Beyond Day-Ahead

Multi-Day Electricity Price Forecasting for Smarter Electric Vehicle Charging

Karl Kjellergren,
Johan Lindqvist Sabel

Master's thesis
2026:E99

Lund Institute of Technology
Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

CENTRUM SCIENTIARUM MATHEMATICARUM

Abstract

Smart charging of electric vehicles can reduce cost by scheduling against time-varying electricity spot prices, but commercial systems generally rely on day-ahead prices and capture only within-day variation. This thesis investigates whether forecasting spot prices beyond the day-ahead horizon adds value for vehicle-integrated smart charging, and how that value varies across charging contexts. A three-stage pipeline is built for the Swedish SE3 bidding zone. Probabilistic hourly forecasts of wind production, solar production and electricity consumption for up to one week ahead are produced from weather forecasts together with standard calendar and historical features. These upstream forecasts, along with day-ahead price history and supply-side market data, then feed an hourly spot price model running over the same horizon. The forecasts are evaluated in an EV charging simulation under two specific frameworks: continuous deferral and a weekly deadline. This evaluation is conducted across two use cases representative of vehicle-manufacturer-controlled charging, specifically home overnight and workplace daytime.

Forecast accuracy is dominated by tree-based model families (Random Forest and XGBoost), which outperform neural sequence families (Long Short-Term Memory and Gated Recurrent Unit) by 15–18 percentage points in normalised mean absolute error at short leads, narrowing to roughly 3 points beyond five days. Accuracy declines markedly across the horizon, driven mostly by the loss of weather forecast skill at multi-day leads. Specifically, the coefficient of determination drops from 0.83 day-ahead to between 0.27 and 0.33 at one week. Despite this decline, the broader price pattern remains clear over the first two to four days, with electricity consumption being the most reliably forecast upstream input. The marginal annual saving from acting on multi-day forecasts, above and beyond within-day optimisation, falls in the range 44–56 EUR per vehicle per year across all four framework $\times$ use-case combinations at a reference 50 km/day driver on a 7.4 kW home wallbox, scaling from roughly 20–28 EUR/year at 25 km/day to 72–78 EUR/year at 100 km/day. The total annual savings against a do-nothing baseline are larger (103–168 EUR/year), but the variation between configurations comes almost entirely from within-day scheduling, which requires no forecast at all. The residual gap to a perfect-information benchmark stays within 3.6 % of the do-nothing cost, indicating that further gains would have to come from a richer decision problem rather than better forecasts. The value is real but dependent on a customer relationship to charging that does not require a fully charged battery every morning.

Acknowledgements

First and foremost, we wish to extend our gratitude to our supervisors at Volvo Cars, Niklas Åkerblom and Viktor Larsson, for being valuable resources throughout this project. Their excellent expertise in electric vehicles has been instrumental in shaping the direction of our work and grounding it in a real-world context. We are grateful for the generous time they have dedicated to guiding us, for the many discussions, and for their continued support and encouragement from start to finish.

Of course, we also extend our thanks to Professor Andreas Jakobsson at Lunds Tekniska Högskola for his excellent supervision. His deep expertise in mathematical statistics has been invaluable in guiding the methodological foundations of this thesis.

Contents

List of Abbreviations

Abbreviation	Meaning
API	Application Programming Interface
BPTT	Backpropagation Through Time
CET	Central European Time
CPO	Charge Point Operator
DA	Day-Ahead (market)
ENTSO-E	European Network of Transmission System Operators for Electricity
GFS	Global Forecast System
GRU	Gated Recurrent Unit
KDE	Kernel Density Estimation
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MET Nordic	Norwegian Meteorological Institute Nordic forecast system
MfRR	Manual Frequency Restoration Reserve
ML	Machine Learning
MW	Megawatt
MWh	Megawatt-hour
nMAE	Normalised Mean Absolute Error
nRMSE	Normalised Root Mean Squared Error
NWP	Numerical Weather Prediction
OEM	Original Equipment Manufacturer
OOS	Out-of-Sample
PV	Photovoltaic
RF	Random Forest
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SDAC	Single Day-Ahead Coupling
SE3	Swedish electricity bidding zone 3
SHAP	Shapley Additive Explanations
SMHI	Swedish Meteorological and Hydrological Institute
SoC	State of Charge
TPE	Tree-structured Parzen Estimator
TSO	Transmission System Operator
UTC	Coordinated Universal Time
V2G	Vehicle-to-Grid

Chapter 1

Introduction

1.1 Problem Statement

As the transition to electric vehicles (EVs) accelerates, charging is emerging as a central element of EV ownership. In its simplest form, a vehicle begins charging as soon as it is connected, drawing electricity at the maximum rate until the battery is full. This uncontrolled approach ignores power market conditions and exposes owners on dynamic contracts to direct price risk, forcing them to pay whatever the electricity costs at the exact moment they plug in.

EV smart charging introduces deliberate scheduling into this process. By distributing charging demand across a window of vehicle inactivity, rather than requiring immediate delivery, smart charging allows the timing and rate of energy delivery to respond to external signals such as electricity prices, grid congestion indicators, or renewable energy availability, while still guaranteeing that the vehicle is charged before the next departure.

The defining constraint of current smart charging implementations is temporal. Many major platforms and vehicle manufacturers base their scheduling decisions on confirmed hourly electricity prices from the day-ahead market, which are published by Nord Pool at 12:45 Central European Time (CET) for the following 24-hour period. The planning horizon is therefore bounded by the availability of confirmed price information, and commercial smart charging systems generally do not optimise charging across multiple future nights based on price expectations beyond that window

This horizon is an information constraint, not a physical one. A vehicle regularly parked over a time-bound session possesses energy flexibility that extends over multiple consecutive sessions. Whether exploiting that multi-night flexibility is economically meaningful depends both on whether electricity prices vary sufficiently across days and whether those variations can be anticipated. Spot prices in the Nordic market are driven by temperature-driven demand and weather-dependent supply factors, such as wind and solar generation and hydropower reservoir levels. All of these variables all evolve on timescales from days to weeks and are themselves the subject of operational weather forecasts.

This thesis addresses the question of whether predictive models can produce electricity price forecasts of sufficient quality to support charging decisions across a multi-day horizon, and quantifies the cost savings such a capability would deliver relative to day-ahead-bounded optimisation.

1.2 Purpose and Project Objectives

The objective of this thesis is to assess how different classes of predictive models can be applied to electricity spot price forecasting beyond the day-ahead horizon. By generating a continuous 168-hour prediction to identify the time horizon at which model efficiency degrades, this study examines how such extended forecasts enable the optimization of EV smart charging strategies over longer temporal horizons.

To this end, the work is structured around a three-step pipeline. In the first step, models for forecasting wind power production, solar power production, and electricity consumption are developed and compared, providing the supply and demand fundamentals that drive spot prices. In the second step, these fundamental forecasts serve as inputs to a range of spot price models spanning tree-based ensemble and deep learning approaches. Additional inputs include historical price signals, nuclear production, hydro reservoir levels, and flows between price areas in Sweden. These models are then evaluated against each other and against price-lag-only baselines. In the third step, the best-performing price forecast is used to implement and evaluate a set of EV smart charging decision strategies across a multi-night horizon. The evaluation covers a range of practically relevant use cases, such as home overnight charging and other scenarios representative of real charging behaviour. This step assesses each strategy's ability to identify low-cost charging windows and the cost savings it achieves relative to a naïve benchmark of always charging immediately.

The three steps are deliberately modular: the input signal models in step one can be improved independently of the price model, and the charging evaluation framework in step three can accommodate any price forecast as its upstream input. This structure also allows the contribution of each stage to final charging performance to be examined in isolation.

1.3 Research Questions

Through discussions with EV smart charging experts, the field of smart charging coupled with electricity spot price forecasts was translated into the following research questions:

1. How accurately can different classes of predictive models forecast electricity spot prices beyond the day-ahead horizon, and how does forecast performance vary across model type and temporal horizon?

2. To what extent can extended-horizon electricity price forecasts improve the cost efficiency of EV smart charging, and how does it vary across charging contexts that differ in connection window length and scheduling flexibility?

1.4 Scope and Limitations

The empirical analysis is limited to the SE3 bidding zone. The forecasting pipeline transfers naturally to the other Swedish zones (SE1, SE2, SE4) by re-training the same model structure on zone-specific data; application to non-Nordic markets would require more substantial rework to accommodate differences in market structure, supply mix, and cross-border flow dynamics.

The study takes the perspective of a vehicle manufacturer (OEM) delivering extended-horizon smart charging as a vehicle-integrated feature. Home and workplace charging are therefore the relevant settings, while public and fast-charging contexts (controlled by the charge point operator) and platform-level fleet coordination fall outside the scope. The evaluation is correspondingly restricted to two use cases: home overnight charging (UC1) and workplace daytime charging (UC2).

The charging direction is charge-only. The price forecasting models apply equally to discharge timing. Vehicle owner could schedule discharging to the grid against price peaks rather than imports against troughs, but evaluating a customer who both charges and discharges falls outside the scope. Such vehicle-to-grid (V2G) operation introduces a substantially more dynamic battery-usage pattern that makes the cost evaluation materially more complex.

On 1 October 2025 the European market transitioned to 15-minute pricing intervals. This thesis nonetheless models hourly prices, aggregating quarter-hourly settlement values to hourly averages where applicable. This keeps training-data requirements tractable while providing sufficient resolution for the research questions, which concern multi-day rather than intra-hour decision-making.

The cost component is the SE3 day-ahead spot price only. Grid fees, energy taxes, network capacity charges, retail markups, time-of-use tariffs, congestion pricing, carbon-intensity pricing, and demand-response incentives are not modelled; these omissions scale absolute cost levels but do not affect the relative ordering of charging strategies, which is the headline output.

The forecast modelling considers four Machine learning (ML) models: Random Forest, XGBoost, Gated Recurrent Unit (GRU), and Long Short-Term Memory (LSTM). These were selected to span tree-based ensembles and recurrent sequence architectures. Alternatives such as transformer-based models and classical econometric specifications are referenced in the related work but not implemented. The maximum forecast horizon is 168 hours (seven days), beyond which model efficiency is not investigated.

The evaluation employs a reference vehicle with stylized assumptions, such as constant consumption and zero battery degradation among others. Detailed in Section 4.4.2, these simplifications isolate the algorithm's economic impact from vehicle-specific variations, though actual driver savings will naturally vary.

Chapter 2

Background

This chapter provides the domain foundation necessary to understand the problem addressed in this thesis. It begins by introducing EV smart charging: the relevant charging contexts, the actors who provide smart charging services today, and the temporal boundary at which current implementations operate. The chapter then describes the Nordic electricity spot market, covering its structure and price formation mechanism, the drivers of spot price variation in the SE3 bidding zone, and the landscape of price information available to an operator at decision time. The chapter closes with a review of related work across the five research streams that together constitute the intellectual background of the thesis, and positions the present study within that literature.

2.1 Electric Vehicle Smart Charging

2.1.1 Electric Vehicle Charging Alternatives

EV owners charge their vehicles across a range of location types that differ in session duration, pricing mechanism, and the degree of scheduling flexibility available. Table 2.1 provides an overview. In all cases, smart charging, the deliberate scheduling of energy delivery over the available connection window, is technically feasible. What differs across contexts is which actor controls the charging process and on what informational basis that control is exercised.

At home and at many workplace installations, the vehicle owner or their chosen platform controls the charge, typically through a wallbox or the vehicle's own software. These settings offer the longest and most predictable connection windows, and the owner is directly exposed to retail spot prices [1], making them the most natural context for price-responsive scheduling. At public and fast charging locations, control over the charging process rests with the Charge Point Operator (CPO), who typically offers customers a fixed per-kilowatt-hour tariff while retaining the ability to manage load across their station portfolio.

This thesis examines the problem from the perspective of an OEM seeking to provide extended-horizon smart charging as a vehicle-integrated feature. In this context, home and workplace charging are the primary relevant settings, as these are the environments where the OEM can directly control charging behaviour through the vehicle's onboard systems. Analogous problem statements can be formulated for other actors, such as CPOs managing portfolio-level load or energy platforms offering multi-vehicle coordination. However, these fall outside the scope of the present study.

Table 2.1: Overview of EV charging alternatives and smart charging control characteristics. Home and workplace dwell ranges are indicative of real-world Swedish charging-session data [2]; public and fast-charging ranges are indicative.

Type	Typical dwell	Price mechanism	Smart charging control
Home (wallbox)	6–10 h (overnight)	Retail spot price	Owner / OEM / platform
Workplace	4–8 h (daytime)	Employer tariff or spot	Owner / OEM / employer
Public (urban)	0.5–3 h	Fixed CPO tariff	CPO
Fast / ultra-fast	0.2–0.5 h	Fixed CPO tariff	CPO

2.1.2 Smart Charging Services and Market Actors

The charging contexts described in Section 2.1.1 are served by three distinct categories of actors, each with a different technical integration point and optimisation basis. Table 2.2 summarises the three categories, their control points, and representative providers.

Smart energy platforms connect to compatible vehicles or home chargers via APIs, retrieving state of charge, connection status, and departure constraints to schedule charging at lowest cost within the available window. Platforms such as Tibber and Gridio operate in the residential and workplace settings where OEM access is highest, but their scheduling is built on confirmed day-ahead prices and does not incorporate information beyond the following 24 hours [3], [4].

Vehicle manufacturers implement smart charging through vehicle-integrated software and proprietary control of the onboard charger, independent of external infrastructure: the OEM controls when and at what rate energy is drawn, based on user schedules, local energy signals, and spot price periods [5]. As with platforms, current OEM implementations operate within the temporal scope of known or user-defined price periods.

Charge point operators manage public and fast-charging infrastructure; their interventions target station-level load management rather than vehicle-level cost minimisation [6], and their retail pricing is commonly fixed or weakly time-differentiated, limiting price-responsive behaviour at the vehicle level.

Table 2.2: Overview of smart charging offerings by actor category.

Category	Control point	Optimisation basis	Examples
Smart energy platforms	Charger-level	Day-ahead spot prices, readiness constraints	Tibber [3], Gridio [4]
Vehicle manufacturers	Vehicle-level	User schedules, local energy	Tesla [5], Volvo Cars [7]
Charge point operators	Infrastructure-level	Fixed or weakly time-differentiated	Vattenfall [8], Mer [9]

Across the three actor categories, smart charging implementations typically share a common temporal boundary: scheduling decisions are generally made within the 24-hour window for which day-ahead prices are confirmed. This reflects the structure of electricity market price formation, in which hourly spot prices beyond the following day are not published and must be estimated [10].

2.1.3 The Case for Extended-Horizon Smart Charging

Current smart charging implementations typically schedule energy delivery within the 24-hour window for which day-ahead prices are confirmed. This horizon reflects the structure of electricity market price formation rather than any physical constraint on vehicle flexibility. A vehicle parked at home overnight on consecutive nights remains connected across multiple sessions, and the energy requirement for each day's driving is typically a fraction of total battery capacity [1]. In principle, charging can therefore be deferred across nights as long as sufficient charge is available for each departure.

Extending the planning horizon to cover multiple nights introduces the possibility of inter-session scheduling: rather than optimising within a single overnight window, a charging strategy can select which night within a multi-day window to charge, or distribute energy across several nights. The value of such flexibility depends on the degree to which electricity spot prices differ across consecutive days. In the Nordic market, spot prices are driven by weather-dependent supply and demand factors (wind generation, hydropower reservoir levels, and temperature) which evolve on timescales of days to weeks, producing day-to-day variation in the price level [11], [12].

Wieberneit et al. evaluate multi-session EV charging control using planning horizons of up to seven days and report improvements of 19–26 % over single-session optimisation and 40–46 % over uncontrolled charging, measured against carbon intensity [13]. Whether comparable savings are available when the optimisation criterion is electricity price, which differs from carbon intensity in that it also reflects demand-side and cross-border drivers, is one of the questions this thesis quantifies. Multi-day price forecasts exist commercially for portfolio management (Section 2.2.3) but are not currently incorporated into consumer-facing smart charging products.

2.2 The Nordic Electricity Spot Market

2.2.1 Market Structure

Sweden participates in the liberalised Nordic electricity market, where electricity has been traded competitively since the market opened in 1996 [14]. Trading takes place on the Nord Pool exchange, where producers, suppliers, industrial consumers, traders, and transmission system operators (TSOs) submit buy and sell orders, with prices determined by the intersection of aggregated supply and demand curves, as illustrated in Figure 2.1.

The primary trading venue is the day-ahead (spot) market, where all electricity for physical delivery on the following day is auctioned in a single daily session [10]. Gate closure is 12:00 CET; results, including a price for each delivery period of the next day,

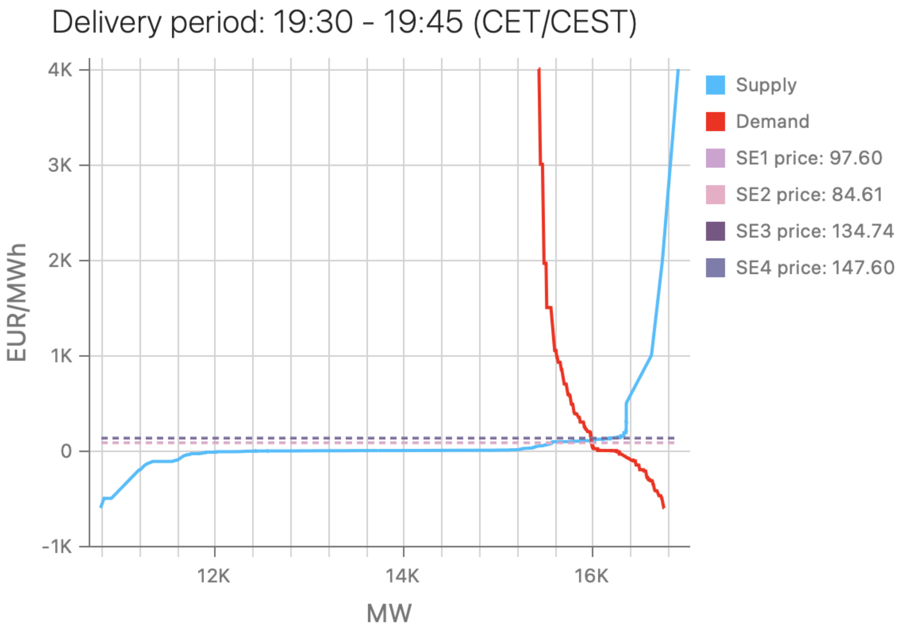

Figure 2.1: Example aggregated bid curves for a delivery period in June 2026. Supply and demand intersect to determine the hourly clearing price per SE bidding zone [15].

are published at 12:45 CET. All accepted orders clear at the same market price within each bidding zone, making the day-ahead auction the reference price for retail contracts, smart charging platforms, and operational dispatch decisions.

Because the physical grid imposes transmission constraints between regions, Nord Pool calculates separate area prices for each bidding zone when cross-border capacity is congested [16]. Sweden is divided into four bidding zones (SE1–SE4); this thesis focuses on SE3, which covers central Sweden including Stockholm and Gothenburg, and is the most populous zone. Prices across zones are equal when the grid is uncongested and diverge when transmission limits bind.

Beyond the day-ahead market, trading also occurs in forward, intraday, and balancing markets [17], but the day-ahead auction is the dominant price-setting venue and the reference for all downstream applications considered in this thesis. From 1 October 2025 its delivery resolution switched from hourly to quarter-hourly under the Single Day-Ahead Coupling (SDAC) [18]; the models and data here use the hourly resolution that prevailed before this change.

2.2.2 Drivers of Spot Price Variation

Because electricity cannot be stored at scale, production and consumption must be balanced in real time, and the day-ahead spot price is highly sensitive to short-run supply and demand conditions in each bidding zone [12]. The principal drivers of price variation in SE3 are described below.

Supply-side drivers.

- **Hydrological balance.** Hydropower dominates Nordic generation [11]. When reservoir levels are low, available hydro output is reduced and higher-cost thermal generation sets the marginal price more frequently, raising spot prices. When reservoirs are full, hydro can displace thermal generation, depressing prices.
- **Wind.** Wind power output is a nonlinear function of wind speed; at high wind speeds, large volumes of near-zero-marginal-cost generation enter the market, pushing higher-cost units out of the merit order and depressing the spot price [11].
- **Nuclear availability.** Sweden's nuclear fleet provides substantial baseload capacity at low marginal cost, primarily in SE3 and SE4 [11]. Planned and unplanned outages reduce this supply, increasing reliance on more expensive alternatives.
- **Thermal and CHP fuel costs.** When combined heat and power or other thermal units set the marginal price, spot prices co-move with fuel and CO_2 prices [14].
- **Solar.** Solar power provides near-zero-marginal-cost generation that is strictly bound to diurnal and seasonal cycles. During sunny periods, particularly midday hours in the summer, this influx of solar generation pushes more expensive units out of the merit order and exerts strong downward pressure on spot prices [19].

Demand-side drivers.

- **Temperature.** Space-heating and direct electric heating make Swedish electricity demand strongly temperature-sensitive [11]. Demand in winter can be approximately twice that in summer [14], with stronger price effects in the southern zones [11].
- **Calendar effects.** Demand follows systematic intraday and intraweek patterns, with weekday working hours producing higher load than nights and weekends, independent of temperature [20].

Cross-border flows. Sweden's bidding zones are connected to Norway, Denmark, Finland, Germany, Poland, and the Baltic states via high-voltage interconnectors [17]. When neighbouring areas experience high prices, electricity flows out of Sweden, tightening domestic supply and raising SE3 prices; the reverse effect occurs when neighbouring areas have surplus generation. Day-ahead prices are quoted in EUR/MWh, and EUR/SEK exchange rate movements introduce an additional source of variation for stakeholders operating in SEK [10].

2.2.3 Price Information Availability at Decision Time

An EV charging decision made at the start of a parking session, such as at 22:00 on a weekday evening, draws on several categories of information, each with a distinct availability horizon.

Confirmed day-ahead spot prices are available for all delivery periods of the following calendar day from 12:45 CET onward [10]. A charging decision made in the evening therefore has access to confirmed prices covering the overnight period into the following afternoon. Prices for the day after tomorrow and beyond are not published by Nord Pool.

Numerical weather prediction (NWP) forecasts are issued operationally at sub-daily intervals and provide hourly projections of wind speed, solar irradiance, temperature, and related variables at lead times of up to approximately 10 days [21]. These forecasts are publicly available and are the primary input to operational wind and solar power production models. Forecast skill degrades with lead time; for synoptic-scale wind patterns, the most rapid degradation typically occurs beyond approximately 3–5 days [21].

Historical price and generation data are available with short delay through market operators (Nord Pool), TSOs (Svenska kraftnät), and the Nordic settlement system eSett, which publishes metered hourly production volumes per balance area [22].

Commercial spot price forecasts covering multi-day horizons are available for the Nordic market from specialist providers. AleaSoft offers price forecasts for all Nordic bidding zones from D+1 to several years using a combination of statistical and machine-learning methods [23]. DNV provides a similar service oriented towards power market modelling and investment decisions [24]. Both services are designed primarily for portfolio management and long-term planning rather than real-time EV charging optimisation.

The result is a two-tier information structure: confirmed prices are available for the next 24 hours, while prices at longer horizons must be estimated from weather forecasts and historical patterns. This structure defines the boundary at which smart charging typically operates.

2.3 Related Work

Research related to this thesis spans five areas: wind power forecasting, solar power forecasting, electricity demand forecasting, electricity spot price forecasting, and EV smart charging under price-based decision rules. The first four areas concern the estimation of the supply and demand fundamentals described in Section 2.2.2, while the fifth concerns how such signals can be translated into charging decisions. Existing studies frequently address one stage at a time; the integration of all stages within a single pipeline targeting a specific bidding zone remains limited.

2.3.1 Wind Power Production Forecasting

Wind power forecasting is a well-established field given its importance for power system balancing and market operation. A recurring finding is that numerical weather prediction data becomes increasingly important as the horizon extends, while hybrid approaches combining physical conversion with data-driven correction often perform well at operational horizons [25].

Aggregated forecasting at system or regional level is studied less than turbine- or farm-level forecasting but is especially relevant for market applications. Belenguer et al. show aggregated renewable generation can be forecast directly from weather-driven inputs without modelling each installation separately [26], placing bidding-zone-level forecasting within an established line of work despite differences in market setting, weather regime, and modelling choices across studies.

2.3.2 Solar Power Production Forecasting

Solar power forecasting differs from wind in that the weather-to-power relationship depends strongly on irradiance, cloud cover, panel orientation, and temperature. The literature typically treats photovoltaic forecasting as a model-chain problem in which irradiance-related inputs are transformed into power output, often aided by clear-sky modelling or other physics-informed steps [27].

Reviews commonly distinguish physical, statistical, and hybrid approaches, often favouring hybrid methods at operational horizons because physical models are sensitive to local cloud uncertainty while statistical models are less robust when conditions change [28]. In market applications, solar's relevance lies not only in energy volume but in its concentration in daytime hours, when it can influence price formation.

2.3.3 Electricity Consumption Forecasting

Short-term electricity demand forecasting is one of the most mature forecasting problems in power systems. The literature consistently identifies weather variables, especially temperature, together with calendar effects, as the main drivers of demand at short and medium horizons. Reviews of load forecasting document the importance of combining weather-driven and calendar-based information when demand is sensitive to heating needs and daily routines [20].

A second common finding is that electricity demand contains multiple seasonal structures, especially intraday and intraweek variation, and that forecast accuracy at longer horizons depends strongly on the quality of upstream weather forecasts [29], [30]. This is particularly relevant in Nordic settings, where temperature sensitivity is pronounced and weather-related uncertainty propagates directly into load forecasts.

2.3.4 Electricity Spot Price Forecasting

Electricity price forecasting is commonly divided into structural, statistical, multi-agent, and computational intelligence approaches [12], where structural models link prices to physical fundamentals and statistical and machine-learning models estimate price dynamics directly from historical prices and exogenous variables.

Within the statistical literature, autoregressive and volatility-based models remain standard benchmarks, and recent reviews identify regularised linear approaches such as LEAR-type models as strong interpretable baselines in European electricity price forecasting [31], [32]. Machine-learning methods are increasingly prominent: Lago et al. show deep learning can perform competitively in day-ahead forecasting when market fundamentals are included [33], while Biro et al. find tree-based methods highly competitive across European day-ahead markets [34]. Kılıç et al. report the opposite ranking for a DK1 intraday task where LSTM leads, showing that performance is sensitive to both market setting and forecast horizon [35].

Renewable generation and demand variables improve price forecasting when the horizon extends beyond the very short run. Belenguer et al. demonstrate this in a framework combining renewable and demand forecasts [26], and for Sweden, Renefalk finds model-combination approaches particularly useful in the more volatile southern zones [36].

2.3.5 Electric Vehicle Smart Charging with Price Signals

The EV smart charging literature is often distinguished by the planning horizon of the charging decision and by whether prices over that horizon are assumed known or forecast. A large share of existing work assumes confirmed day-ahead prices and frames charging as a deterministic optimisation problem over a single 24-hour window, using linear programming, mixed-integer optimisation, or related scheduling methods [37].

More recent work incorporates predictive models into the charging decision. Zhao et al. show forecast information can improve EV scheduling within a reinforcement-learning framework [38]. Wieberneit et al. extend the horizon beyond a single overnight session and show inter-session flexibility improves outcomes when decisions respond to forecasted conditions, using carbon intensity rather than price as signal [13]. In the Swedish context, Fakhry et al. combine machine-learning availability forecasts with optimisation of residential EV flexibility for the manual Frequency Restoration Reserve (mFRR) market, using Swedish residential charging data [39].

Cross-context comparisons across charging settings remain limited in the literature, as do studies that extend scheduling decisions beyond the information available from the day-ahead market.

2.3.6 Positioning of This Thesis

The studies reviewed above treat forecasting of renewable generation, electricity demand, electricity prices, and EV charging decisions as largely separate problems, or connect only two stages at a time. This thesis addresses all three stages in a single SE3-oriented pipeline, spanning from weather-driven input forecasts through spot price prediction to charging evaluation across a multi-night horizon. This approach is motivated by the information structure described in Section 2.2.3 and the research questions in Section 1.3.

Chapter 3

Theory

The primary objective of this thesis is to apply established machine learning methodologies to the practical challenge of electricity price forecasting, rather than to develop novel theoretical frameworks. Consequently, this chapter provides a deliberately scoped, high-level overview of the model families evaluated in the experiments. An extensive exposition of the underlying supervised learning mathematics, training algorithms, and neural architectures is provided in Appendix A for the interested reader.

3.1 Supervised Learning for Price Forecasting

At its core, machine learning replaces hand-coded decision logic with a data-driven estimation procedure. The task of predicting an electricity spot price (or wind/solar generation) from historical and meteorological inputs is framed as a supervised regression problem.

The joint effect of wind availability, temperature-driven load, hydro storage levels, and cross-border flows on hourly spot prices is simultaneously nonlinear, non-stationary, and too difficult to derive analytically. By providing a model with a dataset of historical input-output pairs, the algorithm infers the mapping automatically. The central challenge in this process is navigating the bias-variance trade-off: the model must be flexible enough to capture genuine price-driving patterns (low bias) without memorising the inevitable noise intrinsic to electricity markets, such as unannounced outages or strategic bidding behaviour (low variance).

To ensure models can be trusted operationally, predictive accuracy is augmented with SHapley Additive exPlanations (SHAP). SHAP assigns each feature a fair share of the credit for any given prediction, allowing us to interpret exactly which weather or historical variables are driving the model's outputs.

3.2 Model Families

Two distinct model families are evaluated in this thesis, each chosen to handle specific challenges inherent to weather-driven power markets.

3.2.1 Tree-Based Ensembles (Random Forest and XGBoost)

Linear models impose a globally additive structure that cannot capture the regime-dependent responses characteristic of electricity markets. For instance, the price sensitivity to an additional megawatt of wind generation differs drastically between a system operating near capacity and one with ample reserve.

Tree-based methods naturally learn these nonlinear interactions by partitioning the feature space into distinct regions. Because a single decision tree is prone to overfitting, we evaluate two ensemble techniques:

- **Random Forest:** Builds hundreds of independent trees in parallel on randomised subsets of the data, averaging their predictions to drastically reduce variance.
- **XGBoost (Gradient Boosting):** Builds trees sequentially, where each new tree is specifically trained to correct the residual errors of the current ensemble.

Both tabular models excel at finding complex, non-linear interactions across large, cross-sectional feature matrices.

3.2.2 Sequence Models (GRU and LSTM)

While tree-based models treat each hourly input as a self-contained snapshot, electricity prices are fundamentally sequential. The wind ramp earlier in the night, the temperature trajectory through the evening, and the trend in cross-border flows all carry momentum that a static feature vector cannot fully capture.

To explicitly model this temporal context, we evaluate two Recurrent Neural Network (RNN) architectures: the Gated Recurrent Unit (GRU) and the Long Short-Term Memory (LSTM) network. Standard RNNs struggle to remember information over long sequences due to vanishing gradients during training. GRUs and LSTMs solve this by introducing internal "gating" mechanisms-learned filters that decide what past information to remember, what to forget, and what new information to write to memory. This allows them to maintain a running summary of the system's state, making them theoretically well-suited for capturing multi-day weather patterns and weekly demand cycles.

Chapter 4

Methodology

4.1 Model Development

This chapter describes the full implementation of the three-step forecasting and evaluation pipeline introduced in Section 1.2. The pipeline is modular: each stage produces an output that feeds the next, and the models at each stage can be developed, tuned, and evaluated independently (Figure 4.1).

In the first stage, historical weather forecasts (wind speed and direction, solar irradiance, air temperature across locations in and around SE3) train models predicting hourly wind power, solar power, and electricity consumption, supply and demand fundamentals underpinning spot price formation. In the second stage, these forecasts combine with historical price signals, nuclear and hydro production, cross-border flows, reservoir levels, and calendar features to drive a range of spot price models. In the third stage, the best-performing price forecast is used to implement and evaluate EV smart charging strategies over a multi-night horizon, across practically relevant use cases (home overnight and workplace daytime charging), with performance measured as cost savings relative to a naïve always-charge-tonight benchmark.

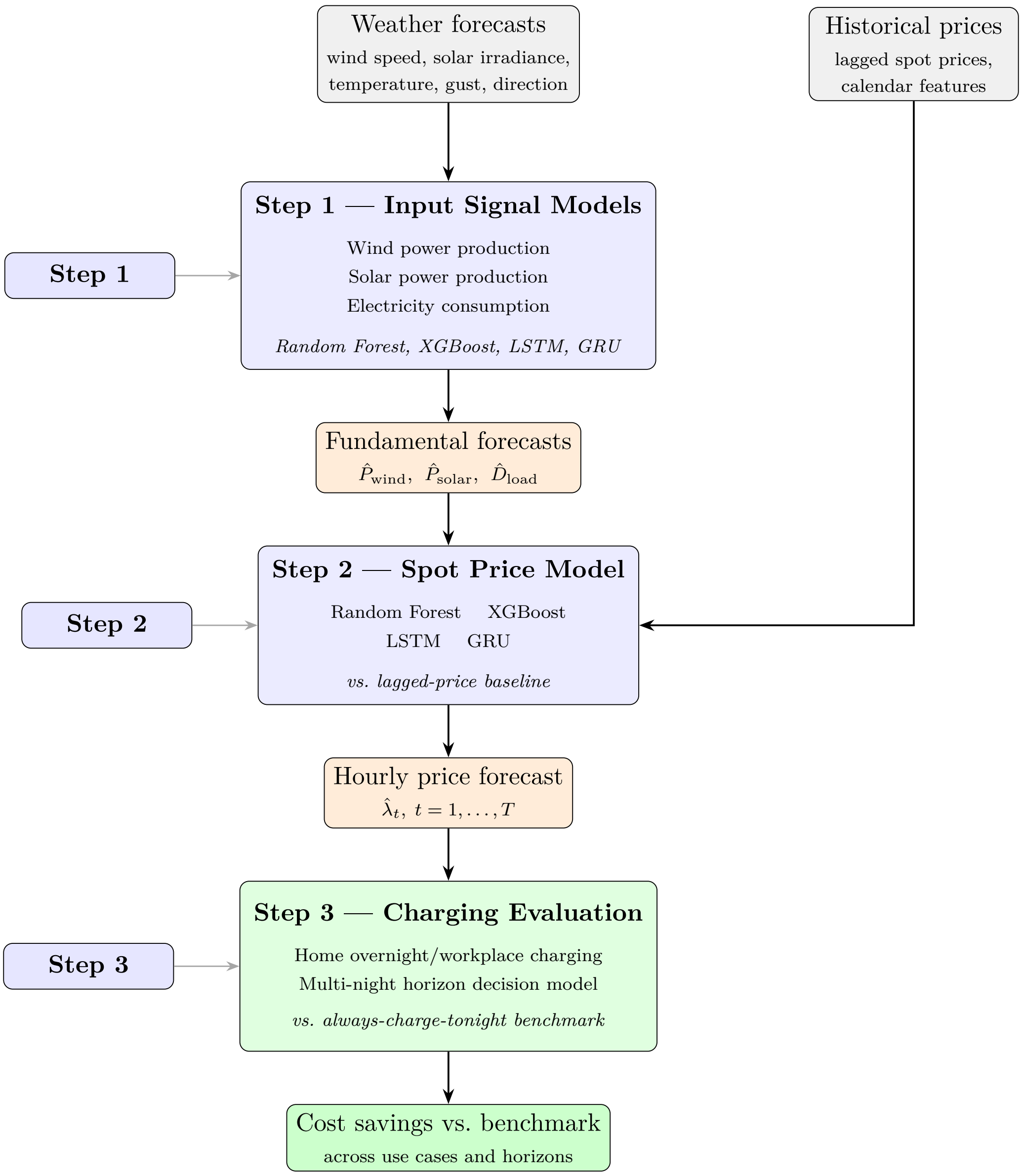

Figure 4.1: Three-step modelling pipeline. Weather forecasts and historical electricity production and consumption feed into stage one input signal models, whose outputs combine with lagged price features in stage two to produce hourly spot price forecasts. Stage three uses these forecasts to evaluate EV charging decision strategies across a multi-night horizon.

The remainder of this chapter is organised as follows. Section 4.2 covers the first pipeline stage: data, preprocessing, and model development for the three input signal forecasts. Section 4.3 covers the second stage, developing and evaluating the spot price models. Section 4.4.1 covers the third stage, implementing and assessing the EV charging decision strategies.

4.2 Input Signals

The first pipeline stage produces three hourly forecasts: wind power production, solar power production, and electricity consumption. Together, these characterise supply and demand fundamentals driving spot prices in SE3. Each forecast is treated as a supervised regression task: a set of meteorological features are mapped to an observed hourly quantity, and models from each of the four models introduced in Chapter 3 are trained and evaluated.

4.2.1 Shared Data and Preprocessing

The three input-signal sub-models share a common data infrastructure, panel layout, training protocol, and evaluation framework. This subsection describes those shared components once; signal-specific modelling choices are deferred to Sections 4.2.3-4.2.5.

Data sources

The primary inputs for all three sub-models are historical weather forecasts rather than observed measurements, ensuring the training feature distribution matches what the model receives at inference. Wind features are sourced from the 72 largest SE3 turbine clusters ($\approx 85\,\%$ of installed SE3 wind capacity), with per-cluster hourly forecasts of wind speed, gust, and direction; the cluster list is in Table B.4 (Appendix B) [40]. Solar and consumption features are sourced from the 24 SE3 municipalities of Table B.3 (Appendix B), selected as the largest population clusters dominating regional solar generation and demand [41], [42], [43], [44]: solar uses hourly short-wave irradiance and cloud area fraction, consumption uses hourly 2 m air temperature. Observed temperatures at the same municipalities are also retrieved from SMHI [45] to fit a temperature–consumption response of Section 4.2.5.

Metered SE3 wind production, solar production, and total consumption are sourced from the eSett open data API [46], which publishes Nordic settlement data on a 24-hour cycle. To guarantee data availability during live forecasting, no actual values more recent than 48 hours are used as lagged features. Day-ahead spot prices for SE3 are retrieved from the European Network of Transmission System Operators for Electricity (ENTSO-E) Transparency Platform [47] and feed the price model in Section 4.3.

