, 2026.

Interviewee 3, Senior Electrical Engineer at WSP Sweden, Industry and Substation 1. 2026. Interviewed on the February 13, 2026.

Interviewee 4, Experienced Electrical Engineer at WSP Sweden, Power Systems 4. 2026. Interviewed on the February 5, 2026.

Plant Manager, chief of operations at the Facility. 2026. Interviewed on the February 3, 2026.

Plant Manager, chief of operations at the Facility. 2026. Interviewed on March 17, 2026.