Solution 7. Concept

Solution 7: AI-based Electricity Market Prediction Tool.

Funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor CINEA can be held responsible for them.

List of Acronyms

AGR	Avoid Grid Reconstruction
AI	Artificial Intelligence
ARIMA	Auto Regressive Integrated Moving Average
BESS	Battery Energy Storage System
CM	Capacity Market
DT	Digital Twins
EMS	Energy Management System
FR	Frequency Regulation
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GRU	Gated Recurrent Unit
IoT	Internet of Things
LSTM	Long Short-Term Memory
ML	Machine Learning
MSC	Maximization of Self-Consumption
PS	Peak Shaving
RNN	Recurrent Neuronal Network
SARIMA	Seasonal ARIMA
TOU	Time of Use

List of Tables

Table 1. Summary of forecasting models in terms of their strengths and weaknesses. 8

Table 2. Summary of forecasting models in terms of their scope and characteristics. 11

Table 3. Summary of the forecasting models depending on the application time horizon. 13

Table 4. Proposal for data to collect 14

Table 5. Proposal for stochastic parameters 15

Table 6. Summary of the target values for the KPIs described. 17

Keywords list

- Simulation
- Control
- Forecasting
- Sun tracking
- Batteries

Table of Contents

List of Acronyms **1**
List of Tables **2**
Keywords list **2**
1. Executive summary **4**
2. Introduction **5**
3. Chapter B: Electricity Market Forecast **6**
3.1.1. Introduction 6
3.1.2. PVOP proposal: concept 9
3.1.2.1. Conceptual aspects 9
3.1.2.2. Stochastic analysis of forecasts 11
3.1.3. MODELING, ANALYSIS AND KEY PERFORMANCE INDICATORS (KPIs) 12
3.1.3.1. Modelling and experimentation process 13
3.1.3.2. KPIs definition 16
3.1.3.3. Integration with the BESS System 17
3.1.4. Conclusions 19

1. Executive summary

This document presents the concepts at the root of WP3 solutions, all of them about the Control of PV plants to optimise their performance. This WP3 is divided into two differentiated tasks:

- Task 3.1: Development and demonstration of a simulation tool and control system to maximize the performance of PV plants with sun tracking systems.

- Task 3.2: Development and demonstration of a control system to maximize the performance and energy trading of PV plants with batteries. In turn, this task is divided into two:

    - Chapter A: Battery control system
    - Chapter B: Electricity Market Forecast

Regarding T3.2, PVOP will develop a PV plant control system that will integrate a Smart Energy Management System (EMS) based on predictions of weather conditions (including extreme events) and an AI-based electricity market predictor. The concept at the root of this solution is the following: the EMS will receive inputs from sensors, electricity price databases, meteorological databases and Electricity Market Analysts predictor that will allow the EMS to select, in a hourly basis, the best strategy to maximize the profitability of the PV asset. Regarding, the forecast prices in the electricity market, a new concept will be implemented that is thought to be decisive: the stochastic analysis of the forecasts.

For the validation of these solutions, experiments and specific KPIs for both, the implementation of control strategies for batteries and for the electricity market forecast tool, have been proposed.

2. Introduction

This document presents the first deliverable of WP3 (Control of PV plants to optimize their performance), and it explains the main objectives to be reached, the techniques used for that and the KPIs defined for the validation of the different solutions.

This WP3 is divided into two differentiated tasks, and one of them is as well separated into two main lines of work. The deliverable is structured accordingly:

- Task 3.1: Development and demonstration of a simulation tool and control system to maximize the performance of PV plants with sun tracking systems.

- Task 3.2: Development and demonstration of a control system to maximize the performance and energy trading of PV plants with batteries.

This document is focused on Task 3.2 and describes the Solution 7: AI-based Electricity Market Prediction Tool.