Forecast data: MET Nordic and GFS

Historical weather forecasts are gathered from two NWP datasets: Norwegian Meteorological Institute Nordic forecast system (MET Nordic) [48], [49], and the Global Forecast System (GFS), produced by NCEP [50]. MET Nordic provides a 1 km grid over the Nordic region with hourly leads from 1 to 58 h; Global Forecast System (GFS) provides a global 2.5 km grid with 3-hourly leads out to 240 h [51]. Both systems issue four runs per day at 00, 06, 12, and 18 UTC. The specific variables retrieved from each source are listed in Tables B.1 and B.2 (Appendix B).

Spatial aggregation

Each input signal is a weighted aggregate of high-resolution local forecasts (municipality for solar and consumption, cluster for wind) that are sequentially scaled to the regional level, and finally to a single SE3 figure. The weighting scheme is signal-specific, relying on installed capacity for wind and solar and historical consumption share for load. However, the structure remains identical: a weighted sum dynamically renormalised over the subset of locations with valid (non-missing) data at each timestamp, ensuring robustness against missing data. The signal-specific weights and aggregation formulas are in the corresponding subsection.

Data partitioning and fold structure

The three input signals share an identical *fold structure*, quarterly expanding-window retraining covering the out-of-sample (OOS) period from 2023-01-01 through 2026-05-01, but differ in how far back their training data extends. Wind-related data are available from 2022-01-01, while solar and consumption data extend further back to 2021-02-24, the earliest date at which MET Nordic and GFS historical forecasts data are simultaneously available.

Where the MET dataset contains intermittent gaps, the corresponding GFS forecast is substituted at that timestamp. This cross-filling preserves continuous time series without sacrificing data volume. Because GFS provides forecasts at 3-hourly steps while MET Nordic is hourly, linear interpolation between adjacent GFS steps is used to align the two streams on a common hourly cadence, applied at the feature level for solar and consumption, and at the model-output level for the wind GFS extension (Section 4.2.3).

The fold schedule is given in Table 4.1. Fold k trains on all data from the signal-specific anchor date through the end of the quarter preceding the prediction window, and produces forecasts for the next quarter; fold $k+1$ then re-trains on a window extended by exactly one quarter. The expanding design replicates live operational conditions and supplies the price model in Section 4.3 with a continuous OOS upstream feature history without temporal leakage. Concatenating the prediction blocks across all 14 folds yields a single continuous OOS panel against which all model families are compared.

Feature construction

Raw hourly weather forecast values are supplemented with engineered features capturing temporal patterns and interaction effects. Calendar variables such as hour of day, day of week, month (cyclic sine/cosine encoding where appropriate), and Swedish public holiday indicators, see Appendix C, are included across all three sub-models. Lagged target observations are included where they are available at inference time without introducing future information; metered eSett actuals are subject to the 48-hour cutoff noted above. Signal-specific features (turbine physics outputs, solar seasonality proxies, temperature-elasticity baselines, etc.) are described in the corresponding subsection.

Table 4.1: Quarterly expanding-window fold schedule shared across all three input-signal models. Training begins at the signal-specific anchor date, 2022-01-01 for wind, 2021-02-24 for solar and consumption, and the training window is extended by one quarter at each refit. Fold 14 covers April 2026 only.

Fold	Training window end	Prediction window
1	2022-12-31	2023-01-01 to 2023-03-31
2	2023-03-31	2023-04-01 to 2023-06-30
3	2023-06-30	2023-07-01 to 2023-09-30
4	2023-09-30	2023-10-01 to 2023-12-31
5	2023-12-31	2024-01-01 to 2024-03-31
6	2024-03-31	2024-04-01 to 2024-06-30
7	2024-06-30	2024-07-01 to 2024-09-30
8	2024-09-30	2024-10-01 to 2024-12-31
9	2024-12-31	2025-01-01 to 2025-03-31
10	2025-03-31	2025-04-01 to 2025-06-30
11	2025-06-30	2025-07-01 to 2025-09-30
12	2025-09-30	2025-10-01 to 2025-12-31
13	2025-12-31	2026-01-01 to 2026-03-31
14	2026-03-31	2026-04-01 to 2026-05-01

Per-lead modelling and panel structure

To capture the unique bias structure of each forecast distance, an independent model is trained for every distinct lead hour: 1–58 h for the MET branch and 60–168 h for the GFS branch. All three input signals are modelled using the same four families introduced in Chapter 3: Random Forest, XGBoost, GRU, and LSTM. Tabular models (Random Forest, XGBoost) consume the panel rows directly; sequence models (GRU, LSTM) reshape the rolling window of lagged feature columns into a time-ordered tensor, with non-lagged features supplied as static context.

Probabilistic outputs

Alongside point forecasts, prediction intervals are produced at the 10th, 25th, 50th, 75th, and 90th percentile levels. For tree-based models this is achieved through a native multi-quantile objective that fits all five levels simultaneously in a single pass. For sequence models the loss is a sum of pinball losses across the five levels. Monotonicity of the predicted quantiles is enforced post-prediction by taking the cumulative maximum across quantile outputs, ensuring no lower quantile exceeds a higher one. These intervals propagate uncertainty into the downstream price model in Section 4.3.

Hyperparameter tuning.

Hyperparameter tuning is performed using Bayesian optimization via the Optuna framework [52]. Specifically, we employ the Tree-structured Parzen Estimator (TPE) algorithm [53]. Because TPE is a non-parametric method that relies on Kernel Density

Estimation (KDE) to model the objective function, it does not utilise classical conjugate priors. Instead, independent prior distributions are defined to establish the search space for each hyperparameter.

For the XGBoost models, the search space consists of a log-uniform distribution for the learning rate (η), discrete uniform distributions for the maximum tree depth and minimum child weight, and uniform distributions for the subsample and column-sample ratios. For the Random Forest models, discrete uniform distributions are used for the number of estimators and maximum tree depth, while the maximum features fraction is drawn from a continuous uniform distribution.

Evaluation framework

Point forecast performance on the fully out-of-sample data is assessed using MAE, RMSE, nMAE, nRMSE, and R^2, as defined in Appendix A.1 Section A.1.2. Errors are reported across the canonical lead bands in 5.1, with each signal evaluated on the bands applicable to its forecast horizon. The best-performing models are further analyzed using SHAP to identify key prediction drivers and improve interpretability for the selected model.

4.2.2 Shared Modeling Methodology

While the physical drivers for wind, solar, and consumption differ, all three input signals are modeled using a two-step architectural framework, as summarized in Figure 4.2.

Step 1: Physical Domain Modeling. The first stage involves transforming raw meteorological variables into a theoretical energy baseline. This "Physics Pipeline" uses deterministic equations, such as transforming wind speeds to estimate turbine-level production, to create an initial estimate of the target variable. This baseline is physically grounded but carries systematic biases due to idealized assumptions (e.g., neutral atmospheric stability or generic efficiency curves).

Step 2: Machine Learning Framework. The second stage trains ML models to map the engineered features, including the physical baseline, to metered actuals. This serves two purposes: it applies statistical bias correction to the physics output and learns the error characteristics of the NWP sources (MET Nordic and GFS) to yield a seamless 1-168 h probabilistic forecast.

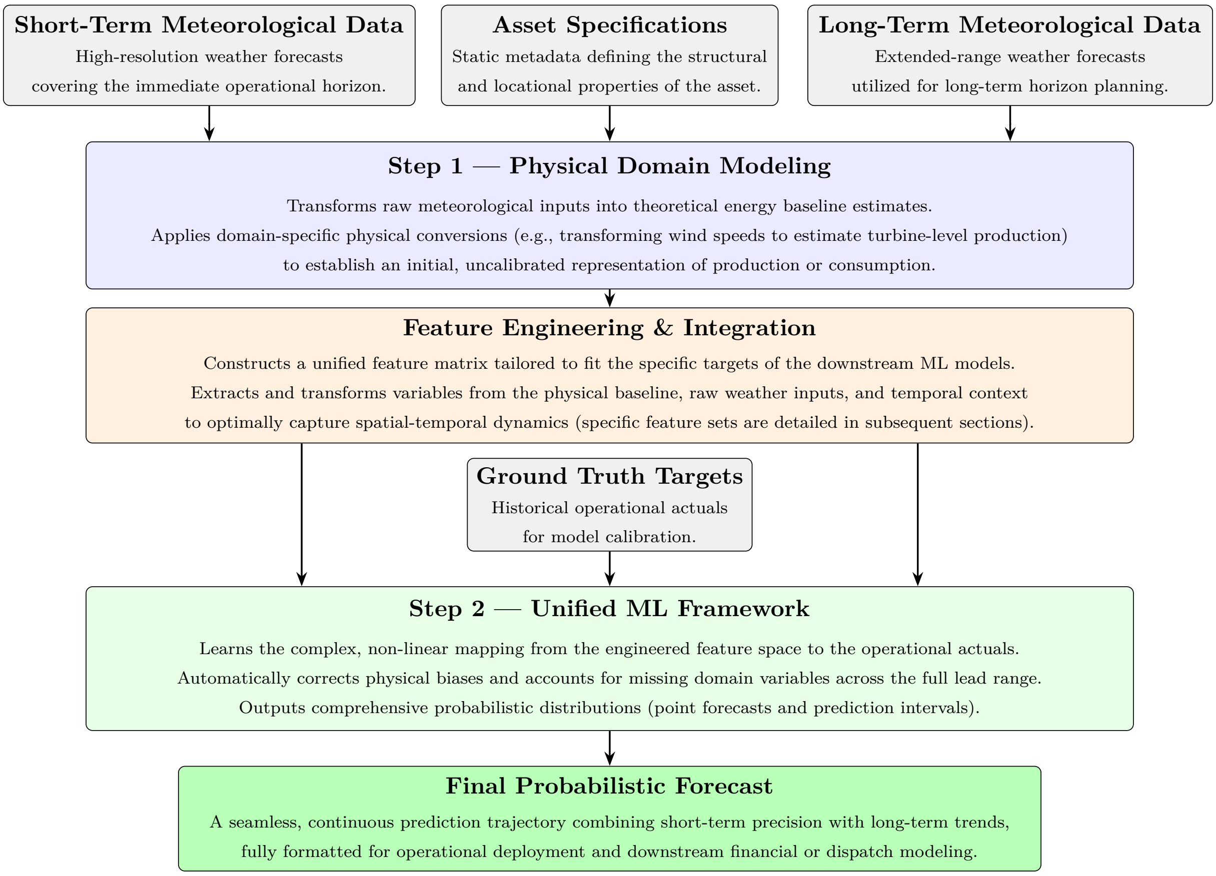

Figure 4.2: Generalized conceptual architecture of the energy asset prediction pipeline. The framework abstracts the process into a domain-specific physical modeling stage (Step 1) followed by a unified ML calibration framework (Step 2), allowing for flexible feature engineering and robust probabilistic forecasting across varying horizons.

4.2.3 Wind Power Production

Hourly metered SE3 wind production serves as the regression target. Wind power output is highly non-linear and sensitive to local turbine characteristics and atmospheric conditions.

Step 1 — Physical Domain Modeling

The physical pipeline converts raw NWP wind variables into megawatt estimates at the cluster level before aggregating to a regional SE3 baseline. Turbines across the 72 clusters are categorised into four generations from metadata (installed capacity, location, hub height). For each cluster, raw NWP wind speeds undergo a sub-hourly gust correction, are extrapolated to hub height via a logarithmic wind profile, and are mapped through family-specific power curves to net capacity factors, which are adjusted for availability and wake losses before aggregation. The complete formulations and turbine power curves are given in Appendix D.1.

Feature Engineering & Integration

To let the ML models correct systematic biases in the physical baseline (unmodeled stability transitions, generic power-curve mismatches), the baseline is embedded in a multi-group feature matrix. Alongside the aggregated baseline and per-cluster MW estimates, it includes raw meteorological aggregates (capacity-weighted mean speeds, standard deviations, cyclic wind directions), cyclic calendar variables, and spatial-coherence metrics (e.g. the fraction of clusters above cut-in/cut-out speed) that proxy for synoptic stability. For the GFS horizon (60–168 h), explicit lead identifiers and handoff features from the MET tail (58 h predictions and ramps) are appended to preserve cross-model continuity.

Step 2 — Unified Machine Learning Framework

The combined feature matrix is passed into the ML framework. Separate models are fitted for each hourly lead on the MET Nordic branch (1–58 h) to handle horizon-specific errors. On the GFS branch (60–168 h), observations are pooled across leads to counter the lower native 3-hourly data cadence. At inference time, the calibrated MET forecast and the GFS extension are concatenated into a seamless 1–168 h output.

4.2.4 Solar Power Production

Targeting hourly metered SE3 solar production, we apply the two-stage calibration framework from Section 4.2.2. The weather feature set is restricted to shortwave irradiance and cloud cover; temperature is excluded to avoid overfitting on a secondary effect [54].

Step 1 — Physical Domain Modeling

Rather than using explicit panel-level thermodynamic equations, the physical baseline for solar production is constructed by defining a synthetic proxy for solar availability and spatially aggregating the relevant meteorological drivers.

Astronomical seasonality profiles. To capture the deterministic cycles of solar availability over the year and within the day, we construct an artificial input signal. This theoretical proxy is formed by taking the product of two min-max scaled Gaussian profiles: an annual profile centered on the summer solstice, and an intraday profile centered on solar noon. To account for seasonal variations in daylight length, the width of the intraday Gaussian scales dynamically with the time of year. The mathematical formulations for these profiles are detailed in Appendix D.2.

Spatial aggregation. The meteorological inputs are aggregated following the two-stage framework of Section 4.2.1 using solar-specific weights. To reflect the urban concentration of rooftop installations, municipality-to-region weights are proportional to historical production shares. Across SE3, region-to-SE3 weights are proportional to installed solar capacity. The explicit formulation of these capacity weights, alongside the dynamic renormalisation for missing data, is provided in Appendix D.2.

Feature Engineering & Integration

Capacity build-out adjustment. Due to the rapid, non-stationary build-out of SE3 solar capacity, historical production features must be scaled to current levels using a year-over-year multiplier. This is computed as the ratio of total production between consecutive years across shared valid dates (formulated in Appendix D.2). For 2026, the multiplier is fixed at 1.15 based on Swedish Energy Agency projections [55].

Seasonal historical baselines. To capture seasonal trends, we compute a historical average for the target hour across a rolling calendar window. This mean is then scaled by the capacity growth multiplier. To smooth out weather anomalies, this growth-adjusted baseline is computed individually for the four preceding years. These four baselines, along with their unweighted mean, provide five distinct features. The explicit definition of the historical window mean is detailed in Appendix D.2.

Short-term autoregressive and time features. Day of the year ($d \in [1, 366]$) provides a non-cyclic complement to the calendar variables from Section 4.2.1. Immediate generation trends are captured by two autoregressive features: lagged production over the preceding week (accounting for a 48-hour data delay) and the week's average production at the target hour.

Step 2 — Unified Machine Learning Framework

The combined feature matrix is passed into the ML framework defined in Section 4.2.1. By combining the deterministic astronomical curves (S_{combined}), stochastic meteorological aggregates, and scaled historical actuals, the model learns the nonlinear mapping required to translate weather attenuation into accurate SE3 megawatt solar generation.

4.2.5 Electricity Consumption

To forecast hourly SE3 consumption (eSett API), we rely on air temperature as our sole meteorological input, as it is the dominant, nonlinear driver of electricity demand. SMHI observations calibrate a dedicated temperature-elasticity model, while weather forecasts drive the final prediction.

Step 1 — Physical Domain Modeling

Spatial aggregation. Temperature forecasts are aggregated following the two-stage framework of Section 4.2.1, but utilizing historical-consumption-share weights. At the intra-regional level, municipality weights are proportional to municipal consumption (skipping this step for regions represented by a single primary city). At the SE3 level, regional weights are proportional to their share of total SE3 consumption. This produces a single consumption-weighted SE3 temperature forecast, $\hat{T}_{SE3,h}$. The mathematical formulation for these weights is detailed in Appendix D.3.

Temperature-elasticity baseline. To capture the non-linear relationship between temperature and demand while preserving complex calendar seasonality, we build a baseline around a quadratic elasticity fit. For a target timestamp, the algorithm identifies matched historical observations from the preceding five years using calendar- and holiday-aware matching criteria. The historical consumption for these matched dates is then synthetically "swapped" from its historical weather onto the forecasted weather using the elasticity model. The unweighted mean of these weather-adjusted historical loads forms a theoretical baseline. The explicit equations for this quadratic swap and adjustment are provided in Appendix D.3.

Feature Engineering & Integration

While the elasticity baseline establishes a strong temperature-dependent anchor, multi-year structural trends and short-term behavioural patterns (e.g., shifting schedules) require additional features.

Multi-year historical baselines. To capture pure seasonal and time-of-day patterns without the temperature swap, we compute a rolling mean around the matched date in each of the four preceding years. These four yearly baselines and their unweighted mean enter the feature set as five separate variables, preserving year-specific structure while stabilising against single-year weather anomalies. The mathematical definition of this rolling window is deferred to Appendix D.3.

Day-adjusted short-term autoregressive baseline. Because electricity consumption profiles vary significantly across the week, utilizing recent historical observations directly introduces a weekday-mismatch bias (e.g., using a Sunday observation to predict a Tuesday). To correct this, we establish a global historical average for every weekday and hour to compute a relative scaling factor. The three most recent valid observations are multiplied by this factor, dynamically projecting them onto the target weekday's profile. The explicit scaling equations are detailed in Appendix D.3.

Lagged hourly averages and direct features. A second short-term feature computes the average consumption over a fixed multi-day lag set (spanning two to nine days prior) at the target hour. Two direct features complete the unified matrix: day of the year ($d \in [1, 366]$) to complement the cyclic calendar variables, and the raw consumption-weighted SE3 temperature $\hat{T}_{SE3,h}$.

Step 2 — Unified Machine Learning Framework

The features above form the complete matrix for the ML models (Section 4.2.1). By capturing both physical thermodynamics and momentum, the framework dynamically weights weather impacts against human habits to produce a hourly SE3 consumption forecast across all time horizons.

4.3 Spot Price Prediction

4.3.1 Overview and Modelling Approach

The second pipeline stage produces hourly SE3 day-ahead spot price forecasts for the EV charging evaluation in Section 4.4.1. Forecasts span 95 distinct lead hours, covering $H \in \{18, \ldots, 58\} \cup \{60, 63, \ldots, 168\}$. The lower cutoff of $H = 18$ reflects the 12:45 CET Nord Pool gate (detailed in Section 4.3.3), ensuring models only predict prices that are not yet confirmed.

Following the NWP-calibration approach established in Section 4.2.3, an independent model is trained for each lead hour. For every lead, two models are trained in parallel: a point model optimised for mean absolute error, and a quantile model predicting $(q_{10}, q_{25}, q_{50}, q_{75}, q_{90})$. The point forecast drives the price level comparisons (Section 4.3.5), while the 80 % central interval supplies the uncertainty bounds required by the risk-adjusted scoring rule (Section 4.4.2). This point-and-quantile framework is applied to evaluate four model families: Random Forest, XGBoost, GRU, and LSTM.

As summarised in Figure 4.3, the pipeline proceeds from feature integration to per-lead training, followed by a six-fold expanding-window cross-validation over the 2025-2026 out-of-sample period (Section 4.3.5). Model performance is ultimately assessed via standard metrics and SHAP interpretability (Sections 4.2.1 and 5.2.3).

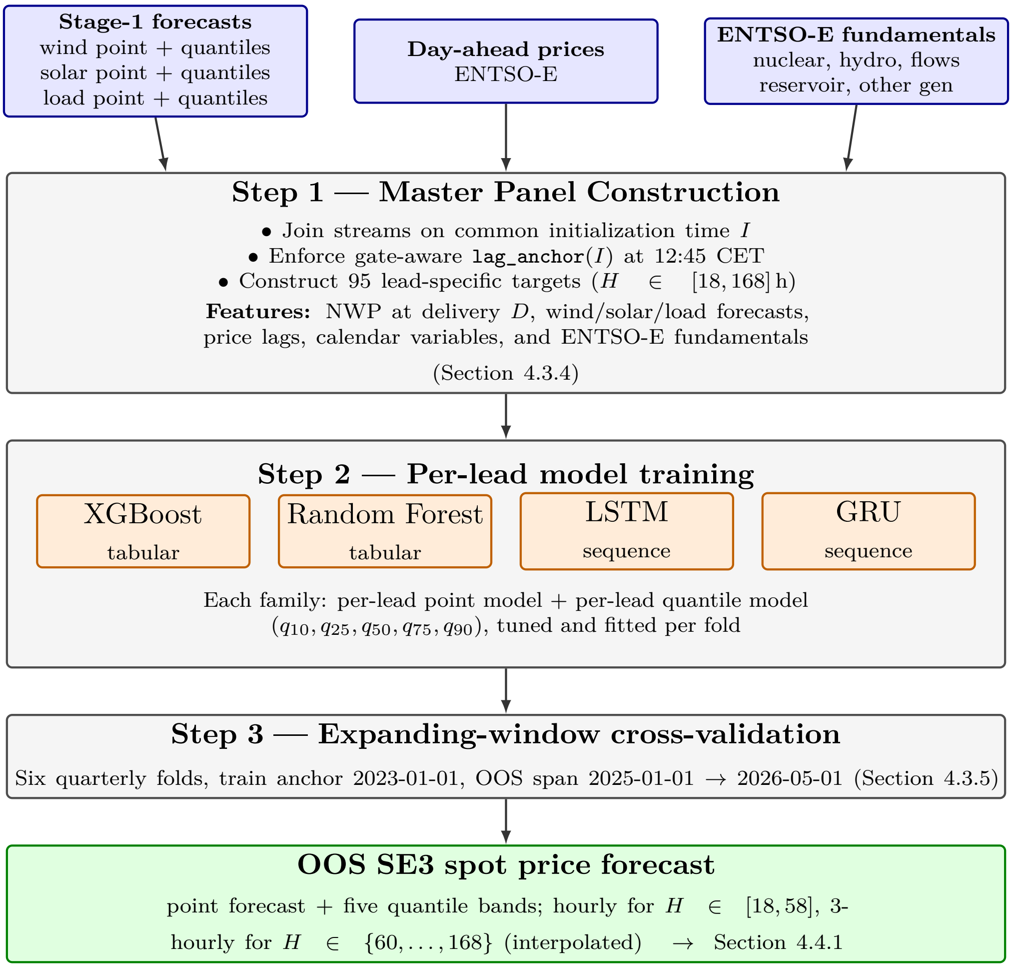

Figure 4.3: Spot price prediction pipeline. Using upstream forecasts and historical actuals to train per-lead models for each family.

4.3.2 Data Sources

The price model draws on three data categories, each with a distinct role in the feature matrix assembled in Section 4.3.4.

Day-ahead spot prices. Hourly SE3 day-ahead prices in EUR/MWh are retrieved from the ENTSO-E Transparency Platform.[1] The price series covers the period from January 1, 2023, onward, and is indexed by the UTC delivery timestamp.

Upstream forecasts. The three Stage-1 forecasts enter the model as wide panels indexed by initialization time I. Each lead step K provides a point forecast and five quantile intervals $(q_{10}, q_{25}, q_{50}, q_{75}, q_{90})$ for the delivery hour $I + K$. Furthermore, wind and load are supplemented with historical hourly actuals from the previous 168 h, derived from the same eSett settlement feed used for Stage-1 calibration.

[1] Cross-checked at model build time against Energi Data Service (Elspotprices dataset, published by Energinet); both sources republish the same Nord Pool auction outcomes and agree to rounding.

ENTSO-E fundamentals. Supplementary supply-side signals are drawn from the ENTSO-E Transparency Platform: per-unit nuclear generation, hydro generation, other generation, cross-border scheduled and physical flows, and weekly hydropower reservoir levels. These are aggregated to hourly SE3 totals (reservoir levels remain at weekly resolution and are forward-filled) and structured as a separate panel indexed by initialization time. This panel is joined to the main feature matrix in Section 4.3.4 whenever the ENTSO-E feature group is included in the configuration.

Temporal coverage. All sources are harmonised to the window 2023-01-01 through 2026-05-01, with 2023–2024 reserved as training data so every fold has at least two full years before its prediction window; 2025-01-01 to 2026-05-01 forms the out-of-sample span used in Section 4.3.5.

4.3.3 Gate-Aware Information Structure

The Nord Pool day-ahead auction for delivery on day $d+1$ closes at 12:00 CET on day d, with hourly clearing prices for all 24 delivery periods published by 12:45 CET [10]. Let $g = 13{:}00$ CET denote the hour by which these prices are reliably available; g is the *gate*.

For an NWP initialisation time I, the *lag anchor* $\tau(I)$ is the earliest Coordinated Universal Time (UTC) delivery timestamp whose price is not yet confirmed at I:

$$\tau(I) \;=\; \mathrm{midnight}_{\mathrm{CET}}(I) \;+\; \begin{cases} 1 \text{ day} & \text{if } I_{\mathrm{CET}} < g, \\ 2 \text{ days} & \text{if } I_{\mathrm{CET}} \geq g, \end{cases} \tag{4.1}$$

returned in UTC. Pre-gate, only prices for the current CET day are confirmed; post-gate, prices for the following day are available as well.

The anchor $\tau(I)$ strictly partitions the feature space to prevent data leakage. All price-based features, including hourly lags up to 168 h, rolling 24 h and 168 h statistics, and same-hour weekly lags, are evaluated strictly prior to $\tau(I)$. Conversely, the Stage-1 forecasts (wind, solar, and load) and their historical actuals are indexed to the initialization time I, as their availability follows the NWP cadence rather than the Nord Pool gate. Finally, calendar features depend solely on the delivery timestamp $D = I + H$.

Equation (4.1) bounds the meaningful lead hours, as predicting any delivery period where $I + H \leq \tau(I)$ is redundant because the price is already confirmed. Across the four NWP cycles (00, 06, 12, and 18 UTC), the shortest valid lead is $H = 18$. This occurs during the 06 UTC initialization: because it is pre-gate, the anchor $\tau(I)$ falls at 23 UTC, making $06 + 18 = 24$ UTC the first unconfirmed hour. Consequently, leads shorter than 18 h are globally excluded. Because the validity threshold depends on the initialization hour, invalid (I, H) pairs are dropped dynamically. As a result, training panels for early leads contain fewer observations per day than mid-range leads.

4.3.4 Feature Set

Each row in the feature matrix represents an initialization-lead pair (I, H), targeting the spot price at delivery hour $D = I + H$. Features are organized into six groups following the gate-aware logic of Section 4.3.3:

- **NWP based forecasts at D:** A point forecast and five quantile predictions (q_{10} through q_{90}) for each of wind, solar, and load (18 features total).
- **Historical production actuals:** Hourly lags (1–168 h) for wind, solar, and load production.
- **Price history:** Hourly price lags (1–168 h) drawn strictly from the confirmed window prior to $\tau(I)$.
- **Anchor-relative price statistics:** Summary metrics computed over the confirmed window up to $\tau(I)$. These include the anchor price p_τ; 24 h and 168 h rolling means and standard deviations; 24 h min/max; the previous day's price spread; and same-hour lags (1, 2, and 4 weeks prior) to capture diurnal recurrence.
- **Calendar features at D:** Hour, weekday, and month ; binary weekend and peak-hour indicators; forecast distance $D - \tau(I)$; and the Gaussian summer-solstice profile ($\sigma = 78$ days) capturing midday price suppression.
- **ENTSO-E fundamentals:** Supply-side anchor values and rolling means (24 h and 168 h) for nuclear, hydro, other generation, and cross-border flows. Hydropower reservoir states are included as both total MWh volume and percentage of the seasonal median.

Panel schema. To minimize the computational overhead of joins and rolling statistics, all lead-specific data is materialized once into a single master panel. During training, a helper routine extracts the valid rows for each lead hour H and standardizes the feature names. This provides a consistent input schema across all leads, which the tabular and sequence models then consume according to their respective architectures.

4.3.5 Expanding-Window Cross-Validation

To maximize available training data, models are evaluated via an expanding-window scheme that simulates quarterly retraining. The out-of-sample period (2025-01-01 to 2026-05-01) is split into six quarterly folds. For each fold, models train on all data from a fixed anchor (2023-01-01) up to the beginning of that quarter, guaranteeing at least two years of historical data per fold. Models are refit from scratch at each step, and their out-of-sample predictions are concatenated into a single evaluation panel. Table 4.2 summarises the fold structure.

Table 4.2: Quarterly expanding-window fold schedule for the spot price model. Training begins at the common anchor 2023-01-01, two full calendar years before the first prediction quarter, and the training window is extended by one quarter at each refit. Fold 6 covers April 2026 only.

Fold	Training window end	Prediction window
1	2024-12-31	2025-01-01 to 2025-03-31
2	2025-03-31	2025-04-01 to 2025-06-30
3	2025-06-30	2025-07-01 to 2025-09-30
4	2025-09-30	2025-10-01 to 2025-12-31
5	2025-12-31	2026-01-01 to 2026-03-31
6	2026-03-31	2026-04-01 to 2026-05-01

4.3.6 Evaluation

The spot price model follows the same evaluation framework as the input modelsdescribed in Section 4.2.1.

4.4 Electric Vehicle Charging Evaluation

4.4.1 Overview and Scope

The third stage of the pipeline translates the spot price forecasts produced in Stage 2 into quantified economic value for EV owners. Rather than evaluating forecast quality in isolation, this stage asks a downstream question: given a particular price forecast, how much can a vehicle owner reduce their charging cost by acting on it, and how does that value vary across charging contexts, driving behaviours, and forecast quality levels?

The evaluation is conducted under two complementary frameworks that differ in how the algorithm operates on the same upstream forecasts. **Framework 1** is opportunistic: at each plug-in session, the algorithm decides whether to charge or defer to a future session, with no external commitment to any particular charging schedule. **Framework 2** introduces a weekly deadline: the driver declares once a week that the vehicle should reach a target state of charge by a specified time (for example, 80 % by Monday 07:00), and the algorithm must allocate the required energy across the week's available charging windows to honour that target. Both frameworks share the same vehicle, driver profiles, wallbox configurations, baselines, and forecast pipeline; only the structure of the algorithm's optimisation problem differs. Table 4.3 summarises the structural differences at a glance; the formal specifications follow in Sections 4.4.4 and 4.4.5.

The two frameworks are evaluated across two use cases that span the residential charging contexts. **Use Case 1** (Section 4.4.6) examines home overnight charging, where the vehicle is parked in a multi-hour window each night. **Use Case 2** (Section 4.4.7) examines workplace daytime charging for a driver without home access, e.g., an individual living in an apartment with no possibility of having their own charger. The two cases

Table 4.3: Comparison of the two evaluation frameworks for EV smart charging.

	Framework 1: Continuous	**Framework 2: Weekly Target**
Driver behaviour	Plugs in each session; the algorithm decides whether to charge or defer.	Declares a target State of Charge (SoC) and deadline once per week (e.g. 80 % by Monday 07:00); no further input.
Algorithm decision	Per-session: charge tonight or wait for a cheaper future session.	Per-week: allocate the required energy across the week's available charging windows.
Optimisation horizon	Rolling, N sessions ahead.	Fixed forward window from now to the deadline.
Binding constraint	SoC floor and connection-window capacity.	Same, plus a hard deadline-energy constraint that must be reachable from every state.
What it measures	Maximum savings extractable from multi-day forecasts under fully flexible charging.	Savings extractable when the optimisation operates under a fixed weekly energy target.

differ structurally in their charging window, decision timing, and information available at the point of decision; these differences interact with each framework's optimisation problem to produce materially distinct savings profiles.

Within each framework × use case combination, value is reported as S_{defer}, the marginal saving from acting on multi-day forecasts, and $S_{\text{sched}} + S_{\text{defer}}$, the cumulative saving against a do-nothing driver, with the diagnostic G_{oracle} alongside; these are defined in Section 4.4.3. The cross-framework comparison is the sharpest test of whether multi-day forecasts hold up when the algorithm must commit ahead of time. Stylised assumptions are introduced at the level they apply to, from vehicle and driver (Section 4.4.2) through to use case (Sections 4.4.6 and 4.4.7).

4.4.2 Vehicle, Driver and Charging Setup

Vehicle Specification

All use cases are evaluated for the Volvo EX60 P10 AWD [56], a battery electric SUV. Table 4.4 summarises the relevant specifications, and Table E.3 (Appendix E) lists the modelling assumptions applied to the vehicle's energy behaviour throughout the evaluation.The manufacturer's energy-consumption figure follows the Worldwide Harmonised Light Vehicles Test Procedure (WLTP), the standardised laboratory cycle used for type-approval consumption ratings.

Table 4.4: Volvo EX60 P10 AWD specifications used in the charging evaluation, based on public information [56].

Parameter	Value	Source
Nominal battery capacity	95 kWh	[56]
WLTP energy consumption	16.2 kWh/100 km	[56]

Driver Profiles

Four synthetic driver profiles are evaluated, differing only in daily driving distance. The reference of 50 km/day is anchored in empirical Swedish travel data: GPS measurements of recent private BEV buyers report a mean of 54 km/day, while register data place the typical annual mileage at 12,000–14,000 km, or 33–38 km per calendar day. The same study finds Swedish BEV users drive 4–6 km/day more than comparable internal-combustion drivers [57]. The 50 km/day reference therefore sits between the two central estimates and represents a plausible average for a Swedish BEV commuter.

Let d (km/day) denote the daily driving distance and $e = d \cdot \varepsilon/100$ (kWh/day) the corresponding daily consumption, with ε (kWh/100 km) the modelled energy consumption defined in Table E.3. Table 4.5 summarises the four profiles, including the maximum SoC-floor-feasible deferral horizon in days $N_{\max} = \lfloor \Delta s/e \rfloor$ for UC1 under Framework 1, where sessions occur every calendar day.

Table 4.5: Driver profiles evaluated across both use cases.

Distance d	Energy e	$N_{\max}$ (UC1, F1)	Slots at 7.4 kW	Note
25 km/day	5.0 kWh	10	1	Short commuter
50 km/day	10.0 kWh	5	2	Reference profile
75 km/day	15.0 kWh	3	3	Long commuter
100 km/day	20.0 kWh	2	3	High-mileage driver

The 25 km/day and 75 km/day profiles bracket the reference within the day-to-day variation reported in [57]. The 100 km/day profile captures the upper tail of long-distance commuters and is included to expose the interaction between high consumption and the SoC floor.

The driver-side behavioural assumptions applied across both frameworks and both use cases are listed in Table E.4 (Appendix E).

Wallbox Configurations

Three alternating current (AC) wallbox configurations are evaluated, corresponding to the single-phase and three-phase installations available to residential and workplace EV owners in Sweden. The 7.4 kW (1-phase 32 A), 11 kW (3-phase 16 A), and 22 kW (3-phase 32 A) levels correspond to the single-phase and three-phase AC configurations supported by standard residential wallboxes [58], [59].

For an energy delivery of e kWh at charging rate P, the required number of full one-hour slots is $H = \lceil e/P \rceil$, with the final slot used fractionally. Table 4.6 lists the three configurations and the slot count for the 50 km/day reference profile.

Table 4.6: Wallbox configurations and slot requirements for the 50 km/day reference profile ($e = 10.0$ kWh).

Configuration	Power P	Full slots	Partial fraction	Standard
1-phase 32A	7.4 kW	1	0.351	Common residential [58], [59]
3-phase 16A	11.0 kW	1	0 (exact)	Workplace standard [58], [59]
3-phase 32A	22.0 kW	1	0 (exact)	Fast AC [58], [59]

The 7.4 kW wallbox requires two slots for the reference profile and presents the richer slot-selection problem; results in the main text correspond to this configuration unless otherwise stated.

Window-capacity constraint. A nine-hour charging window (Sections 4.4.6 and 4.4.7) bounds the energy that can be delivered in a single charging event at $9 \cdot P$ kWh. Under Framework 1 the energy required at the eventual charging event grows with the calendar distance to the previous charge, so combinations of high e and low P, such as a 100 km/day driver on a 7.4 kW wallbox, can be forced to charge earlier than the SoC floor alone would allow because deferring further would push the required delivery beyond what the window can supply. Under Framework 2 the same constraint applies session-by-session during weekly slot allocation. Both frameworks enforce the window-capacity bound jointly with the SoC floor in their feasibility checks.

4.4.3 Shared Evaluation Components

This subsection collects the elements common to both frameworks: the benchmarks against which each algorithm is measured, the cost decomposition that reports the savings, the annual scaling that converts per-event costs into the headline numbers carried into Chapter 5, and the market and contract assumptions that determine what a charged kWh costs the driver.

Benchmarks

Three benchmarks are defined against which the algorithm is measured. The naive and smart baselines represent strategy-free or forecast-free reference behaviours and are framework-independent: their per-session logic is the same regardless of which framework's algorithm they are being compared against. The oracle is framework-specific. It runs the same algorithm as the framework it is paired with, but using confirmed prices in place of forecasts. The oracle's purpose is to measure forecast error within each framework on its own terms.

Naive baseline. The driver plugs in and charges immediately, with no scheduling or deferral. The naive cost is computed from the first H slots of the charging window in chronological order at confirmed prices. This is the do-nothing reference.

Smart baseline. The driver schedules optimally within the current session by selecting the cheapest H confirmed hours, but never defers to a future day or optimises across sessions. The smart baseline cost is C_0, computed via (4.9) on the current session. The difference $C_{\text{naive}} - C_0$ isolates the value of intra-day scheduling alone, a component requiring no price forecast. This corresponds to the behaviour of current-generation smart charging products [3], [4].

Oracle. The oracle runs the same algorithm as its framework but on confirmed prices in place of forecasts: under Framework 1 it selects $K^*_{\text{oracle}} = \arg\min_{K \in \mathcal{F}} C_K$ at each session (re-evaluating if no charge occurs); under Framework 2 it solves the same weekly slot-allocation problem with true prices across the remaining-week window. Both variants share the algorithm's horizon and feasibility constraints, so the resulting gap isolates forecast error rather than a difference in information structure. Note that the oracle is optimal *per session*, not over the full window: because it greedily minimises each session's confirmed cost across the feasible candidate nights, its locally optimal choices can occasionally lock in a more expensive cumulative trajectory than the algorithm's, so G_{oracle} is a forecast-quality diagnostic and not a strict lower bound. This per-event design is deliberate. By restricting the oracle to the same decision-by-decision structure as the algorithm, differing only by using true prices instead of forecasts, G_{oracle} strictly measures forecast quality. This ensures a fair comparison rather than evaluating a completely different, globally optimal policy.

