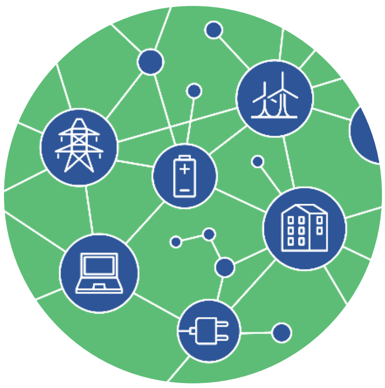
OSMOSE

OPTIMAL SYSTEM-MIX OF FLEXIBILITY SOLUTIONS FOR EUROPEAN ELECTRICITY

Models for market mechanisms simulation taking into account space-time downscaling and novel flexibility technologies

D2.3



Contact: www.osmose-h2020.eu

The project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 773406

Document properties

Project Information

Programme	**O**ptimal **S**ystem-**M**ix **O**f Flexibility **S**olutions For **E**uropean Electricity
Project acronym	OSMOSE
Grant agreement number	773406
Number of the Deliverable	2.3
WP/Task related	WP2: Task 2.3 and Task 2.4

Document information

Document Name	Models for market mechanisms simulation taking into account space-time downscaling and novel flexibility technologies
Date of delivery	09/02/2022
Status and Version	1.0
Number of pages	54

Responsible

Document Responsible	Florian Boehnke (UDE)
Author(s)	Florian Boehnke (UDE), Benjamin Botor (UDE), Sven Kolkmann (UDE), Hendrik Kramer (UDE), Christoph Weber (UDE), Gabriel Kasmi (RTE), Marie Girod (RTE), Clément Brenas (RTE), Maxime Laasrii (RTE), Sandrine Bortolotti (RTE), Virginie Dussartre (RTE), Thomas Heggarty (RTE), Giuditta Pisano (UNICA),
Reviewer(s)	Jens Weibezahn (TUB), Leonardo Petrocchi (TERNA)
Approver	Nathalie Grisey (RTE)

Dissemination Level

Type (distribution level)	☒ PU, Public ☐ CO – full consortium, Confidential, only for members of the consortium (including the Commission Services) ☐ CO – some partners, Confidential, only for some partners (list of partners to be defined)

Review History

Version	Date	Reviewer	Comment
1.0	09/02/2022	Nathan Grisey (RTE)	

Table of content

List of figures

List of tables

0 Executive summary

The electricity system is subject to significant structural changes on both the supply and demand side. In particular, the increasing decarbonization with the associated massive expansion of decentralized renewable energies is leading to the challenge of ensuring the balance of supply and demand with increasing uncertainty in terms of both time and space. These changes are affecting the electricity market's ability to perform their different tasks: determining the short-term behaviour of power system assets, sending appropriate investment signals and ensuring assets can recover their investment costs.

The aim of work package two is to evaluate the ability of different market designs to lead to an optimal mix of flexibility solutions, as well as operate it effectively; this deliverable D2.3 describes the modelling used for this evaluation. Different model frameworks were used that complement each other in terms of their functionalities or use cases, see Figure 1.

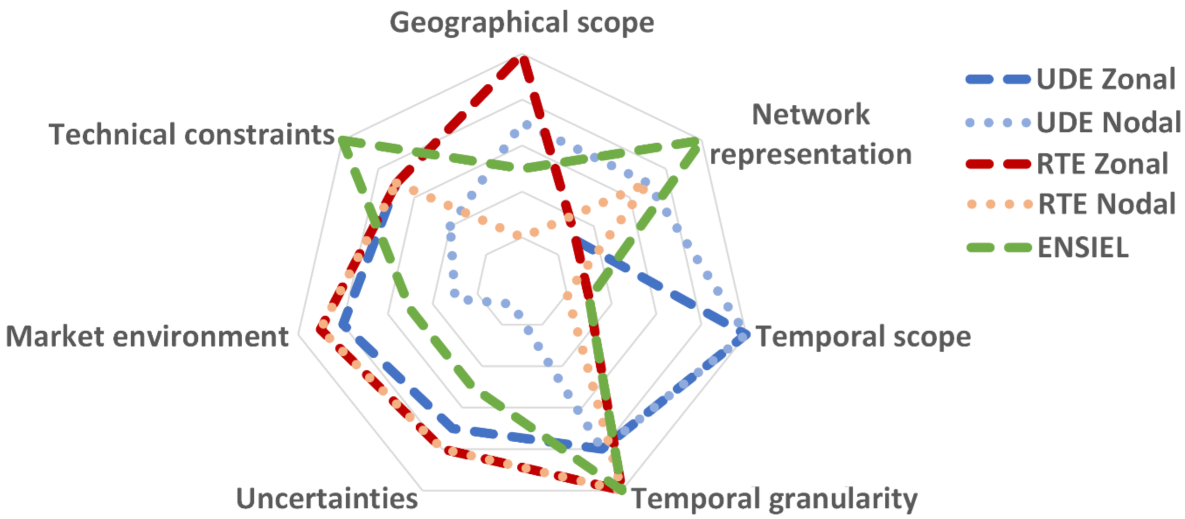


Figure 1: Comparison of different complementary model frameworks of UDE, RTE and ENSIEL

UDE's bottom-up model JMM, described in section 3, uses a rolling horizon planning approach to model a day-ahead, an intraday and a reserve market. Extensions were made to the model to allow the consideration of a nodal market design. RTE's agent-based model ATLAS, described in section 4, is designed to simulate a sequence of electricity markets, gradually revealing information to market participants on the day-ahead, intra-day and real time settings. It is to be used to study both a nodal and a zonal market design.