3. Chapter B: Electricity Market Forecast

3.1.1. Introduction

Background

The objective of this project is to maximize the potential of photovoltaic installations in combination with Battery Energy Storage Systems (BESS). One of the main factors influencing this optimization is economic profitability, as without it, no project can be feasible. Therefore, the correct operation of the plants, maximizing the value of their generation and the charge-discharge cycles of the battery systems, is key.

To ensure this operation is carried out correctly, it is necessary to evaluate the different operational scenarios that are viable for the specific plant. For this, predicting the different selling prices of the energy, whether produced or fed into the system from the batteries, is crucial. For this reason, PVOP includes among its development objectives the creation of forecasting models, based on various existing tools, to enable the most efficient and profitable operation of the plants.

In recent years, the modelling and forecasting of electricity prices have seen significant advances, largely driven by the emergence of artificial intelligence (AI) and self-learning regression models. Traditionally, electricity price forecasting relied on statistical and econometric models that used historical data and exogenous variables such as demand, supply, and weather conditions. However, these models often faced limitations in their ability to capture the complexity and dynamics of the electricity market.

With the advent of AI, more sophisticated models have been developed that can handle large volumes of data and learn complex patterns. Machine learning algorithms, such as neural networks and decision trees, have proven particularly effective in forecasting electricity prices. These models can adapt and continuously improve as they are fed with new data, allowing them to anticipate market changes with greater accuracy. Models like ARIMA (AutoRegressive Integrated Moving Average) and LSTM (Long Short-Term Memory) are capable of capturing both long-term trends and short-term fluctuations in electricity prices. Additionally, they can incorporate a wide range of variables, from weather data to information on renewable energy generation, further enhancing their accuracy.

Nevertheless, there is still much progress to be made in this field. Modelling and influencing chaotic events (unpredictable a priori), as well as rare or infrequent events that affect prices, remain challenges and are foundational aspects of this part of the PVOP project. The goal is not so much to obtain a fixed price, which is almost impossible to achieve from a technical standpoint, for each hour or day for the next 10-15 years, but to generate models that provide price ranges to facilitate decisions regarding productive operation, maintenance, or predictive measures.

Generic characteristics of initial models

Regarding the most commonly used models, the objective of this document is not to provide detailed information on each one, but it is necessary to give an overview to understand the uses, weaknesses, and strengths of each chosen as a first option from those available.

ARIMA

The ARIMA (AutoRegressive Integrated Moving Average) model is a statistical technique used for time series analysis and forecasting. It consists of three components: autoregressive (AR), moving average (MA), and differencing (I). The AR component uses dependencies between past observations, the MA component models the prediction error as a linear combination of past errors, and differencing is used to make the time series stationary.

- Technical characteristics:
    - AR(p): Number of autoregressive terms.
    - I(d): Number of differences needed to make the series stationary.
    - MA(q): Number of moving average terms.
- Use cases:
    - Best: Time series without seasonality.
    - Worst: Series with strong seasonal components.
- Computational complexity: Moderate, as it requires the estimation of several parameters and can be computationally intensive for long series.

Neural Networks (LSTM and GRU)

Recurrent neural networks (RNN) such as LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) are advanced deep learning techniques designed to handle sequential data.

- Technical characteristics:
    - LSTM: Uses memory cells and three gates (input, output, and forget) to maintain and update information over time.
    - GRU: Similar to LSTM but with a simpler architecture, using only two gates (update and reset).
- Use cases:
    - Best: Time series with long-term dependencies and complex patterns.
    - Worst: Series with few data points or without clear patterns.
- Computational complexity: Very high, due to the need to train multiple layers and parameters, requiring significant computational resources.

GARCH

The GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model is primarily used to model and forecast volatility in financial time series. This model assumes that the variance of the error is autocorrelated and follows an autoregressive moving average process.

- Technical characteristics:
    - GARCH(p, q): p is the number of autoregressive terms and q is the number of moving average terms in the variance.
- Use cases:

  - Best: Financial time series with variable volatility.
  - Worst: Series with constant variance or without heteroscedasticity.
- Computational complexity: Moderate to high, depending on the number of terms included in the model.