Cost Decomposition

Recall from Section 4.4.3 that C_{naive} is the naive-baseline cost on the current session, and C_0 the smart-baseline cost on the same session. Let C_{K^*} denote the confirmed cost on the session the algorithm eventually charges, and C_{oracle} the corresponding cost for the oracle under the same framework. Three components are defined per charging event in Equations (4.2), (4.3), and (4.4), illustrated for a single charging sequence in Figure 4.4.

$$S_{\text{sched}} = C_{\text{naive}} - C_0, \tag{4.2}$$

$$S_{\text{defer}} = C_0 - C_{K^*}, \tag{4.3}$$

$$G_{\text{oracle}} = C_{K^*} - C_{\text{oracle}}. \tag{4.4}$$

Two answers to research question 2 are reported per use case and per framework. The marginal answer is S_{defer}, the additional cost reduction obtained from acting on multi-day forecasts over and above scheduling within the current session. This is the component that Stage 2 forecasts directly enable, and it isolates the contribution of the multi-day forecasts on top of behaviour already available to current-generation smart charging products. The cumulative answer is $S_{\text{sched}} + S_{\text{defer}} = C_{\text{naive}} - C_{K^*}$, the total cost reduction relative to a do-nothing driver, which captures both the forecast-free scheduling component and the forecast-driven deferral component. Annual versions of both quantities (Equation 4.5) are the headline numbers carried into Chapter 5.

The arithmetic is identical across frameworks; only the interpretation of S_{defer} shifts. Under Framework 1 it is the value of choosing the right *day* to charge; under Framework 2, the value of allocating the weekly energy budget to the forecast-cheaper *hours across the week*.

G_{oracle} is reported as a diagnostic, not a headline answer: it measures the headroom perfect-information operation could capture beyond the algorithm. A small G_{oracle} indicates the forecasts are doing nearly all they can; a large one indicates material room for forecast improvement to add savings.

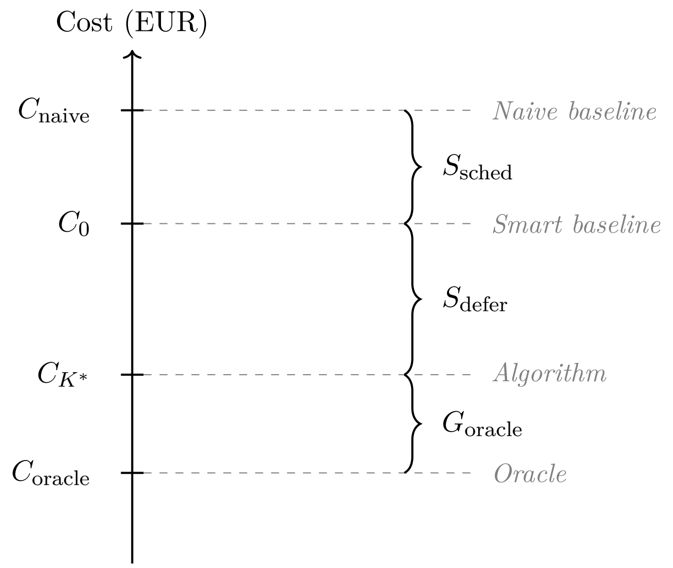

Figure 4.4: Conceptual cost decomposition for a single charging sequence, applicable to both frameworks.The four cost levels (naive, smart baseline, algorithm, and oracle) define the three savings components: S_{sched} (scheduling saving, forecast-free), S_{defer} (deferral saving, enabled by Stage 2 forecasts), and G_{oracle} (oracle gap, the headroom remaining under perfect-information operation). Magnitudes are illustrative.

Annual Scaling

Annual costs are computed as the sum of all event costs over the evaluation period, divided by the number of evaluation years:

$$\bar{C}_{\text{annual}} = \frac{1}{Y} \sum_{t \in \mathcal{E}} C_t, \tag{4.5}$$

where $\mathcal{E}$ is the set of charging events and Y the period length in years. Annual savings follow from the corresponding cost differences between strategies. This formulation makes no assumption about the number of charging events per year, which depends endogenously on each framework's operating logic and is reported in the results.

Market and Contract Assumptions

The driver is assumed to be exposed to the SE3 day-ahead spot price through a pass-through variable-rate contract. The cost-side assumptions are listed in Table E.5 (Appendix E) and apply identically across both frameworks and both use cases.

4.4.4 Framework 1: Continuous Deferral

Under Framework 1, the driver plugs in at every session and the algorithm decides, session by session, whether to charge or defer to a future session. The algorithm is opportunistic: it commits to no charging schedule, makes its decision using the current forecast, and re-evaluates on the next plug-in if no charge occurred. The benchmarks and cost decomposition of Section 4.4.3 apply directly to the resulting cost trajectory.

Decision Algorithm

At each plug-in event the algorithm selects the candidate day $K^* \in \{0, 1, \ldots, N\}$ on which to charge, where $K = 0$ denotes the current night (UC1) or working day (UC2) and $K = k$ the k-th future candidate. The decision is made in two stages: an inter-day choice of which day to charge, and an intra-day choice of which H slots within the chosen day to use.

Intra-day slot pre-selection. At decision time, H candidate slots are pre-selected for each candidate day by choosing the H slots with the lowest predicted price $\hat{p}_h^{(0.50)}$. For confirmed days, $\hat{p}_h^{(0.50)} = p_h$, the published spot price.

Inter-day scoring. Each candidate day K is assigned a risk-adjusted score based on its pre-selected slots. Define the risk-adjusted price for slot h as

$$r_h(\lambda) = (1 - \lambda)\, \hat{p}_h^{(0.50)} + \lambda\, \hat{p}_h^{(0.90)}, \tag{4.6}$$

where $\lambda \in [0, 1]$ is the risk-appetite parameter and $\hat{p}_h^{(0.50)}$, $\hat{p}_h^{(0.90)}$ are the predicted median and 90th-percentile prices for slot h. The score for candidate day K is the energy-weighted average across its H pre-selected slots

$$\sigma_K(\lambda) = \frac{1}{e} \left[P \sum_{i=1}^{H-1} r_{h_i}(\lambda) + \Big(e - (H-1)P\Big)\, r_{h_H}(\lambda) \right], \tag{4.7}$$

where $h_1, \ldots, h_H$ are the pre-selected slot indices ordered as fitted in the energy budget, P the charging power, and e the energy required to top the vehicle up from the SoC floor to the SoC target (Section 4.4.2). Three values $\lambda \in \{0,\ 0.5,\ 1\}$ are evaluated; primary results are reported at $\lambda = 0$.

Decision rule and feasibility. The algorithm selects the feasible candidate with the lowest score, as in Equation (4.8),

$$K^* = \underset{K \in \mathcal{F}}{\arg\min}\ \sigma_K(\lambda), \tag{4.8}$$

where the feasible set $\mathcal{F} \subseteq \{0, \ldots, N\}$ contains those K that satisfy both

(a) the SoC floor: cumulative depletion to candidate day K must not breach $\underline{s}$, and

(b) the window-capacity bound: the energy that would have to be delivered at the eventual charging event must not exceed $9 \cdot P$ kWh (Section 4.4.2).

Both constraints are evaluated against the vehicle's current SoC and the calendar distance to the previous charge, so $\mathcal{F}$ is recomputed at every re-evaluation step.

Confirmed-price re-optimisation. Once K^* is selected, the H cheapest slots on that day are re-selected using confirmed prices at charging time: the inter-day decision uses forecasts, the intra-day assignment waits for the day-ahead publication. The actual charging cost is (with the 1/1000 factor converting kWh · EUR/MWh to EUR)

$$C_K = \frac{1}{1000}\left[P \sum_{i=1}^{H-1} p_{(i),K} + \Big(e - (H-1)P\Big)\, p_{(H),K}\right], \tag{4.9}$$

where $p_{(1),K} \leq \cdots \leq p_{(H),K}$ are the H lowest confirmed prices on day K. The complete step-by-step procedure is summarised as pseudocode Algorithm 2 (Appendix E).

Rolling re-evaluation. If $K^* > 0$ the vehicle does not charge and the algorithm re-runs the following session with fresh predictions, an updated SoC, and a rolling N-step horizon; previous-session decisions carry no commitment. By the second iteration of a deferral sequence the preferred future day has moved closer to the confirmed range and forecast quality has improved, and as the battery depletes $\mathcal{F}$ contracts until the algorithm is forced to charge.

The modelling choices specific to Framework 1's algorithm, forecast inputs, the fixed risk parameter, the rolling $N = 3$ horizon, full re-evaluation, median-based slot pre-selection, and confirmed-price intra-day re-optimisation, are listed in Table E.1 (Appendix E). These shape what Framework 1 measures and distinguish it from a more general dynamic-programming formulation.

4.4.5 Framework 2: Weekly Target

Under Framework 2, the driver declares once a week that the vehicle should reach the target SoC $\bar{s}$ by a specified deadline d_{target} (taken throughout this work as 80 % by Monday 07:00 CET). The declaration is repeated identically every Monday in perpetuity, and the driver provides no further input during the week. The algorithm must therefore deliver the required energy across the available connection windows in the remaining-week horizon, allocating it to the slots forecast to be cheapest while honouring the deadline as a hard constraint.

This framework changes the algorithm's optimisation problem in two ways relative to Framework 1. First, the energy delivered across the week is not endogenous: it is fixed by the difference between the current SoC and the deadline target, plus depletion between now and the deadline. Second, the deadline imposes a forward-feasibility constraint that must hold at every state: the remaining connection capacity must be sufficient to deliver the remaining energy by d_{target}, even if every remaining slot were used. The benchmarks and cost decomposition of Section 4.4.3 apply directly; only the algorithm differs.

Weekly Declaration and Energy Budget

At the start of each weekly cycle (Monday 07:00 CET), the driver's declaration fixes the target SoC $\bar{s}$ and deadline d_{target}. At any subsequent session t within the week, the algorithm computes the energy still to be delivered before the deadline

$$E_{\text{rem}}(t) = (\bar{s} - s_t) \cdot B + e \cdot d_{\text{rem}}(t), \tag{4.10}$$

where s_t is the current SoC fraction at session t, B is the usable battery capacity, e is the daily consumption (Section 4.4.2), and $d_{\text{rem}}(t)$ is the number of calendar days from t to the deadline. The first term accounts for restoring the battery from its current state to the target; the second term accounts for ongoing consumption between t and the deadline.

Forward-Feasibility Constraint

Let $\mathcal{R}(t) = \{j : t \leq j < d_{\text{target}}\}$ denote the sessions remaining between t and the deadline, and W_j the connection window length of session j (taken throughout as $W_j = 9$ hours). The remaining charging capacity at t is

$$C_{\text{rem}}(t) = P \cdot \sum_{j \in \mathcal{R}(t)} W_j, \tag{4.11}$$

where P is the wallbox charging power and $C_{\text{rem}}(t)$ the maximum energy the vehicle could deliver between t and the deadline if every remaining slot were used at full power. Forward-feasibility requires that this capacity is sufficient to deliver the remaining energy, so that

$$C_{\text{rem}}(t) \geq E_{\text{rem}}(t), \tag{4.12}$$

with $E_{\text{rem}}(t)$ the remaining energy requirement from (4.10). Equivalently, the slack expressed in hours,

$$\text{slack}(t) = \frac{C_{\text{rem}}(t) - E_{\text{rem}}(t)}{P}, \tag{4.13}$$

where slack(t) is the number of remaining connection-hours the algorithm is free to skip while still satisfying (4.12); it must be non-negative at every state. Large slack gives substantial flexibility to defer charging to forecast-cheaper slots; small slack approaches forced charging across most remaining hours. Slack contracts deterministically across the week as connection windows are consumed, regardless of forecasts, so by the final session the algorithm typically has either zero slack (forced charging) or a small positive value depending on how much energy was delivered earlier.

Slot Allocation

Let $\mathcal{T}_j$ denote the set of hourly slots in the connection window of session j, of cardinality $|\mathcal{T}_j| = W_j$. The candidate slot pool at session t is the union of remaining sessions' slot sets

$$\mathcal{P}(t) = \bigcup_{j \in \mathcal{R}(t)} \mathcal{T}_j, \tag{4.14}$$

as given in Equation (4.14). Each slot $h \in \mathcal{P}(t)$ is scored using the risk-adjusted price $r_h(\lambda)$ of equation (4.6). Slots in confirmed days have $r_h = (1-\lambda)p_h + \lambda p_h = p_h$; slots in predicted days use the forecast quantiles directly. The number of slots required to deliver $E_{\text{rem}}(t)$ at power P is $H_{\text{rem}}(t) = \lceil E_{\text{rem}}(t)/P \rceil$, with the final slot used fractionally. The algorithm selects the $H_{\text{rem}}(t)$ cheapest slots from $\mathcal{P}(t)$ by ascending $r_h(\lambda)$. The decision rule for the current session is then

> The vehicle charges at session t if and only if at least one of the selected slots falls within $\mathcal{T}_t$, the connection window of session t.

If the selected slots include hours from $\mathcal{T}_t$, those slots are charged at session t at confirmed prices (forecasts for those hours are now confirmed because t is the current session). If no selected slot falls in $\mathcal{T}_t$, the vehicle does not charge and the algorithm re-evaluates at session $t+1$ with an updated state.

Forward-Feasibility Binding

When deferring at session t would yield slack$(t+1) < 0$, the algorithm overrides the cheapest-slot allocation and charges the minimum hours $h_{\text{min}}(t) = -\text{slack}(t+1)$ needed to restore non-negativity, selecting the $h_{\text{min}}(t)$ cheapest slots in $\mathcal{T}_t$ at confirmed prices. This is the simplest defensible response to a binding deadline: the algorithm charges only as much as forward-feasibility demands, leaving the remaining pool $\mathcal{P}(t) \setminus \mathcal{T}_t$ free to optimise at later re-evaluations. The complete step-by-step decision procedure for a single session within a weekly cycle is detailed as pseudocode in Algorithm 3 (Appendix E).

Weekly Cycle Reset

The algorithm operates on rolling weekly cycles. Each Monday at 07:00 CET the deadline is met (or, in the rare case of week-long forward-feasibility binding, the algorithm has charged the maximum it could) and a new cycle begins immediately: the deadline advances to the following Monday, the remaining-week horizon resets to seven calendar days, and allocation begins over the new $\mathcal{P}(t)$. The transition is continuous, the current cycle's deadline and the next cycle's origin are the same moment, so the declaration is implicitly renewed every Monday with no planning gap and no driver intervention.

The modelling choices specific to Framework 2's algorithm are listed in Table E.2 (Appendix E). Forecast inputs, the fixed risk parameter, and the full per-session re-evaluation are as in Framework 1; the Framework-2-specific choices are the once-weekly driver declaration with no mid-week revision, the hard $\bar{s} = 80\,\%$ deadline, and the forward-feasibility override of Section 4.4.5.

4.4.6 Use Case 1: Home Overnight Multi-Night Deferral

Context and Motivation

UC1 represents a driver charging at home on a private wallbox with a regular overnight parking schedule. This is the natural deployment context for OEM-integrated smart charging, offering the longest and most predictable connection window (nine hours, 22:00–06:00 CET) alongside maximum scheduling flexibility [1]. The driver is exposed to the SE3 day-ahead spot price through the variable-rate contract of Section 4.4.3.

Application under Framework 1

Under Framework 1, the algorithm decides at 22:00 CET each plug-in night whether to charge tonight or defer. The decision uses the 18:00 UTC NWP forecast init, which is *post-gate*: it falls after the Nord Pool day-ahead auction gate closure at 12:45 CET [10], so the following calendar day's prices have already been published. The lag anchor is therefore midnight CET plus two calendar days, converted to UTC. Delivery periods at or before the anchor are confirmed prices; those after are model predictions.

Table 4.7: Price information available at decision time (22:00 CET, 18:00 UTC init) for each candidate charging night in UC1 under Framework 1. MET refers to MET Nordic 1-hour predictions; GFS to 3-hour predictions for leads ≥ 60 h. The H=59 gap denotes the missing lead between MET and GFS coverage.

Night	Candidate	Slots	Source	Note
$K = 0$	Tonight	9/9	Confirmed day-ahead	All slots published
$K = 1$	Tomorrow night	9/9	3 confirmed + 6 MET 1h	22:00–00:00 confirmed
$K = 2$	Night +2	8–9/9	MET 1h	H=59 missing in winter
$K = 3$	Night +3	3/9	GFS 3h	H=75, 78, 81 only

The forecast resolution falls off sharply across candidate nights (Table 4.7; visualised in Figure E.1, Appendix E): $K = 3$ lies entirely in the 3-hour GFS range and so offers much poorer intra-night hour selection than $K = 1$ or $K = 2$. This matters less than it appears, because re-evaluation moves the preferred night toward the confirmed range, a night seen as $K = 2$ becomes $K = 1$ the next day, gaining three confirmed slots.

The primary horizon is $N = 3$. At the reference profile all four candidate nights remain feasible (the buffer stays above the floor across three deferrals); for the 100 km/day profile the buffer is exhausted after two deferrals and $\mathcal{F}$ contracts accordingly, consistent with the $N_{\max}$ column of Table 4.5.

Application under Framework 2

Under Framework 2, the driver declares each Monday at 07:00 CET that the vehicle should reach 80 % by the following Monday 07:00. The intervening week contains seven overnight connection windows, each from 22:00 to 06:00 CET, so the algorithm's slot pool $\mathcal{P}(t)$ at the start of the cycle contains $7 \times 9 = 63$ candidate hourly slots. Figure E.2 (Appendix E) illustrates the cycle: at the Monday 07:00 origin the full seven-night pool is visible, with the current evening confirmed and progressively deeper forecast horizons for later nights, and as the week progresses past sessions leave the pool while remaining slots move toward the confirmed range.

Across the 63 connection-hours in the week, the reference 50 km/day profile on a 7.4 kW wallbox needs roughly 9.5 charging-hours ($E_{\text{week}} = 7e = 70$ kWh), leaving slack of about 53 hours; the 100 km/day profile needs about 19 hours, still leaving roughly 44. UC1's nightly schedule keeps slack well above zero across all reasonable driver $\times$ wallbox combinations, so forward-feasibility binding is expected to be rare here.

Session Definition and Annual Scaling

The evaluation iterates over every calendar night in the out-of-sample period 2025-01-01 to 2026-05-01, matching the Stage 2 window. A charging event is recorded whenever the algorithm charges (Framework 1: $K^* = 0$; Framework 2: at least one slot from $\mathcal{T}_t$ assigned); annual costs follow (4.5), with the endogenous event count reported in the results. The use-case-specific assumptions for UC1 are listed in Table E.6 (Appendix E).

4.4.7 Use Case 2: Workplace Daytime Multi-Day Deferral

Context and Motivation

UC2 represents a driver who lacks home charging and instead relies on a workplace installation as their primary energy source. This reflects residents of multi-family dwellings without private parking, a constrained but substantial demographic in Sweden [60]. The vehicle is connected during working hours, Monday to Friday. The setting differs from UC1 in three structural ways: the connection window falls in the daytime, where Nordic prices tend to be higher and more volatile than late evening; the deferral candidates are working days, requiring explicit handling of weekend gaps during which the vehicle is driven but cannot charge; and the decision is made on morning arrival rather than the previous evening, which shifts the relationship between decision time and the day-ahead gate closure.

Application under Framework 1

Under Framework 1, the algorithm decides at 08:00 CET on each working-day arrival, using the 06:00 UTC forecast init. This init is *pre-gate*: it falls before the Nord Pool gate closure at 12:45 CET [10], so the following day's prices are not yet published, and the lag anchor is midnight CET plus one calendar day. The current day's prices are confirmed, but *all* candidate future working days are fully predicted, unlike UC1 under Framework 1, where $K = 1$ has three confirmed slots.

The interaction between working-day deferral and weekend calendar gaps produces a day-of-week-dependent resolution pattern, mapped in Table 4.8.

Table 4.8: Prediction resolution for UC2 under Framework 1, by session day-of-week and deferral step K. MET refers to hourly predictions; GFS to 3-hour predictions.

Session day	$K = 0$	$K = 1$	$K = 2$	$K = 3$	Last MET K
Monday	Confirmed	MET (Tue)	MET (Wed)	GFS (Thu)	2
Tuesday	Confirmed	MET (Wed)	MET (Thu)	GFS (Fri)	2
Wednesday	Confirmed	MET (Thu)	MET (Fri)	GFS (Mon)	2
Thursday	Confirmed	MET (Fri)	GFS (Mon)	GFS (Tue)	1
Friday	Confirmed	GFS (Mon)	GFS (Tue)	GFS (Wed)	0

The number of MET-resolution candidates shrinks toward the end of the week as deferral steps cross the weekend, leaving Friday sessions with all three candidates in GFS range. Since Fridays are roughly 20% of sessions, a non-trivial share of UC2 deferrals is decided on 3-hourly rather than hourly predictions, though re-evaluation recovers resolution: a deferred Friday decision is re-evaluated on Monday with the target day back in MET range. Figure E.3 (Appendix E) shows the full pattern.

Deferral to the K-th next working day crosses a variable number of *calendar* days, not just K working days, because the vehicle is driven on weekends. The energy required at the eventual charging event is therefore proportional to the calendar distance d_K, and feasibility requires the cumulative depletion to stay above the SoC floor, Equation (4.15):

$$d_K \cdot e \leq \Delta s_{\text{now}} = \text{SoC}_{\text{now}} - \underline{s}. \tag{4.15}$$

On the first working day the buffer is the full $\Delta s = 52.2$ kWh; on later re-evaluations it is reduced by the intervening calendar-day depletion. Table 4.9 shows the maximum SoC-feasible K from a fresh 80% starting charge.

The window-capacity constraint of Section 4.4.2 also applies, binding dynamically as the calendar distance to the previous charge grows. For high e and low P, most notably the 100 km/day profile on the 7.4 kW wallbox, it is tighter than the SoC floor on mid-deferral

Table 4.9: Maximum SoC-feasible deferral K from a fresh 80% SoC for each driver profile and session day in UC2 under Framework 1, determined by the floor constraint $d_K \cdot e \leq 52.2$ kWh.

Driver (km/day)	Monday	Tuesday	Wednesday	Thursday	Friday
25	$N \leq 3$	$N \leq 3$	$N \leq 3$	$N \leq 3$	$N \leq 3$
50	$N \leq 3$	$N \leq 3$	$N \leq 3$	$N \leq 3$	$N \leq 3$
75	$N \leq 3$	$N \leq 3$	$N \leq 2$	$N \leq 1$	$N \leq 1$
100	$N \leq 2$	$N \leq 2$	$N \leq 2$	$N \leq 1$	$N = 0$

steps and forces earlier charging than Table 4.9 alone would suggest. Both constraints are enforced jointly in the feasibility check $\mathcal{F}$ at every re-evaluation.

Application under Framework 2

Under Framework 2, the driver declares each Monday at 07:00 CET that the vehicle should reach 80 % by the following Monday 07:00. The intervening week contains five working-day connection windows, each 08:00 to 17:00 CET, totalling 45 connection-hours. Weekend days contribute none but consume battery: the vehicle depletes by $2 \cdot e$ kWh between Friday's disconnect and Monday's plug-in, so forward-feasibility must absorb this unbroken weekend depletion within $E_{\text{rem}}(t)$. The 45-slot pool $\mathcal{P}(t)$ and its weekly contraction are shown in Figure E.4 (Appendix E).

UC2 under Framework 2 is materially tighter than UC1, as both the weekly connection-hours and the depletion-coverage flexibility are reduced. Across the 45 connection-hours, the reference 50 km/day profile on a 7.4 kW wallbox needs roughly 9.5 charging-hours (about 35 of slack); the 100 km/day profile needs about 19, leaving roughly 26, positive but materially less flexible than UC1. The 100 km/day $\times$ 7.4 kW corner is where forward-feasibility binding becomes a real concern: by Friday, with one working day left and weekend depletion still ahead, $E_{\text{rem}}(t)$ may approach $C_{\text{rem}}(t)$ closely enough to leave little deferral flexibility.

Session Definition and Annual Scaling

The evaluation iterates over every calendar day in the out-of-sample period 2025-01-01 to 2026-05-01. Weekend days deplete the battery silently; working days trigger the algorithm. Annual costs follow (4.5), with the period length Y measured in working years on a 251-day calendar that accounts for Swedish public holidays. The use-case-specific assumptions for UC2 are listed in Table E.7 (Appendix E).

Chapter 5

Results

This chapter reports the empirical results of the three-stage pipeline developed in Chapter 4. Section 5.1 evaluates the upstream input signal forecasts that feed the price model. Section 5.2 reports the spot price forecasting results and addresses RQ1. Section 5.3 evaluates the EV smart charging strategies built on those forecasts and addresses RQ2. All out-of-sample results in Sections 5.2 and 5.3 cover the period 2025-01-01 to 2026-05-01; the input signal evaluation in Section 5.1 covers the wider window 2023-01-01 to 2026-05-01, from which the price-model OOS span is drawn. Lead-band definitions are shared across Sections 5.1 and 5.2 and are introduced at the start of Section 5.1.

5.1 Input Signals

This section reports the performance of the three upstream forecasting models: wind power production, solar power production, and electricity consumption, that provide the supply- and demand-side fundamentals to the spot price model in Section 5.2. For each signal, three elements are reported: a point-forecast accuracy table broken down by ML model family and lead band, a set of exemplary forecasts with quantile intervals overlaid on observed values, and a SHAP-based feature attribution table identifying the inputs the selected model leans on most heavily. Throughout, the accuracy tables report nMAE, MAE, nRMSE, RMSE, and R^2 per model family and lead band, with absolute errors in MWh and normalised metrics computed against the mean observed quantity over the test period.

Performance is evaluated across fourteen lead bands defined in Table 5.1. The MET-resolution range (1–58 h) is divided into nine 6 h bands, with one (h=49–58) absorbing the boundary at the MET/GFS handoff. The GFS-resolution range (60–168 h) consists of one 12 h bridge band followed by four 24 h day-aligned bands. Lead-range labels (e.g. "h=19–24") are used as column headers throughout the chapter; the day-window context is provided as an annotation in Table 5.1 and need not be repeated in subsequent tables. The fourth column indicates which bands are active in the price-model evaluation in Section 5.2; the price model does not operate below $h = 18$, so the first three bands are dashed there.

Table 5.1: Lead band definitions used throughout Sections 5.1 and 5.2. MET-resolution bands (1–58 h) are 6 h wide, with the final band absorbing the boundary near the MET/GFS handoff. GFS-resolution bands (60–168 h) consist of one 12 h bridge band followed by four 24 h day-aligned bands.

Lead range	Width	Day window	Active in 5.2?
MET-resolution bands			
h=1–6	6 h	Day 0	—
h=7–12	6 h	Day 0	—
h=13–18	6 h	Day 0	h=18 only
h=19–24	6 h	Day 0	✓
h=25–30	6 h	Day +1	✓
h=31–36	6 h	Day +1	✓
h=37–42	6 h	Day +1	✓
h=43–48	6 h	Day +1	✓
h=49–58	10 h	Day +2 (MET)	✓
GFS-resolution bands			
h=60–72	12 h	Day +2 (GFS)	✓
h=73–96	24 h	Day +3	✓
h=97–120	24 h	Day +4	✓
h=121–144	24 h	Day +5	✓
h=145–168	24 h	Day +6	✓

5.1.1 Wind Power Prediction

Point Forecast Accuracy by Lead Band

Table 5.2 reports point forecast accuracy for hourly SE3 wind power production across the four model families. The best-performing family shifts with horizon: XGBoost leads at the shortest leads (h=1–6), LSTM across the rest of the MET range (h=7–58), and GRU through the early GFS range (h=60–120); at the longest horizons GRU and Random Forest split the absolute and squared-error metrics. The three leading families (XGBoost, LSTM, GRU) stay within one to two nMAE percentage points of each other and converge to within roughly 1.3 points by h=145–168, with Random Forest trailing by a further two to three points in the MET range. Forecast skill falls sharply across the GFS range, the best R^2 dropping from 0.72 at h=60–72 to 0.10 at h=145–168.

Table 5.2: Wind production point forecast accuracy per model and lead band. Normalized metrics are presented in percentage and absolute metrics in MWh. Best in bold.

	Random Forest					XGBoost					GRU					LSTM				
Band	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2
MET																				
1–6	22.47	284.30	30.68	388.30	0.81	**19.40**	**245.41**	**26.20**	**331.65**	**0.87**	22.56	285.40	31.01	392.60	0.81	19.89	251.60	27.23	344.70	0.85
7–12	22.66	286.20	30.83	390.20	0.81	19.70	248.79	26.60	336.63	0.86	20.47	258.50	27.75	351.20	0.85	**19.23**	**242.90**	**26.28**	**332.60**	**0.86**
13–18	23.12	292.20	31.15	393.80	0.81	20.20	255.25	27.10	342.64	0.86	20.79	262.70	27.97	353.70	0.85	**19.69**	**248.80**	**26.84**	**339.40**	**0.86**
19–24	23.34	295.20	31.41	396.40	0.81	20.50	259.23	27.40	345.80	**0.85**	20.89	264.20	28.14	355.10	0.84	**19.90**	**251.70**	**27.15**	**342.70**	**0.85**
25–30	24.12	304.20	32.22	406.70	0.80	21.50	271.17	28.60	361.01	0.84	21.38	269.70	28.54	360.30	0.84	**20.53**	**258.90**	**27.81**	**351.10**	**0.85**
31–36	24.85	313.80	33.27	419.90	0.78	22.60	285.39	30.10	379.86	0.82	22.32	281.80	29.88	377.10	0.82	**21.40**	**270.20**	**28.95**	**365.40**	**0.83**
37–42	25.99	329.40	34.72	440.00	0.76	23.70	300.43	31.60	400.43	0.80	23.22	294.40	31.16	394.90	0.81	**22.10**	**280.10**	**30.14**	**381.90**	**0.82**
43–48	26.95	341.90	35.90	455.50	0.74	24.90	315.92	33.20	421.23	0.78	24.39	309.40	32.66	414.40	0.79	**23.31**	**295.80**	**31.62**	**401.20**	**0.80**
49–58	28.09	357.20	37.00	470.10	0.73	26.40	335.75	35.10	445.94	0.76	25.47	323.90	34.29	435.70	0.77	**24.97**	**317.60**	**34.00**	**432.00**	**0.77**
GFS																				
60–72	30.56	390.30	40.48	518.40	0.67	30.20	385.69	39.90	510.93	0.68	**27.88**	**356.00**	**37.02**	**474.00**	**0.72**	28.14	359.40	37.52	480.40	**0.72**
73–96	36.08	460.40	47.41	605.70	0.55	35.90	458.07	47.30	604.24	0.55	**35.28**	**450.10**	**46.51**	**594.20**	**0.56**	36.07	460.20	48.17	615.30	0.53
97–120	43.47	554.70	56.36	718.00	**0.37**	43.50	555.10	56.80	723.66	0.36	**43.11**	**550.10**	**56.29**	**717.10**	**0.37**	44.31	565.40	58.68	747.60	0.31
121–144	50.44	643.40	**63.56**	**811.00**	**0.19**	50.50	644.20	64.10	817.95	0.18	**49.73**	**634.40**	63.80	814.10	0.18	51.22	653.40	66.24	845.30	0.12
145–168	54.30	691.70	**67.39**	**858.20**	**0.10**	55.20	703.20	69.00	878.68	0.05	**53.90**	**686.70**	68.98	878.40	0.05	54.72	697.10	70.41	896.60	0.01

Note:

Exemplary Forecasts with Quantile Intervals

To visualise the probabilistic forecasts, two 168-hour windows are presented: Figure 5.1 (summer; prediction made 2025-07-09 12:00) and Figure 5.2 (winter; prediction made 2026-01-15 12:00), selected to span contrasting seasonal regimes.

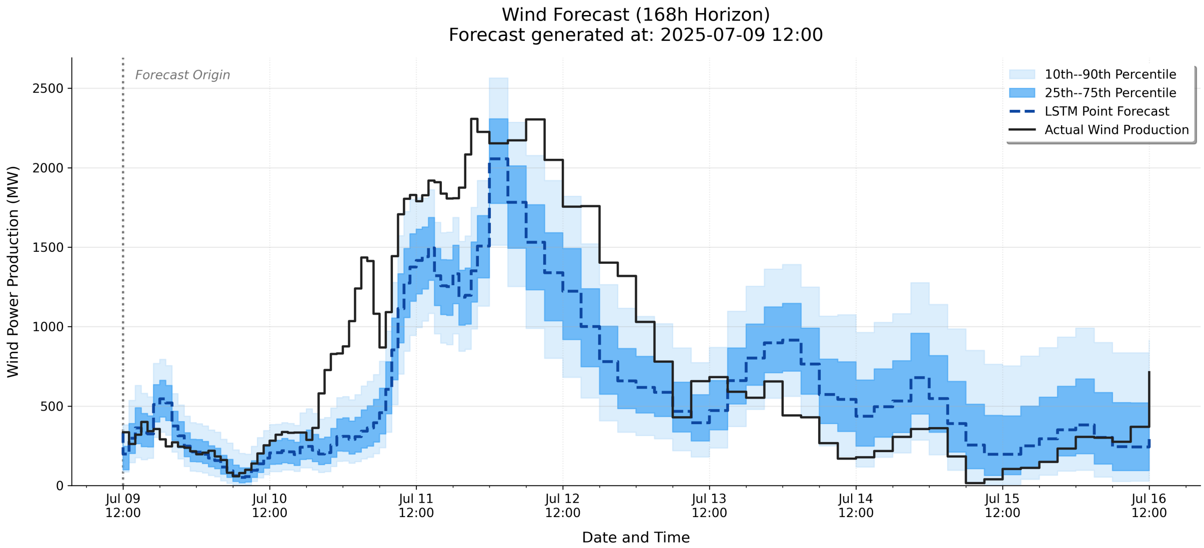

Figure 5.1: Wind power forecast summer example. Prediction made at 2025-07-09 12:00 for the upcoming week showing point forecast and quantile intervals.

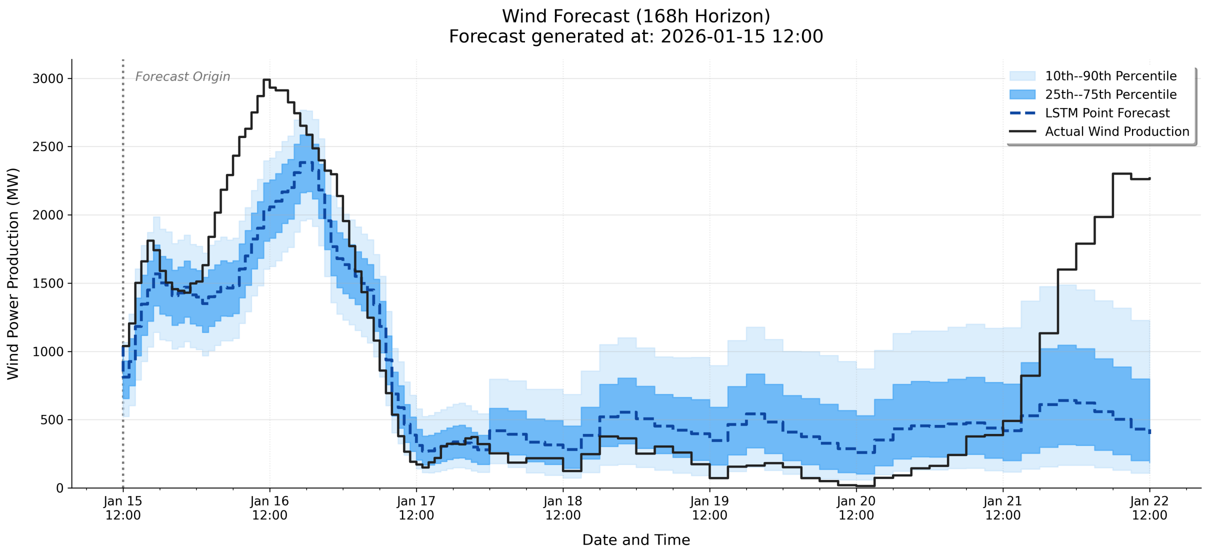

Figure 5.2: Wind power forecast winter example. Prediction made at 2026-01-15 12:00 for the upcoming week showing point forecast and quantile intervals.

Within the MET-resolution portion of the horizon (leads up to 58 hours), the point forecast tracks the major observed production peaks in both examples. Beyond the MET/GFS handoff, the point forecast becomes progressively smoother and oscillates around a seasonal baseline, and several observed high-wind episodes in the later part of Figure 5.2 are not reflected in the point forecast. The 10–90 % quantile band widens mostly with lead time and tends to widen further during observed peak-production episodes, even at long leads.

Feature Attribution

XGBoost was selected for SHAP analysis as the model family for which exact Shapley contributions are tractable. Table 5.3 presents a heatmap of mean absolute SHAP values (MWh) for the most impactful wind features from h=24 to h=168 (top 25 of 160 inputs). Two naming conventions appear: `ws_*` are raw NWP wind-speed inputs, while `mw_*` are megawatt estimates from the physics pipeline of Section 4.2.3.

The attribution shifts cleanly with horizon. At the shortest leads the physics-pipeline raw MW estimate (`mw_raw`), current wind-speed aggregates, and mean gust excess dominate; `mw_raw` declines steeply with lead time, and the gust feature is unavailable beyond the MET range as GFS provides no gust data. At longer horizons calendar and seasonality features take over, with `month_cos` the single largest contributor at the longest lead and `month_sin` and `day_of_week` following the same growth. Cluster-level physics and raw-wind-speed features appear from rank 12 onward, each individually small but collectively substantial across the 135 omitted inputs.

Table 5.3: Heatmap of SHAP feature importance for Wind Production across forecast horizons (Impact in MWh). Features are ordered by their total aggregate importance. Top 25 features shown (out of 160).