While describing UDE´s and RTE´s model frameworks, special attention is given to the work carried out to allow space and time downscaling, as well as to model specific flexibility solution behaviour. Note that both these models use the forecast data obtained based on the methodology described in D2.1 "Methodology for error forecasts at a European scale".

A third model was proposed by ENSIEL, described in section 5, to study the interactions between the TSO and the DSO. In the process, a quantitative assessment of the market potential of distributed flexibility solutions is carried out.

This document solely describes the modelling effort; the final simulation results and the resulting conclusions are respectively presented in deliverable D2.4 and D2.5.

1 List of acronyms and abbreviations

In the following table are listed the acronyms and abbreviations used in this document.

Acronym	Meaning
AC	Alternate Current
aFRR	automatic Frequency Restoration Reserve
ATLAS	Agent-based short-Term eLectricity mArkets Simulation
ATC	Available Transfer Capacity
CHP	Combined heat and power
D	Deliverable
DA	Day Ahead
DC	Direct Current
DER	Distributed Energy Sources
DG	Distributed Generation
DPS	Distribution Primary Station
DSM	Demand Side Management
DSO	Distribution System Operator
EPEC	Equilibrium Problems with Equilibrium Constraints
EV	Electric Vehicle
FBMC	Flow Based Market Coupling
FCR	Frequency Control Reserve
FO	Flexibility Options
GDP	Gross Domestic Product
GIS	Geographical Information Systems
HV	High Voltage
HVDC	High Voltage Direct Current
JMM	Joint Market Model
mFRR	manual Frequency Restoration Reserve
MPEC	Mathematical Problems with Equilibrium Constraints
MV	Medium Voltage
NTC	Net Transfer Capacity
OPF	Optimal Power Flow

PCC	Point of common coupling
PF	Power Flow
PS	Primary Station
PTDF	Power Transfer Distribution Factor
PtG	Power to Gas
RAM	Remaining Available Margin
RES	Renewable Energy Sources
TSO	Transmission System Operator
TYNDP	Ten-Year Network Development Plan
UPS	User Primary Station
WP	Work Package

2 Introduction

The overarching objective of the work package 2 is to simulate the market outcome under different market designs considering novel flexibility options and space-time downscaling for the future European electricity system. Deliverable D2.3 contains the applied modelling efforts in the process of the project work in WP2. It is the third in a series of five deliverables in this work package.

Figure 2 outlines the structure of the planned case studies to reach this objective.

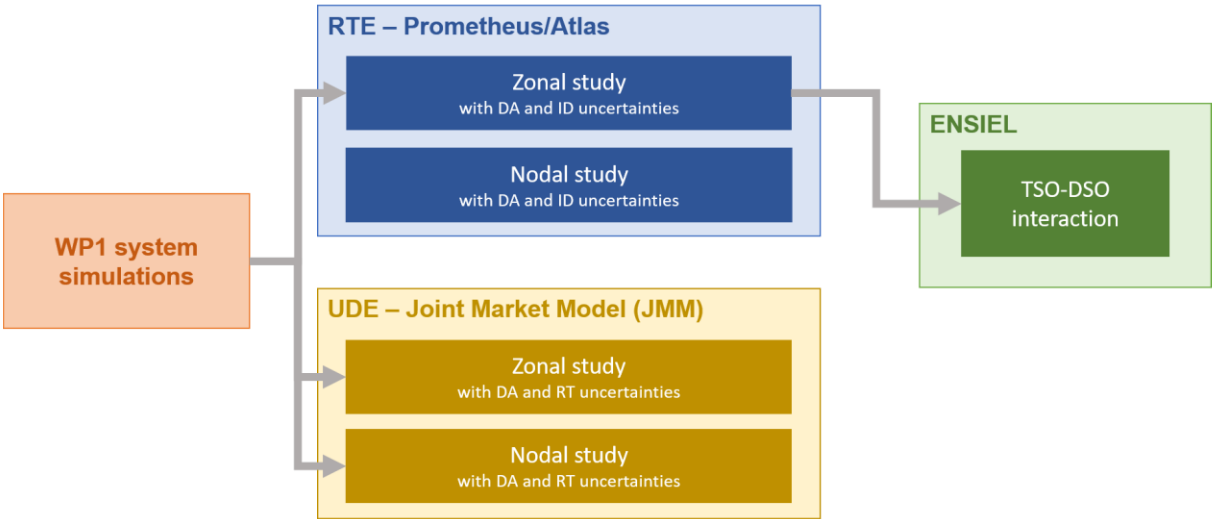


Figure 2: Case Study Roadmap for RTE, UDE and ENSIEL

The base for the work of WP2 are the two electricity market models of UDE (JMM) and RTE (PROMETHEUS/ATLAS). The consideration of uncertainties due to forecast errors bases on preliminary work (D2.1 Methodology for error forecasts at a European scale).