Model	Strength	Weakness	Computational Complexity
ARIMA	Time series without seasonality	Series with strong seasonal components	Moderate
SARIMA	Time series with clear seasonal patterns	Series without seasonality or with irregular seasonality	High
LSTM and GRU	Time series with long-term dependencies and complex patterns	Series with few data points or without clear patterns	Very high
GARCH	Financial time series with variable volatility	Series with constant variance or without heteroscedasticity	Moderate to high

Table 1. Summary of forecasting models in terms of their strengths and weaknesses.

These models offer different approaches and levels of complexity for time series forecasting, and their choice depends on the specific characteristics of the data and the objectives of the analysis.

Regarding the temporal use of the aforementioned models, we must consider that not all models will be equally accurate, a priori, for all simulation scopes. Based on the three simulation ranges to be used in the project (hourly, monthly, and annual), we can conduct an initial analysis to indicate which of these previously mentioned models may be a priori, more suitable for each of them.

Hourly forecasting

- Time series models: ARIMA (AutoRegressive Integrated Moving Average) and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models are widely used to predict hourly prices. These models consider seasonal patterns and historical trends.

- Neural networks: Neural networks, such as LSTM (Long Short-Term Memory), have proven to be effective in forecasting electricity prices. These networks can capture nonlinear relationships and complex patterns in the data.

Monthly forecasting

- Seasonal analysis: The monthly forecast is based on seasonal patterns and climatic factors. Variables such as demand, availability of renewable sources and special events (holidays, vacations, etc.) are considered. So, models such as ARIMA and SARIMA are appropriate.

- Econometric models: Models such as the multiple linear regression model are used to relate prices to explanatory variables such as demand, supply and fuel prices.

Annual forecast

- Macroeconomic scenarios: The annual forecast involves considering long-term economic, political and technological factors. Scenarios such as economic growth, infrastructure investment and the transition to clean energy are evaluated based, among others, on government strategic plans. Structural models (SSM) could be the most accurate for this work.

Hybrid models

Combined approaches, such as the integration of econometric models with neural networks (LSTM), allow for more accurate and robust forecasting.

All of the above allows us to outline a starting scenario from which to begin the necessary work to achieve the project's objectives.

3.1.2. PVOP proposal: concept

3.1.2.1. Conceptual aspects

The purpose of this PVOP project is clear: to push PV generation to the next level of efficiency, safety and profitability to reach the level of excellence that will allow this technology to lead the energy revolution in the coming years. To optimize profitability to the maximum, it is necessary to be able to operate these installations, and more specifically, those with storage that discharge energy into the system.

To achieve this goal, it is necessary to be able to create energy management strategies that maximize the economic return on each MWh generated. When to store, when to feed the associated facilities in self-consumption or when to discharge to the grid, and how much, will be one of the keys to the success of these systems. And for this, it is necessary to know as accurately as possible the grid energy prices in the necessary time frames as well as the foreseeable balance of generation and demand with the necessary anticipation. It is important to keep in mind that the forecast domains depend very much on the purpose of the forecast. We can reduce these forecast blocks to 4 scenarios:

- **Pluriannual model**: The objective of these models is to forecast over a period of 5 - 10 - 15 years both the expected average prices and the possible fluctuations of these prices. Their objective is the realization of the business plan of each facility, as well as the analysis of its payback period and profitability. As it is logical, these forecasts cannot pretend to be of a high accuracy and must be more oriented to create scenarios of not exceeding or lowering the price, giving a range of prices that allow, with the necessary reliability, to create business scenarios that facilitate the financing of the facilities. Although their calculation should include price weightings according to each hour, to be able to adjust the data correctly to PV installations, which have a large component of seasonality both monthly and hourly, they should be based on global annual data.