Feature Name	**h=24**	**h=48**	**h=72**	**h=96**	**h=120**	**h=144**	**h=168**
mw_raw	210.24	149.66	231.94	161.89	61.35	39.39	6.85
month_cos	49.95	61.47	96.45	107.14	144.80	166.45	194.23
ws_p90	178.45	197.87	52.69	46.21	47.56	31.82	19.29
dir_sin	43.80	55.16	65.48	81.61	89.44	84.44	56.83
ws_mean	31.77	14.26	137.56	145.52	74.31	27.82	32.28
hour_cos	68.84	78.56	61.56	49.72	44.69	47.56	31.72
dir_cos	34.03	47.52	20.44	23.32	23.70	39.12	36.29
frac_clusters_strong	3.20	51.29	20.12	20.62	62.49	58.73	7.04
hour_sin	46.73	34.30	15.36	27.61	24.32	25.43	30.66
month_sin	20.81	17.21	19.53	15.55	28.47	47.77	51.19
ws_spatial_std	51.87	44.30	9.10	13.05	24.38	25.34	23.81
mw_loc___gummarsen	5.83	19.49	13.79	13.98	22.04	32.23	39.44
day_of_week	7.85	10.98	13.51	21.51	22.38	33.14	28.18
mw_loc___kyrkberget	22.23	16.40	25.89	16.33	17.78	9.08	22.24
ws_loc___kyrkberget	19.41	19.13	17.79	18.54	17.67	23.74	9.85
mw_loc___vindpark_rnsliden	36.93	26.76	13.79	11.78	7.38	10.19	15.84
ws_loc___vindpark_lursng	16.50	13.14	8.61	9.11	30.56	15.33	28.37
ws_loc___varsvik	5.02	11.06	18.71	14.09	21.07	26.82	23.64
mw_loc___vindpark_lursng	16.85	13.85	11.07	8.86	24.06	29.85	15.75
mw_loc___nsudden	4.23	8.42	10.17	18.60	35.88	28.28	14.64
mw_loc___vettsenfinnberget	15.07	6.07	23.62	20.25	15.02	23.19	15.23
mw_loc___varsvik	6.20	9.75	20.23	28.42	12.19	18.29	21.16
mw_loc___svartns	30.42	14.66	8.05	9.73	18.52	17.35	16.32
ws_loc___hedbodberget	3.86	24.16	22.13	8.99	13.79	18.15	20.53
gust_excess_mean	34.82	50.05	—	—	—	—	—

Impact (MWh)

0 240

5.1.2 Solar Power Prediction

Point Forecast Accuracy by Lead Band

Table 5.4 reports point forecast accuracy for hourly SE3 solar power production. Solar production is structurally nighttime-zero, so all metrics are computed over daylight hours only, avoiding artificially inflated R^2 from trivially-correct nighttime predictions. The tree-based families dominate the MET range: Random Forest and XGBoost are essentially tied (within 0.1 nMAE points), while GRU and LSTM trail by 5–7%. At the MET/GFS handoff the spread narrows to within roughly one point, and from h=97 onward LSTM takes the lead in every GFS band with GRU close behind, the tree-based models falling back by 2–3 points. The best-model R^2 declines from 0.95 at h=1–6 to 0.84 at h=145–168.

Table 5.4: Solar production point forecast accuracy per model and lead band. Normalized metrics are presented in percentage and absolute metrics in MWh. Best in bold.

	Random Forest					XGBoost					GRU					LSTM				
Band	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2
MET																				
1–6	19.48	39.55	44.81	90.97	0.95	**19.48**	**39.54**	**44.60**	**90.55**	**0.95**	26.65	54.10	57.80	117.34	0.91	26.88	54.56	59.04	119.86	0.91
7–12	20.01	41.28	45.71	94.31	0.94	**19.86**	**40.97**	**45.11**	**93.07**	**0.94**	26.39	54.46	57.10	117.82	0.91	26.89	55.48	58.70	121.12	0.91
13–18	**19.99**	**41.08**	**46.16**	**94.85**	**0.94**	20.26	41.63	46.30	95.14	0.94	27.02	55.52	58.25	119.69	0.91	26.74	54.94	57.66	118.48	0.91
19–24	20.90	42.74	48.47	99.13	0.94	**20.81**	**42.55**	**48.09**	**98.35**	**0.94**	27.32	55.86	58.86	120.38	0.91	27.02	55.26	58.67	119.99	0.91
25–30	**20.88**	**42.62**	**48.21**	**98.37**	**0.94**	20.93	42.70	48.25	98.45	0.94	26.98	55.05	57.96	118.28	0.91	26.74	54.56	58.27	118.90	0.91
31–36	21.97	45.64	50.31	104.53	0.93	**21.90**	**45.50**	**50.12**	**104.14**	**0.93**	27.63	57.41	59.19	122.99	0.90	27.35	56.83	59.50	123.64	0.90
37–42	22.41	46.59	51.14	106.34	0.93	**22.06**	**45.86**	**50.30**	**104.57**	**0.93**	27.52	57.22	59.41	123.52	0.90	28.06	58.33	60.43	125.64	0.90
43–48	23.55	48.83	53.51	110.95	0.92	**23.37**	**48.46**	**53.32**	**110.56**	**0.92**	28.55	59.20	61.66	127.86	0.90	28.46	59.02	61.62	127.77	0.90
49–58	**24.20**	**50.39**	**55.29**	**115.15**	**0.92**	24.22	50.44	55.52	115.62	0.91	28.71	59.79	61.72	128.53	0.89	29.20	60.81	62.53	130.22	0.89
GFS																				
60–72	28.99	60.25	66.50	138.22	0.88	**28.13**	**58.46**	64.20	133.44	**0.89**	28.69	59.62	**63.37**	**131.71**	**0.89**	28.92	60.11	63.78	132.57	**0.89**
73–96	30.95	64.54	70.16	146.29	0.86	**30.38**	**63.35**	68.70	143.24	**0.87**	30.71	64.03	**67.85**	**141.46**	**0.87**	30.87	64.36	68.38	142.58	**0.87**
97–120	33.83	71.11	76.24	160.27	0.84	33.67	70.77	76.44	160.69	0.84	33.23	69.85	73.06	153.58	0.85	**33.11**	**69.61**	**72.67**	**152.77**	**0.85**
121–144	35.82	75.83	80.34	170.09	0.82	35.66	75.51	80.75	170.95	0.82	34.36	72.74	**75.04**	**158.88**	**0.84**	**34.22**	**72.44**	75.10	159.00	**0.84**
145–168	37.74	80.06	84.49	179.27	0.80	37.71	80.01	85.38	181.15	0.80	35.37	75.04	**76.82**	**162.99**	**0.84**	**35.33**	**74.97**	76.86	163.06	**0.84**

Note: Normalized errors for solar production are systematically higher than consumption due to the high variability of cloud cover and daylight availability.

Exemplary Forecasts with Quantile Intervals

To visualise the probabilistic forecasts, two 168-hour windows are presented: Figure 5.3 (summer; prediction made 2025-07-15 12:00) and Figure 5.4 (winter; prediction made 2026-01-15 12:00), selected to span contrasting seasonal regimes, peak production at roughly 1 500–1 800 MWh in summer against restricted production peaking below 150 MWh in winter.

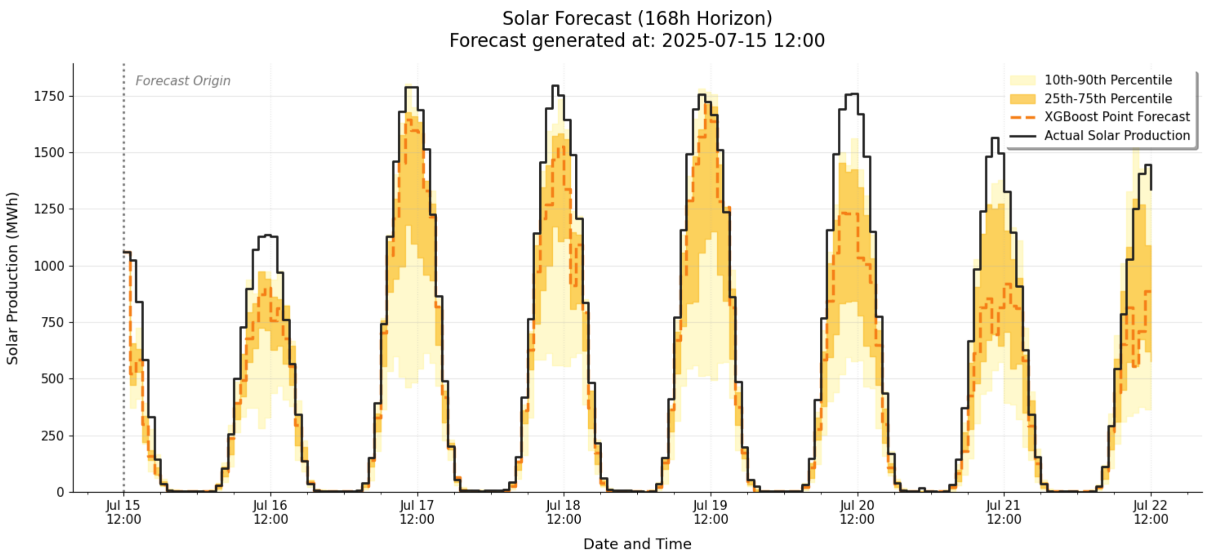

Figure 5.3: Solar power forecast summer example. Prediction made at 2025-07-15 12:00 for the upcoming week showing point forecast and quantile intervals.

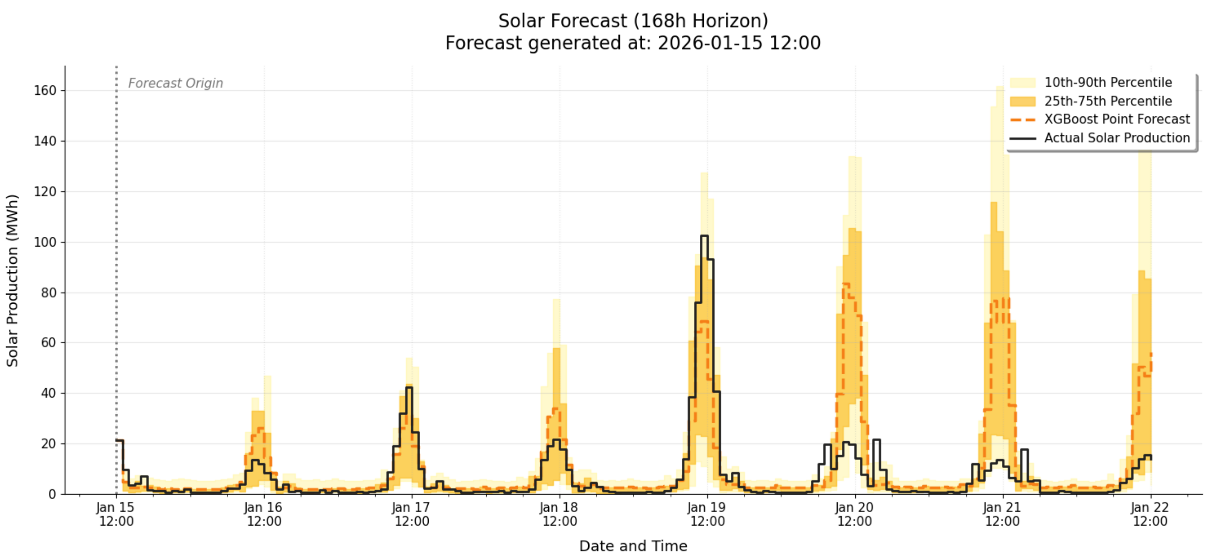

Figure 5.4: Solar power forecast winter example. Prediction made at 2026-01-15 12:00 for the upcoming week showing point forecast and quantile intervals.

The two windows show qualitatively different behaviour. In the summer example, the generation profile exhibits significant volatility during the morning and evening hours and the point forecast tracks the daily rise and fall through the full 168-hour horizon, but it systematically underestimates the peak daytime production from around Jul 18 onward, with observed peaks reaching 1 700-1 800 MWh while the point forecast peaks at 1 200-1 400 MWh. The actual values nonetheless remain within the 10-90 % quantile band across almost the entire window.

In the winter example, where production peaks are an order of magnitude smaller and more irregular, the model exhibits the opposite tendency: the point forecast overestimates production on the early days of the horizon (Jan 16–18), where observed peaks are 20-50 MWh but the predicted envelope reaches 30-50 MWh, and the 10-90 % quantile band widens markedly on the later days (Jan 20-22). Across both horizons, the model retains accurate timing of the daily rise and fall throughout, even as the point forecast smooths and the quantile bands expand at long leads.

Feature Attribution

XGBoost was selected for SHAP analysis as the strongest tree-based model across the short- and medium-term range and the family for which exact Shapley contributions are tractable. Table 5.5 presents a heatmap of mean absolute SHAP values (MWh) for the most impactful solar features from h=24 to h=168.

Historical window means dominate the profile across the entire horizon, structurally different from the wind case, with the 3-year window mean the single largest contributor at every lead and the other multi-year and combined means adding a steady climatological share. The meteorological forecast features (solar radiation, cloud cover) behave as horizon-decaying corrections on top of this baseline, contributing most at the shortest lead and declining steadily; day of year shows the opposite trend, growing as the meteorological signal weakens. Recent production lags and hour of day are negligible throughout.

Table 5.5: Heatmap of SHAP feature importance for Solar Production across forecast horizons (Impact in MWh). Features are ordered by their total aggregate importance.

Feature Name	h=24	h=48	h=72	h=96	h=120	h=144	h=168
3-Year Hist. Window Mean	98.05	120.09	67.25	69.32	72.62	65.53	60.59
4-Year Hist. Window Mean	34.87	9.14	57.39	63.28	36.25	41.37	59.04
1-Year Hist. Window Mean	29.72	30.45	45.77	36.17	37.43	37.61	41.57
Combined Hist. Mean	26.44	28.82	25.92	26.85	48.83	47.56	35.58
SE3 Solar Radiation Forecast	36.12	31.22	19.32	22.69	15.71	13.28	10.46
SE3 Cloud Cover Forecast	21.43	21.17	24.75	16.27	14.17	11.39	10.76
2-Year Hist. Window Mean	9.90	7.09	11.49	10.26	7.78	8.53	10.50
Day of Year	3.39	4.35	4.67	4.50	7.71	8.85	11.42
Combined Solar Profile	4.84	5.18	5.53	3.97	5.00	3.55	3.73
Actual Prod. (t-48)	2.05	2.19	2.58	4.16	5.86	6.62	5.93
Actual Prod. (t-72)	1.50	1.47	2.50	4.63	0.00	0.00	0.00
Actual Prod. (t-96)	1.43	1.67	2.97	0.00	0.00	0.00	0.00
Actual Prod. (t-120)	1.76	3.46	0.00	0.00	0.00	0.00	0.00
Hour of Day	0.35	0.99	0.79	0.56	0.37	0.55	0.46
Last Week Hourly Avg.	1.47	1.84	0.00	0.00	0.00	0.00	0.00
Actual Prod. (t-144)	2.23	0.00	0.00	0.00	0.00	0.00	0.00

Impact (MWh)

0 150

5.1.3 Consumption Prediction

Temperature–Demand Relationship

The temperature-demand regression yields a convex relationship between ambient temperature and electricity consumption (Equation 5.1, Figure 5.5). The negative linear coefficient and positive quadratic coefficient together imply a minimum baseline consumption at approximately 26.0°C, above and below which demand rises through cooling and heating loads respectively.

$$\text{Consumption} = 11041.6144 - 296.0429 \cdot \text{Temp} + 5.7026 \cdot \text{Temp}^2 \tag{5.1}$$

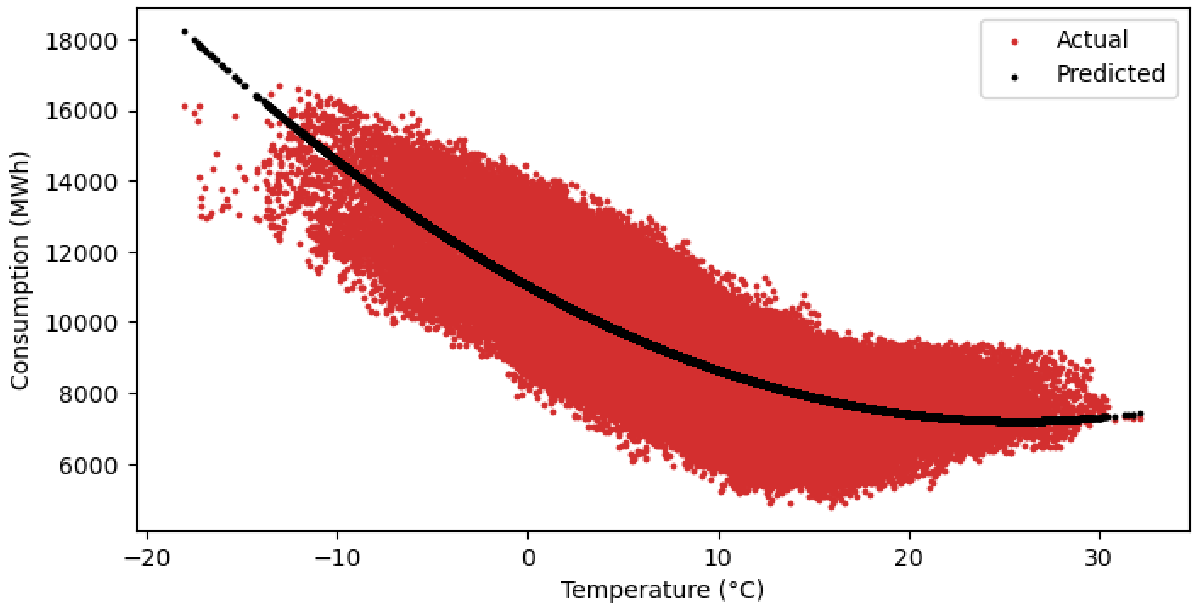

Figure 5.5: Relationship between temperature and consumption. Red dots represent the observed historical temperature–consumption data pairs; the solid black line shows the fitted quadratic regression model.

Point Forecast Accuracy by Lead Band

Table 5.6 reports point forecast accuracy for hourly SE3 electricity consumption. XGBoost leads in every band from h=1–6 through h=121–144 (nMAE rising smoothly from 2.68 % to 4.66 %, R^2 at or above 0.88 throughout), with GRU edging it only fractionally at h=145–168 and Random Forest trailing by 0.2–0.4 points across the horizon; the sequence models sit 0.3–1.0 points behind the tree-based families through most of the MET range and converge only at the longest leads. The absolute error level is dramatically lower than for either renewable signal: the worst band of the worst model (LSTM at h=121–144, nMAE 5.22 %) is still tighter than the best band of any wind or solar model.

Table 5.6: Electricity consumption point forecast accuracy per model and lead band. Normalized metrics are presented in percentage and absolute metrics in MWh. Best in bold.

	Random Forest					XGBoost					GRU					LSTM				
Band	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2
MET																				
1–6	2.82	266.0	3.79	357.0	0.97	**2.68**	**252.6**	**3.58**	**337.3**	**0.97**	2.99	281.6	3.90	367.5	0.97	3.09	290.8	4.03	379.2	0.96
7–12	3.18	299.2	4.30	405.4	0.96	**2.78**	**261.7**	**3.73**	**351.2**	**0.97**	3.65	343.4	4.75	448.0	0.95	3.65	343.8	4.75	447.3	0.95
13–18	3.28	309.1	4.41	416.0	0.96	**2.85**	**268.8**	**3.84**	**361.8**	**0.97**	3.79	357.4	4.93	464.6	0.95	3.74	352.7	4.86	458.1	0.95
19–24	3.21	302.2	4.32	406.6	0.96	**2.95**	**277.9**	**3.99**	**376.2**	**0.96**	3.49	328.8	4.59	432.6	0.95	3.52	331.6	4.62	435.6	0.95
25–30	3.43	323.4	4.60	433.5	0.95	**3.19**	**300.5**	**4.27**	**401.9**	**0.96**	3.87	364.5	5.08	478.7	0.94	3.93	370.1	5.15	484.7	0.94
31–36	3.53	332.3	4.76	448.3	0.95	**3.18**	**299.7**	**4.28**	**402.8**	**0.96**	4.14	390.1	5.44	512.2	0.93	4.18	393.8	5.47	515.4	0.93
37–42	3.58	336.7	4.83	454.9	0.95	**3.25**	**306.1**	**4.38**	**412.5**	**0.96**	4.17	392.7	5.45	513.4	0.93	4.34	408.4	5.64	531.2	0.93
43–48	3.57	336.3	4.83	454.2	0.95	**3.34**	**314.1**	**4.50**	**423.6**	**0.96**	4.11	387.2	5.38	506.6	0.94	4.12	387.4	5.36	504.2	0.94
49–58	3.81	358.4	5.11	480.4	0.94	**3.57**	**335.2**	**4.79**	**450.2**	**0.95**	4.47	419.6	5.84	549.1	0.92	4.43	416.0	5.77	541.8	0.93
GFS																				
60–72	3.77	354.4	5.09	479.3	0.94	**3.56**	**335.4**	**4.81**	**452.9**	**0.95**	4.29	403.9	5.65	531.7	0.93	4.35	410.0	5.70	536.5	0.93
73–96	4.04	379.7	5.46	512.8	0.93	**3.87**	**363.5**	**5.25**	**493.3**	**0.94**	4.55	428.0	5.98	561.8	0.92	4.56	428.9	5.99	562.7	0.92
97–120	4.41	413.2	6.76	634.1	0.90	**4.27**	**400.7**	**6.61**	**620.4**	**0.91**	4.88	457.7	7.15	670.5	0.89	4.91	460.4	7.17	672.4	0.89
121–144	4.74	444.1	7.47	700.1	**0.88**	**4.66**	**436.8**	**7.39**	**692.5**	**0.88**	5.13	480.4	7.73	723.9	0.87	5.22	488.9	7.79	729.8	0.87
145–168	5.08	475.2	7.83	732.9	0.87	5.02	469.5	7.79	728.8	0.87	**4.96**	**464.1**	**7.62**	**713.7**	**0.88**	5.07	474.8	7.71	721.5	0.87

Note: Normalized errors for consumption are systematically lower than wind- and solar production due to significantly larger absolute load magnitudes and high cyclical predictability.

Exemplary Forecasts with Quantile Intervals

To visualise the probabilistic forecasts, two 168-hour windows are presented: Figure 5.6 (summer; prediction made 2025-07-15 12:00) and Figure 5.7 (winter; prediction made 2026-01-15 12:00), selected to span contrasting seasonal load regimes.

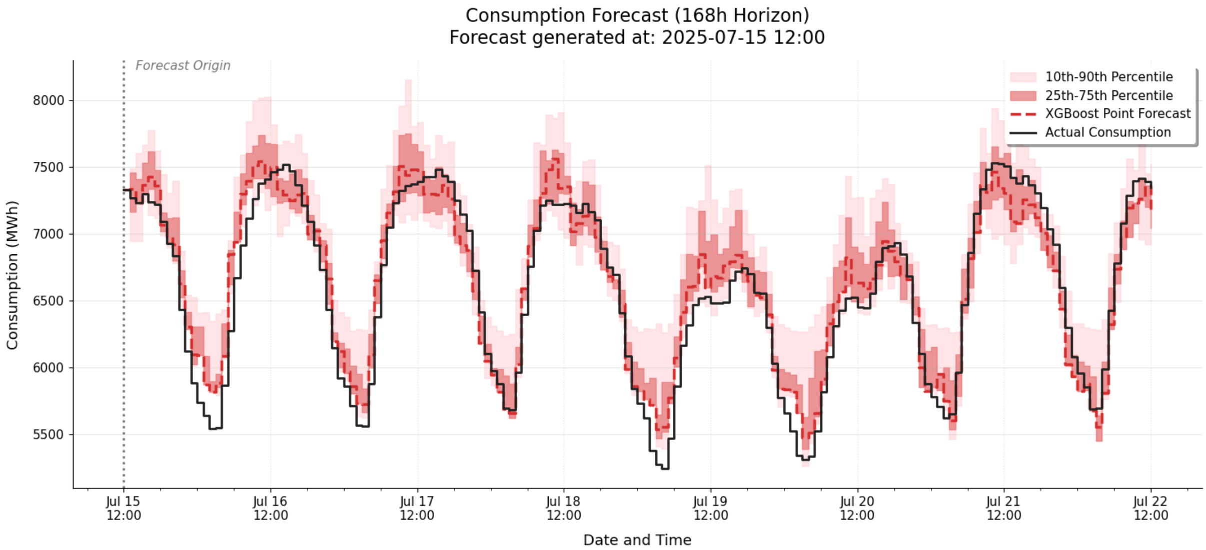

Figure 5.6: Consumption forecast summer example. Prediction made at 2025-07-15 12:00 for the upcoming week showing point forecast and quantile intervals.

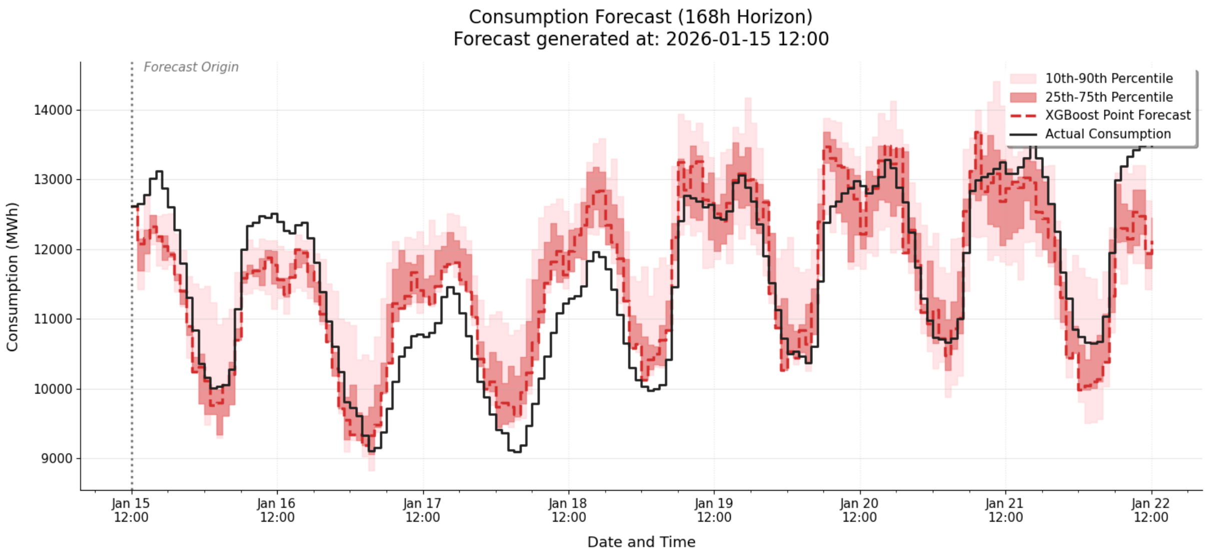

Figure 5.7: Consumption forecast winter example. Prediction made at 2026-01-15 12:00 for the upcoming week showing point forecast and quantile intervals.

The point forecast captures the dominant structural features of consumption across both seasons: the daily peak-trough cycle, the weekly pattern with reduced weekend demand (visible at Jul 19–20 in summer and Jan 17–18 in winter), and the absolute load level appropriate to the season. In the summer example, the forecast tracks the observed series closely throughout the 168-hour horizon, with a mild tendency to overestimate the daytime peaks of the early-week working days. In the winter example, the early part of the horizon (Jan 15–17) shows a more pronounced overestimation of daytime levels, after which the forecast aligns well with the observed series from Jan 18 onward.

The 10–90 % quantile band does not widen materially with lead time in either example. The model's stated uncertainty is broadly stable across the horizon, in contrast to the wind and solar cases. The dominant variation in band width is seasonal rather than temporal: winter bands are wider in absolute terms throughout, reflecting the larger absolute amplitude of winter consumption. The observed values remain within the 10–90 % band across most of both windows, with occasional excursions during the deepest weekend troughs.

Feature Attribution

XGBoost was selected for SHAP analysis as the best-performing family across the operational horizon and the model for which exact Shapley contributions are tractable. Table 5.7 presents a heatmap of mean absolute SHAP values (MWh) for the most impactful features from h=24 to h=168.

Two features dominate the profile, and their relative weight inverts across the horizon: the day-adjusted baseline is the single largest contributor from h=24 through h=120, then decays, while the temperature-adjusted consumption feature moves oppositely, dipping mid-horizon before rising sharply to overtake the baseline as the largest contributor by h=168. SE3 mean temperature and the historical and calendar features (combined and multi-year window means, day of year) carry a secondary but non-trivial share, while recent load lags are present only at the shortest leads and negligible at long horizons.

Table 5.7: Heatmap of SHAP feature importance across forecast horizons (Impact in MWh). Features are ordered by their total aggregate importance.

Feature Name	**h=24**	**h=48**	**h=72**	**h=96**	**h=120**	**h=144**	**h=168**
Day Adjusted baseline	1070.74	1198.05	1156.93	1079.52	839.36	631.46	267.68
Temp. Adjusted Consumption	509.90	194.34	116.37	57.84	74.63	887.20	1138.55
SE3 Mean Temperature	206.61	318.62	355.70	402.99	351.70	191.11	31.11
Combined Hist. Mean	9.56	53.16	238.49	265.70	291.00	136.86	46.20
Day of Year	69.67	83.02	90.72	86.64	89.34	115.44	114.58
4-Year Hist. Window Mean	63.48	87.16	89.29	99.57	99.33	69.71	92.44
3-Year Hist. Window Mean	39.47	53.36	61.20	51.29	88.58	98.77	143.81
Actual Load (t-48)	13.81	28.96	36.31	12.37	254.51	19.46	30.85
Hour of Day	44.47	73.70	66.30	83.71	62.47	21.13	15.88
Actual Weekly Mean	16.23	21.89	38.75	56.94	56.44	51.47	29.55
Actual Load (t-72)	37.14	37.64	10.33	52.96	0.00	0.00	0.00
2-Year Hist. Window Mean	16.07	14.15	18.59	16.49	20.12	19.18	25.86
1-Year Hist. Window Mean	11.37	13.36	14.74	19.45	17.06	21.52	31.44
Actual Load (t-120)	15.58	65.09	0.00	0.00	0.00	0.00	0.00
Actual Load (t-144)	65.13	0.00	0.00	0.00	0.00	0.00	0.00
Actual Load (t-96)	14.67	9.57	28.42	0.00	0.00	0.00	0.00

Impact (MWh)

0 1200

5.2 Electricity Price Prediction

This section reports the spot price forecasting results that address RQ1: how accurately can different classes of predictive models forecast SE3 electricity spot prices beyond the day-ahead horizon, and how does forecast performance vary across model type and temporal horizon. All results use the six-fold expanding-window cross-validation defined in Section 4.3.5 over the OOS period 2025-01-01 to 2026-05-01, the six folds' out-of-sample predictions concatenated into a single evaluation panel before metric computation. Lead bands match Table 5.1; the first three bands ($h = 1$–6, $h = 7$–12, $h = 13$–18) are inactive because the price model does not operate below $h = 18$, except that the third reports a single value at $h = 18$, following the publication of Nord Pool day-ahead prices at 13:00 CET.

The section follows the same overall structure as the input signals in Section 5.1: point forecast accuracy by lead band, exemplary forecasts with quantile intervals, and SHAP-based feature attribution. A final subsection identifies the forecast used as input to the EV charging evaluation in Section 5.3.

5.2.1 Point Forecast Accuracy by Lead Band

Table 5.8 reports point forecast accuracy for hourly SE3 day-ahead spot prices across the four model families.

Table 5.8: Electricity price point forecast accuracy per model and lead band. Normalized metrics are presented in percentage and absolute metrics in EUR/MWh. Best in bold.

	Random Forest					XGBoost					GRU					LSTM				
Band	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2	nM.	MAE	nR.	RMSE	R^2
MET																				
1–6	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7–12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
13–18	**22.70**	**7.64**	**31.32**	**10.53**	**0.835**	23.54	7.92	32.07	10.79	0.827	37.98	12.78	50.01	16.82	0.579	41.40	13.93	54.72	18.41	0.495
19–24	**28.98**	**12.79**	**44.27**	**19.55**	**0.725**	30.29	13.37	44.59	19.69	0.721	35.54	15.69	51.03	22.53	0.635	38.87	17.16	55.03	24.30	0.575
25–30	**36.79**	**17.20**	53.37	24.95	0.664	36.94	17.27	**52.16**	**24.38**	**0.679**	38.11	17.81	53.99	25.24	0.656	40.89	19.11	57.00	26.64	0.617
31–36	**36.94**	**19.44**	**54.46**	**28.66**	**0.630**	38.02	20.01	54.70	28.79	0.627	38.50	20.26	55.54	29.23	0.615	40.60	21.37	57.86	30.45	0.582
37–42	**36.68**	**19.15**	52.98	27.67	0.627	37.21	19.43	**52.27**	**27.29**	**0.637**	39.17	20.45	55.33	28.89	0.593	41.24	21.53	57.61	30.08	0.559
43–48	**38.53**	**20.10**	54.72	28.54	0.605	39.11	20.40	**54.00**	**28.17**	**0.615**	41.04	21.41	57.03	29.75	0.571	42.29	22.06	58.83	30.69	0.543
49–58	**40.17**	**20.72**	56.52	29.16	0.578	40.69	21.00	**56.15**	**28.97**	**0.583**	43.74	22.57	60.39	31.16	0.518	43.17	22.28	60.25	31.09	0.520
GFS																				
60–72	**41.66**	**22.11**	57.90	30.73	0.547	42.51	22.56	**57.47**	**30.50**	**0.554**	44.43	23.58	60.64	32.19	0.503	44.40	23.56	61.32	32.55	0.492
73–96	**44.44**	**23.47**	**61.28**	**32.36**	**0.491**	46.72	24.68	62.28	32.89	0.474	46.59	24.60	63.35	33.46	0.456	46.26	24.43	63.57	33.58	0.452
97–120	**47.89**	**25.47**	**65.38**	**34.77**	**0.416**	48.85	25.98	65.72	34.95	0.410	50.38	26.79	67.71	36.01	0.373	49.53	26.34	67.26	35.77	0.382
121–144	**50.33**	**26.82**	**67.95**	**36.21**	**0.368**	51.71	27.55	69.38	36.97	0.341	53.33	28.42	71.40	38.05	0.302	52.08	27.75	70.58	37.61	0.318
145–168	**52.29**	**27.72**	**70.10**	**37.16**	**0.333**	54.74	29.02	72.96	38.68	0.278	55.87	29.62	74.17	39.32	0.254	54.88	29.09	73.61	39.02	0.265

Note: Normalized errors are reported as percentages. — indicates that final results are not available for the corresponding model and lead band.

The tabular models clearly outperform the sequence models across the whole horizon: at h=13–18 the tree-based nMAE (Random Forest 22.70 %, XGBoost 23.54 %) sits 15–18 percentage points below GRU and LSTM, and Random Forest still leads by 2.6–3.6 points at h=145–168. Random Forest and XGBoost are practically tied, Random Forest shows the lowest nMAE and MAE in every band (margin typically below one point), while XGBoost wins the squared-error metrics from h=25–30 onward, the standard divergence between MAE- and RMSE-optimal tree ensembles. Accuracy degrades smoothly across the MET–GFS handoff rather than discontinuously, and the leading-model R^2 declines from 0.835 at h=13–18 to 0.27–0.33 at h=145–168, with the steepest losses in the late MET range and across the boundary into the early GFS range.

5.2.2 Exemplary Forecasts with Quantile Intervals

To visualise the probabilistic forecasts, two 168-hour windows are presented: Figure 5.8 (summer; prediction made 2025-04-15 00:00) and Figure 5.9 (winter; prediction made 2025-11-19 12:00), selected to span contrasting seasonal price regimes.

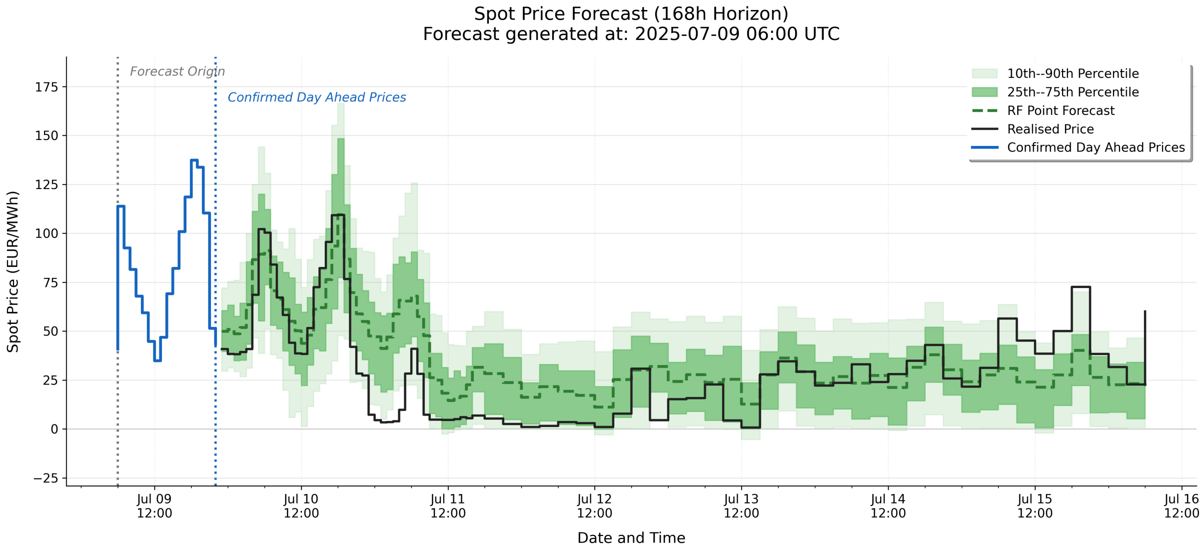

Figure 5.8: Electricity price forecast summer example. Prediction made at 2025-04-15 00:00 for the upcoming week showing point forecast and quantile intervals.