Regarding the JMM (see chapter 3), this implies performing multiple computations for zonal market design, e.g., considering NTCs or FBMC. Additionally, the extensions made to model a European nodal electricity market design will be applied using the same input parameters but scaling them down to high-voltage grid nodes. For both market designs, the consideration of DA uncertainties is an option (see D2.1).

RTEs has performed similar case studies, using an agent-based electricity market model (see chapter 4).

Additionally, the modelling framework of the distribution grid level of ENSIEL is used to analyse the market potential, availability and occurrence of distribution grid level flexibility and serves as an input to inform the large-scale models (see chapter 5).

3 European Market Modelling with JMM

The Joint Market Model is part of a larger model chain as depicted in Figure 3. It is specifically designed to model the outcomes of the current and future interconnected electricity markets, including also related markets like the reserve and district heating markets and reflecting the interplay with the European electricity grid.

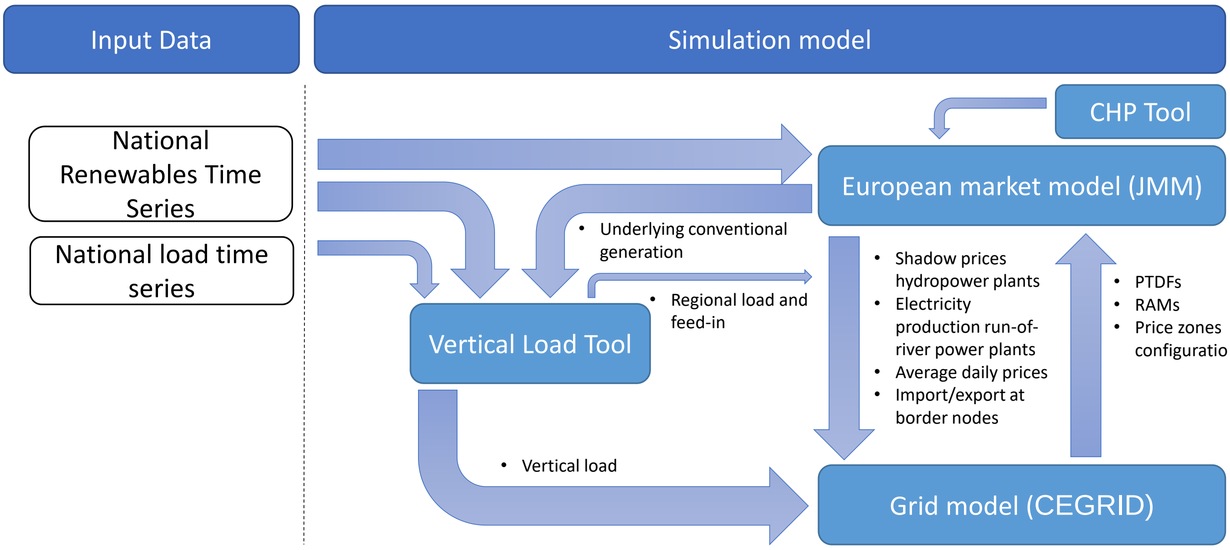


Figure 3: Model chain scheme

3.1 Overview of JMM

The Joint Market Model[1] (JMM) is a model for operational system optimization, which in particular performs dynamic power plant and plant dispatch planning. It was originally developed as part of the EU-funded Wind Power Integration in Liberalised Electricity Markets (WILMAR) project and has been in use in research and industry for over ten years. An LP and an MIP formulation exist, which are largely identical, can access the same database, and can be selected via a software switch. The model is written in the General Algebraic Modelling System (GAMS) programming language and is usually solved with the CPLEX solver.

The JMM is based on the assumption of system cost minimization but models each hour of the considered year. The models' objective function includes fuel costs, CO_2 costs, start-up costs and other variable costs for maintenance and insurance. Technical restrictions such as start-up times, minimum operating and shutdown times, part-load efficiencies, minimum and maximum generation, and reserve requirements are considered. In addition to electricity demand, heat demand must be met for all 8,760 hours of the year. Modelled market prices reflect marginal generation costs. The input data for the model are the university's own Europe-wide power plant database as well as detailed load and weather data, which are available in regional

[1] https://openenergy-platform.org/factsheets/models/61/

resolution. The data is stored and maintained in a structured database and can be specifically compiled for different scenarios. The power plant database comprises around 7,500 power plant units. In addition to conventional power plants, the model can also represent various storage technologies and demand-side flexibilities (demand side management). The following power plant types are implemented in the model:

- Biomass plants
- Lignite and hard coal power plants
- Gas-fired power plants, notably combined cycle gas turbines (CCGT)
- Nuclear power plants
- Waste incineration plants
- Oil-fired power plants
- Pumped storage power plants, water reservoirs and run-of-river power plants
- Thermal boilers
- Wind and PV plants

The conventional power plant types are also divided into technology classes, e.g. to distinguish different turbine types.