- **Annual model**: This pricing model should be able to, with greater predictability than the multi-year model, present price scenarios with tighter reliability. With the objective of planning a year ahead from the operational point of view of the plant (looking for maintenance windows, knowing the expected monthly cash flows, foreseeing advanced amortizations or other actions of a more operational or financial nature), this model should present better results on a monthly basis than the previous one, and should be able to be weighted by generation hours, although the hourly nature is not its main objective.

- **Monthly model**: For this model, the temporal scope should be of several months, being the chosen term discussed later during the PVOP project with the objective of adjusting more to the needs of the plants from the point of view of technical shutdowns, maintenance operations or revenue optimization. Its character is more operative than the previous ones, since it will allow making decisions not so much in terms of plant financing, but more in terms of its technical effectiveness, minimizing the energy losses in usual operations. Its hourly forecast must be more adjusted, and the seminal scope should be the basic calculation unit from which to start. It should allow a more global planning of the load/discharge cycles of the installations, which will then be optimized in the following model.

- **Daily model:** the daily model is the one that should be the basis for the short-term operation of each facility. With a horizon of several days (maximum 10 days), it must reach high efficiency quotas in the hourly forecast, since it will be the basis for short-term sales/storage strategies. Starting from the daily forecasting unit, which must have very high reliability values in the forecast price, it must perform an efficient hourly price forecast in all the days of analysis, and especially in the short period of time of 5 days ahead. In addition to this hourly forecast, this daily model has to be able to reach a quarter-hourly forecast resolution on the shortest time horizon (from a few hours, to act intraday up to 1 or 2 days), to allow the efficient operation of the asset in regulation, secondary, adjustment markets, etc.

As mentioned in the previous point, there are several methodologies in use, with better or worse results, for each area of analysis of the above models. Moreover, the objective is not the same for all of them, and therefore, their approach to the calculation cannot be similar. To simplify the understanding of the objective, scope and characteristics of each model, Table 4 is attached.

MODEL	Multi-year	Annual	Monthly	Daily
Time scope	5 to 15 years	12 months	3 or 4 months	10 days
Base forecasting unity	Annual	Monthly	Monthly	Daily
Accuracy required	High	High	Very high	Very high
Maximum granularity	Hourly	Hourly	Hourly	Quarter Hour

Accuracy required	Medium	High	High	Very high
Model objective	Financial	Financial/operative	Operative/Financial	Operative
Reliability interval	High	High	Very high	Very high
Sensitivity to stochastic analysis	Very high	High	High	Medium

Table 2. Summary of forecasting models in terms of their scope and characteristics.

The selection of the mathematical forecasting model will be carried out throughout the experimentation process, starting for each block from the most suitable ones according to the scope of analysis. This process does not differ much from the standard analyses that are already being carried out, but in PVOP project, a new concept will be implemented that is thought to be decisive: the stochastic analysis of the forecasts.

3.1.2.2. Stochastic analysis of forecasts

Stochastic analysis is a statistical technique, widely used in the financial sector, which aims to determine **the strength and trend of a price based on external events**. In this way, it is possible to try to anticipate trading operations to sudden changes or price caps/lows.

In more specific terms, the stochastic indicator compares the current closing price with previous closing prices over a chosen period. It was created by George Lane in the 1950s and is popular in Forex, indices and stock trading. Some common uses of this indicator are as follows:

- Divergences: It can help to identify divergences between the indicator and the price, which can signal possible trend changes.

- Day trading and scalping: Intraday traders can use it to make quick decisions.

- Buy and sell confirmations: Provides signals to enter or exit the market.

- Overbought and oversold: Helps to identify extreme price levels.

Currently, most electricity market price analyses are based on technical criteria such as the degree of penetration of renewable energies in the market or the overall demand forecast for a given period. Under stable conditions, being these data well known, this type of analysis gives good results, especially in the short term. But experience and observation allow us to confirm that external events not directly associated with actual consumption affect the price in a critical way. As an example:

- We can appreciate slight price increases in the electricity market in the order of 2-3 months before an electoral process, not being this increase related to objective parameters.