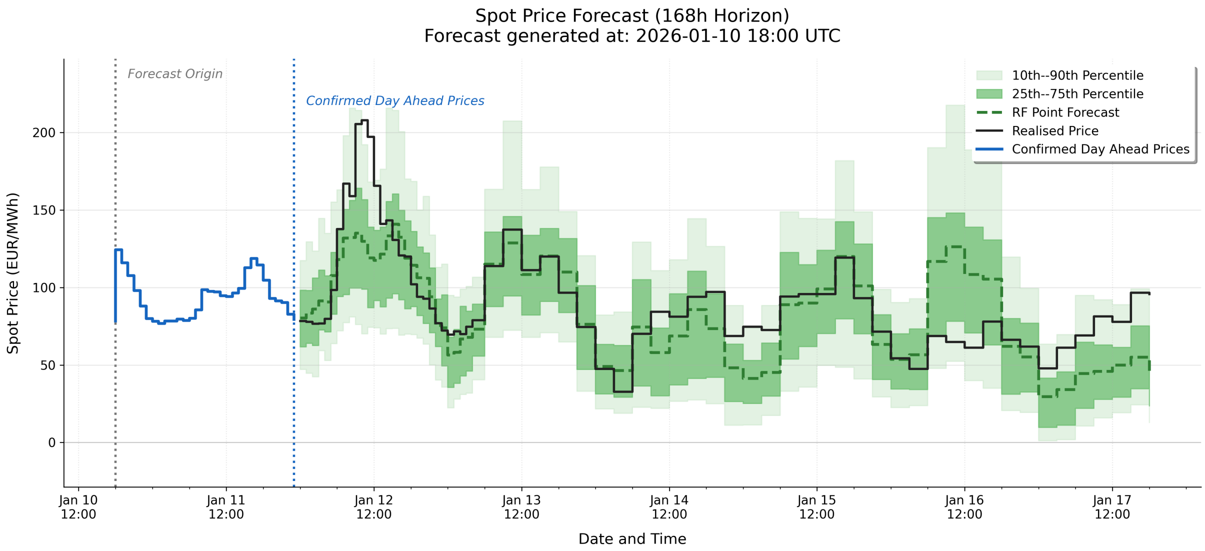

Figure 5.9: Electricity price forecast winter example. Prediction made at 2025-11-19 12:00 for the upcoming week showing point forecast and quantile intervals.

The diurnal price cycle is visible in both windows, with the point forecast capturing the morning and evening peaks against the overnight troughs throughout the 168-hour horizon. The point forecast tracks the broad price level in both seasons, though individual peak hours, particularly the evening peaks in the winter window, are not always reflected in the median prediction and instead surface in the upper quantile bound. The 10–90 % quantile band widens with lead time in both examples; the rate of widening is steeper in the winter window, where absolute price variability is larger throughout the forecast.

5.2.3 Feature Attribution

Random Forest was selected for SHAP analysis as the model family for which exact Shapley contributions are tractable on the price feature matrix. Table 5.9 presents a heatmap of the mean absolute SHAP values (in EUR/MWh) for the most impactful features from $h = 24$ to $h = 168$.

Table 5.9: Mean absolute SHAP values for the spot price prediction model across selected forecast horizons. Darker green shading indicates higher feature impact.

Feature Name	h=24	h=48	h=72	h=96	h=120	h=144	h=168
Cons_Pred_h017	7.256	12.385	9.126	7.344	7.160	2.927	4.427
day_of_week	3.891	3.974	3.225	3.512	4.080	3.854	3.245
Cons_q25_h017	1.331	1.540	2.345	4.985	4.520	5.261	4.034
Wind_q10_h017	3.290	3.592	3.662	3.035	3.958	2.817	2.956
price_same_h_1w	3.309	4.923	3.590	2.467	1.747	1.414	3.425
Cons_q75_h017	1.133	2.975	2.056	3.466	2.096	2.183	5.217
Price_Actual_-h001	10.298	5.516	0.985	0.326	0.706	0.847	0.440
Wind_q25_h017	0.740	2.921	4.631	3.627	1.794	1.320	1.178
Cons_q50_h017	0.234	0.561	3.853	2.977	2.497	2.153	1.680
Wind_q90_h017	1.735	1.068	1.457	0.851	1.645	1.810	4.326
Cons_q10_h017	0.646	0.896	0.591	0.558	2.769	5.036	2.311
price_mean_168h	0.410	1.003	2.225	2.507	2.351	2.047	1.817
Wind_q75_h017	0.568	1.423	2.142	3.175	2.548	1.732	0.605
reservoir_mwh	0.163	0.849	1.786	2.261	1.654	2.564	2.774
flow_total_mean_168h	0.766	0.754	0.924	1.582	2.361	3.199	2.420
Cons_q90_h017	1.104	1.059	2.988	1.331	1.200	2.100	2.221
price_same_h_2w	0.419	0.886	1.775	1.906	1.584	3.106	1.354
reservoir_pct_of_median	0.332	1.578	1.892	1.834	1.590	1.947	1.481
Wind_q50_h017	0.768	2.232	2.042	1.959	1.674	0.741	0.814
solar_seasonality	0.455	0.877	1.148	1.431	1.962	1.636	2.721
days_ahead	2.540	1.801	1.046	0.743	1.006	0.550	0.553
price_same_h_4w	0.331	1.005	0.949	2.457	1.987	1.045	0.456
hydro_gen_mean_168h	0.396	1.342	0.859	1.144	1.467	1.346	1.369
dow_sin	0.558	0.517	1.659	1.452	1.673	0.792	0.978
nuclear_gen_at_anchor	0.850	2.158	1.409	1.011	0.658	0.433	0.796
price_std_168h	0.479	1.171	1.015	1.002	1.425	0.976	0.988
other_gen_at_anchor	0.413	0.733	0.779	0.772	0.730	1.684	1.447
Price_Actual_-h005	1.959	1.349	0.407	0.641	0.851	0.555	0.619
flow_total_at_anchor	0.362	0.967	1.873	0.896	0.769	0.743	0.446
Solar_Actual_-h001	0.364	1.134	1.352	0.347	0.853	0.893	1.088
Wind_Pred_h017	0.577	1.145	0.517	0.994	1.014	1.107	0.471
Price_Actual_-h004	3.136	0.599	0.539	0.192	0.388	0.712	0.241
hydro_gen_mean_24h	0.141	0.388	1.152	1.450	0.568	0.720	1.105
Solar_Actual_-h012	3.648	0.276	0.172	0.418	0.282	0.285	0.327
Price_Actual_-h006	1.547	1.324	0.531	0.240	0.658	0.342	0.333
nuclear_gen_mean_168h	0.698	0.518	0.607	0.446	0.668	0.857	1.135
Price_Actual_-h003	1.156	0.631	0.848	0.603	0.454	0.523	0.512
price_spread_prev_day	1.018	0.632	0.328	0.450	0.645	0.789	0.803

Mean absolute SHAP value

0 12.5

The dominant pattern is a shift in feature reliance across the horizon. At h=24 the most recent confirmed price is the single largest contributor, followed by the upstream consumption point forecast, but the recent price lags decay sharply with lead time and are effectively zero from h=96 onward. From h=48 the upstream forecast group becomes the dominant input: the consumption point and quantile forecasts lead, with the upstream wind quantiles contributing persistently throughout. Calendar features (day of week, same-hour-of-week price lags) and ENTSO-E fundamentals (reservoir levels, hydro and nuclear generation, cross-border flows) carry a stable secondary share, gaining slightly in relative weight at the longest leads as the dominant features fade.

5.2.4 Selected Forecast for Downstream Evaluation

At horizons relevant to the EV charging evaluation in Section 5.3 ($h = 24$ to $h = 168$), Random Forest and XGBoost report essentially equivalent point accuracy. Random Forest holds a consistent but marginal lead in nMAE and MAE in every band; XGBoost holds a similarly consistent lead in nRMSE, RMSE, and R^2 from $h = 25$–30 onward. The split reflects the standard divergence between MAE-optimal and RMSE-optimal tree ensembles.

The full point-and-quantile forecast set from Random Forest is used as the price input to the EV charging evaluation in Section 5.3. Its consistent lead in median-error metrics across every band is aligned with the role of the point forecast in indexing candidate charging slots in the scoring rule of Section 4.4.4, where the median predicted price is the primary driver of slot selection and the quantile band enters as a secondary risk adjustment.

5.3 Electric Vehicle Smart Charging Evaluation

This section reports the EV smart charging evaluation that addresses RQ2: to what extent extended-horizon electricity price forecasts improve the cost efficiency of EV smart charging, and how outcomes vary across charging contexts that differ in connection window length and scheduling flexibility. All results use the Random Forest point and quantile forecasts selected in Section 5.2.4. The reference configuration is 50 km/day daily distance, 7.4 kW wallbox, risk parameter $\lambda = 0$. Results span the 2×2 grid of Framework 1 and Framework 2 crossed with Use Case 1 (home overnight) and Use Case 2 (workplace daytime). The OOS evaluation window is 2025-01-01 to 2026-05-01.

Two answers to RQ2 are reported throughout. S_{defer} is the marginal saving from forecast-driven operation on top of intra-day scheduling. $S_{\text{sched}} + S_{\text{defer}}$ is the cumulative saving relative to a do-nothing driver. The oracle gap G_{oracle} is reported alongside as a forecast-quality diagnostic.

5.3.1 Cost Decomposition for the Reference Configuration

This subsection reports the cost decomposition for the reference configuration across the four framework $\times$ use-case cells. Tables 5.10–5.13 give the per-cell annual costs and savings components; Figure 5.10 visualises the same numbers as a 2×2 panel of cost-ladder bar charts. The cross-framework comparison is reported separately at the end of the subsection.

Framework 1, UC1: Home Overnight

At the reference configuration, intra-day scheduling alone delivers $S_{\text{sched}} = 58.7$ EUR/year (32.5 % of naive cost) and forecast-driven deferral adds $S_{\text{defer}} = 55.6$ EUR/year (30.8 %), for a cumulative annual saving of 114.3 EUR/year. The algorithm cost (66.4 EUR/year) falls below the per-event oracle (72.9 EUR/year), producing a negative oracle gap $G_{\text{oracle}} = -6.5$ EUR/year. The algorithm charges on 150 of 365 days, deferring on the remaining 215.

Table 5.10: Framework 1 × UC1 reference configuration cost decomposition (50 km/day, 7.4 kW, $\lambda = 0$). Annual costs in EUR/year; savings components in EUR/year and as percentage of the naive baseline. Charging events per year are endogenous.

Quantity	EUR/year	% of naive	Notes
Naive baseline cost (C_{naive})	180.7	100.0	
Smart baseline cost (C_0)	122.0	67.5	
Algorithm cost (C_{K^*})	66.4	36.8	
Oracle cost (C_{oracle})	72.9	40.4	
Scheduling saving (S_{sched})	58.7	32.5	Forecast-free
Deferral saving (S_{defer})	55.6	30.8	Marginal answer to RQ2
Cumulative ($S_{\text{sched}} + S_{\text{defer}}$)	114.3	63.2	Total vs. naive
Oracle gap (G_{oracle})	-6.5	-3.6	Headroom diagnostic
Charging events / year (algorithm)	150		Endogenous
Charging events / year (naive)	365		

Framework 1, UC2: Workplace Daytime

Naive cost in UC2 is 289.8 EUR/year, 60 % above UC1's 180.7 EUR/year. Intra-day scheduling extracts $S_{\text{sched}} = 118.3$ EUR/year (40.8 %), roughly twice the corresponding UC1 figure. The deferral saving $S_{\text{defer}} = 49.6$ EUR/year is close to UC1 in absolute terms but a smaller fraction of the naive baseline (17.1 % vs 30.8 %). G_{oracle} is tight at 2.0 EUR/year (0.7 %).

Table 5.11: Framework 1 × UC2 reference configuration cost decomposition. Same row structure as Table 5.10.

Quantity	EUR/year	% of naive	Notes
Naive baseline cost (C_{naive})	289.8	100.0	
Smart baseline cost (C_0)	171.5	59.2	
Algorithm cost (C_{K^*})	121.9	42.1	
Oracle cost (C_{oracle})	119.8	41.4	
Scheduling saving (S_{sched})	118.3	40.8	Forecast-free
Deferral saving (S_{defer})	49.6	17.1	Marginal answer to RQ2
Cumulative ($S_{\text{sched}} + S_{\text{defer}}$)	167.9	57.9	Total vs. naive
Oracle gap (G_{oracle})	2.0	0.7	Headroom diagnostic
Charging events / year (algorithm)	139		Endogenous
Charging events / year (naive)	251		

Framework 2, UC1: Home Overnight

Under the weekly 80 %-SoC deadline, the algorithm cost rises from 66.4 EUR/year (F1) to 77.8 EUR/year, reducing S_{defer} to 44.2 EUR/year (24.5 %) and the cumulative saving to 102.9 EUR/year (57.0 %). G_{oracle} is positive at 4.0 EUR/year, the algorithm no longer falling below the per-event oracle once weekly commitment is required. The algorithm charges on 215 of 365 days, 43 % more events than F1 in the same use case.

Table 5.12: Framework 2 × UC1 reference configuration cost decomposition. Same row structure as Table 5.10.

Quantity	EUR/year	% of naive	Notes
Naive baseline cost (C_{naive})	180.7	100.0	
Smart baseline cost (C_0)	122.0	67.5	
Algorithm cost (C_{K^*})	77.8	43.0	
Oracle cost (C_{oracle})	73.8	40.8	
Scheduling saving (S_{sched})	58.7	32.5	Forecast-free
Deferral saving (S_{defer})	44.2	24.5	Marginal answer to RQ2
Cumulative ($S_{\text{sched}} + S_{\text{defer}}$)	102.9	57.0	Total vs. naive
Oracle gap (G_{oracle})	4.0	2.2	Headroom diagnostic
Charging events / year (algorithm)	215		Endogenous
Charging events / year (naive)	365		

Framework 2, UC2: Workplace Daytime

Under the weekly deadline, UC2 outcomes are essentially identical to F1 × UC2 at the cost level: $S_{\text{defer}} = 48.8$ EUR/year (49.6 EUR/year under F1) and cumulative saving 167.1 EUR/year (167.9). G_{oracle} is 5.6 EUR/year (1.9 %).

Table 5.13: Framework 2 × UC2 reference configuration cost decomposition. Same row structure as Table 5.10.

Quantity	EUR/year	% of naive	Notes
Naive baseline cost (C_{naive})	289.8	100.0	
Smart baseline cost (C_0)	171.5	59.2	
Algorithm cost (C_{K^*})	122.7	42.3	
Oracle cost (C_{oracle})	117.1	40.4	
Scheduling saving (S_{sched})	118.3	40.8	Forecast-free
Deferral saving (S_{defer})	48.8	16.8	Marginal answer to RQ2
Cumulative ($S_{\text{sched}} + S_{\text{defer}}$)	167.1	57.7	Total vs. naive
Oracle gap (G_{oracle})	5.6	1.9	Headroom diagnostic
Charging events / year (algorithm)	182		Endogenous
Charging events / year (naive)	251		

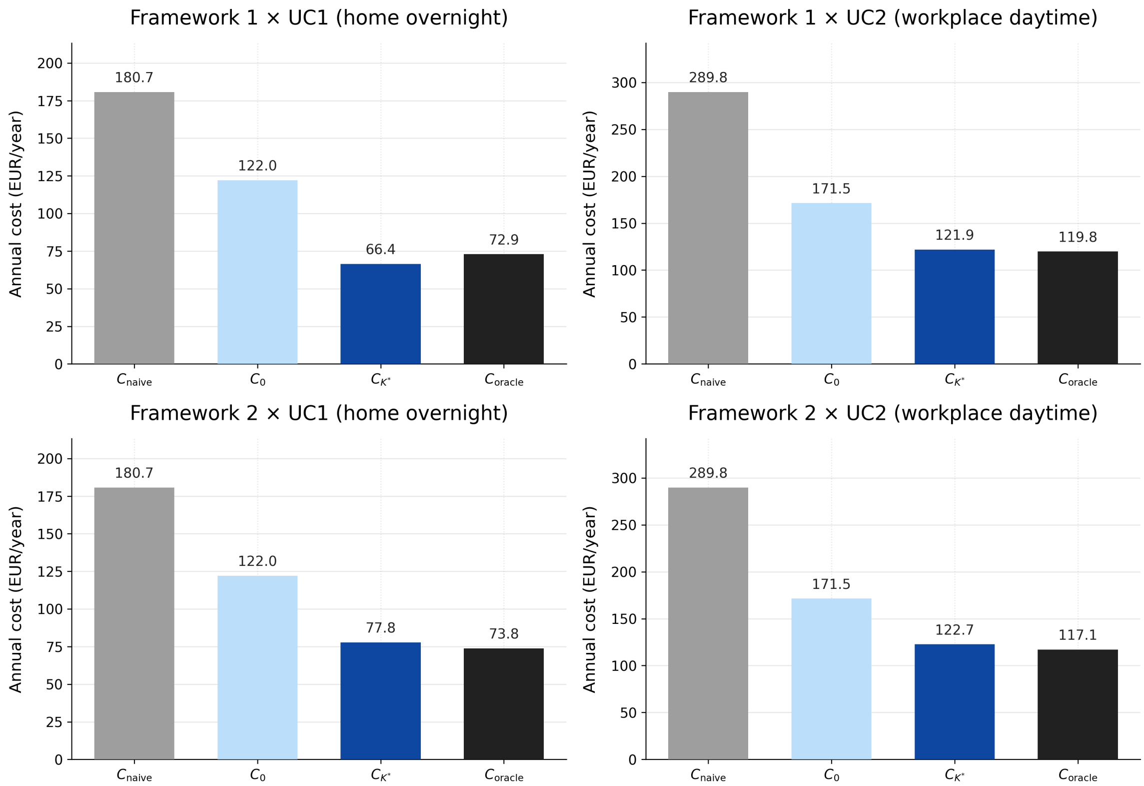

Figure 5.10: Cost decomposition bar charts at the reference configuration (50 km/day, 7.4 kW, $\lambda = 0$), populated with annual cost levels for the four framework × use-case cells. Bars from left to right in each panel: naive baseline (C_{naive}), smart baseline (C_0), algorithm (C_{K^*}), per-event oracle (C_{oracle}). Y-axes differ between UC1 and UC2 panels because of the different absolute scale (UC1 reference cost is ≈181 EUR/year, UC2 is ≈290 EUR/year); relative cost-ladder shapes are comparable across panels.

Cross-Framework Comparison

Within UC1, switching from continuous deferral (F1) to weekly deadline-bound operation (F2) costs 11.4 EUR/year in both S_{defer} and cumulative savings (the two gaps are identical because S_{sched} is forecast-free and does not depend on framework choice). Within UC2, the same switch costs 0.8 EUR/year, under 2 % of the F1 cumulative saving. The deadline constraint therefore binds asymmetrically across use cases at the reference configuration, as summarised in Table 5.14.

Table 5.14: Cross-framework comparison at the reference configuration: S_{defer} and $S_{\text{sched}} + S_{\text{defer}}$ under F1 versus F2 within each use case. Gap = F1 − F2, quantifying the cost of the deadline constraint (positive gap = F2 saves less than F1).

	S_{defer} (EUR/yr)			$S_{\text{sched}} + S_{\text{defer}}$ (EUR/yr)		
Use case	F1	F2	Gap	F1	F2	Gap
UC1 (home overnight)	55.6	44.2	11.4	114.3	102.9	11.4
UC2 (workplace daytime)	49.6	48.8	0.8	167.9	167.1	0.8

Framework-Specific Diagnostics

Forward-feasibility binding never triggered under Framework 2 in either use case across the full sensitivity grid; the deadline rather than the safety override was the operative constraint in every cycle. The Framework 1 K^*-agreement diagnostic, the rate at which the algorithm's chosen K matches the oracle's, is reported in Appendix E.

5.3.2 Sensitivity to Charging Context

This subsection sweeps the driver profile (25, 50, 75, 100 km/day) and wallbox capacity (7.4, 11, 22 kW) at $\lambda = 0$. Each of the four framework × use-case cells is reported as a 12-row sensitivity table (Tables 5.15–5.18). The 50 km/day × 7.4 kW row in each table reproduces the reference configuration reported in Section 5.3.1. The rightmost flag column marks cells where the SoC-floor (F1) or forward-feasibility (F2) constraint binds on more than 5 % of sessions.

Framework 1, UC1

Cumulative savings scale monotonically with daily distance, from 60.0–65.2 EUR/year at 25 km/day to 161.4–223.2 EUR/year at 100 km/day. The constraint flag (forced charging > 5 % of sessions) activates from 50 km/day × 11 kW onward and is set in all cells from 75 km/day. The oracle gap turns negative in most cells from 50 km/day, reaching −30.6 EUR/year at 100 km/day × 7.4 kW.

Table 5.15: Framework 1 × UC1 sensitivity to charging context. Annual savings (EUR/year) and oracle headroom across 4 driver profiles × 3 wallbox configurations. Flag: F = binding constraint observed (forced-charging rate > 5% of sessions); – = unconstrained.

Driver	Wallbox	S_{defer}	$S_{\text{sched}} + S_{\text{defer}}$	G_{oracle}	Flag
25 km/day	7.4 kW	26.0	60.0	6.2	–
25 km/day	11 kW	29.2	63.2	4.5	–
25 km/day	22 kW	31.2	65.2	4.2	–
50 km/day	7.4 kW	55.6	114.3	-6.5	–
50 km/day	11 kW	57.4	125.4	-4.1	F
50 km/day	22 kW	45.3	113.3	11.5	F
75 km/day	7.4 kW	70.4	144.9	-18.1	F
75 km/day	11 kW	74.1	161.8	-13.2	F
75 km/day	22 kW	49.4	151.4	17.2	F
100 km/day	7.4 kW	77.9	161.4	-30.6	F
100 km/day	11 kW	83.6	187.4	-25.0	F
100 km/day	22 kW	87.2	223.2	-24.4	F

Framework 1, UC2

Cumulative savings range from 93.2 EUR/year (25 km/day × 7.4 kW) to 349.6 EUR/year (100 km/day × 22 kW). The constraint flag is set in all cells from 50 km/day onward. The 100 km/day × 7.4 kW cell is excluded: the 9-slot daily window capacity (66.6 kWh) cannot cover four-day depletion under the $N=3$ deferral horizon (80 kWh), and the algorithm enters a degenerate regime with approximately 36 charging events per year.

Table 5.16: Framework 1 × UC2 sensitivity to charging context. Same column structure as Table 5.15. The 100 km/day × 7.4 kW cell is excluded: the wallbox is structurally underpowered for daily consumption under the $N=3$ deferral horizon and the algorithm enters a degenerate operating regime in this corner.

Driver	Wallbox	S_{defer}	$S_{\text{sched}} + S_{\text{defer}}$	G_{oracle}	Flag
25 km/day	7.4 kW	19.9	93.2	22.2	–
25 km/day	11 kW	28.9	105.3	15.5	–
25 km/day	22 kW	39.9	119.6	6.3	–
50 km/day	7.4 kW	49.6	167.9	2.0	F
50 km/day	11 kW	53.4	192.3	4.3	F
50 km/day	22 kW	43.3	196.2	19.4	F
75 km/day	7.4 kW	61.8	201.3	-20.0	F
75 km/day	11 kW	67.3	243.8	-18.0	F
75 km/day	22 kW	26.7	246.3	28.0	F
100 km/day	7.4 kW	*excluded (structurally degenerate)*			
100 km/day	11 kW	70.8	270.4	-34.5	F
100 km/day	22 kW	71.8	349.6	-32.3	F

Framework 2, UC1

Cumulative savings range from 60.0 EUR/year (25 km/day × 7.4 kW) to 249.1 EUR/year (100 km/day × 22 kW), scaling monotonically with both axes. The forward-feasibility binding never triggers anywhere in the grid. The oracle gap stays within ±6 EUR/year across all cells, materially tighter than the F1 × UC1 counterpart.

Table 5.17: Framework 2 × UC1 sensitivity to charging context. Same column structure as Table 5.15. Flag: F = forward-feasibility binding triggered (> 5% of sessions); – = unconstrained.

Driver	Wallbox	S_{defer}	$S_{\text{sched}} + S_{\text{defer}}$	G_{oracle}	Flag
25 km/day	7.4 kW	26.0	60.0	6.2	–
25 km/day	11 kW	27.0	61.0	5.9	–
25 km/day	22 kW	29.1	63.1	4.9	–
50 km/day	7.4 kW	44.2	102.9	4.0	–
50 km/day	11 kW	47.1	115.1	4.8	–
50 km/day	22 kW	51.1	119.1	4.0	–
75 km/day	7.4 kW	60.0	134.5	-0.9	–
75 km/day	11 kW	70.0	157.7	-1.5	–
75 km/day	22 kW	78.9	180.9	0.9	–
100 km/day	7.4 kW	71.8	155.4	3.6	–
100 km/day	11 kW	93.1	196.9	-2.7	–
100 km/day	22 kW	113.1	249.1	-1.8	–

Framework 2, UC2

Cumulative savings range from 101.5 EUR/year (25 km/day × 7.4 kW) to 439.7 EUR/year (100 km/day × 22 kW), the largest cumulative figure in the section. Forward-feasibility never binds. Above 25 km/day, however, the 80 %-SoC deadline is structurally unachievable in UC2 because weekend depletion exceeds the battery-cap headroom; the algorithm reaches the physics-floor minimum miss across these cells with negligible algorithmic slack, and per-driver deadline outcomes are not reported in this chapter.

Table 5.18: Framework 2 × UC2 sensitivity to charging context. Same column structure as Table 5.15. Flag: F = forward-feasibility binding triggered (> 5% of sessions); – = unconstrained.

Driver	Wallbox	S_{defer}	$S_{\text{sched}} + S_{\text{defer}}$	G_{oracle}	Flag
25 km/day	7.4 kW	28.1	101.5	4.5	–
25 km/day	11 kW	30.4	106.9	3.0	–
25 km/day	22 kW	31.9	111.6	2.2	–
50 km/day	7.4 kW	48.8	167.1	5.6	–
50 km/day	11 kW	54.5	193.4	5.1	–
50 km/day	22 kW	64.3	217.2	0.5	–
75 km/day	7.4 kW	67.0	206.4	2.6	–
75 km/day	11 kW	84.5	261.0	0.3	–
75 km/day	22 kW	110.7	330.3	-3.8	–
100 km/day	7.4 kW	75.6	215.5	10.7	–
100 km/day	11 kW	115.0	314.7	-3.8	–
100 km/day	22 kW	161.9	439.7	-17.3	–

5.3.3 Risk-Appetite Sensitivity

Sweeping the risk parameter λ over $\{0, 0.5, 1\}$ at the reference configuration in all four framework $\times$ use-case combinations shows that $\lambda = 0$ is empirically near-optimal in three of four cells, with risk-averse scoring ($\lambda \to 1$) costing 4–18 EUR/year relative to $\lambda = 0$. The full λ sweep is reported in Appendix E.2.

Chapter 6

Discussion

6.1 Input Signals

6.1.1 Cross-signal patterns

Two findings hold across all three input signals and are worth stating once before turning to the signal-specific patterns.

First, performance at long leads is set by the underlying NWP skill, not by the ML architecture: the four model families converge as the horizon grows, from their largest spread in the MET range to roughly one nMAE percentage point apart across all three signals in the GFS range. Model-family choice therefore becomes secondary as the horizon extends.

Second, as lead time grows the point forecast smooths toward a signal-specific baseline, climatology for wind, the historical-mean daily profile for solar, the temperature-demand regression for consumption, but the probabilistic output retains information the point forecast loses, with quantile bands either widening or staying seasonally stable as forecast-driven information decays. This matters downstream: the price model receives the full probabilistic structure of each upstream forecast, so volatility and uncertainty invisible in the smoothed long-lead point forecasts is still passed on.

6.1.2 Wind Power Prediction

At long leads the wind model expresses unreliability as widening uncertainty rather than false-confident peaks. Within the MET range it identifies individual high-wind episodes; beyond the handoff the point forecast reverts to a smoothed, mean-reverting trajectory (the late-week peaks in Figure 5.2 are not anticipated) while the 10–90 % quantile band widens during observed peaks, the model has learned that elevated long-lead inputs are unreliable peak indicators and routes them into uncertainty. The SHAP attribution supports this: the deterministic wind-speed feature `mw_raw` collapses across the horizon while calendar features rise to replace it.

The cluster-level features confirm the two-stage architecture works as designed. No single cluster feature ranks among the dominant contributors, but collectively the raw-wind-speed and physics-pipeline cluster signals let the ML stage correct residual cluster-level biases the generic physics pipeline cannot remove.

6.1.3 Solar Power Prediction

Tree-based families beat the sequence models by 5–7 percentage points across the entire MET range, a wider gap than in the wind case, because solar has stronger deterministic structure: the daily cycle, seasonal level, and cloud-cover response are captured efficiently by axis-aligned splits on calendar and weather features, and the temporal dependence the sequence models add gives little extra once hour-of-day and seasonality are present.

Solar prediction leans on a climatological baseline across the entire horizon, unlike wind. The SHAP profile shows historical window means as the dominant contributor at every lead, with the meteorological forecasts acting only as horizon-decaying corrections on top. The dominant feature never collapses the way wind's does. The exemplary forecasts match this: the daily cycle is preserved across the full 168-hour horizon, and the summer-peak underestimation and winter overestimation are scale errors with correct timing, leaving little of the regression-toward-baseline drift seen in wind.

6.1.4 Consumption Prediction

The most striking feature of the consumption results is the much lower absolute error level. The deterministic structure of consumption means a large share of the variance is predictable from features that are themselves known or well-forecast at long lead times. This predictability is driven by temperature dependence, weekly seasonality, and the daily cycle. The quantile bands in Figures 5.6 and 5.7 are narrow in relative terms and do not widen meaningfully with lead time, in marked contrast to wind and solar; what variation they do show is seasonal rather than temporal, reflecting the larger absolute amplitude of winter demand.

Consumption depends on temperature throughout, but switches between two forms of it across the horizon, a pattern qualitatively different from wind and solar. The SHAP attribution shows the explicit temperature-regression baseline dominating at short and medium leads and a residual temperature form taking over at long leads, rather than the switch-to-climatology of wind or the single-baseline reliance of solar.

Consumption is the most informative of the three upstream signals at every horizon, and this is its key downstream consequence. Its point forecasts are tighter, its quantile bands narrower and roughly flat across the 168-hour horizon, and its feature dependence never collapses to climatology. Therefore the price model receives a near-deterministic consumption input even at the longest leads, unlike the wind and solar inputs whose uncertainty widens considerably.

6.2 Electricity Price Prediction

Mirroring the solar results, tree-based algorithms surpass the sequence models at every lead time in the price horizon. The margin of outperformance begins at 15–18 percentage points for the $h = 13$–18 band and persists at roughly 3 points at the extended leads. The underlying cause is that SE3 prices depend strongly on structured inputs, specifically recent price lags, weekday indicators, and fundamental levels. Tree-based architectures isolate these axis-aligned variables efficiently, meaning the complex temporal dependencies modeled by sequence architectures are effectively redundant when explicit lags are already provided. Accuracy degrades smoothly across the MET–GFS handoff, since price does not depend purely on weather and the NWP source change matters less than the gradual shift in the model's information mix.

The price model shifts cleanly from confirmed price lags to upstream Stage 1 forecasts as leads grow, exactly as the two-stage pipeline intends. Recent price lags dominate near the gate but decay to negligible by $h = 96$, at which point the upstream consumption forecasts take over the leading role. Consumption dominates the upstream group throughout, with wind quantiles persistent but secondary, a dominance that reads two ways: consumption may genuinely be the strongest driver of multi-day SE3 price variation, or the model may simply lean on it because it is the most accurately forecast upstream signal and attribution rewards the cleanest input. The data cannot separate the two, and both effects likely operate.

The price model is accurate exactly where the charging algorithm needs it, which is what makes the downstream evaluation work. Within about a day of the gate the point forecast resolves individual hours; further out it no longer pins down specific peaks but still captures the broader price regime, with the quantile band widening to absorb the rest. That regime-level skill is sufficient because the deferral algorithm acts on relative price differences between candidate nights, not absolute hour-level predictions, meaning correct night choices follow even where individual-hour R^2 has fallen to 0.27–0.33. Read across Sections 5.3.1 and 5.3.2, the price model proves "good enough" precisely where it matters: highly accurate at the gate to schedule the first night, and regime-informative further out to defer over multi-day horizons.

6.3 Electric Vehicle Smart Charging Evaluation

6.3.1 Cross-Cell findings

Extended-horizon forecasting delivers a comparable marginal benefit across every deployment context studied. The marginal saving S_{defer} varies only from 44.2 to 55.6 EUR/year across the four framework × use-case cells, representing a narrow 25% range despite the structural differences between continuous and deadline-bound operation and between overnight and daytime charging. The cumulative saving varies far more (102.9 to 167.9 EUR/year), but almost entirely through S_{sched}, the forecast-free intra-day component.

Which metric is reported changes the conclusion. Only S_{defer} isolates the marginal value of extended-horizon information; the cumulative figure also includes intra-day savings the smart baseline already captures without any forecast. UC2 has the largest cumulative saving but the smallest deferral fraction, so a cumulative-based comparison would wrongly cast it as the strongest case for forecasting when its marginal contribution matches UC1's.

The algorithm captures nearly all the value available in its decision space. The oracle gap stays within ±3.6 % of naive cost at the reference and the algorithm matches the oracle's night choice on 62–80 % of sessions, so the residual is mostly close calls rather than forecast bias, meaning further gains must come from a richer decision problem, not better forecasts.

6.3.2 The Negative Oracle Gap

The algorithm's cost falls below the per-event oracle's in a recurring pattern, scaling monotonically negative with daily distance (from −6.5 EUR/year at the UC1 × F1 reference to −30.6 EUR/year at 100 km/day), a monotone trend that rules out statistical noise. The cause is that the per-event oracle is myopic: at each decision it picks the cheapest of the several feasible candidate charging nights in that window, but this locally optimal night-choice is not the same as minimising cumulative cost across the full horizon, so its greedy selections can occasionally lock the trajectory into a more expensive overall path than the algorithm's q_{50}-averaged scoring follows. The decomposition identity holds to machine precision in every cell, so the negative gap is a property of the per-event oracle definition, not a flaw. As set out in Section 4.4.3, this per-event design is deliberate: it makes G_{oracle} a measure of forecast quality under the evaluation framework actually used, the fair comparison, rather than a contrast with a different, globally optimal decision policy.

6.3.3 From Simulation to a Real Customer

The savings reported here describe an idealised driver, and a real customer will typically realise less. The evaluation, scoped to the OEM perspective (home and workplace scheduling; other actors are addressed in Chapter 8), treats the driver as bound to one operating framework, but a real customer is bound to neither: plug-in regularity, departure timing, target SoC, and willingness to defer over several nights are all decisions made on factors the algorithm cannot see. Continuous-deferral operation in particular asks the customer to trust the algorithm to wait several nights, which most drivers will override at the first sign of low SoC even when the decision was correct in expectation.

Real consumption is not constant either. It varies with season, weekday, and occasional long trips and both behavioural override and consumption variability soften the value of extended-horizon forecasting, since the algorithm's night choice depends on a future state trajectory that becomes less predictable once session-level uncertainty enters. The reported S_{defer} figures should therefore be read as the marginal value available to an idealised driver: in expectation a real customer realises less, and in any given period the realised saving is a random variable that can land on either side of the idealised value.

Chapter 7

Conclusions

This thesis investigates whether multi-day electricity spot price forecasts can improve the cost efficiency of electric vehicle smart charging. A three-stage pipeline was built for the Swedish SE3 bidding zone: probabilistic seven-day forecasts of wind production, solar production, and consumption from weather inputs; an hourly spot price forecast combining those upstream forecasts with price history and supply-side market data; and an EV charging evaluation across two frameworks (continuous-deferral and weekly-deadline) and two use cases (home overnight, workplace daytime), seen from the perspective of a vehicle manufacturer offering smart charging as a connected feature. Two research questions structure the analysis.

Research Question 1. *How accurately can different classes of predictive models forecast electricity spot prices beyond the day-ahead horizon, and how does forecast performance vary across model type and temporal horizon?*

Tree-based families (Random Forest, XGBoost) clearly outperform sequence families across the full 168-hour horizon, a 15–18 percentage-point nMAE gap at short leads narrowing to roughly 3 points beyond five days; Random Forest is selected as the production model on its consistent median-error advantage. Accuracy falls as the horizon grows. R^2 falls from 0.83 at the day-ahead edge to 0.27–0.33 at one week, tracking the loss of weather-forecast skill. However within the first two to four days the model still captures the broader price pattern (the daily cycle and the high- versus low-price periods) even where hour-by-hour accuracy weakens. Of the three upstream signals, consumption is the most reliably forecast across the horizon, held up by behavioural and seasonal regularity, while wind and solar decline more sharply with their weather inputs. Feature attribution confirms the pipeline design: the price model leans on confirmed price history at short leads and shifts to upstream consumption forecasts once that information is out of date.

Research Question 2. *To what extent can extended-horizon electricity price forecasts improve the cost efficiency of EV smart charging, and how does it vary across charging contexts that differ in connection window length and scheduling flexibility?*

The marginal annual saving from acting on multi-day price forecasts falls in a narrow 44–56 EUR-per-vehicle-per-year band across all four framework × use-case combinations at the reference 50 km/day, 7.4 kW configuration. This demonstrates a comparable benefit regardless of deployment context, achieved on top of the within-day optimisation that current smart charging products already perform. Total savings against a do-nothing baseline are larger (103–168 EUR/year), but that variation is driven almost entirely by the forecast-free within-day component, so total figures overstate the marginal value of extended-horizon information. The algorithm captures most of the value available in its decision space, with the gap to a perfect-information benchmark remaining within 3.6 % of the do-nothing cost at the reference configuration. Consequently, further financial gains would require a richer decision problem rather than more accurate price forecasts. Absolute savings scale with mileage, from roughly 20–28 EUR/year at 25 km/day to 72–78 at 100 km/day where the configuration remains workable. These figures describe an idealised driver with regular plug-in and a willingness to defer over several nights; a real customer realises less, depending on how closely their behaviour follows that pattern.

Scope and broader applicability. The evaluation is bounded to the vehicle manufacturer's perspective and to the Swedish SE3 market. The forecasting capability itself is general: Chapter 8 outlines extensions to charge-point operators managing public charging infrastructure and to commercial fleets coordinating across multiple vehicles, both of which face different decision problems but use the same probabilistic price information. Within the manufacturer's scope, the work suggests that multi-day price forecasting delivers a marginal contribution that is real, consistent across deployment contexts, and dependent on a customer relationship to charging that does not require a fully charged battery every morning.

Chapter 8

Future Work

Several extensions to this thesis are worth pursuing.

15-Minute Resolution. On 1 October 2025, the European day-ahead market transitioned to 15-minute settlement intervals. Extending the pipeline from hourly to 15-minute resolution would require rebuilding the panel structure at the finer cadence but would not change the overall methodology, and would let intra-hour price variation become actionable for the downstream charging algorithm.