- Steam turbines, separate classes for extraction condensing and backpressure turbines
- Gas turbines, separate classes for fixed or flexible waste heat utilization
- Combined cycle gas turbines, separate classes for extraction condensing and back-pressure turbines
- Internal combustion engines, separate classes for fixed or flexible waste heat utilization

A further differentiation of power plant types is made according to plant age. Overall, the power plants can thus be differentiated by fuel, technology and age, which leads to a broad differentiation with respect to their parameters.

For the LP runs, the power plants are aggregated by power plant type and age class. These aggregated units are used for Europe-wide runs to determine exchange flows and prices. In the MIP runs, the power plants are modelled on a unit basis. These runs are calculated to determine the power plant dispatch under fixed exchange flows at market area level and determine the hourly power plant and storage schedules.

In order to be as close to reality as possible, the interactions between conventional power plant dispatch, feed-in of renewable energies, available reserve capacity (FCR, aFRR and mFRR) and available transmission capacities are represented. By modelling regional heat markets, potential constraints of heat-led CHP plants in the electricity market can also be mapped. In addition, international electricity trading is mapped via Flow-Based Market Coupling (FBMC) using a PTDF approach. The mapping via a linear transport model using the NTC approach is also possible for regions that do not participate in the FBMC. The geographical coverage of the JMM currently includes the 28 EU countries (except Malta and Cyprus), Norway, Switzerland and the Balkans.

The basic concepts for spatial resolution in the JMM are shown in Figure 4. Different bidding (or price) areas in electricity markets are represented by different regions, which are interconnected. Usually, a country is divided into one or more regions. The market in each region is modelled under the assumption of no internal congestions. Different areas further may be used to subdivide these regions and to allow detailed modelling of the district heating grid.

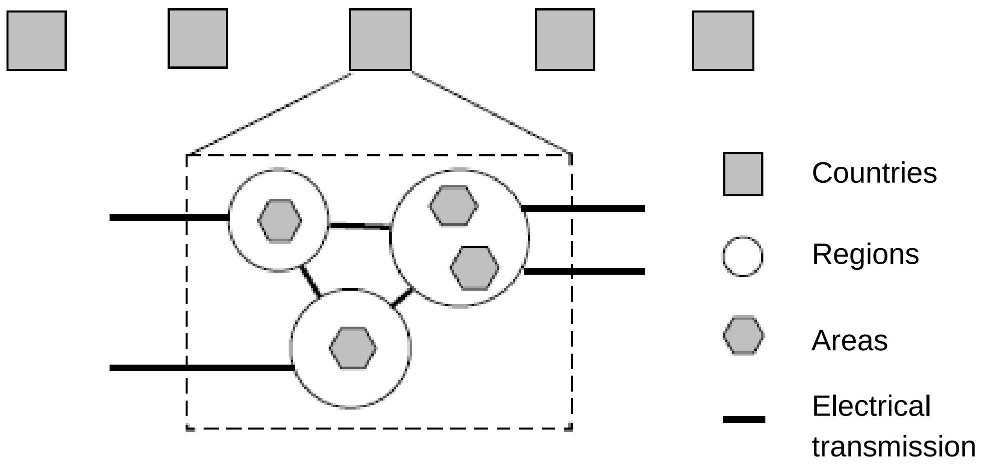


Figure 4: Spatial structure in the JMM

An important functionality of the JMM is the rolling planning, which maps the interaction of day-ahead and intraday markets and allows to model the arrival of new information. Here, day-ahead and intraday loops alternate. In the standard specification, the DA loop covers 36 hours and the ID loop 24 hours. The start of the optimization period advances by 12 hours after each loop. In the DA loop, the trading volumes of the following day are determined for the period from 0 to 24 hours. The reserve market is also optimized for the following day in this step. The following ID loop then starts at 0 o'clock of the next day and determines the intraday market result of the first twelve hours taking into account the day-ahead market result. In addition, the power plant dispatch of hours 13 to 24 can be corrected if deviations in feed-ins or load require it. The following figure shows this process:

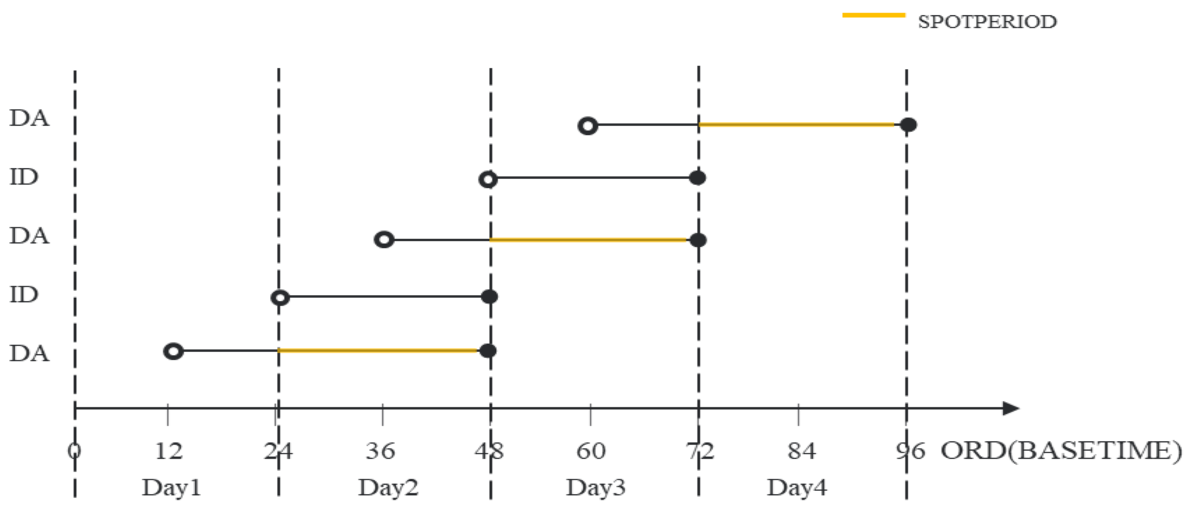


Figure 5: Example of the used rolling planning approach, here for day-ahead/intra-day planning

This approach is illustrated in Figure 5 and does not only allow to model repeated information updates but it also reduces computation times. The time resolution of the JMM is strongly connected to that rolling planning structure. The rolling planning structure is also essential to cope with uncertainties due to renewable power generation. During the day ahead allocation the optimal unit commitment decision is being computed subject to technical characteristics and an initial RES forecast. Updates on this RES forecast can be fed in the model by inserting an information update time series. The unit commitment adaptions and price changes can be interpreted as the intraday market results.

Obviously reducing the time intervals between information updates leads to an increase in computation time needed to receive the model outcomes. Therefore, this function has been used in the past mostly in a 12-h looping structure as it is displayed in figure 3. In the context of OSMOSE however we implement and apply a 3-h looping structure to reflect a more realistic setup. This has resulted in some changes in the model structure, mainly adjusting the output and input structure as well as adjustments regarding the data outputs of the connected tool chain (Forecast Error Tool).

The JMM provides a comprehensive hourly resolved output that includes:

- Electricity prices
- System costs and welfare effects
- Electricity and heat production per plant
- Provision of control power (reserve) per plant
- Storage levels
- Utilization rates
- RE balancing
- Electricity exchange between market areas
- Transmission losses
- fuel consumption
- CO_2 emissions

The input data is pre-processed via a chain of two computation tools which were developed at the Chair of Energy Economics of the University of Duisburg-Essen:

- The chair's CEGRID model is used to simulate the operation of the power grid on a nodal level and to derive PTDFs and RAMs for the calculation in the JMM. To be able to obtain FBMC parameters (zonal PTDF and RAM parameters), a meaningful base case market run without any grid information, thus only NTC information, is required as input for the CEGRID model. Therefore, the JMM must be executed once priorly without any grid information. Based on the dispatch information and prices, the FBMC parameters can be calculated and be used in another FMBC market run by the JMM.
- The chair's Vertical Load Tool is used to disaggregate country-level data to obtain demand and RES time series for each node in the network. Due to the huge amount of data for the nodal model, it is not stored in the chair's data base, which usually does the conversion of input data into input files for the JMM, but some new code has been

elaborated in order to short-cut the data base for most of the nodal input data, leading to less storage need and faster processing.

3.2 Model Description Regionalization Model Vertical Load Tool

The Vertical Load Tool is a versatile tool used for detailed mapping of both regional loads and regional generation for Germany and the CORE (CWE + CEE) region. The main component of the tool is the simulation of wind feed-in as well as PV feed-in using available official power plant data and detailed weather data. Furthermore, the tool is able to adequately distribute future power plant capacities regionally.

In addition to installed capacities and feed-in quantities for renewable energies, load time series of private households, trade and industry are derived by means of standard load profiles and regionalized using suitable methods.

With the help of data from the European market model (JMM), the generation from non-volatile renewable energies is distributed regionally, thus completing the load regionalized from the Vertical Load Tool. Finally, the regionalized load is assigned to the nodes of the CEGRID model.

The Vert Load Tool determines the load curves of individual regions based on publicly available socio-economic parameters. In particular, data on economic performance and the number of inhabitants are relevant, which are used to derive the energy demand of individual regions. In summary, the following input data are used:

- Current total population per regional (Germany and Europe)
- Regional gross value added of industry and trade (Germany and Europe)
- Regional annual electricity demand for industry including electricity demand of individual industrial sites for electricity-intensive processes
- Standard load profiles for trade and private households optionally also real measured time series of various distribution network operators of non power-metered customers
- Current and future national annual electricity demand of individual sectors

The tool follows a bottom-up approach: The country-wide electricity demand for the sectors households, trade and industry is specified and allocated proportionally to the respective sectors. The distribution key is based on the socio-economic parameters mentioned above. For the industrial sector, known regional electricity demands are additionally used and only the remaining industrial electricity demand is distributed on the basis of the industrial gross value added.