- The world pandemic, and more specifically its confinement, produced a critical drop in prices. In this case, it is related to the drop in demand, but previous analyses indicate that this drop was not correlated and had an indeterminate component related more to bearish market expectations than to fundamentals.

- The Ukrainian war generated a price increase that, although it can be attributed to the gas price increase, should not have responded abruptly given the level of storage filling or supply conditions.

These examples are just a sample of how external processes, or "black swans", which often do not have a direct effect on supply or do not affect price formation parameters with the weight that justifies the price change they produce, generate critical impacts on these prices.

The analysis, modelling and implementation of these external elements is defined as key to understand and forecast the price in the medium and long term, mainly. In the short term, only major events can affect these forecasts, which, as indicated above, respond more to technical parameters, but in a monthly or longer-term analysis, their presence is critical.

That is why in PVOP project, the stochastic analysis of external events and their effect on the market price is defined as one of the mainstays of the work to be developed.

3.1.3. MODELING, ANALYSIS AND KEY PERFORMANCE INDICATORS (KPIs)

The objective of the modelling and analysis processes of this project will be to determine which mixed models (data-driven and stochastic) respond best to the analysis in each time domain, and to select their priority for use in the implementation, development and calculation of PV plant operation strategies. These models will be able to exploit their full potential in installations that have a certain level of management when hybridized with associated BESS, but for some applications, especially for issues of participation in regulation and secondary markets, this does not necessarily have to be the case.

First, we must define what we consider as success. It must be considered that, as in any practical project, the objectives must be feasible from the theoretical point of view and achievable from the technical point of view. The objective, therefore, cannot be to determine with high precision the outcome in each time range of a given strategy. Except for very short-term predictions (from quarter-hourly to a couple of days), accurately predicting the outcome of such a strategy by thinking of a closed value of € is impossible, and pursuing such a goal would be unrealistic.

The objective, therefore, should be that the chosen models achieve a range of results for each of the strategies as close to the real one as possible, with sufficient reliability to allow them to:

- Develop firm facility business plans that improve their bankability and facilitate their actual execution.

- Allow medium-term operational decisions to be made, such as in which month it is more appropriate to stop for corrective work or similar actions.

- Allow to define short-term strategies for plant operation. These strategies will have as main objective the maximization of plant revenues, considering, in addition, where appropriate, the possible technical restrictions of the associated storage.

On the other hand, the general management strategies of the BESS associated to the PV plant will be defined and agreed jointly with WATT partner, initially proposing those already mentioned in Chapter A of this deliverable.

3.1.3.1. Modelling and experimentation process

To perform this analysis, the following actions should be carried out:

A) First, some models should be chosen for each temporal scope of analysis, as indicated in Table 5 of this document (Chapter B). The state of the art at the time of initiating this work will determine which mathematical models are most appropriate.

The models initially proposed based on the application time horizon are the following

MODEL	Pluriannual	Annual	Monthly	Daily
Time scope	5 to 15 years	12 months	3 or 4 months	10 days
Proposed Models	• ARIMA • Structural models SSM • Neural Networks LSTM • Econometric models (VAR/VECM)	• SARIMA (Seasonal ARIMA) • Neural Networks (LSTM o GRU) • GARCH Models	• ARIMA y SARIMA • Neural Networks (LSTM o GRU) • Machine learning models (random forest, gradient boosting)	• ARIMA y SARIMA • Neural Networks (LSTM o GRU) • Machine learning models (random forest, gradient boosting) GARCH models

Table 3. Summary of the forecasting models depending on the application time horizon.

In this phase, it will be necessary to collect data that is known to directly affect the price of the electricity market, such as the degree of penetration of renewable energy sources, weather forecasts, the availability of hydroelectric power, economic indicators, the price of natural gas or the price of emission rights.