Bidirectional Operation and Grid Services. The price forecasting models developed here are equally applicable to discharge timing and to bidding in markets beyond the day-ahead spot, such as the Swedish mFRR balancing market [39]. Both extensions introduce a substantially more dynamic battery usage pattern that would benefit from being evaluated alongside empirical battery-degradation curves.

Additional Charging Contexts. Within the OEM perspective, the evaluation covers home overnight (UC1) and workplace daytime (UC2) charging. Natural extensions include mixed home-and-workplace access for a single driver and explicit modelling of weekend long-trip events that fall outside the regular weekly rhythm.

Charging Point Operators. CPOs operating public charging stations currently set fixed per-kWh tariffs and absorb the full price risk between their wholesale procurement and retail delivery. An extended-horizon forecast opens the door to dynamic CPO pricing: lifting prices during high-price windows to protect margins, passing through low-price windows to attract demand, and informing site-level scheduling of co-located battery storage assets where present. The probabilistic forecasts produced here are particularly well suited to the margin-protection use case, where the upper quantile band is the relevant signal rather than the median.

Commercial Fleets. Commercial vehicle fleets with centralised dispatch, such as utilities, last-mile delivery, and mobility services, face a deferral problem similar to UC1 but with the added dimension of cross-vehicle coordination. Aggregated deferral over N vehicles reaches scales of price exposure that an individual driver cannot, and the deferral decision becomes a portfolio problem across heterogeneous duty cycles rather than a per-session optimisation. The framework developed here applies per-vehicle; extending it to fleet level requires modelling the joint decision and the operational constraints that bind across vehicles.

Geographic Expansion. The methodology transfers naturally to the other Swedish bidding zones (SE1, SE2, SE4) with no structural change: only the underlying data differs. Extension beyond the Nordic market requires more substantial rework to accommodate different price-formation structures, generation mixes, and cross-border flow dynamics, but the conceptual framework remains applicable with market-specific modifications.

Commercial Framing. The evaluation quantifies value as annual cost savings against a do-nothing baseline. Cost savings are unlikely to be the only driver of customer adoption: extending the visible price curve from the confirmed 24-hour window to a multi-day forecast horizon also changes what the customer can see and act on, introducing a sense of control and informed decision-making that current smart charging products do not offer.

High-Frequency Weather Forecast Integration. The current setup relies on weather predictions restricted to a six-hour resolution, with updates initialized only at 00:00, 06:00, 12:00, and 18:00. Furthermore, these updates from GFS are subject to a significant publication lag, typically becoming available four to six hours after their initialization. As a result of this information lag, the pipeline cannot operate in real time. Because the latest data is severely delayed, a spot price prediction generated at 16:00 remains bound to the 12:00 weather initialization, forcing the model to rely on past meteorological inputs. Future work should explore implementing a solution from an alternative weather forecast provider that issues high-frequency updates with minimal delay. Transitioning to a continuous, low-latency data feed would resolve this operational bottleneck and enable the pipeline to produce truly live, hourly spot price forecasts.

Bibliography

[1] International Energy Agency, “Global EV outlook 2024,” IEA, Tech. Rep., 2024. [Online]. Available: https://www.iea.org/reports/global-ev-outlook-2024.

[2] P. Huang and Z. Ma, “Unveiling electric vehicle (EV) charging patterns and their transformative role in electricity balancing and delivery: Insights from real-world data in Sweden,” *Renewable Energy*, vol. 236, p. 121 511, 2024. DOI: 10.1016/j.renene.2024.121511.

[3] Tibber, *Smart charging*, Accessed: 2026-03-18, 2024. [Online]. Available: https://tibber.com/se/smartladdning.

[4] Gridio, *Smart EV charging*, Accessed: 2026-03-18, 2024. [Online]. Available: https://www.gridio.io/sv.

[5] Tesla, *Scheduled precondition and charge*, Accessed: 2026-03-18, 2026. [Online]. Available: https://www.tesla.com/ownersmanual/modely/en_us/GUID-76995CEC-6402-4BFF-99FA-CEFA36E64A19.html.

[6] Rechargeability, “Smart charging for charge point operators,” Rechargeability, Tech. Rep., 2023. [Online]. Available: https://www.rechargeability.com/knowledge/smart-charging-for-cpos.

[7] Volvo Cars, *Smart charging*, Accessed: 2026, 2026. [Online]. Available: https://www.volvocars.com/se/support/topic/0ef7c0498845e1cdc0a801514899a3be/.

[8] Vattenfall, *Kostnad att ladda elbil*, Accessed: 2026, 2026. [Online]. Available: https://www.vattenfall.se/fokus/eldrivna-transporter/kostnad-ladda-elbil/.

[9] Mer, *Priser för laddstationer*, Accessed: 2026, 2026. [Online]. Available: https://se.mer.eco/laddstationer/priser/.

[10] Nord Pool, *Day-ahead market*, Accessed: 2026-03-18. [Online]. Available: https://www.nordpoolgroup.com/en/the-power-market/Day-ahead-market/.

[11] R. Sandberg, “Wind and electricity prices in Sweden — A statistical analysis,” Energiforsk, Tech. Rep. 2024-975, 2024. [Online]. Available: https://energiforsk.se/media/33100/2024-975-wind-and-electricity-prices-in-sweden-a-statistical-analysis.pdf.

[12] R. Weron, “Electricity price forecasting: A review of the state-of-the-art with a look into the future,” *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014. DOI: 10.1016/j.ijforecast.2014.08.008.

[13] F. Wieberneit, E. Crisostomi, A. Quinn, H. Hamedmoghadam, and R. Shorten, “Predictive control of EV overnight charging with multi-session flexibility,” *arXiv preprint arXiv:2505.05087*, 2025.

[14] Energimarknadsinspektionen, “Sweden’s electricity and natural gas market 2023,” Swedish Energy Markets Inspectorate (Ei), Tech. Rep. Ei R2025:07, 2025. [Online]. Available: https://ei.se/download/18.3a4ca3ae195b23dd0051daa2/1742815664184/Swedens-electricity-and-natural-gas-market-2023-Ei-R2025-07.pdf.

[15] Nord Pool, *Day-ahead aggregated bidding curves*, Accessed: 2026-03-18. [Online]. Available: https://www.nordpoolgroup.com/en/services/power-market-data-services/aggregatedbiddingcurves/.

[16] Svenska kraftnät, *The control room*, Accessed: 2026-03-18, 2024. [Online]. Available: https://www.svk.se/en/national-grid/the-control-room/.

[17] Svenska kraftnät, *Operations and electricity markets*, Accessed: 2026-03-18, 2024. [Online]. Available: https://www.svk.se/en/national-grid/operations-and-electricity-markets/.

[18] EPEX SPOT, *15-minute products in market coupling*, Accessed: 2026-03-18, 2025. [Online]. Available: https://www.epexspot.com/en/15-minute-products-market-coupling.

[19] IEA PVPS Task 1, “National survey report of PV power applications in Sweden 2024,” International Energy Agency Photovoltaic Power Systems Programme, Tech. Rep., 2025. [Online]. Available: https://iea-pvps.org/national_survey/nsr-sweden-2024/.

[20] H. S. Hippert, C. E. Pedreira, and R. C. Souza, “Neural networks for short-term load forecasting: A review and evaluation,” *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44–55, 2001. DOI: 10.1109/59.910780.

[21] P. Bauer, A. Thorpe, and G. Brunet, “The quiet revolution of numerical weather prediction,” *Nature*, vol. 525, no. 7567, pp. 47–55, 2015. DOI: 10.1038/nature14956.

[22] eSett Oy, *Nordic imbalance settlement*, Accessed: 2026-03-18, 2024. [Online]. Available: https://www.esett.com.

[23] AleaSoft, *Short-term electricity market price forecasting*, Accessed: 2026-03-18, 2026. [Online]. Available: https://aleasoft.com/services/energy-price-forecasting/short-term/.

[24] DNV, *Power price forecasts*, Accessed: 2026-03-18, 2024. [Online]. Available: https://www.dnv.com/services/power-price-forecasts/.

[25] P. Piotrowski, I. Rutyna, D. Baczyński, and M. Kopyt, “Evaluation metrics for wind power forecasts: A comprehensive review and statistical analysis of errors,” *Energies*, vol. 15, no. 24, p. 9657, 2022. DOI: 10.3390/en15249657.

[26] E. Belenguer, J. Segarra-Tamarit, E. Pérez, and R. Vidal-Albalate, "Short-term electricity price forecasting through demand and renewable generation prediction," *Mathematics and Computers in Simulation*, vol. 229, pp. 350–361, 2025. DOI: 10.1016/j.matcom.2024.10.004.

[27] P. Ineichen and R. Perez, "A new airmass independent formulation for the linke turbidity coefficient," *Solar Energy*, vol. 73, no. 3, pp. 151–157, 2002. DOI: 10.1016/S0038-092X(02)00045-2.

[28] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martínez-de-Pisón, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136, pp. 78–111, 2016. DOI: 10.1016/j.solener.2016.06.069.

[29] J. W. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," *Journal of the Operational Research Society*, vol. 54, no. 8, pp. 799–805, 2003. DOI: 10.1057/palgrave.jors.2601589.

[30] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016. DOI: 10.1016/j.ijforecast.2015.11.011.

[31] H. Liu and J. Shi, "Applying ARMA–GARCH approaches to forecasting short-term electricity prices," *Energy Economics*, vol. 37, pp. 152–166, 2013. DOI: 10.1016/j.eneco.2013.02.006.

[32] D. O'Connor, W. Bahloul, S. Prestwich, and L. Visentin, "A review of electricity price forecasting models in the day-ahead, intra-day, and balancing markets," *Energies*, vol. 18, no. 12, p. 3097, 2025. DOI: 10.3390/en18123097.

[33] J. Lago, F. D. Ridder, and B. D. Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, vol. 221, pp. 386–405, 2018. DOI: 10.1016/j.apenergy.2018.02.069.

[34] B. Biró, C. Kiss, R. Molontay, and A. Aszódi, "Industry-adaptable explainable AI based methodology for forecasting electricity prices," *Energy Conversion and Management: X*, vol. 30, p. 101 583, 2026. DOI: 10.1016/j.ecmx.2026.101583.

[35] D. K. Kılıç, P. Nielsen, and A. Thibbotuwawa, "Intraday electricity price forecasting via LSTM and trading strategy for the power market: A case study of the west denmark DK1 grid region," *Energies*, vol. 17, no. 12, p. 2909, 2024. DOI: 10.3390/en17122909.

[36] V. Renefalk, "Forecasting day-ahead electricity market prices using ML and rule-based ensemble models: A case study on the swedish electricity markets," Master's thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2025.

[37] E. Sortomme and M. A. El-Sharkawi, "Optimal charging strategies for unidirectional vehicle-to-grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 1, pp. 131–138, 2011. DOI: 10.1109/TSG.2010.2090910.

[38] Z. et al., "Electric vehicle cluster charging scheduling optimization: A forecast-driven multi-objective reinforcement learning method," *Energies*, vol. 19, no. 3, p. 647, 2026.

[39] S. Z. A. Fakhry, F. Golzar, and J. Ardelius, "Forecasting and optimizing residential ev flexibility for the swedish mfrr market using machine learning," *International Journal of Electrical Power and Energy Systems*, vol. 174, p. 111 558, 2026. DOI: 10.1016/j.ijepes.2025.111558.

[40] Swedish Energy Agency, *Vindbrukskollen*, Accessed: 2026, 2026. [Online]. Available: https://www.energimyndigheten.se/energisystem-och-analys/elproduktion/vindkraft/vindbrukskollen/.

[41] Energimyndigheten, *Över 250 000 installerade solcellsanläggningar i sverige*, Apr. 2024. Accessed: Apr. 4, 2026. [Online]. Available: https://www.energimyndigheten.se/nyhetsarkiv/2024/over-250-000-installerade-solcellsanlaggningar-i-sverige/.

[42] Energimyndigheten, *Nätanslutna solcellsanläggningar*, 2025. Accessed: Apr. 4, 2026. [Online]. Available: https://www.energimyndigheten.se/statistik/officiell-energistatistik/tillforsel-och-anvandning/natanslutna-solcellsanlaggningar/?currentTab=0.

[43] Statistiska centralbyrån, *Slutlig energianvändning efter sektor*, Statistikdatabasen, 2024. Accessed: Apr. 4, 2026. [Online]. Available: https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__EN__EN0203__EN0203A/SlutAnvSektor/.

[44] P. Holmberg and T. P. Tangerås, "The swedish electricity market – today and in the future," *Sveriges Riksbank Economic Review*, no. 1, pp. 5–36, 2023. [Online]. Available: https://www.riksbank.se/globalassets/media/rapporter/pov/artiklar/engelska/2023/230512/2023_1-the-swedish-electricity-market--today-and-in-the-future.pdf.

[45] Swedish Meteorological and Hydrological Institute (SMHI), *Open data portal, introduction*, https://opendata.smhi.se/metobs/introduction, Data retrieved programmatically via API. Accessed: 2026-05-04, 2026.

[46] eSett, *eSett open data API*, https://api.opendata.esett.com/, Accessed: 2026-05-05, 2026.

[47] European Network of Transmission System Operators for Electricity (ENTSO-E), *Transparency Platform Restful API Integration Guide*, https://transparencyplatform.zendesk.com/hc/en-us/sections/12783116987028-Restful-API-integration-guide, Accessed: 2026-05-05, 2026.

[48] Norwegian Meteorological Institute (MET Norway), *MET Nordic: 1 km high-resolution post-processed weather forecast*, https://thredds.met.no/thredds/catalog/metpplatest/catalog.html, Data retrieved programmatically via OPeNDAP. Accessed: 2026-05-04, 2026.

[49] Norwegian Meteorological Institute (MET Norway), *MET Nordic Dataset*, https://github.com/metno/NWPdocs/wiki/MET-Nordic-dataset, Accessed: 2026-05-05, 2026.

[50] National Centers for Environmental Prediction (NCEP), National Oceanic and Atmospheric Administration (NOAA), *Global Forecast System (GFS)*, https://nomads.ncep.noaa.gov/, Accessed: 2026-05-05, 2026.

[51] National Center for Athmospheric Research (NCAR), *NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive*, https://data.ucar.edu/dataset/ncep-gfs-0-25-degree-global-forecast-grids-historical-archive, Accessed: 2026-05-05, 2026.

[52] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2019, pp. 2623–2631. DOI: 10.1145/3292500.3330701.

[53] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Advances in neural information processing systems*, vol. 24, 2011.

[54] Waaree Energies Ltd., *Irradiance and temperature: The dynamic duo behind every solar panel's performance*, Accessed: 2026-05-04, 2025. [Online]. Available: https://www.waaree.com/blog/irradiance-temperature-in-solar-panel-performance/.

[55] Energimyndigheten, *Kortsiktsprognos i siffror vinter 2026*, Excel spreadsheet, Accessed: 2026-05-04, 2026. [Online]. Available: https://www.energimyndigheten.se/4948fd/globalassets/energisystem-och-analys/kortsiktiga-prognoser/kortsiktsprognos-i-siffror-vinter-2026.xlsx.

[56] Volvo Cars, *EX60 electric: Specifications and dimensions*, Accessed: 2026, 2025. [Online]. Available: https://www.volvocars.com/mt/cars/ex60-electric/specifications/.

[57] E. Lång, E. Adell, and L. Winslott Hiselius, "Reser elbilister annorlunda? Verkligt resbeteende som underlag för att styra mot transporteffektivitet," Swedish National Road and Transport Research Institute (VTI), VTI rapport 1258, 2026, Accessed: 2026. DOI: 10.65151/vti240576. [Online]. Available: https://doi.org/10.65151/vti240576.

[58] Volvo Cars, *Volvo wallbox charger with socket*, Accessed: 2026, 2026. [Online]. Available: `https://www.volvocars.com/mt/shop/additionals/volvo-wallbox-with-socket/`.

[59] Easee, *Charge up — electric car charger for charging at home*, Accessed: 2026, 2026. [Online]. Available: `https://easee.com/en/products/ev-chargers/charge-up/`.

[60] M. Xylia and S. Joshi, "A three-dimensional view of charging infrastructure equity," Stockholm Environment Institute, Tech. Rep., 2022. DOI: `10.51414/sei2022.020`.

[61] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

[62] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd. Springer, 2009.

[63] A. Lindholm, N. Wahlström, F. Lindsten, and T. B. Schön, *Machine Learning: A First Course for Engineers and Scientists*. Cambridge University Press, 2026, Online draft version, March 4, 2026, `http://smlbook.org`.

[64] I. T. Jolliffe and J. Cadima, "Principal component analysis: A review and recent developments," *Philosophical Transactions of the Royal Society A*, vol. 374, no. 2065, 2016.

[65] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of Artificial Intelligence Research*, vol. 4, pp. 237–285, 1996.

[66] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[67] C. Bergmeir and J. M. Benítez, "On the use of cross-validation for time series predictor evaluation," *Information Sciences*, vol. 191, pp. 192–213, 2012. DOI: `10.1016/j.ins.2011.12.028`.

[68] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.

[69] I. Gitman, H. Lang, P. Zhang, and L. Xiao, "Understanding the role of momentum in stochastic gradient methods," in *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[70] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[71] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[72] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.

[73] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[74] P. J. Werbos, "Backpropagation through time: What it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990. DOI: `10.1109/5.58337`.

[75] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.

[76] R. Pascanu, T. Mikolov, and Y. Bengio, “On the difficulty of training recurrent neural networks,” in *International Conference on Machine Learning*, 2013, pp. 1310–1318.

[77] K. Cho et al., “Learning phrase representations using RNN encoder-decoder for statistical machine translation,” *arXiv preprint arXiv:1406.1078*, 2014.

[78] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” *arXiv preprint arXiv:1412.3555*, 2014.

[79] C.-J. Chen, F.-I. Chou, and J.-H. Chou, “Temperature prediction for reheating furnace by gated recurrent unit approach,” *IEEE Access*, vol. 10, pp. 33 362–33 369, 2022.

[80] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[81] Z. Su and Z. Zhou, “Modeling the impacts of climate change on monthly commercial sector electricity consumption: A case of suzhou, china,” *Energy Science & Engineering*, vol. 11, no. 4, pp. 1370–1381, 2023. DOI: `https://doi.org/10.1002/ese3.1333`. eprint: `https://scijournals.onlinelibrary.wiley.com/doi/pdf/10.1002/ese3.1333`. [Online]. Available: `https://scijournals.onlinelibrary.wiley.com/doi/abs/10.1002/ese3.1333`.

Appendix A

Extended Machine Learning Theory

This appendix provides the rigorous mathematical formulations, algorithmic procedures, and architectural details underpinning the machine learning models summarized in Chapter 3.

A.1 Machine Learning Fundamentals

Formally, we observe a dataset of n labelled pairs $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$, where $\mathbf{x}_i \in \mathbb{R}^p$ collects p measured or forecasted input signals for observation i and $y_i \in \mathbb{R}$ is the quantity to be predicted. The task is to find a function $\hat{f}_{\boldsymbol{\theta}} : \mathbb{R}^p \to \mathbb{R}$, parameterised by $\boldsymbol{\theta}$, whose predictions generalise beyond the training sample. Defining $\ell(\cdot, \cdot)$ as a loss that penalises discrepancies between predictions and observations, and letting $\mathcal{P}(\mathbf{x}, y)$ denote the data-generating distribution, the ideal parameters satisfy:

$$\boldsymbol{\theta}^\star = \arg\min_{\boldsymbol{\theta}} \ \mathbb{E}_{(\mathbf{x},y)\sim\mathcal{P}}\Big[\ell\Big(y,\ \hat{f}_{\boldsymbol{\theta}}(\mathbf{x})\Big)\Big]. \tag{A.1}$$

Since $\mathcal{P}$ is unknown, the expectation in (A.1) is approximated empirically over $\mathcal{D}$ during training [61], [62]. The notation adopted throughout the thesis is as follows

- $\mathbf{x}_i \in \mathbb{R}^p$: input feature vector for observation i — for example, forecasted wind speeds, solar irradiance, or air temperatures;
- $y_i \in \mathbb{R}$: scalar target — hourly spot price (EUR/MWh), wind power production (MW), solar generation (MW), or electricity consumption (MW) depending on the sub-model;
- $\hat{y}_i = \hat{f}_{\boldsymbol{\theta}}(\mathbf{x}_i)$: the model's predicted output;
- n, p: number of training observations and input features respectively;
- $\boldsymbol{\theta}$: parameters learned by minimising the training objective;
- λ: regularisation hyperparameter, selected via a held-out validation set rather than gradient optimisation.

Two categories of model settings must be distinguished. The *parameters* $\boldsymbol{\theta}$, including regression coefficients, neural network weights, and tree split thresholds are determined automatically during training. *Hyperparameters* such as regularisation strength, tree depth, learning rate, and network width define the structure within which that training takes place; their values are chosen by evaluating generalisation on data withheld from the fitting procedure.

Bias, variance, and the cost of complexity. Two competing sources of prediction error arise whenever a model is fitted to finite data. A model that is too simple fails to capture genuine structure in the inputs, exhibiting high *bias*. A model that is too expressive memorises the specific noise present in the training sample and performs poorly on new observations, exhibiting high *variance*. For squared-error loss, this tension is formalised by decomposing the expected prediction error at any fixed input $\mathbf{x}$ into three additive components, as in Equation (A.2)

$$\mathbb{E}\left[\left(y - \hat{f}(\mathbf{x})\right)^2\right] = \underbrace{\left(\mathbb{E}[\hat{f}(\mathbf{x})] - f^\star(\mathbf{x})\right)^2}_{\text{squared bias}} + \underbrace{\mathbb{V}\left[\hat{f}(\mathbf{x})\right]}_{\text{variance}} + \underbrace{\mathbb{V}[y \mid \mathbf{x}]}_{\text{irreducible noise}}, \tag{A.2}$$

where $\mathbf{x} \in \mathbb{R}^p$ is the feature vector, $y \in \mathbb{R}$ the response variable, $\hat{f}$ a model trained on a random training set, $f^\star(\mathbf{x}) = \mathbb{E}[y \mid \mathbf{x}]$ the theoretically optimal (Bayes) predictor, and $\mathbb{E}[\cdot]$ and $\mathbb{V}[\cdot]$ expectation and variance taken jointly over the training data and the new input–response pair $(\mathbf{x}, y)$ [62], [63]. The third term is irreducible: no model can explain randomness intrinsic to the data-generating process, which in electricity markets includes unobserved cross-border flows, unannounced nuclear outages, and strategic bidding behaviour. Figure A.1 illustrates how total prediction error is minimised at an intermediate model complexity where neither bias nor variance dominates. Effective model selection navigates these first two terms: enough flexibility to represent genuine price-driving patterns, yet sufficient constraint to avoid absorbing noise.

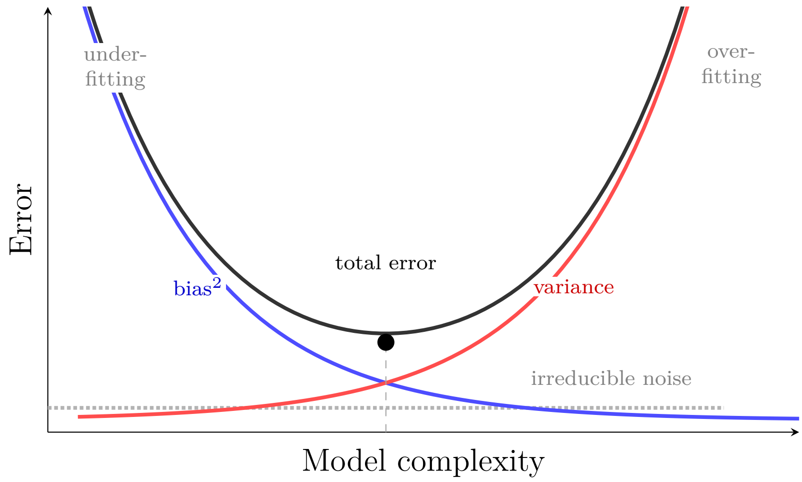

Figure A.1: Bias–variance trade-off: total prediction error is minimised at an intermediate complexity where the decrease in bias^2 is offset by increasing variance.

A.1.1 Types of Machine Learning

Four broad paradigms organise the ML literature, distinguished by the nature of the available supervision signal.

Supervised learning. Every model evaluated in this thesis falls under supervised learning, where each training observation $(\mathbf{x}_i, y_i)$ carries an observed target. When the target is a continuous real value (such as spot prices, wind power output, solar generation, or electricity demand), the task is regression. The model is fitted to minimise a training loss over $\mathcal{D}$, and its ultimate quality is judged on chronologically held-out test observations that represent the practical forecasting setting.

Unsupervised learning. When no output labels exist, unsupervised methods search for latent structure within the input distribution. One widely used technique is Principal Component Analysis (PCA) [64], which finds the directions of greatest variance in the feature space. For a mean-centred data matrix $\mathbf{X} \in \mathbb{R}^{n \times p}$, PCA computes the eigendecomposition of the sample covariance $\mathbf{C} = \frac{1}{n}\mathbf{X}^\top \mathbf{X}$ and projects observations onto the k leading eigenvectors, discarding directions dominated by noise [64]. In this thesis PCA is optionally applied as a preprocessing step to reduce redundancy across the correlated weather signals before model fitting.

Reinforcement learning. Rather than learning from labelled data, a reinforcement learning agent discovers a policy through repeated environmental interaction: actions are taken, outcomes are observed, and scalar reward signals gradually shape the policy towards higher cumulative return [65]. The EV charging scheduling decision of whether to charge during the current period or defer to a predicted lower-cost future window is structurally a reinforcement learning problem. The scope here, however, is the upstream price forecasting component that would supply the state signal for such an agent.

Generative models. Generative approaches learn to synthesise new observations consistent with the training distribution, most visibly in the form of large language models. While synthetic scenario generation for energy systems is an active research direction, these methods are not used here: their data and computational requirements exceed what the present forecasting task demands [66].

A.1.2 Model Training and Evaluation

Fitting a supervised model involves three sequential decisions. The training loss is first minimised over $\mathcal{D}$ to produce parameter estimates $\hat{\boldsymbol{\theta}}$. Hyperparameters are then chosen by comparing variants on a validation set that was withheld during fitting. Finally, the selected configuration is evaluated on a test set that has played no role in any prior modelling decision. A critical constraint throughout this thesis is temporal ordering: because electricity prices exhibit strong serial dependence, all data partitions are performed chronologically. Shuffling observations before splitting would permit the model to implicitly condition on future information, yielding evaluation results that would not reflect real forecasting performance [67].

Gradient descent. Minimising a differentiable training loss $\mathcal{L}(\boldsymbol{\theta})$ requires an optimisation algorithm. Gradient descent exploits the fact that the negative gradient $-\nabla_{\boldsymbol{\theta}}\mathcal{L}$ points in the direction of steepest local decrease, producing iterates

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \alpha \nabla_{\boldsymbol{\theta}}\mathcal{L}\Big(\boldsymbol{\theta}^{(k)}\Big), \tag{A.3}$$

where $\boldsymbol{\theta} \in \mathbb{R}^p$ is the model's parameter vector, $\boldsymbol{\theta}^{(k)}$ its value at iteration $k \in \mathbb{N}_0$, $\alpha > 0$ the learning rate, $\mathcal{L}(\boldsymbol{\theta})$ the differentiable training loss, and $\nabla_{\boldsymbol{\theta}}\mathcal{L}$ its gradient with respect to $\boldsymbol{\theta}$ [68]. For convex objectives a small enough α guarantees convergence; for the non-convex landscapes of neural networks no such guarantee holds. Computing the exact gradient over all n training observations at each step is often impractical. Mini-batch stochastic gradient descent (SGD) instead estimates the gradient from a randomly drawn subset, trading precision for speed while introducing noise that can help escape shallow local minima [68]. Augmenting SGD with a momentum term, an exponentially weighted accumulation of past gradient directions, smooths the update trajectory and accelerates convergence along consistent descent directions [69]. These ideas underpin the optimisers used when training the recurrent models in Section A.3.

Evaluation metrics. Several complementary metrics are computed to assess regression performance. Let $\{(y_i, \hat{y}_i)\}_{i=1}^{n}$ be the test-set collection of true and predicted values, and $\bar{y} = n^{-1}\sum_{i=1}^{n} y_i$ the mean observation over the test period.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{A.4}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{A.5}$$

$$\mathrm{nMAE} = \frac{\mathrm{MAE}}{\bar{y}}, \tag{A.6}$$

$$\mathrm{nRMSE} = \frac{\mathrm{RMSE}}{\bar{y}}, \tag{A.7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}. \tag{A.8}$$

The Mean Absolute Error (MAE) defined in (A.4) reports error in the native output unit (EUR/MWh for prices, MW for power quantities) and is robust to occasional large deviations. The Root Mean Square Error (RMSE) defined in (A.5) squares each residual before averaging and is therefore particularly sensitive to price spikes, a defining feature of Nordic spot markets. Because the output variables span very different magnitudes, the normalised variants (A.6)–(A.7) divide by the test-period mean $\bar{y}$, expressing each error as a fraction of typical magnitude and enabling fair cross-variable comparison. The coefficient of determination R^2 (A.8) measures the share of output variance explained relative to a naïve mean forecast: 1 is perfect, 0 matches the mean baseline, and negative values indicate worse-than-mean predictions.

One nuance deserves attention. R^2 weights every observation equally regardless of economic consequences. For wind power, however, the price impact of a forecasting error depends strongly on where in the generation distribution it falls: errors at very high or very low wind propagate far more strongly into price deviations than mid-range errors, where the merit-order curve is flatter. This asymmetry is acknowledged when interpreting wind model performance, and the metrics are examined conditionally across the wind output distribution where appropriate [70].

SHAP feature importance. Predicting accurately is necessary but not sufficient: for a model serving as a component in an operational forecasting pipeline, understanding which inputs actually drive predictions is essential for validation and trust. This thesis uses SHAP (Shapley Additive explanations) [71], which borrows the Shapley value concept from cooperative game theory to assign each feature a fair share of the credit for any given prediction. For model $\hat{f}$, observation $\mathbf{x}$, and feature index j, the Shapley value is given by Equation (A.9)

$$\varphi_j(\hat{f}, \mathbf{x}) = \sum_{\mathcal{S} \subseteq \mathcal{F} \setminus \{j\}} \frac{|\mathcal{S}|!\,(|\mathcal{F}| - |\mathcal{S}| - 1)!}{|\mathcal{F}|!} \Big[\hat{f}_{\mathcal{S} \cup \{j\}}(\mathbf{x}_{\mathcal{S} \cup \{j\}}) - \hat{f}_{\mathcal{S}}(\mathbf{x}_{\mathcal{S}})\Big], \tag{A.9}$$

where $\varphi_j(\hat{f}, \mathbf{x})$ is the Shapley value of feature j for the prediction at observation $\mathbf{x}$, $\mathcal{F}$ the complete feature set, $\mathcal{S} \subseteq \mathcal{F} \setminus \{j\}$ a subset of features not containing j, $|\cdot|$ denotes set cardinality, $\mathbf{x}_{\mathcal{S}}$ the restriction of $\mathbf{x}$ to the features in $\mathcal{S}$, and $\hat{f}_{\mathcal{S}}$ the model restricted to the feature subset $\mathcal{S}$ [71]. The combinatorial weighting averages each feature's marginal contribution across all orderings in which it could enter the model, ensuring that credit is distributed fairly when features are correlated. Averaging $|\varphi_j|$ over the test set yields a global importance ranking that is consistent across all model families evaluated in this thesis.

A.2 Ensemble and Tree-Based Models

Unlike standard linear regression or standard neural networks, which assume a globally smooth and continuous response surface, a regression tree models relationships as a piecewise-constant function. This means the resulting predictions are inherently discontinuous. This step-like, discontinuous behavior is theoretically well-suited for modeling electricity spot prices, as the market's underlying merit-order curve frequently exhibits abrupt, non-smooth price jumps when system transmission constraints bind or when the marginal pricing technology shifts entirely.

The basic unit is the *regression tree*, which partitions the feature space $\mathbb{R}^p$ into a set of axis-aligned rectangular regions through a sequence of binary splits. At every node, the split variable and threshold are selected to maximise the reduction in mean squared prediction error, as in Equation (A.10)

$$\Delta\,\mathrm{MSE} = \mathrm{MSE}(\mathcal{N}) - \frac{|\mathcal{N}_L|}{|\mathcal{N}|}\mathrm{MSE}(\mathcal{N}_L) - \frac{|\mathcal{N}_R|}{|\mathcal{N}|}\mathrm{MSE}(\mathcal{N}_R). \tag{A.10}$$

where $\mathcal{N}$ denotes the current node, $\mathcal{N}_L$ and $\mathcal{N}_R$ its proposed left and right children under the candidate split, $|\cdot|$ the number of training observations in a node, $\mathrm{MSE}(\mathcal{N})$ the mean squared prediction error of the responses falling in node $\mathcal{N}$, and $\Delta\,\mathrm{MSE}$ the resulting weighted reduction in MSE that the split achieves. Growth stops when a maximum depth or minimum node size is reached. The tension between shallow trees that underfit and deep trees that overfit is managed by aggregating many trees into an ensemble. The two ensemble strategies evaluated in this thesis (Random Forest and XGBoost) differ fundamentally in how they construct and combine their trees, as illustrated in Figures A.2 and A.3.

A.2.1 Random Forest

Random Forest [72] builds B regression trees in parallel, each on an independently drawn bootstrap resample $\mathcal{D}_b$ of the training data, and combines them by averaging

$$\hat{f}_{\mathrm{RF}}(\mathbf{x}) = \frac{1}{B}\sum_{b=1}^{B} \hat{f}_b(\mathbf{x}). \tag{A.11}$$

A bootstrap resample $\mathcal{D}_b$ is generated by drawing n observations from the original training dataset $\mathcal{D}$ uniformly at random with replacement. Because unpruned regression trees are highly sensitive to their training data (i.e., they are high-variance estimators), the bootstrap procedure ensures that each tree is grown on a structurally unique variation of the data. This process decorrelates the individual trees. Consequently, when their independent predictions are averaged in (A.11), the overall variance of the ensemble is drastically reduced while the bias remains largely unchanged [62], [72]. The procedure is stated as Algorithm 1.

Algorithm 1 Random Forest (Regression)

Require: Dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$, number of trees B, features per split m_{try}
1: **for** $b = 1$ **to** B **do**
2: Draw bootstrap resample $\mathcal{D}_b$ from $\mathcal{D}$
3: Grow tree $\hat{f}_b$ on $\mathcal{D}_b$; at each node sample m_{try} features and split on the one giving the greatest impurity reduction
4: **end for**
5: **return** $\hat{f}_{\mathrm{RF}}(\mathbf{x}) = \frac{1}{B}\sum_{b=1}^{B} \hat{f}_b(\mathbf{x})$

A Random Forest requires no feature scaling, and its bootstrap resampling naturally provides out-of-bag estimates of model generalisation at no extra computational cost. Its principal limitation in this setting is that predictions cannot exceed the range of training targets: the model has no mechanism to extrapolate to price levels not encountered historically, which constrains performance during extreme scarcity events. Key hyperparameters are B, maximum tree depth, minimum samples per leaf, and m_{try}. Figure A.2 illustrates the parallel construction strategy, which contrasts with the sequential approach of XGBoost shown in Figure A.3.

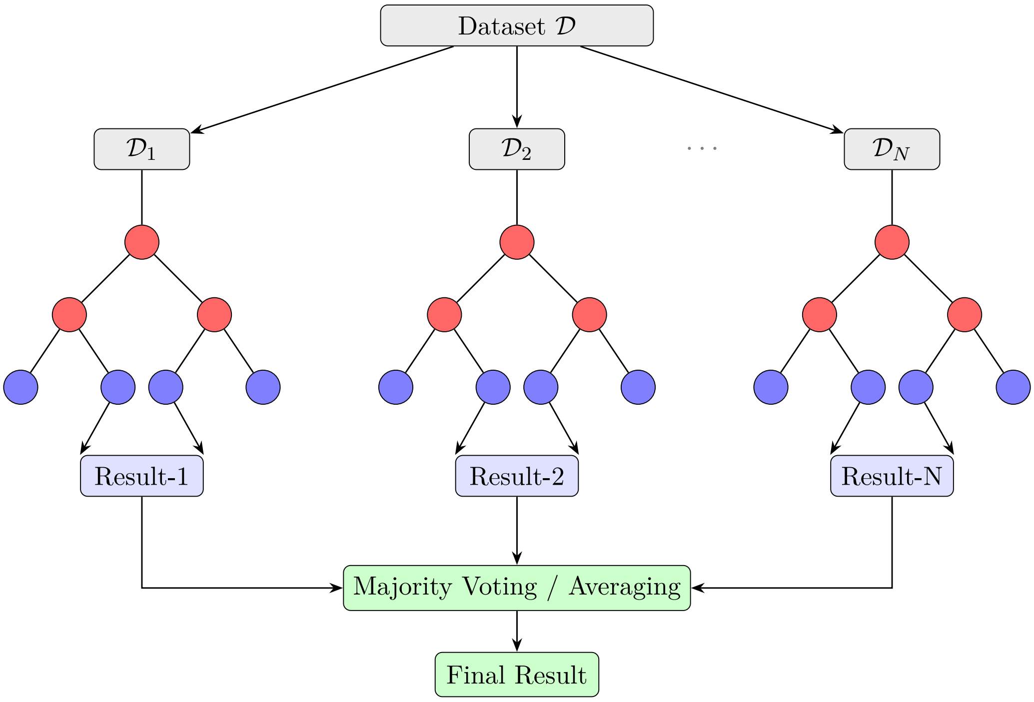

Figure A.2: A Random Forest trains B independent trees on bootstrap resamples and aggregates their predictions by averaging.