Standard load profiles are used to generate the load curves, with which both the household load and the commercial load are mapped and extrapolated on an hourly basis. The industrial electricity demand is calculated by means of the difference between the nationwide hourly load and the aggregated hourly load from trade and households. With the help of the regional total electricity demand from industry, the previously described load profile and the regional load profiles from trade and households, regional hourly load profiles are generated.

Regional time series for variable generation are determined using real production data, weather data and plant data. To determine the hourly production of volatile generation units, various factors are relevant - the generation capacities of a region (from the Marktstammdaten-register), the plant data (power curve, height of a plant) as well as the available weather data of the German Weather Service (DWD), and national feed-in time series for the base year to adjust modelling inaccuracies.

The hourly feed-in time series are modelled with the known plant types and using the DWD data for each individual region. The turbine data (wind and PV) for Germany are assigned to the respective regions by means of latitude and longitude lines using the publicly accessible market master data register (for Europe: WindPower database) and scaled up according to the national total installed capacity. The short-term addition of generation capacity is based on the current distribution of wind turbines. In the medium and long term, generation capacities are distributed considering regional saturation effects, land restrictions and site quality.

3.3 Forecast modelling – spatial component

The application of the European market model framework requires appropriate forecast data. These are determined based on the forecast methodology presented in Deliverable 2.1. Yet to cover several countries, the forecast modelling approach has to be extended by a spatial component. Given the limited availability of data sources that allow to assess combined temporal and spatial dependencies, we develop a method suitable for sparsely available, but coherent data.

Our existing methodology assesses dependency structures by transposing forecast trajectories, namely $P_{t,T}$, into forecast updates $\Delta_t P_{t,t+k} = P_{t,t+k} - P_{t-1,t+k} = x_k$. Based on the marginal distribution function $u_k = F_k(x_k) \Leftrightarrow x_k = F_k^{-1}(u_k)$ we elaborate the, in a first step solely temporal, joint distribution function of all updates published at the same point in time t. To avoid the complex process of finding a joint distribution function, we take advantage of Sklar's theorem and apply a copula on the marginals to replicate the joint function: $F(x_1, x_2, \dots, x_n) = C\big(F_1(x_1), \dots, F_n(x_n)\big)$.

Though the copula parameters above can be extended and fitted to assess both temporal and spatial dependencies at the same time, this demands the availability of a full data set. Namely, hourly updated forecast trajectories simultaneously available for all involved spatial entities (regions) S. As this matrix grows in size vastly, we increase the practicableness of the approach by implementing two restrictions in our modelling. First, we restrict ourselves to the special case of a Gaussian Copula, $C = F_K(x_1, \dots, x_K, R)$ with R being a covariance matrix. Second, we constrain the spatial and temporal component of R to be independent. This allows us to estimate a global temporal covariance matrix R^K used for all involved spatial entities $s \in S$ based on quality monitored data provided by RTE. Next, we fit the weaker spatial dependency component with data from various sources including the ENTSO-E transparency platform:

$$\mathrm{R^S} = \begin{pmatrix} \rho_{\mathrm{s_1,s_1}} & \cdots & \rho_{\mathrm{s_1,s_n}} \\ \vdots & \ddots & \vdots \\ \rho_{\mathrm{s_m,s_1}} & \cdots & \rho_{\mathrm{s_m,s_n}} \end{pmatrix}$$

Given the simplification of independent temporal and spatial correlation, we can compose the estimate of the joint matrix as

$$\hat{R}^{S,K} = \begin{pmatrix} \rho_{s_1,s_1} \cdot R^K & \cdots & \rho_{s_1,s_n} \cdot R^K \\ \vdots & \ddots & \vdots \\ \rho_{s_n,s_1} \cdot R^K & \cdots & \rho_{s_n,s_n} \cdot R^K \end{pmatrix}$$

In a last step, we need to adjust the covariance matrix obtained through the estimation procedure sketched above, to restore the properties of a Gaussian copula, namely the correlation matrix being positive semidefinite. We therefore apply the approach presented by Cheng & Higham (1998)[2]. The resulting matrix $R^{S,K}$ is then incorporated in the copula of the approach presented in Deliverable 2.1.

3.4 Spatial Downscaling

3.4.1 From zonal to nodal market modelling

For nodal market modelling, the regions are redefined as nodes of the extra-high voltage grid. Thus, the PTDF-based approach to flow-based market coupling leads to a nodal DC optimal power flow. In this case, areas and regions are decoupled from each other, which means that none of them needs to be a sub-division of the other one.