Data	Description	Granularity	Data Source
Electricity Price	Hourly electricity price in the wholesale market.	Hourly	Electricity market operators (e.g., OMIE in Spain)
Electricity Demand	Hourly electricity consumption.	Hourly	System operators (e.g., REE in Spain)

Renewable Production	Electricity generation from renewable sources (wind, solar, hydro).	Hourly	System operators, energy agencies
Non-renewable Production	Electricity generation from non-renewable sources (coal, gas, nuclear).	Hourly	System operators, energy agencies
Temperature	Ambient temperature, which affects electricity demand.	Hourly	Meteorological services (e.g., AEMET in Spain)
Natural Gas Price	Natural gas price, influencing electricity generation costs.	Daily	Commodity markets (e.g., Henry Hub)
CO2 Price	Price of CO2 emission allowances.	Daily	Carbon markets (e.g., EU ETS)
International Exchanges	Electricity imports and exports.	Hourly	System operators, energy agencies

Table 4. Proposal for data to collect

B) Secondly, we will proceed to determine a base set of stochastic parameters to be implemented in the model. At this point, we must define what they are and how much they affect the price in relative terms, what is their time cycle (time in which they surely occur), their probability of occurrence in time...... If, for example, we think of an electoral process, the probability of occurrence in 5 years is 100%, and as time passes since the last process, the probability of repetition increases (the quarter after the electoral process, the probability of repetition is close to zero, but after 3 years it is much higher).

A starting proposal for the stochastic parameters to be considered in these models would be the following:

Stochastic parameters
Extreme heat events
Extreme cold events
Droughts
Natural disasters (earthquakes, hurricanes of extreme magnitude)
Elections
Events affecting demand (World Cups, Olympics, etc)
Epidemics
Economics crises

Military Conflicts / Terrorist Attacks / Cyber Attacks
Disruptions / supply reductions Gas / Oil
Legislative changes

Table 5. Proposal for stochastic parameters

After these first two points, the mixed mathematical models that combine both analysis philosophies will be generated. Specifically, the initial idea is that stochastic analysis will be applied to technical analysis, modifying its forecast as effects may or may not appear.

C) Finally, the validity of the models determined will be analysed by choosing past time periods and determining whether those models are appropriate based on the chosen KPIs.

The data to be used for the validation of the different models, and for their subsequent training, will be extracted from the different European NEMOS (Nominated Electricity Market Operator) on the one hand and from the TSOs (Transmission System Operators) on the other hand. Both NEMOS and TSOs are subject to the different European directives that establish the need to publish information related to electricity generation, consumption, capacity, balancing energies etc.

It is important to keep in mind that in these experiments, the boundary conditions will be very important. It is difficult to know a priori how the weather will behave or if there will be a global pandemic in the next five years. Therefore, in order to launch the models and validate the results, Monte Carlo simulations will be applied.

Also known as the Monte Carlo method or multiple probability simulation, this technique is used to predict outcomes in uncertain situations. It was invented by John von Neumann and Stanislaw Ulam during World War II to improve decision making under uncertainty. Unlike a normal forecasting model, Monte Carlo simulation predicts a set of outcomes based on an estimated range of values rather than fixed input values. It uses probability distributions (such as uniform or normal) to model uncertain variables. It then calculates the results repeatedly using different random numbers between the minimum and maximum values. This process is repeated thousands of times to generate many probable outcomes.

This simulation allows sensitivity analyses to be performed and correlations between inputs to be calculated. In addition, it is accurate over the long term, projecting results with greater accuracy as the number of inputs increases.

Taking as inputs all our variables, stochastic or technical, determined as a normal distribution or more complex ones, but allowing us to generate random scenarios in a given time frame, we can simulate different scenarios that generate different expected market prices, and determine which models are in the desired ranges.

It is important to indicate that the analyses will be carried out on past dates, in order to be able to validate the model. It is unfeasible in the early stages of the project to validate the modelling without reference prices as a result, so we will proceed to model past price scenarios, starting with more stable periods and ending in periods of great changes, to study how the model acts against these imponderables. Future forecasts will only be studied in the final stages of the project, once the possible final candidates have been chosen, to carry out real tests in the second period of the project.