A.2.2 XGBoost

As illustrated in Figure A.3, the distinction between Random Forest and gradient boosting is one of construction strategy: where the former grows all trees independently and averages their predictions, the latter builds them one at a time, each tasked with correcting the systematic errors of the current ensemble. After K rounds, the prediction is the sum of K shallow trees, given by Equation (A.12)

$$\hat{y}_i^{(K)} = \sum_{k=1}^{K} \eta \, t_k(\mathbf{x}_i), \tag{A.12}$$

where $\hat{y}_i^{(K)}$ is the boosted prediction for observation i after K rounds, $\mathbf{x}_i \in \mathbb{R}^p$ its feature vector, $K \in \mathbb{N}$ the total number of boosting rounds, t_k the regression tree fitted at round k, and $\eta \in (0, 1]$ a shrinkage coefficient that scales each tree's contribution. Each t_k maps an input to the weight w_j at the leaf node $j = q(\mathbf{x}_i)$ it reaches under tree structure q. The training objective penalises both prediction error and model complexity

$$\mathcal{L}(\mathcal{D}) = \sum_{i=1}^{n} \ell(y_i, \, \hat{y}_i) + \sum_{k=1}^{K} \Big(\gamma \, T_k + \tfrac{1}{2} \mu \, \|\mathbf{w}_k\|_2^2 \Big), \tag{A.13}$$

where $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ is the training set of n observations, $\mathcal{L}(\mathcal{D})$ the regularised training objective, $\ell(y_i, \hat{y}_i)$ a differentiable per-example loss between target y_i and prediction $\hat{y}_i$, K the number of boosting rounds, T_k the leaf count of tree k, $\mathbf{w}_k$ the vector of leaf weights of tree k, $\|\cdot\|_2$ the Euclidean norm, and $\gamma \geq 0$, $\mu \geq 0$ penalise tree size and leaf-weight magnitude respectively [63], [73]. Rather than minimising (A.13) over all trees jointly, at each round k only the new tree t_k is fitted to the current residuals. Substituting $\hat{y}_i^{(k)} = \hat{y}_i^{(k-1)} + \eta\, t_k(\mathbf{x}_i)$ and applying a second-order Taylor approximation of ℓ around $\hat{y}_i^{(k-1)}$ converts the problem into a weighted least-squares objective in t_k, yielding the closed-form optimal leaf weight for node j

$$w_j^\star = -\frac{\displaystyle\sum_{i\in\mathcal{I}_j} g_i}{\displaystyle\sum_{i\in\mathcal{I}_j} h_i + \mu}, \tag{A.14}$$

where $w_j^\star$ is the closed-form optimal weight assigned to leaf j of the new tree t_k, g_i and h_i are the first- and second-order gradients of ℓ evaluated at the current prediction $\hat{y}_i^{(k-1)}$, $\mathcal{I}_j = \{i \mid q(\mathbf{x}_i) = j\}$ is the set of training observations reaching leaf j, and $\mu \geq 0$ is the leaf-weight regularisation parameter from (A.13) [73]. Substituting (A.14) back yields a scoring function for any candidate tree structure q, shown in Equation (A.15)

$$\widetilde{\mathcal{L}}^{(k)}(q) = -\frac{1}{2}\sum_{j=1}^{T} \frac{\left(\sum_{i\in\mathcal{I}_j} g_i\right)^2}{\sum_{i\in\mathcal{I}_j} h_i + \mu} + \gamma T. \tag{A.15}$$

Since exhaustive search over tree structures is intractable, XGBoost grows each tree greedily from a single root node. A split of node $\mathcal{I}$ into children $\mathcal{I}_L$ and $\mathcal{I}_R$ is accepted when the gain, defined in Equation (A.16)

$$\mathcal{G}_{\text{split}} = \frac{1}{2}\left[\frac{\left(\sum_{i\in\mathcal{I}_L} g_i\right)^2}{\sum_{i\in\mathcal{I}_L} h_i + \mu} + \frac{\left(\sum_{i\in\mathcal{I}_R} g_i\right)^2}{\sum_{i\in\mathcal{I}_R} h_i + \mu} - \frac{\left(\sum_{i\in\mathcal{I}} g_i\right)^2}{\sum_{i\in\mathcal{I}} h_i + \mu}\right] - \gamma \tag{A.16}$$

is positive, providing a natural pruning criterion [73]. Column and row subsampling, cache-aware computation, and parallel split search make XGBoost practical on the feature matrices assembled in this thesis. The key hyperparameters are summarised in Table A.1.

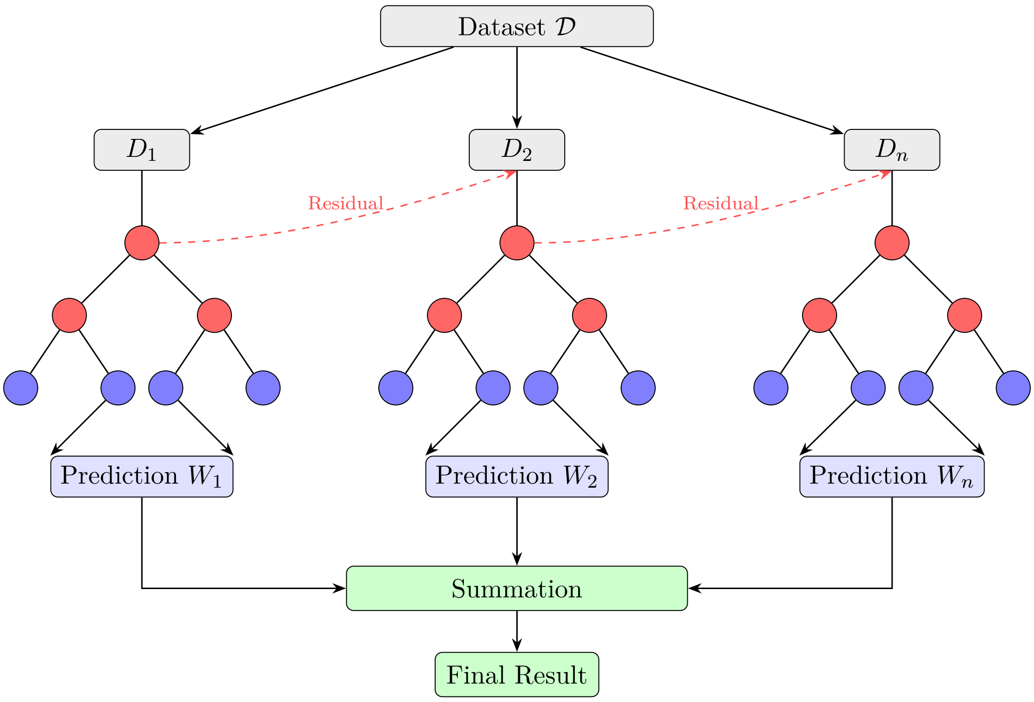

Figure A.3: XGBoost constructs trees sequentially: each tree fits the residuals of the previous one, with all predictions summed into the final output.

Table A.1: XGBoost hyperparameters and their roles.

Hyperparameter	**Effect**
`n_estimators`	Total boosting rounds K; more trees reduce training error but risk overfitting without adequate regularisation.
`learning_rate`	Shrinkage η per tree; smaller values generalise better but require more rounds to converge.
`max_depth`	Maximum tree depth; limits the order of feature interactions the model can represent.
`subsample`	Fraction of rows sampled per round; values below 1 add stochasticity and reduce overfitting.
`colsample_bytree`	Fraction of features considered per tree; reduces correlation between successive trees.
`reg_alpha`	ℓ_1 penalty on leaf weights (γ in (A.13)); encourages sparse tree structures.
`reg_lambda`	ℓ_2 penalty on leaf weights (μ in (A.13)); shrinks large leaf values.

A.3 Sequence Models

The tree-based models in Section A.2 treat each input feature vector $\mathbf{x}_i$ as a self-contained summary of an observation, with no explicit representation of how the system evolved to reach that state. Electricity prices are fundamentally sequential, however: the wind ramp earlier in the night, the temperature trajectory through the evening, and the trend in cross-border flows all carry information that a static feature vector cannot fully capture. This section introduces the two recurrent architectures evaluated in this thesis: the Gated Recurrent Unit (GRU) and the Long Short-Term Memory (LSTM). Both of these architectures are designed to exploit this kind of sequential structure. The vanilla RNN is presented first in Section A.3.1 as a foundation, since the gated mechanisms in Sections A.3.2 and A.3.3 are extensions of its core recurrence and are most easily understood as solutions to its limitations.

A.3.1 Recurrent Neural Architectures

The recurrent neural network (RNN) is presented here not as a candidate model in the empirical comparison but as the architectural baseline that the GRU and LSTM extend. Understanding why their gating mechanisms exist requires understanding what they fix.

A standard feedforward neural network composes affine transformations and elementwise nonlinearities to map an input vector to an output, with no memory of previous inputs. Its parameters are trained by backpropagation, which propagates the gradient of the loss back through each layer via the chain rule, updating each weight by a step in the negative gradient direction consistent with (A.3) [63], [66]. Two nonlinearities appear later in this section. The Rectified Linear Unit, Equation (A.17)

$$\mathrm{ReLU}(u) = \max(0,\, u), \tag{A.17}$$

is the standard hidden-layer activation: computationally cheap and effective at sustaining usable gradient magnitudes in deep architectures [66]. The logistic sigmoid,

$$\sigma(u) = \frac{1}{1 + e^{-u}}, \tag{A.18}$$

maps any real input to $(0, 1)$ and appears inside the gating mechanisms of the GRU and LSTM [66].

A *recurrent* neural network extends the feedforward architecture to sequential inputs by maintaining a hidden state vector $\mathbf{s}_t \in \mathbb{R}^d$ that accumulates information across time steps. Given input $\mathbf{x}_t$ at step t, the state update is given by Equation (A.19)

$$\mathbf{s}_t = \phi(\mathbf{V}\,\mathbf{x}_t + \mathbf{U}\,\mathbf{s}_{t-1} + \mathbf{c})\,, \tag{A.19}$$

where $\mathbf{s}_t \in \mathbb{R}^H$ is the hidden state at time step t, $\mathbf{s}_{t-1}$ its value at the previous step, $\mathbf{x}_t \in \mathbb{R}^p$ the input at step t, $\mathbf{V} \in \mathbb{R}^{H\times p}$ the input-to-state weight matrix, $\mathbf{U} \in \mathbb{R}^{H\times H}$ the recurrent state-to-state weight matrix, $\mathbf{c} \in \mathbb{R}^H$ the bias vector, and $\phi : \mathbb{R} \to \mathbb{R}$ a nonlinear activation applied element-wise. The recurrent weight matrix $\mathbf{U}$ connects each state to the previous one, so that $\mathbf{s}_t$ encodes a running summary of the entire input history $\mathbf{x}_1, \ldots, \mathbf{x}_t$. The same cell is applied repeatedly across time steps, passing its hidden state forward at each step. In the forecasting context of this thesis, a sequence of hourly weather observations and lagged prices forms the input, with a single price prediction being read from the final hidden state, representing the many-to-one configuration used throughout.

Training requires propagating the loss gradient backwards through the unrolled recurrence using an algorithm known as *backpropagation through time* (BPTT). Each backward step involves a multiplication by $\mathbf{U}$, and when the singular values of $\mathbf{U}$ are less than one these products shrink to zero exponentially with sequence length, a phenomenon known as the *vanishing gradient* problem [74]. When they exceed one, gradients grow without bound, leading to the *exploding gradient* problem [75], [76]. Both pathologies make it practically impossible for vanilla RNNs to learn dependencies spanning more than a few dozen time steps. In electricity forecasting, relevant dependencies easily span 24 hours or more: yesterday's wind pattern, the previous weekend's demand profile, or a multi-day anticyclone suppressing generation all carry information about current prices.

The two gated architectures evaluated in this thesis (the Gated Recurrent Unit and the Long Short-Term Memory) both address this limitation, with different trade-offs in parameter count and expressivity. They are introduced in turn below, and the choice between them is treated as a modelling decision resolved empirically (Section 5.1 and Section 5.2).

A.3.2 Gated Recurrent Unit

The Gated Recurrent Unit (GRU) [77] addresses gradient vanishing by augmenting the vanilla RNN state update with two multiplicative gates that learn to control what information is preserved and what is discarded at each step. At time t, given input $\mathbf{x}_t$ and previous state $\mathbf{s}_{t-1}$, the GRU computes

$$\mathbf{r}_t = \sigma(\mathbf{W}_r\mathbf{x}_t + \mathbf{U}_r\mathbf{s}_{t-1} + \mathbf{b}_r)\,, \tag{A.20}$$

$$\mathbf{z}_t = \sigma(\mathbf{W}_z\mathbf{x}_t + \mathbf{U}_z\mathbf{s}_{t-1} + \mathbf{b}_z)\,, \tag{A.21}$$

$$\tilde{\mathbf{s}}_t = \tanh(\mathbf{W}_s\mathbf{x}_t + \mathbf{U}_s(\mathbf{r}_t \odot \mathbf{s}_{t-1}) + \mathbf{b}_s)\,, \tag{A.22}$$

$$\mathbf{s}_t = (\mathbf{1} - \mathbf{z}_t) \odot \mathbf{s}_{t-1} + \mathbf{z}_t \odot \tilde{\mathbf{s}}_t, \tag{A.23}$$

where $\mathbf{r}_t \in (0,1)^H$ is the reset gate controlling how much of the previous state contributes to the candidate update, $\mathbf{z}_t \in (0,1)^H$ the update gate (A.21) controlling the interpolation between previous and candidate state, $\tilde{\mathbf{s}}_t \in \mathbb{R}^H$ the candidate hidden state (A.22), $\mathbf{s}_t \in \mathbb{R}^H$ the new hidden state at step t, $\mathbf{W}_{(\cdot)} \in \mathbb{R}^{H\times p}$ and $\mathbf{U}_{(\cdot)} \in \mathbb{R}^{H\times H}$ learned weight matrices with subscripts r, z, s indicating the reset, update, and candidate-state components respectively, $\mathbf{b}_{(\cdot)} \in \mathbb{R}^H$ the corresponding bias vectors, $\mathbf{1} \in \mathbb{R}^H$ the all-ones vector, σ the logistic sigmoid (A.18) applied element-wise, $\odot$ the element-wise (Hadamard) product, and tanh the hyperbolic tangent applied element-wise, Equation (A.24)

$$\tanh(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}}. \tag{A.24}$$

The internal structure is illustrated in Figure A.4.

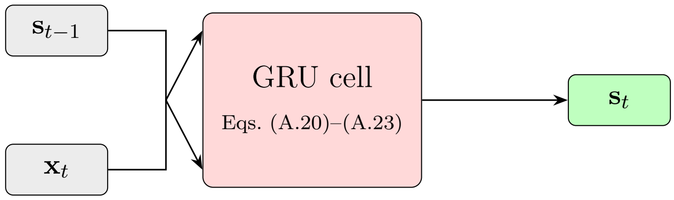

Figure A.4: GRU cell. The previous hidden state $\mathbf{s}_{t-1}$ and current input $\mathbf{x}_t$ are combined through the gating mechanism defined by Equations (A.20)–(A.23) to produce the new hidden state $\mathbf{s}_t$.

Each gate answers a specific question about what the state update should do. The *reset gate* $\mathbf{r}_t$ asks: how much of the past state is relevant when computing a new candidate? When $\mathbf{r}_t \approx \mathbf{0}$, the candidate $\tilde{\mathbf{s}}_t$ is formed almost entirely from the current input $\mathbf{x}_t$, effectively reinitialising the state which is useful when an abrupt event such as a sudden price spike or an unexpected generator outage breaks the continuity of the time series. The *update gate* $\mathbf{z}_t$ then asks: how much of the old state should carry forward versus be replaced by the candidate? Values near one adopt the fresh candidate while values near zero preserve the existing state, allowing the model to maintain slowly evolving context (such as a persistent anticyclone suppressing wind generation) across many time steps without overwriting it at each hour. The additive interpolation in (A.23) creates a direct gradient path from $\mathbf{s}_t$ back to $\mathbf{s}_{t-1}$, substantially alleviating gradient vanishing relative to the vanilla RNN [77], [78].

The GRU has no separate cell state and one fewer gate than the LSTM introduced below, resulting in a smaller parameter count. This parsimony reduces overfitting risk when labelled sequences are limited and lowers training time, while retaining the ability to capture the medium-range temporal dependencies most relevant to day-ahead price forecasting. Prior work on energy-related time series has found GRUs to be competitive with LSTMs in data-constrained industrial settings [78], [79].

A.3.3 Long Short-Term Memory

The Long Short-Term Memory network [80] takes the idea of learned gating further by introducing a second state variable alongside the hidden state $\mathbf{s}_t$. The *cell state* $\mathbf{c}_t \in \mathbb{R}^d$ acts as a long-term memory channel: it flows through time with only targeted, gate-controlled modifications, shielded from the repeated nonlinear transformations that erode gradient signals in vanilla RNNs. The forget, input, and output gates collectively determine what is written to, preserved in, and read from this channel, as illustrated in Figure A.5. At time t, given $\mathbf{x}_t$, $\mathbf{s}_{t-1}$, and $\mathbf{c}_{t-1}$

$$\mathbf{f}_t = \sigma(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f \mathbf{s}_{t-1} + \mathbf{b}_f)\,, \tag{A.25}$$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i \mathbf{s}_{t-1} + \mathbf{b}_i)\,, \tag{A.26}$$

$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c \mathbf{x}_t + \mathbf{U}_c \mathbf{s}_{t-1} + \mathbf{b}_c)\,, \tag{A.27}$$

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t, \tag{A.28}$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o \mathbf{s}_{t-1} + \mathbf{b}_o)\,, \tag{A.29}$$

$$\mathbf{s}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t), \tag{A.30}$$

where $\mathbf{f}_t, \mathbf{i}_t, \mathbf{o}_t \in (0,1)^d$ are the forget, input (A.26), and output (A.29) gates respectively, $\tilde{\mathbf{c}}_t \in \mathbb{R}^d$ the candidate cell update (A.27), $\mathbf{c}_t \in \mathbb{R}^d$ the cell state, and $\mathbf{s}_t \in \mathbb{R}^d$ the hidden state at step t. The weight matrices $\mathbf{W}_{(\cdot)} \in \mathbb{R}^{d\times p}$ and $\mathbf{U}_{(\cdot)} \in \mathbb{R}^{d\times d}$, together with the bias vectors $\mathbf{b}_{(\cdot)} \in \mathbb{R}^d$, carry subscripts f, i, c, o indicating the forget, input, candidate-cell, and output components respectively. The activations σ, tanh, and the Hadamard product $\odot$ are as introduced in Section A.3.2.

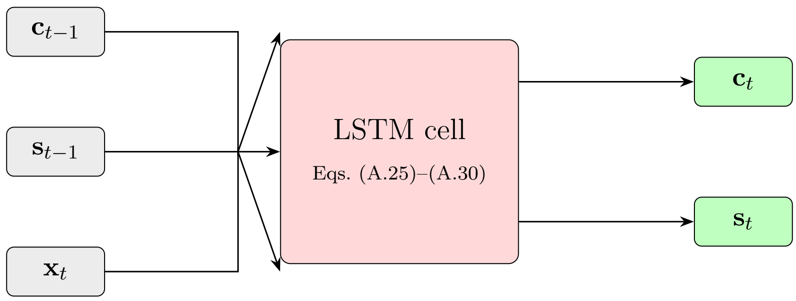

Figure A.5: LSTM cell. The previous cell state $\mathbf{c}_{t-1}$, previous hidden state $\mathbf{s}_{t-1}$, and current input $\mathbf{x}_t$ are processed through the gating mechanism defined by Equations (A.25)–(A.30) to produce the new cell state $\mathbf{c}_t$ and hidden state $\mathbf{s}_t$.

The roles of the three gates map naturally onto the temporal structure of electricity markets. The *forget gate* $\mathbf{f}_t$ multiplies the previous cell state element-wise, selectively erasing content that is no longer relevant . For example, discarding the memory of a cold spell once temperatures recover, or dropping the context of a prior wind drought once generation has normalised. The *input gate* $\mathbf{i}_t$ controls how much of the newly computed candidate $\tilde{\mathbf{c}}_t$ is written into the cell, governing how strongly the current hour's observations update the long-term memory. Finally, the *output gate* $\mathbf{o}_t$ determines which aspects of the cell state are surfaced as the hidden state $\mathbf{s}_t$ passed forward to the prediction head or subsequent layers.

The key theoretical property is the additive cell update in (A.28): the gradient of the loss with respect to $\mathbf{c}_t$ flows back to $\mathbf{c}_{t-1}$ through only a multiplicative rescaling by $\mathbf{f}_t$, rather than through repeated nonlinear squashing. This creates a near-unobstructed gradient highway through time [66], [80], thereby enabling the LSTM to learn dependencies spanning hundreds of time steps. Consequently, the model successfully captures multi-day weather patterns, weekly demand cycles, and seasonal generation regimes, all of which are highly relevant to Nordic spot price dynamics.

Relative to the GRU, the LSTM's added expressivity comes at the cost of a larger parameter count: three gate matrices per layer instead of two, plus the separate cell state. This increases both training time and the risk of overfitting when training sequences are short. In practice, GRU and LSTM achieve comparable accuracy on medium-sized datasets [78], and the choice between them is treated as a modelling decision resolved by cross-validation across both architectures.

Appendix B

Weather Forecast Variables and Data Gathering Locations

This appendix lists the meteorological variables retrieved from each forecast system and the geographic locations at which historical weather forecasts and actual weather observations are collected for the three input signals modelled: wind power production, solar power production, and electricity consumption. For every location, hourly forecasts are pulled from two numerical weather prediction systems: the MET Nordic NWP, which provides hourly leads from 0 to 58 h, and the Global Forecast System (GFS) operated by NOAA, which provides 3-hourly leads from 60 to 240 h. The two streams are stitched into a single continuous forecast as described in Section 4.2.3.

B.1 Forecast Variables

Tables B.1 and B.2 list the variables retrieved from the MET Nordic and GFS systems respectively, as referenced in Section 4.2.1.

Table B.1: Meteorological variables from MET Nordic weather forecast.

Application Programming Interface (API) variable ID	Description	Unit
`air_temperature_2m`	Air temperature at 2 m above ground level	K
`cloud_area_fraction`	Fraction of the sky covered by clouds	0–1
`integral_of_surface_downwelling_shortwave_flux_in_air_wrt_time`	Hourly integrated incoming solar radiation at the surface	$\mathrm{W\,s\,m^{-2}}$
`wind_speed_10m`	Average wind speed at 10 m height	$\mathrm{m\,s^{-1}}$
`wind_direction_10m`	Average wind direction at 10 m	degree
`wind_speed_of_gust`	Maximum wind gust speed	$\mathrm{m\,s^{-1}}$

B.2 Data Gathering Locations

The locations are organised by input signal. Solar power production and electricity consumption are both modelled at municipality level. The 12 regions and 24 municipalities representing SE3 listed in Table B.3. The cities selected to capture the major urban centres

Table B.2: Meteorological variables from GFS weather forecast.

`API variable ID`	**Description**	**Unit**
`TMP`	Temperature	K
`TCDC`	Total Cloud Cover	%
`DSWRF`	Downward Shortwave Radiation Flux	$\mathrm{W\,m^{-2}}$
`UGRD`	U-Component of Wind (Eastward)	$\mathrm{m\,s^{-1}}$
`VGRD`	V-Component of Wind (Northward)	$\mathrm{m\,s^{-1}}$

and residential areas that account for the bulk of solar generation and electricity demand in the bidding zone, following the population- and capacity-weighted aggregation described in Sections 4.2.4 and 4.2.5. Wind power production is instead modeled at turbine-cluster level, since the relevant atmospheric variables — 10 m wind speed, gust, and direction — exhibit substantial sub-municipality variation driven by terrain and proximity to the coast. The 72 SE3 clusters listed in Table B.4 are sourced from the Swedish Energy Agency's Vindbrukskollen register [40] and together cover approximately 85 % of total installed onshore wind capacity in SE3.

Table B.3: Complete aggregation structure. Column "Län Demand" is the average electricity demand over 2020 and 2021 per region which are the most recent years when confirmed consumption data is available for all regions. The data in the column "Kommun Demand" are from 2024, the most recent when the data was downloaded. City fractions are normalized within each region; single-city regions have fraction equal to 1.

Län	**Län Demand (MWh)**	**Län frac.**	**Kommun**	**Kommun Demand (MWh)**	**Kommun frac.**	**Lat. (°N)**	**Lon. (°E)**
Västra Götaland	18,628,313	0.230	Göteborg[1]	8,314,135	0.748	57.69460	11.92045
			Trollhättan/Vb.	1,193,908	0.107	58.30418	12.33562
			Borås	859,189	0.077	57.72277	12.94568
			Skövde	743,480	0.067	58.39777	13.83382
Jönköping	4,459,363	0.055	Jönköping	1,150,873	0.675	57.77800	14.20884
			Eksjö/Nässjö	554,516	0.325	57.64281	14.90141
Örebro	3,959,415	0.049	Örebro	1,130,362	0.673	59.27608	15.19751
			Karlskoga[2]	550,550	0.327	59.32624	14.52629
Östergötland	6,324,844	0.078	Linköping	1,179,043	0.237	58.40400	15.62396
			Norrköping	3,152,918	0.634	58.60009	16.19443
			Motala	353,587	0.071	58.54316	15.04277
			Mjölby	286,760	0.058	58.32844	15.11620
Södermanland	3,321,908	0.041	Nyköping	430,419	0.250	58.75366	17.01555
			Eskilstuna	925,879	0.538	59.37734	16.50096
			Strängnäs	364,508	0.212	59.37252	17.02731
Uppsala	3,269,024	0.040	Uppsala	1,556,003	0.809	59.86213	17.64376
			Enköping	366,693	0.191	59.64870	17.10332
Västmanland	2,824,300	0.035	Västerås	1,252,260	0.725	59.61394	16.54244
			Arboga/Köping	476,150	0.275	59.47171	15.91921
Stockholm	20,219,301	0.250	Stockholm	–	1.000	59.36317	18.11892
Gotland	925,326	0.011	Visby	–	1.000	57.62902	18.32727
Dalarna	6,218,628	0.077	Falun/Borlänge	–	1.000	60.54034	15.55410
Gävleborg	5,066,453	0.063	Gävle	–	1.000	60.66751	17.14708
Värmland	5,642,296	0.070	Karlstad	–	1.000	59.38721	13.50788

Table B.4: Wind power production: the 72 SE3 turbine clusters at which MET Nordic and GFS forecasts of 10 m wind speed, gust, and direction are collected. Clusters formed by aggregating individual turbines within a radius of 10 km and are ordered by installed capacity (P_{inst}). The Cluster column reproduces the verbatim *Projekteringsområde* field from Vindbrukskollen [40] and lists every underlying wind park grouped into the cluster. Hub heights are mean values across the turbines in each cluster; entries shown as "—" indicate clusters for which the hub-height field is missing in the source register. Geographic coordinates are given in decimal degrees (WGS 84).

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
Bruzaholm Vindkraftpark + Grevekulla	Eksjö	Jönköpings län	57.7063	15.2255	175.8	27	146
Svartnäs + Enviksberget	Falun	Dalarnas län	60.8422	16.1639	152.1	41	111
Bäckhammar + Krontorp	Kristinehamn	Värmlands län	59.1633	14.2610	131.1	31	130
Tormoseröd Vindpark + Tolvmanstegen	Tanum	Västra Götalands län	58.8761	11.4283	116.6	33	112
Grönhult Vindkraftpark + Grimsås	Tranemo	Västra Götalands län	57.4199	13.5045	114.0	25	120
Målarberget	Norberg	Västmanlands län	60.0970	16.1182	113.4	27	125

continued on next page

[1]Göteborg aggregation includes the municipalities: Härryda, Partille, Stenungsund, Tjörn, Ale, Lerum, Göteborg, Mölndal, Kungälv, and Alingsås.
[2]Karlskoga aggregation includes Degerfors.

Table B.4 continued from previous page

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
Näsudden Väst N1 + Stugylparken S1 + Näsudden Vindkraftpark + Gansparken 1 + Kulle vindpark + Skåls vindpark + Skålsparken Väst + Grötlingbo Uddvide Domerarve + Näs Levide Selma + Havdhem Bols Ryftes Hulda [Kulle 2] + Näs Amfunds Snaigsto + Grötlingbo Kattlunds + Näs Sigsarve Sladdkvenni	Gotland	Gotlands län	57.0989	18.2509	112.5	68	65
Årjäng SV + Årjäng NV etapp 2 + Årjäng NV	Årjäng	Värmlands län	59.3985	12.0442	98.0	30	128
Älgkullen	Smedjebacken	Dalarnas län	60.0150	15.1677	93.0	15	119
Norra vedbo (Jönköping/Aneby)	Aneby	Jönköpings län	57.8240	14.6189	86.0	20	130
Rödene	Alingsås	Västra Götalands län	57.9829	12.5170	85.8	13	123
Lyrestad + Lyrestad II + Fåleberg + Börstorp	Mariestad	Västra Götalands län	58.7823	13.9889	81.9	26	129
Zinkgruvan + Gärdshyttan + Södra Kärra + Alva af Hulta	Askersund	Örebro län	58.7502	15.0450	81.0	26	105

continued on next page

Table B.4 continued from previous page

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
Velinga Vindkraftpark + Råmmarehemmet	Tidaholm	Västra Götalands län	58.0948	14.0248	79.8	15	106
Vindpark Kingebol + Fröskog Vindpark + Söderbodane + Hässelbacka	Åmål	Västra Götalands län	58.9714	12.5653	79.7	18	99
Skallberget/Utterberget	Avesta	Dalarnas län	60.3342	16.4885	79.1	12	115
Sättravallen + Långmarken	Kristinehamn	Värmlands län	59.3216	14.2091	74.4	24	125
Hultema + Björkö	Motala	Östergötlands län	58.7935	15.3289	73.4	13	98
Häjsberget och södra Länsmansberget + Norra Länsmansberget	Sunne	Värmlands län	59.7839	12.9723	71.9	16	126
Riskebo + Tjärnäs	Hedemora	Dalarnas län	60.5176	16.1607	65.2	11	118
Skaveröd/Gurseröd + Mungseröd/Gurseröd/Skaveröd + Vindpark Tågeröd + Naverstads-Tyft Edsäm	Tanum	Västra Götalands län	58.7641	11.4230	64.3	24	95
Kronoberget	Lekeberg	Örebro län	59.2123	14.6842	60.8	16	117
Vindpark Femstenaberg + Vettebergets Vindpark + Neanberg/VIK	Strömstad	Västra Götalands län	58.8862	11.2845	59.1	12	109

continued on next page

Table B.4 continued from previous page

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
Lyckås	Jönköping	Jönköpings län	57.8488	14.3682	58.5	13	97
Vindkraftpark Fågelås + Prästbolet + Nyskog + Stakahemmet	Hjo	Västra Götalands län	58.2393	14.1666	51.6	11	138
Varsvik	Norrtälje	Stockholms län	60.0798	18.5579	51.0	17	129
Knöstad	Säffle	Värmlands län	59.2622	12.8871	49.6	8	115
Töftedalsfjället	Dals-Ed	Västra Götalands län	58.8545	11.8338	48.3	21	100
Stöllsäterberget	Torsby	Värmlands län	60.4557	13.4739	47.0	8	118
Lervik	Västervik	Kalmar län	57.6463	16.2484	46.2	7	115
Hornamossen + Ekekullen Dintestorp	Habo	Jönköpings län	58.0016	13.8836	45.9	12	106
Saxberget + Fjällberget + Paljakoberget	Ludvika	Dalarnas län	60.1064	14.8642	45.8	21	98
Sötterfällan + Gunillaberg	Jönköping	Jönköpings län	57.7793	13.8653	44.8	14	114

continued on next page

Table B.4 continued from previous page

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
Fryksdalshöjden	Sunne	Värmlands län	59.6923	13.0726	43.4	7	115
Rämsberget + Korpfjället	Malung-Sälen	Dalarnas län	60.3672	13.9255	41.7	16	108
Laxåskogen + Norra Hunna	Laxå	Örebro län	58.9374	14.7545	41.6	11	126
Gummaråsen + Sotared + Hässlås Morups-Lunnagård + Munkagård + Tvååker + Bäck-agård + Kärrets Gård II + Kärrets Gård	Varberg	Hallands län	57.0194	12.4366	41.5	24	81
Fredriksdal + Tapplarp + Topperyd 3 och 4 + Hallhult + Topperyd 1 och 2	Nässjö	Jönköpings län	57.6328	14.5561	40.8	24	103
Hästhalla Basebo + KILAGÅRDEN /7306-1 + NYCKELTORP / 7308-2 + KARSTORP / 7044-3 + Stenkulla + Stenbrona-Saleby + Tagelberg 1 + Tagelberg 2 + Lanneholm + Lanna 1, Lidköping + Jung	Skara	Västra Göta-lands län	58.3518	13.1898	39.9	22	92

continued on next page

Table B.4 continued from previous page

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
VIGLUNDA / 7307 + Västermark + MUNSTORP 7043 + Stora lund II + Stockebäck I + NEDERSÅNNA / 7042 + NATTORP 7303 + Erikstorp III + Kollbogården II + Kollbogården III + Västermark I + MÄRENE 1992 + Stora lund I + VIGLUNDA / 4456 + Erikstorp II + ISTRUMS-LYCKE / 5695 + Dalaholm II + Dalaholm I + ÖLANDA / 4391 + Kollbogården I + MUNSTORP 7012 + MUNSTORP 5691 + RANÅKER / 747 + Erikstorp I	Skara	Västra Götalands län	58.4554	13.5041	39.3	28	87
Prässebergens vindpark + Vindkraftpark Högen + Sköllunga + Sannersby	LillaEdet	Västra Götalands län	58.1307	12.0661	38.7	16	102
Anneberg + Källeberg + Orreholmen + Skörstorp + Lammevad + Källeberg 6 + Hångsdala Vindkraftverk (Ann-El-i) + Åsle	Tidaholm	Västra Götalands län	58.1372	13.7370	35.0	17	91
Vindpark Lursäng + Kil + Torseröds Vindkraftpark 5 + Tannam + Torseröds Vindkraftpark 6 + Tannam Smeby	Tanum	Västra Götalands län	58.7993	11.2781	34.7	16	77

continued on next page

Table B.4 continued from previous page

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
Ås + Svarteborgs-Skogen + Hällevadsholm 2 och 3 + Hällevadsholm Väster	Munkedal	Västra Göta-lands län	58.5520	11.5254	34.7	15	107
Iglasjön + Dal	Kungsbacka	Hallands län	57.4007	12.2262	34.4	12	96
Lebo	Västervik	Kalmar län	57.6151	16.4560	33.0	5	—
Orrberget/Stensvedberget	Ludvika	Dalarnas län	60.3586	15.1003	32.4	9	92
Sundby Vindpark	Eskilstuna	Södermanlands län	59.4096	16.6372	32.4	9	87
Jung-Åsa + Uvereds Vindkraftförening + Närefors 1 Norr + Härjevad + Rockagården + Bränneberg + Rytterås + St Levene + Härjevads Vind + Helleberg 1 + Månsagården	Vara	Västra Göta-lands län	58.3346	13.0196	31.8	23	78
Vindpark Vänern + Lucia, Boholmen	Karlstad	Värmlands län	59.2960	13.4087	30.5	11	86
Brahehus + Tuggarpsgruppen + Adelövs Lantbruk	Jönköping	Jönköpings län	58.0494	14.5862	30.2	14	95
Hedbodberget Etapp 1	Rättvik	Dalarnas län	61.1773	15.3849	30.0	15	90
Röbergsfjället + Köjkeberget + Silkomhöjden	Vansbro	Dalarnas län	60.2576	14.2101	28.0	14	88

continued on next page

Table B.4 continued from previous page

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
Vettåsen/Finnberget + Mårtensklack	Sandviken	Gävleborgs län	60.7977	16.6240	27.6	12	108
Näsby/Överberg + Örbacken + Hyppinge Gårdsverk	Mjölby	Östergötlands län	58.2733	15.1793	27.5	12	90
Västergården 1 + Brålanda-Torp + Erikstads-Björnebol + Fågelbacka + Tängelsbol + Erikstads-Mossebol 2 + Erikstads-Mossebol 1 + Vena	Mellerud	Västra Göta-lands län	58.6144	12.4074	27.2	16	96
Ramsnäs + Slottsbol	Laxå	Örebro län	58.8552	14.4457	25.4	13	97
Boarp	Vaggeryd	Jönköpings län	57.4441	14.2326	25.0	4	119
Fågelås-Torp + Bossgården (Galneryd) + Getaryggen, Söderryd, Stora Björstorp + Skalleberg och Dunabolet + Signesbo + Hjällö + Simona Vindkraft Leringen + Granbolet	Hjo	Västra Göta-lands län	58.1698	14.1609	24.5	15	89
Vindpark Ränsliden	Mellerud	Västra Göta-lands län	58.8372	12.3089	24.0	7	135
Slageryd	Vetlanda	Jönköpings län	57.2954	14.9092	23.4	6	114

continued on next page

Table B.4 continued from previous page

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
Fröreda	Hultsfred	Kalmar län	57.4289	15.5392	23.1	7	117
Kyrkberget	Mora	Dalarnas län	60.8888	13.6190	23.0	10	90
Rosenskog + Västtorp + Rosenskog A1 + Rosenskog A2 + Göteve 1	Falköping	Västra Götalands län	58.1168	13.3991	22.9	7	89
Bösjövarden	Mora	Dalarnas län	61.3090	14.3269	22.5	9	100
Mässingberget	Orsa	Dalarnas län	61.3204	14.6060	22.0	11	95
Storugns vindpark + Lärbro Nors Norrvange Nr 1 & 2 + Storugns 10 + Hangvar Kyrkebys, Kyrkebys 3 & 4 + Hangvar Kyrkebys 1 + Tängelgårde + Hangvar Kyrkebys 2 + Norrvange	Gotland	Gotlands län	57.8225	18.7524	22.0	20	72
Hån Vindpark	Årjäng	Värmlands län	59.5065	11.7390	21.0	5	125
Gunnarby + RÅLANDA	Uddevalla	Västra Götalands län	58.3838	11.7091	20.8	11	99
SIMMATORP / 7835 + BLOMBACKA / 7837-1 + HORSHAGA / 7309 + ENTORP / 5690-1 + Sävedalen	Skara	Västra Götalands län	58.3630	13.3497	20.8	10	98
Munkhyttan	Lindesberg	Örebro län	59.6009	15.1141	18.6	3	139

continued on next page

Table B.4 continued from previous page

Cluster (*Projekteringsområde*)	**Kommun**	**Län**	**Lat. (°N)**	**Lon. (°E)**	P_{inst} **(MW)**	N_{turb}	**Hub h. (m)**
Dållebo	Ulricehamn	Västra Göta-lands län	57.8108	13.1670	18.0	4	105

Appendix C

Holidays Included in Consumption Modeling

This appendix presents the specific Swedish public holidays that were incorporated into the consumption modelling in Table C.1 and C.2. Because these dates exhibit significantly different electricity demand patterns compared to standard workdays, explicitly accounting for them was necessary to maintain forecasting accuracy.