A dedicated analysis was performed to identify the impact of different model simplifications on key outcomes and performance. The following simplifications were tested

- Minimal operating and down-time
- Consideration of reserve
- Start-up costs

The total runtime changes significantly with different solution algorithms and especially with different simplifications for our test case (nodal German on transmission network level) from about 2 days to 10 days.

Analyses of the changes in the operation of the various units in the systems show that, in particular, taking start-up costs into account leads to significantly better results for moderate runtime extensions.

3.4.2 Obtention of nodal input data

To obtain nodal input time series for renewable energy technology power infeed as well as load, a three-step bottom-up process is implemented:

- First, capacities for the future scenario are determined by a techno-economic discrete-choice investment model. Restricted by regional upper potential bounds, the model considers local weather information to determine the locations where investment is the most profitable. This approach is used to distribute a national target capacity onto regions.

[2] Higham, N. J., & Cheng, S. H. (1998). Modifying the inertia of matrices arising in optimization. Linear Algebra and its Applications, 275, 261-279.

- Second, a bottom-up time series heuristic is applied (cf. also Section 3.2). Input data for the model are the before-mentioned regional capacities at NUTS3 level, as well as high-resolution weather data (solar irradiation and wind speed). Through technical representation of the decentral units, an aggregated power infeed time series is derived at regional level.
  Regarding load, a national load profile is separated into a national household, commercial and industry load profile. Then every regional load profile is composed based on the local shares of these national time series. Different types of GDP (industry and commercial), as well as population figures are used to derive the relative shares of each region in reference to the national demand figures. In that way, the national demand is distributed onto all regions to obtain regional demand time series.
- Third, the regional power time series are allocated to grid nodes nearby, namely substations. This is implemented by calculating Voronoi polygons around each grid node and assigning the relevant linear combination of the adjacent regional production time series onto the specific grid node according to its surface area.

This detailed bottom-up method requires more data than provided by OSMOSE project partners in working package 1, thus data pre-processing has been performed. This includes particularly the analysis of restrictions regarding areas not suitable for wind turbines or photovoltaic plants. These include lakes, forests, cities, airports and their relevant protection or safety perimeter. Furthermore, depending on the model year, a fraction of the potential area must be subtracted, given the existing turbines and power park modules, thus today's capacities must be known.

After performing the bottom-up simulation based on estimated regional capacities and the underlying weather information, the results are aligned to the scenario data of WP1. Thus, for every hour, the power infeed is adapted linearly to match the given exogenous power value provided by WP1.

The nodal assignment of the time series finally is performed using Voronoi areas. Given the existing power time series on regional level (see blue boundaries in Figure 6), each node is allocated a certain share of the regional time series.

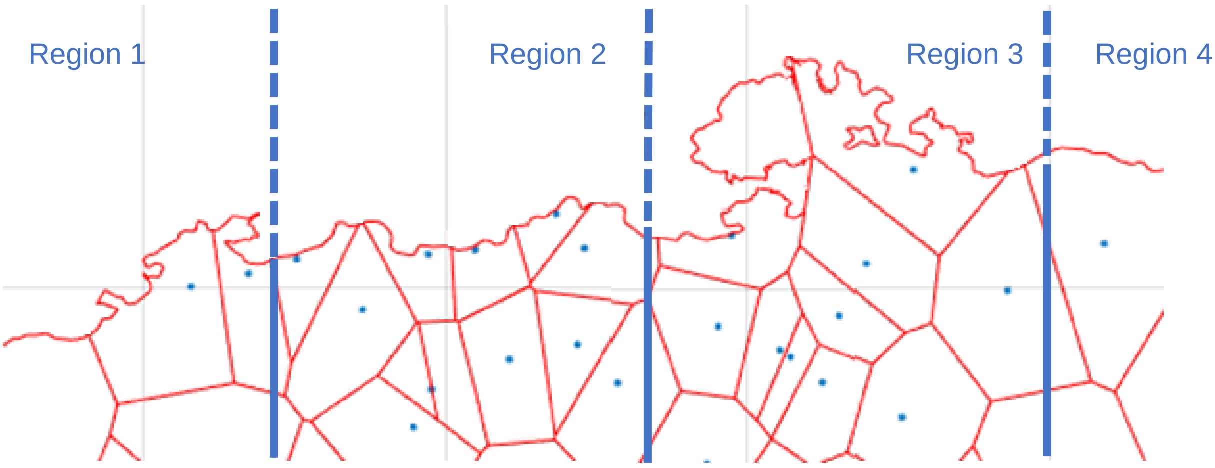


Figure 6: Exemplary boundaries of regions (blue) and Voronoi polygon based on grid nodes (red)

If the nodal Voronoi area is part of two regions, then a share of each of these regions is allocated to the Voronoi area, proportionally to the respective surface area. Likewise, the total production of one region is spread over the nodes according to the Voronoi polygon areas in that region.

3.5 Grid Modelling: CEGRID

The CEGRID model supports power flow (PF) analyses on the extra-high transmission grid level. Besides the grid data, it comprises sanity check routines to fix inconsistent input data and then passes the relevant parameters to the JMM. The required information to incorporate the grid representation into the market model are the zonal and nodal exchange capacities between regions as well as PTDF factors in case of flow-based market coupling or nodal simulations.