Finally, it is important to consider that market prices behave mostly as an elastic system: when an event takes them out of their equilibrium point, they oscillate until they reach that equilibrium again. The amplitude of the wave, the transition period and the distance from that point of equilibrium to the previous one is variable, but what seems clear is that after large declines appears a large rise, which chains lower and lower rises until stabilizing at a point not far from the previous one, and that can be estimated with technical parameters.

3.1.3.2. KPIs definition

The objective of the work previously described is twofold:

- Determine with simulations the minimum and maximum prices for a period of time that are reached with a given reliability interval.

- Determine which models extract results with the above conditions, dispersing those minimum and maximum values at the shortest possible distance from the actual price of the analysis period.

These objectives directly lead to the definition of two fundamental KPIs to evaluate the quality of the results:

1. **Confidence interval:** Probability that the price is in a range of maximum and minimum prices.

2. **Maximum percentage dispersion:** Percentage deviation of the real final price with the minimum and maximum values that comply with the confidence intervals indicated for each model layer. For this, the indicated simulations will be carried out, obtaining the prices that each forecast model generates for each period of analysis, and determining its maximum dispersion with respect to the average price of the period to be analysed.

Table 8 shows the target values for the two KPIs, for each Base Forecasting Unit (BFU), for the Maximum granularity (MG) and for the actual prediction models. Note that the pluriannual analysis is the one where more uncertainty is accepted in the predictions, while the daily analysis must be very precise to be considered valid.

MODEL	Pluriannual	Annual	Monthly	Daily
BFU	Annual	Monthly	Monthly	Diary
Confidence Interval BFU	80 %	80 %	90 %	95 %
Maximum percentage dispersion BFU	20%	15%	10%	5%
MG	Hourly	Hourly	Hourly	Quarter-Hourly

Confidence Interval MG	50%	80%	90%	95%
Maximum percentage dispersion MG	50%	25%	15%	5%
Confidence Interval Model	80%	85%	90%	95%
Maximum percentage dispersion Model	15 %	7,5 %	5 %	2,50 %

Table 6. Summary of the target values for the KPIs described.

The Maximum Granularity (MG) is a forecast control group considered essential for evaluating the model's performance. In other words, an annual model is of no use if, despite meeting its annual average objectives, it generates monthly prices that deviate significantly from the average.

For example, and following the previous table, an annual model should:

- Obtain an annual forecast where the hourly dispersion does not exceed 25% and the confidence interval is 80%.
- Obtain average monthly prices with a confidence interval of 80% and a dispersion of 15%.
- Obtain an overall model with a confidence interval of 85% and a dispersion of 7.5%.

So, it should not only be a model that provides a correct annual result, but the different granularities should also meet these criteria.

3.1.3.3. Integration with the BESS System

The main objective of obtaining models with reliable price predictions in the different markets is the calculation of the potential revenues obtained by the PV plant according to the different possible operating models for a given time horizon.

As mentioned in Chapter A of this document, the results obtained by these models will serve for each moment as input for the calculation of these different operating modes for the BESS EMS, which in turn will take into account technical operating restrictions, both of the PV plant itself, such as generation forecasts based on the weather, as well as restrictions of the BESS system itself (for example, maximum discharge depth allowed).

Development plan

In order to outline an initial roadmap for the development of this solution, a basic plan of the tasks to be carried out is included. The phases indicated for the development of each model are as follows:

Phase 1, Model Use Validation: For each model, the first step will be to perform an analysis of the different calculation options for the model in question. To do this, working with past data, the necessary set of simulations will be carried out to recognize which forecasting models and against which parameters the desired results are obtained. Once this phase is completed, the model or models that will be the basis of the development will be chosen.

Phase 2, Simple Modelling: In this phase, and after selecting the best possible methods in the previous phase, the future forecast will be made for the periods and granularities indicated in each model. The objective will be to obtain a future forecast that is recognized as validatable, eliminating from the analysis the stochastic processes that may negatively affect pure mathematical modelling.