Table C.1: Fixed-Date Holidays Included for Consumption Model

Holiday name	**Calendar date**	**Model identifier**
New Year's Day	January 1	`new_year`
Epiphany	January 6	`epiphany`
May Day	May 1	`may_day`
National Day	June 6	`national_day`
Christmas Eve	December 24	`christmas_eve`
Christmas Day	December 25	`christmas_day`
Boxing Day	December 26	`boxing_day`
New Year's Eve	December 31	`new_year_eve`

Table C.2: Movable Holidays Included in the Model (Easter-Based)

Holiday name	**Date definition**	**Model identifier**
Good Friday	2 days before Easter Sunday	`good_friday`
Easter Eve	1 day before Easter Sunday	`easter_eve`
Easter Sunday	Easter Sunday	`easter_sunday`
Easter Monday	1 day after Easter Sunday	`easter_monday`
Ascension Day	39 days after Easter Sunday	`ascension_day`

Appendix D

Mathematical Formulations for Input Models

D.1 Mathematical Formulations for Wind Features

This appendix details the physical domain modeling and mathematical formulations used to construct the uncalibrated regional wind baseline discussed in Section 4.2.3.

D.1.1 Cluster Metadata and Parameters

The physical pipeline converts raw NWP wind variables into preliminary megawatt estimates at the cluster level. Each of the 72 SE3 clusters is characterized by its installed capacity (MW), geographic county (*Län*), and hub height per turbine. These determine three core physics parameters:

- **Terrain roughness (z_0):** Assigned via a lookup table reflecting the dominant land cover of the cluster's *Län*. Values range from $z_0 = 0.03\,\mathrm{m}$ on Gotland and the southwestern coast to $z_0 = 0.25\,\mathrm{m}$ in the dense boreal forests of northern Sweden.
- **Turbine family:** Inferred from hub height. Turbines below 70 m are classified as `legacy` (pre-2000 era), 70–95 m as `onshore_mid`, 95–120 m as `onshore_modern`, and above 120 m as `onshore_large`.
- **Loss factors:** Availability (η_c^{avail}) and wake-loss (ℓ_c^{wake}) factors are fixed at 96 % and 8 % respectively for modern turbine families, with slightly relaxed values for legacy turbines.

D.1.2 Gust Correction

The NWP 10-metre wind speed is a one-hour mean, which understates the effective aerodynamic loading near cut-out wind speeds where sub-hourly gusts can briefly saturate the power curve. A gust correction is applied prior to hub-height extrapolation, as in Equation (D.1)

$$u_{\text{eff}} = u_{\text{mean}} + \alpha_g \max(0,\, u_{\text{gust}} - u_{\text{mean}}), \tag{D.1}$$

where u_{mean} is the NWP one-hour mean 10-metre wind speed, u_{gust} is the corresponding gust speed, u_{eff} is the gust-corrected effective wind speed, and $\alpha_g = 0.10$ is the gust-correction coefficient. This ensures that 10 % of the gust excess above the hourly mean is added to the effective speed, improving behavior near cut-out conditions without overweighting transient peaks.

D.1.3 Hub-Height Extrapolation

The corrected 10-metre wind speed is extrapolated to hub height using the logarithmic wind profile (IEC 61400), which assumes neutral atmospheric stability. This serves as a reasonable approximation for hourly NWP averages, as in Equation (D.2)

$$U(z_{\text{hub}}) = u_{\text{eff}} \frac{\ln(z_{\text{hub}}/z_0)}{\ln(z_{\text{ref}}/z_0)}, \tag{D.2}$$

where z_{hub} is the turbine hub height, $z_{\text{ref}} = 10\,\text{m}$ is the reference measurement height, z_0 is the terrain roughness length, and $U(z_{\text{hub}})$ is the extrapolated hub-height wind speed. The amplification factor depends heavily on local geography: a modern turbine at a 120 m hub height over coastal terrain ($z_0 = 0.05\,\text{m}$) sees approximately 1.55× the 10-metre wind speed, while the same hub height over northern Swedish forest ($z_0 = 0.25\,\text{m}$) yields only a 1.34× amplification.

D.1.4 Power Curve Lookup and Spatial Aggregation

Hub-height wind speed is converted to a normalized capacity factor $C_p \in [0, 1]$ using empirical power curves tailored to the four turbine families (Figure D.1). All curves share a cut-in wind speed of approximately 3 m/s and a hard cut-out at 25 m/s. However, older turbines reach rated power at lower wind speeds due to their smaller rotor diameters.

The net MW estimate for cluster c at time t is computed as in Equation (D.3)

$$\hat{P}_{c,t} = C_p(U_{c,t}) \times P_c^{\text{inst}} \times \eta_c^{\text{avail}} \times (1 - \ell_c^{\text{wake}}), \tag{D.3}$$

where $C_p(\cdot)$ is the family-specific normalized power curve, P_c^{inst} is the cluster's installed nameplate capacity, η_c^{avail} is its availability factor, ℓ_c^{wake} is the wake loss factor, and $\hat{P}_{c,t}$ is the resulting net power estimate.

Wind production is then aggregated across the 72 SE3 turbine clusters as a direct per-cluster sum to formulate the system-wide raw forecast, Equation (D.4)

$$\hat{P}_t^{\text{raw}} = \sum_c \hat{P}_{c,t}, \tag{D.4}$$

Because each summand carries its cluster's installed capacity directly through the physics step, this per-cluster sum naturally coincides with a capacity-weighted aggregate.

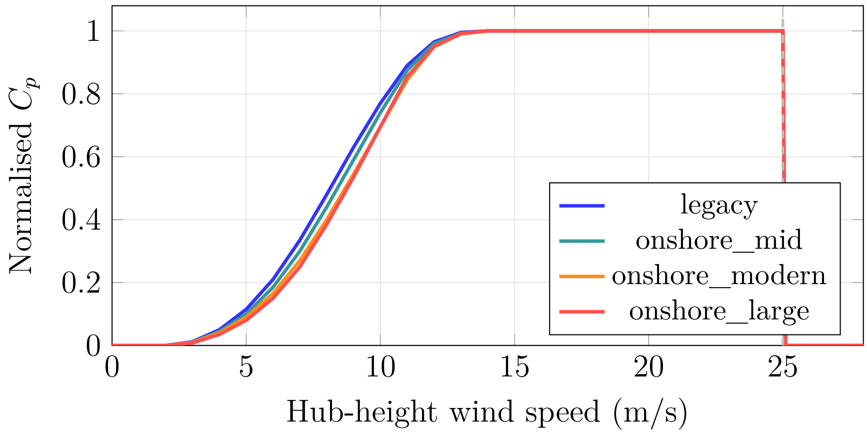

Figure D.1: Normalised power curves for the four Swedish turbine generations. Older turbines reach rated power at lower wind speeds; all families share a hard cut-out at 25 m/s.

D.2 Solar Features

This appendix provides the explicit mathematical definitions underpinning the physical baseline modeling and feature engineering for the solar production forecasts detailed in Section 4.2.4.

D.2.1 Astronomical Seasonality Profiles

Synthetic geometric features are utilized to make the deterministic astronomical availability of solar energy explicit to the model. Two Gaussian profiles are constructed on a circular timeline using the distance function $D(x, p, T) = \min(|x - p|, T - |x - p|)$.

The annual profile is centred on the summer solstice ($p = 172$, $T = 365.25$ days) with an empirically tuned standard deviation of $\sigma_{\text{annual}} = 78$ days:

$$S_{\text{annual}} = \max\left(0, \frac{\exp\left(-\frac{1}{2}\left(\frac{D(d,172,365.25)}{78}\right)^2\right) - \min_{\text{annual}}}{1 - \min_{\text{annual}}}\right), \tag{D.5}$$

where $d \in \{1, \dots, 365\}$ is the day of the year, $\min_{\text{annual}}$ is the empirical minimum of the unscaled Gaussian over the year, and $S_{\text{annual}} \in [0, 1]$ is the resulting min-max scaled annual availability profile.

The intraday profile is centred on solar noon ($p = 12$, $T = 24$ hours) with a dynamic standard deviation $\sigma_{\text{daily}} = H_d/4$. The daylight length H_d (in hours) at latitude $\approx 57.7°$N varies throughout the year according to:

$$H_d = 12.25 + 5.75 \cos\left(2\pi \frac{d - 172}{365.25}\right). \tag{D.6}$$

Using this daylight length, the intraday profile is defined as:

$$S_{\text{intraday}} = \max\left(0,\ \frac{\exp\left(-\frac{1}{2}\left(\frac{D(h,12,24)}{\sigma_{\text{daily}}}\right)^2\right) - \min_{\text{daily}}}{1 - \min_{\text{daily}}}\right), \tag{D.7}$$

where $h \in \{0, \ldots, 23\}$ is the hour of the day, $\min_{\text{daily}}$ is the empirical minimum of the unscaled Gaussian over the day, and $S_{\text{intraday}} \in [0, 1]$ is the resulting min-max scaled intraday availability profile.

Both profiles are min-max scaled to $[0, 1]$ via their respective minimum constants. Their product serves as the final theoretical solar-availability proxy:

$$S_{\text{combined}} = S_{\text{annual}} \times S_{\text{intraday}}. \tag{D.8}$$

D.2.2 Spatial Aggregation Weights

When spatially aggregating meteorological inputs, the region-to-SE3 weights are strictly proportional to installed solar capacity:

$$w_r = \frac{C_r}{\sum_{i \in R} C_i}, \tag{D.9}$$

where C_r is the installed capacity in sub-region r, and R is the set of all SE3 sub-regions.

D.2.3 Capacity Build-out Adjustment

To account for the non-stationary growth of the SE3 solar fleet, historical features are scaled. The year-over-year growth multiplier g_Y is calculated as the ratio of total production between year Y and year $Y - 1$:

$$g_Y = \frac{\sum_{x \in S_Y} y_{Y,x}}{\sum_{x \in S_Y} y_{Y-1,x}}, \tag{D.10}$$

where $y_{Y,x}$ is the total solar production in year Y on date x, and S_Y is the set of calendar dates with valid production data in both year Y and year $Y - 1$.

D.2.4 Seasonal Historical Baselines

The seasonal baseline relies on a historical window mean. For a target timestamp at hour h on target date d_c, this mean is computed as the average production at hour h across days within a $\pm n$ calendar day window:

$$\bar{y}_{\text{hist}}(d_c, h, n) = \frac{1}{|D|} \sum_{d \in D} y_{d,h}, \quad D = \{d : |d - d_c| \leq n\}, \tag{D.11}$$

where n is the half-width of the calendar window in days, D is the set of historical dates within that window, and $y_{d,h}$ is the historical production at hour h on date d. The final growth-adjusted seasonal baseline is defined as $\hat{y}_{\text{seasonal}} = \bar{y}_{\text{hist}} \times g_Y$.

D.3 Consumption Features

This appendix provides the explicit mathematical definitions underpinning the physical baseline modeling and feature engineering for electricity consumption forecasts detailed in Section 4.2.5 and Section 4.2.5.

D.3.1 Spatial Aggregation Weights

Temperature forecasts are aggregated using consumption-based weights. At the intra-regional level, the municipality-to-region weights v_m are proportional to historical municipal consumption:

$$v_m = \frac{E_m}{\sum_{i \in M_r} E_i}, \tag{D.12}$$

where E_m is the historical consumption of municipality m, and M_r is the set of primary municipalities representing region r. Region-to-SE3 weights are similarly proportional to the regional share of total SE3 consumption.

D.3.2 Temperature-Elasticity Baseline

The temperature-dependent base load is defined by a quadratic function:

$$f(T) = \beta_0 + \beta_1 T + \beta_2 T^2, \tag{D.13}$$

where T is the air temperature in degrees Celsius, and β_0, β_1, and β_2 are coefficients fitted by ordinary least squares (OLS) against total SE3 consumption using SMHI observed temperatures.

This quadratic specification was chosen to capture the non-linear, U-shaped relationship between ambient temperature and electricity demand in the SE3 zone, a modeling approach well-supported by previous research [81]. While a linear model is restricted to a monotonic trend, incorporating the squared term ($\beta_2 T^2$) successfully accounts for the inflection point where elevated summer temperatures increase electricity consumption due to cooling requirements.

For a target timestamp, let y_k and $T_{\text{hist},k}$ be the consumption and observed temperature at a matched calendar timestamp k years prior. Given a forecasted temperature T_{fc}, the historical consumption is synthetically projected onto the forecasted weather:

$$\tilde{y}_k = y_k - f(T_{\text{hist},k}) + f(T_{\text{fc}}). \tag{D.14}$$

The final temperature-adjusted consumption baseline $\hat{C}_{\text{adj}}$ is the unweighted mean across the set of successfully matched years $K \subseteq \{1, \dots, 5\}$:

$$\hat{C}_{\text{adj}} = \frac{1}{|K|} \sum_{k \in K} \tilde{y}_k. \tag{D.15}$$

D.3.3 Multi-Year Historical Baselines

To establish seasonal baselines, a centred rolling mean is computed over a $\pm n$-day window around the matched date d_{t-k} in each of the four preceding years ($k \in \{1, 2, 3, 4\}$):

$$\bar{C}_k(d_t, h, n) = \frac{1}{|W|} \sum_{i=-n}^{n} y_{d_{t-k}+i,\, h}, \tag{D.16}$$

where d_t is the target date, h is the target hour, $n = 16$ is the half-width of the rolling window in days, W is the set of valid observations within the window, and $y_{d,h}$ is the consumption at hour h on date d.

D.3.4 Day-Adjusted Short-Term Autoregressive Baseline

To project recent observations across different weekdays, a scaling factor S is calculated using the global historical average consumption $\bar{C}_{w,h}$ for weekday w at hour h:

$$S(w_{\text{base}}, w_{\text{target}}, h) = \frac{\bar{C}_{w_{\text{target}},h}}{\bar{C}_{w_{\text{base}},h}}. \tag{D.17}$$

For a target with hour h_{target} and weekday w_{target}, the three most recent valid observations y_i (occurring on weekday w_i) are rescaled and averaged to form the short-term feature $\hat{C}_{\text{recent},h}$:

$$\hat{C}_{\text{recent},h} = \frac{1}{3} \sum_{i=1}^{3} y_i \cdot S(w_i, w_{\text{target}}, h_{\text{target}}). \tag{D.18}$$

D.3.5 Lagged Hourly Averages

An additional short-term feature averages historical consumption over a fixed multi-day lag set $L = \{48, 72, 96, 120, 144, 168, 192, 216\}$ hours:

$$\bar{C}_{\text{lagged},t} = \frac{1}{|L|} \sum_{l \in L} y_{t-l}, \tag{D.19}$$

where t is the target timestamp and y_{t-l} is the consumption l hours before t.

Appendix E

EV Smart Charging: Risk-Appetite Sensitivity and Algorithmic Diagnostics

This appendix collects supporting analyses for the EV smart charging evaluation reported in Section 5.3. Section E.2 reports the full risk-appetite λ sweep at the reference configuration for each of the four framework $\times$ use-case combinations. Section E.3 reports the Framework 1 K^*-agreement diagnostic, which quantifies how often the algorithm's chosen deferral horizon K matches the oracle's per-event optimum.

E.1 Supplementary Tables and Figures for the Charging Evaluation

This section collects the modelling-assumption tables, the vehicle, driver and market specifications, the decision-structure figures, and step-by-step algorithmic pseudocode supporting the EV charging evaluation of Section 4.4.1. The framework assumptions (Tables E.1 and E.2) record the algorithm-specific choices for the two frameworks of Sections 4.4.4 and 4.4.5; the accompanying step-by-step execution logic is detailed as pseudocode in Algorithms 2 and 3; Tables E.3–E.5 record the vehicle-energetics, driver-behaviour, and market-contract assumptions of Section 4.4.2; and Figures E.1–E.4 visualise the per-night and per-week forecast-resolution structure described in Sections 4.4.6 and 4.4.7.

Table E.1: Framework 1 modelling assumptions.

Aspect	Assumption
Forecast inputs	Algorithm consumes the predicted median ($\hat{p}^{(0.50)}$) and 90th percentile ($\hat{p}^{(0.90)}$) only. Other quantile bands are produced by the price model but not used in Framework 1.
Risk parameter	λ is fixed at decision time and does not adapt within a deferral sequence or based on forecast confidence. Three values ($\lambda \in \{0, 0.5, 1\}$) are evaluated as a sensitivity.
Decision horizon	Rolling N-step, with $N = 3$ as the primary value. The algorithm considers $K \in \{0, \ldots, N\}$ at each session; longer-horizon information is ignored.
Re-evaluation	Full re-evaluation at every plug-in: previous-session decisions carry no commitment beyond their effect on current SoC.
Slot pre-selection	Within each candidate day, the H cheapest slots are selected using the predicted *median* price only. The risk parameter λ applies only to the inter-day score (4.7).
Intra-day re-optimisation	Once $K^* = 0$ is selected, the H cheapest slots are re-picked using *confirmed* prices, not forecasts. Forecast errors at the intra-day level are therefore neutralised on the day of charging.

Algorithm 2 Framework 1 deferral decision — single evaluation step.

Require: Candidate days $K = 0, \ldots, N$; slots $\mathcal{T}_K$ for each candidate; predicted quantiles $\hat{p}_h^{(0.50)}, \hat{p}_h^{(0.90)}$ on $K \geq 1$ and confirmed prices p_h on $K = 0$; energy requirement e; charging power P; risk parameter λ; feasible set $\mathcal{F}$

Ensure: Decision at session: charge at $K = 0$ (with cost C_0) or defer

Phase 1 — Score each candidate day

for each $K \in \{0, \ldots, N\}$ **do**
  $\mathcal{S}_K \leftarrow$ the H slots in $\mathcal{T}_K$ with lowest $\hat{p}_h^{(0.50)}$ $\triangleright$ intra-day pre-selection
  $\sigma_K(\lambda) \leftarrow$ energy-weighted average of $r_h(\lambda)$ over $\mathcal{S}_K$ $\triangleright$ eqs. (4.6)–(4.7)
end for

Phase 2 — Select feasible candidate with lowest score

$K^* \leftarrow \arg\min_{K \in \mathcal{F}} \sigma_K(\lambda)$

Phase 3 — Realise decision

if $K^* = 0$ **then**
  $\mathcal{S}_0^* \leftarrow$ the H cheapest slots in $\mathcal{T}_0$ by *confirmed* price $\triangleright$ intra-day re-optimisation
  $C_0 \leftarrow$ confirmed cost of $\mathcal{S}_0^*$ $\triangleright$ eq. (4.9)
  return charge at $K = 0$ with cost C_0
else
  return defer — re-evaluate at next session
end if

Table E.2: Framework 2 modelling assumptions.

Aspect	Assumption
Forecast inputs	Algorithm consumes the predicted median ($\hat{p}^{(0.50)}$) and 90th percentile ($\hat{p}^{(0.90)}$) only — identical to Framework 1. Forecast horizon spans the entire remaining-week window.
Driver declaration	Identical every Monday in perpetuity: target SoC $\bar{s} = 80\,\%$, deadline d_{target} at the next Monday 07:00 CET. No mid-week revision is allowed.
Risk parameter	λ is fixed at decision time and applied uniformly to all slots in the remaining-week pool. It does not adapt across sessions or based on forecast confidence.
Deadline binding	Hard. The vehicle reaches $\bar{s} = 80\,\%$ at d_{target} in every weekly cycle.
Forward-feasibility	When binding, the algorithm charges the minimum number of slots in the current session's window required to keep slack non-negative at the next session (Section 4.4.5). It does not charge a full session opportunistically.
Re-evaluation	Full re-evaluation at every plug-in within the weekly cycle. Slot allocations from earlier sessions in the same cycle do not commit the algorithm; only the energy already delivered (and thus reflected in s_t) is preserved across re-evaluations.
Intra-day re-optimisation	For sessions where the algorithm charges, the chosen slots are realised at confirmed prices (the current session's window is always in the confirmed range). Future-session slot allocations remain forecast-driven until those sessions become current.

Algorithm 3 Framework 2 weekly-target decision — single session within a weekly cycle.

Require: Current session t; deadline d_{target}; current SoC s_t; predicted and confirmed prices on remaining sessions $j \in \mathcal{R}(t)$; charging power P; daily consumption e; risk parameter λ

Ensure: Decision at session t: charge (with assigned slots) or defer

Phase 1 — Required quantities

$E_{\text{rem}}(t) \leftarrow (\bar{s} - s_t) \cdot B + e \cdot d_{\text{rem}}(t)$ ▷ remaining energy; eq. (4.10)

$C_{\text{rem}}(t) \leftarrow P \cdot \sum_{j \in \mathcal{R}(t)} W_j$ ▷ remaining capacity; eq. (4.11)

$\text{slack}(t) \leftarrow \big(C_{\text{rem}}(t) - E_{\text{rem}}(t)\big)/P$ ▷ eq. (4.13)

$H_{\text{rem}}(t) \leftarrow \lceil E_{\text{rem}}(t)/P \rceil$

Phase 2 — Forward-feasibility check (forced charge)

if deferring at t would yield $\text{slack}(t+1) < 0$ **then**

  $h_{\min}(t) \leftarrow$ minimum slots in $\mathcal{T}_t$ required to restore $\text{slack}(t+1) \geq 0$

  return charge the $h_{\min}(t)$ cheapest slots in $\mathcal{T}_t$ by *confirmed* price

end if

Phase 3 — Cost-optimal slot allocation

$\mathcal{P}(t) \leftarrow \bigcup_{j \in \mathcal{R}(t)} \mathcal{T}_j$ ▷ candidate slot pool

$r_h(\lambda) \leftarrow (1-\lambda)\, \hat{p}_h^{(0.50)} + \lambda\, \hat{p}_h^{(0.90)}$ for every $h \in \mathcal{P}(t)$ ▷ eq. (4.6)

$\mathcal{S}^*(t) \leftarrow$ the $H_{\text{rem}}(t)$ slots in $\mathcal{P}(t)$ with the lowest $r_h(\lambda)$

Phase 4 — Decision at session t

if $\mathcal{S}^*(t) \cap \mathcal{T}_t \neq \emptyset$ **then**

  return charge the slots in $\mathcal{S}^*(t) \cap \mathcal{T}_t$ at *confirmed* prices

else

  return defer — re-evaluate at session $t+1$

end if

Table E.3: Vehicle modelling assumptions applied throughout the charging evaluation.

Aspect	Assumption
Usable battery capacity (B)	Set to 87 kWh. Assumed as the share of the 95 kWh nominal capacity available for use if there is an energy buffer reserved by the manufacturer
Energy consumption (ε)	Set to 20.0 kWh/100 km, rounded up from the WLTP figure of 16.2 kWh/100 km to account for a wider range of real driving conditions. Held constant regardless of season, temperature, driving style, payload, or trip length.
SoC target ($\bar{s}$)	80 % of usable capacity (69.6 kWh). Assumed as a representative everyday charge target.
SoC floor ($\underline{s}$)	20 % of usable capacity (17.4 kWh). Assumed as a departure-readiness constraint.
Usable buffer (Δs)	52.2 kWh, computed as $(\bar{s} - \underline{s}) \cdot B$.
Battery degradation	Battery capacity is constant across the evaluation window (2025-01-01 to 2026-05-01). No calendar or cycle ageing modelled.
Depletion model	Linear in distance driven. No battery curve effects, no preconditioning energy, no rolling-resistance or auxiliary-load variation.
Charging taper	Charging is at constant power P across the entire connection window; no taper at high SoC. The $9 \cdot P$ kWh window-capacity bound is therefore an exact deliverable.
Charging efficiency	Energy delivered to the battery equals grid-side metered hours multiplied by P. AC–DC and ohmic losses are not modelled separately.
SoC limits	The 80 % target ($\bar{s}$) and 20 % floor ($\underline{s}$) are hard constraints, not soft preferences. The vehicle never charges above $\bar{s}$ and never depletes below $\underline{s}$.

Table E.4: Driver behaviour assumptions applied throughout the charging evaluation. These define a stylised, perfectly regular driver against which the algorithms are evaluated. Real driver irregularity (missed sessions, schedule changes, holidays) would soften forecast value but is not modelled.

Aspect	Assumption
Daily driving distance	Constant at d km/day for every day in the evaluation period. No day-to-day variation, no holidays beyond the working-day calendar defined per use case, no occasional long trips.
Plug-in regularity	The driver plugs in at every session: every calendar night under UC1, every working day under UC2. No missed sessions.
Connection window	The full 9-hour connection window (specified per use case in Sections 4.4.6 and 4.4.7) is available at every session, never foreshortened by user behaviour (early disconnect, late arrival).
Departure SoC	Under Framework 1, departure SoC is whatever the algorithm has delivered subject to the SoC floor. Under Framework 2, the driver expects $\bar{s} = 80\,\%$ at the declared deadline; the algorithm must deliver this.
User intervention	Under Framework 1, none — the driver delegates all decisions to the algorithm. Under Framework 2, the driver declares the target SoC and deadline once per week (Monday 07:00) and provides no further input.

Table E.5: Market and contract assumptions applied throughout the charging evaluation. The exclusion of fees scales absolute cost levels but does not change the relative ordering of strategies, which is the headline output.

Aspect	Assumption
Contract type	Pass-through variable-rate spot contract: hourly delivered energy is priced at the SE3 Nord Pool day-ahead clearing price for that hour.
Retail markup	None. The driver pays the SE3 spot price directly, with no retail markup applied. UC1 assumes a customer-owned wallbox; UC2 assumes the workplace charger passes the spot price through without markup.
Excluded charges	Grid fees, energy taxes, network capacity charges, and time-of-use tariffs are not modelled.
Settlement resolution	Hourly throughout. The 15-minute imbalance settlement that became active in October 2025 is not modelled; this constraint is acknowledged as a thesis-wide limitation in Section 1.4.

Table E.6: UC1-specific assumptions, applied under both frameworks.

Aspect	Assumption
Connection window	22:00 to 06:00 CET (9 hours) on every calendar night in the evaluation period. The vehicle is plugged in at 22:00 and disconnected at 06:00.
Decision time	22:00 CET on each plug-in night (Framework 1: per-session decision; Framework 2: per-session re-evaluation within the weekly cycle).
Forecast init	18:00 UTC, the post-gate forecast init available at the decision time. This delivers next-day confirmed prices to the algorithm.
Holiday and weekend handling	The vehicle is plugged in every calendar night including weekends and holidays. There is no calendar gap in UC1.
Naive baseline behaviour	The vehicle starts charging at 22:00 CET on every night, drawing the H chronologically first slots until the daily energy requirement e is met.

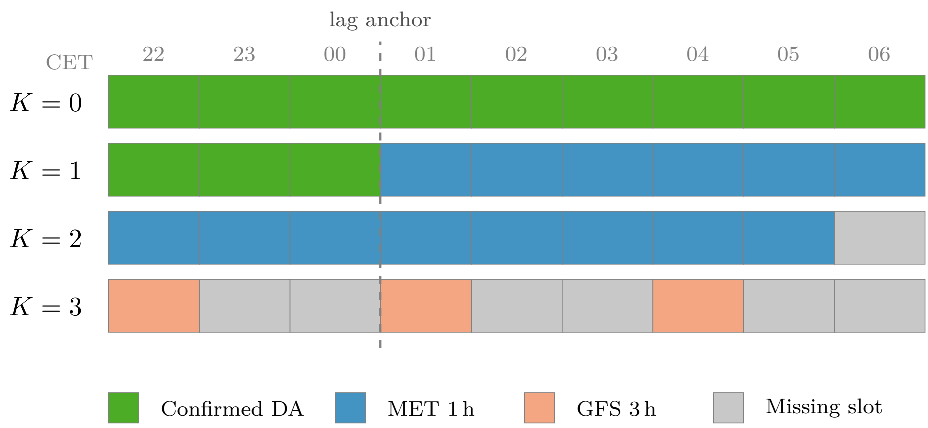

Figure E.1: Decision structure for UC1 under Framework 1. The driver decides at 22:00 CET on $K = 0$ which of four candidate nights to charge, using a mix of confirmed and predicted slots. The lag anchor (dashed) separates confirmed from predicted regions. Tonight ($K = 0$) is fully confirmed; tomorrow ($K = 1$) has its first three slots confirmed; $K = 2$ falls in MET 1-hour range with a winter gap at lead 59 h; $K = 3$ lies in the GFS 3-hour range, leaving only three of nine slots covered. Deferral triggers re-evaluation the following night with fresh predictions.

Table E.7: UC2-specific assumptions, applied under both frameworks.

Aspect	Assumption
Connection window	08:00 to 17:00 CET (9 hours) on every working day in the evaluation period. The vehicle is plugged in at 08:00 and disconnected at 17:00.
Decision time	08:00 CET on each working-day arrival (Framework 1: per-session decision; Framework 2: per-session re-evaluation within the weekly cycle).
Forecast init	06:00 UTC, the pre-gate forecast init available at the decision time. The day-ahead gate at 12:45 CET has not yet closed, so the next day's prices are not confirmed at decision time.
Working calendar	251 working days per year, accounting for Swedish public holidays. Public holidays falling on weekdays are treated as non-working (no plug-in).
Weekend and holiday handling	The vehicle continues to be driven at d km/day on weekends and public holidays but cannot charge. Battery depletion accumulates unbroken across the weekend; the Monday plug-in inherits the reduced SoC.
Naive baseline behaviour	The vehicle starts charging at 08:00 CET on every working day, drawing the H chronologically first slots until the daily energy requirement (adjusted upward at the start of the working week to recover weekend depletion) is met.

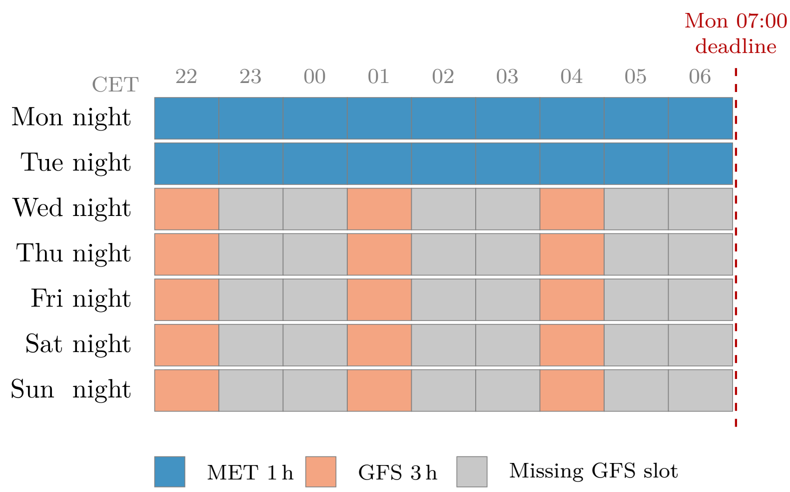

Figure E.2: Slot pool $\mathcal{P}(t)$ for UC1 under Framework 2 at the weekly cycle origin (Monday 07:00 CET). Seven overnight connection windows of nine hours each contribute their hourly slots to the remaining-week pool, colour-coded by forecast source available at the cycle origin. Monday and Tuesday nights fall within the MET 1-hour resolution range; Wednesday onward lies in the GFS 3-hour range, where only every third overnight slot is covered. The Monday 07:00 deadline marks the close of the cycle. As the week progresses, slots in past windows leave the pool and remaining slots move toward the confirmed range, with $\mathcal{P}(t)$ contracting accordingly.

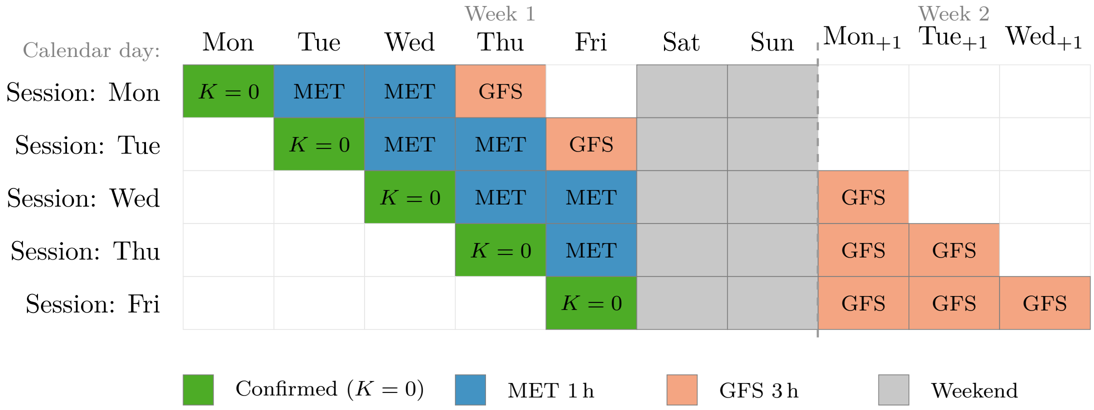

Figure E.3: Prediction resolution for UC2 under Framework 1, by session day-of-week. Each row shows the candidate days $K \in \{0, 1, 2, 3\}$ available to a session on that weekday, colour-coded by forecast source. Friday sessions have no MET-resolution candidate; all three deferred days lie in the GFS range. Weekends consume battery but are never candidate charging days.

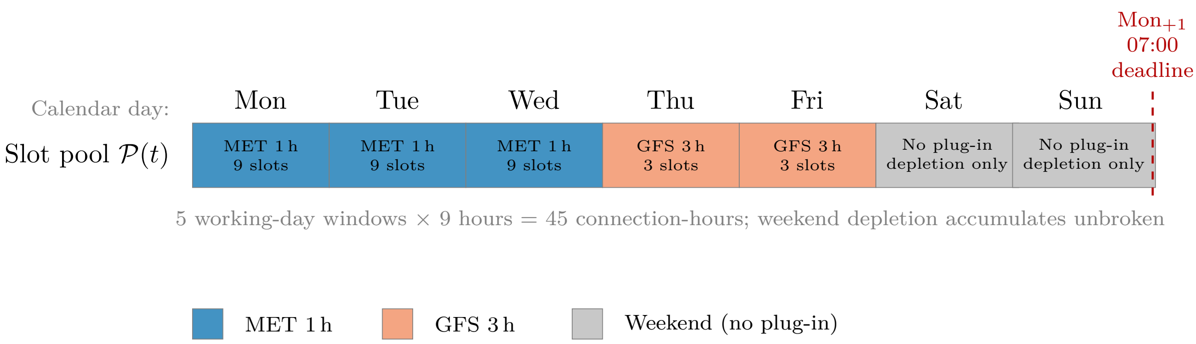

Figure E.4: Slot pool $\mathcal{P}(t)$ for UC2 under Framework 2 at the weekly cycle origin (Monday 07:00 CET). Five working-day connection windows of nine hours each contribute 45 hourly slots to the remaining-week pool, colour-coded by forecast source. Monday through Wednesday windows fall within the MET 1-hour resolution range; Thursday and Friday lie in the GFS 3-hour range. Weekend days carry no connection windows but contribute fixed battery depletion of $2 \cdot e$ that must be covered by $E_{\text{rem}}(t)$ before the Monday 07:00 deadline. As the working week progresses, slots in past windows leave the pool and remaining slots move toward the confirmed range.

E.2 Risk-Appetite Sensitivity at the Reference Configuration

The risk parameter $\lambda \in \{0, 0.5, 1\}$ controls the algorithm's slot-scoring rule (Section 4.4.4). $\lambda = 0$ uses the median predicted price (q_{50}); $\lambda = 1$ uses the 90th-percentile predicted price (q_{90}), penalising slots with high forecast uncertainty; $\lambda = 0.5$ blends the two. Table E.8 sweeps λ at the reference configuration only (50 km/day, 7.4 kW); the $\lambda = 0$ rows reproduce the reference numbers reported in Section 5.3.1.

Table E.8: Risk-appetite sensitivity at the reference configuration (50 km/day, 7.4 kW): annual savings (EUR/year) and oracle headroom for $\lambda \in \{0, 0.5, 1\}$ across the four framework $\times$ use-case cells. The $\lambda = 0$ rows reproduce the reference numbers of Section 5.3.1.

Cell	λ	S_{defer}	$S_{\text{sched}} + S_{\text{defer}}$	G_{oracle}
F1 $\times$ UC1	0	55.6	114.3	-6.5
	0.5	50.3	109.0	-1.2
	1	45.5	104.2	3.6
F1 $\times$ UC2	0	49.6	167.9	2.0
	0.5	50.6	168.9	1.0
	1	44.5	162.8	7.1
F2 $\times$ UC1	0	44.2	102.9	4.0
	0.5	39.4	98.1	8.8
	1	33.8	92.5	14.4
F2 $\times$ UC2	0	48.8	167.1	5.6
	0.5	39.4	157.7	14.9
	1	30.3	148.6	24.0

E.3 Framework 1 K^*-Agreement Diagnostic

Under Framework 1, the algorithm selects a deferral horizon $K^* \in \{0, 1, \ldots, N\}$ at each session by minimising a forecast-quantile-based score; the oracle makes the same selection but scores candidate slots using confirmed (realised) prices. The K^*-agreement rate is the fraction of algorithm charging events where the chosen K^* equals the oracle's per-event optimum. A high agreement rate indicates that the forecast resolves the deferral question correctly at the decision level; the residual gap closes the loop with the oracle headroom G_{oracle} reported in the main results. Table E.9 reports the agreement rates.

Table E.9: Framework 1 K^*-agreement diagnostic: percentage of algorithm charging events where the chosen K^* matches the oracle's. Reference column at 50 km/day $\times$ 7.4 kW, $\lambda = 0$; per-driver columns pool across all wallboxes and all λ values (UC2 excludes the 100 km/day $\times$ 7.4 kW cell consistent with the exclusion applied in Table 5.16).

Use case	Reference (%)	Pooled by driver (%)			
		25 km/day	50 km/day	75 km/day	100 km/day
UC1	69.0	56.6	56.6	63.4	72.3
UC2	61.7	45.3	51.1	64.8	80.1

Master's Theses in Mathematical Sciences 2026:E99
ISSN 1404-6342

LUTFMS-3572-2026

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lu.se/