Phase 3, Stochastic Modelling: Once the base forecast model from the previous point has been validated for each analysis model, the previously indicated stochastic events will be introduced into this model. Starting from past data and modelling these events in specific periods where non-estimable events appear in the calculation, the application of the effect of these unpredictable events in the model will be validated.

Phase 4, Monte Carlo Implementation: With the final model that includes the validated and defined stochastic models, the Monte Carlo analysis will be introduced into these models. In the first phase, work will be done with past values, and then with future values. Both processes must be validated.

Each phase of each model will be validated when the work of the same yield deviation and dispersion data, both for the model in general and for the maximum granularities indicated, within the ranges presented in Table 4 of this chapter.

The development plan will be defined under two axes: On the one hand, each phase of the development within its own model cannot be started until the previous one has been completed, for obvious reasons. Each part of it is built on the results of the previous step, so the strength of each phase of the project must be maximum.

As a second axe, it should be noted that the 4 models will be developed in parallel, with the first starting the hourly modelling and the last the multi-annual modelling, although the phases of the different models cannot be started until that same phase of the previous one is completed. That is, the Validation phase of the use model or Phase 1 of the monthly modelling cannot be started until Phase 1 of the previous modelling or hourly modelling is completed. This is done in this way due to three reasons:

- The complexity of checking and implementing the models increases with the future time horizon being analysed. In this way, it is easier to define a good model for 2 days than for 5 years. This development format enhances continuous learning and makes it so that once the multi-annual modelling is reached, the experience acquired in a repetitive process of increasing complexity simplifies these tasks in the more complex models.
- Minimizes the time to market the most required products in other parts of the project, with more direct application.
- Avoids possible model blockages due to eventual problems in others.

The period from November 1, 2024, to October 31, 2025, will be dedicated to the development of solutions and their validation in controlled environments (e.g., using historical data in a laboratory). Subsequently, from November 1, 2025, to October 31, 2026, the focus will shift to improving the solutions and validating them under real operating conditions in one or more photovoltaic (PV) plants or in laboratory but using real time predictions, due to the work characteristics.

Definition of Validation Environments

- Controlled Environment: This refers to the analysis of historical prices, where both the boundary conditions (events, meteorological data, sector consumption, etc.) and the final results are known. The objective is to obtain results consistent with historical prices based on previous data, simulating price modelling as if it were conducted just before the evaluated dates.

- Real Conditions: In this environment, work will be conducted in real-time, validating the results obtained afterward. This means future modelling will be performed, and subsequently, the accuracy of the forecast modelling will be analysed by comparing it with actual prices. Although ideally, this process would be extended as long as possible, it will be necessary to limit the analysis period to avoid unnecessarily prolonging the project.

Regarding the previous phases, the validation environments for each phase will be as follows:

Phases 1 and 3: Evaluation of models in Controlled Environments, as their objective is to generate the appropriate models.

Phases 2 and 4: Evaluation under Real Conditions, as the objective is to test these models in real use.

It is very likely that some iteration of the verification in Phase 3 will be conducted under real conditions in its final stages, depending on the results obtained.

3.1.4. Conclusions

This document corresponds to Chapter B of the first PVOP deliverable of T3.2 of WP3 which is intended to develop and optimize the AI-based electricity marked predictor for the development of an advanced control system and to truly optimize the battery integration. With the proposed innovations it can be determined that the PVOP project:

- Will implement concepts that are not currently included in forecasting systems, such as the inclusion of stochastic analysis in combination with technical analysis or the modelling of the price as an elastic system.

- The reliability and dispersion objectives indicated will be a clear improvement in the current situation, especially in terms of making financial or operational decisions.

- The price forecast will be launched for several types of markets simultaneously and for several time horizons. The immediate and simultaneous linking of these price signals with the different operating modes of the PV plant and its associated BESS will allow not only to maximize the monetization of the asset but also to establish sufficiently in advance medium- and long-term operating strategies that contribute to a better maintenance and performance of the asset