The CEGRID model is based on the TYNDP dataset, enriched by attached open-source geographic information. The countries of scope (Austria, Belgium, France, Germany, Luxembourg, the Netherlands and Switzerland) are surrounded by an equivalent node for each continental neighbouring country that accounts for the zonal power generation and load in these countries.

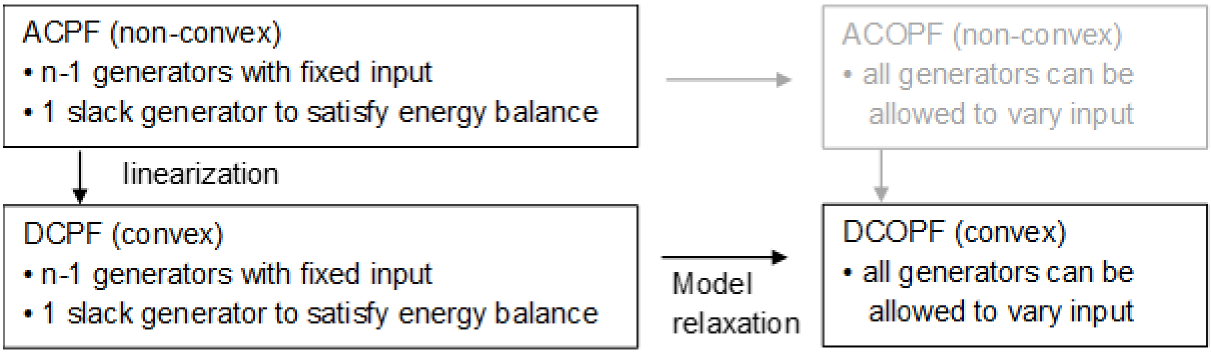


Figure 7: Power flow calculation and simplification procedure

In general, different types of PF computation exist, see Figure 4. The existing power transmission system is based predominantly on alternate current (AC) which comes with non-convex physical principles for ACPF computations. Consider the daily task of a TSO: Estimation of line loading at a known power plant schedules. Non-convex estimation of the line loading has to be performed numerically with several iterations to converge towards a local optimum, defined as least distance to mathematical representation of the grid. Thus, a meaningful starting point close to the final result for this ACPF has to be known to start the numerical iterations.

One option to obtain such a starting point may be the direct current (DC)PF, which linearizes non-convex physical properties of the grid. This may distort the result, yet convex problems can be solved much easier and nice properties exist. In the case of linear models, additionally it can be proven, that if there is a valid solution to the problem, it will be found exactly. This solution then may serve as the starting point for the exact ACPF described above. However, this is just the exact solution of the linearized problem. Furthermore, the DCPF formulation needs to make sure that it will return a valid solution. The linearization yields in disregarding certain physical characteristics such as reactive power and losses. In the new lossless formulation power generation and consumption most likely will not add up and thus there is no solution to the problem. Therefore, one generation unit in the model will be denoted as slack power plant, that compensates the differences in power supply and demand.

When the linearized DCPF model formulation is relaxed furthermore, such that all generation units have a range of operation modes instead of a previously fixed input, the model has a larger set of possible solutions. Given a meaningful objective function such as minimizing operation cost, the model will rather use cheap power plants, if possible, to satisfy all model constraints. This approach is widely used in the research area of economic grid allocation questions – with various alternative settings, e.g., regarding the unit commitment, i.e., whether the on/off status of the units is fixed.

Given the solely linear dependencies within the DCOPF model formulation, the impact of a change of generation output at one node onto each transmission line loading in the grid can be described through linear factors. These are called power transmission distribution factors (PTDF) and may be combined into one data matrix.

Zonal market clearing processes in Europe yet do not account for each single transmission line loading in the first step. Rather the markets are coupled by the flow-based market coupling mechanism. Therefore, CEGRID computes the relevant flow-based parameters, given a previous market base case and then returns this information of relevant transmission line limits to the JMM. Then, the market model JMM estimates the cost-optimal dispatch, such that the relevant flow-based gates will not be overloaded. Subsequently, ex post line overloading must be mitigated if applicable. Thereby a modified OPF formulation is used to model this so-called “redispatch”. It typically fixes the on-

Grid modelling is generally prone to data availability, Power flow consistency checks only return information, if the grid topology is completely consistent. Furthermore, only one grid topology for the whole year is chosen, as hourly new topology information causes large amounts of data that must be transferred between the CEGRID and JMM model.

3.6 Flexibility solution modelling

3.6.1 Storage technologies

The JMM includes various flexibility options (FO). This section aims to provide insights on which technologies are modelled as storage technologies according to their load shifting ability. As the model runs are based on the OSMOSE WP1 results, we focus on the technology types proposed, namely: Stationary Batteries, Electric Vehicles, Heat Pumps, Industrial DSM, Pumped Hydro Storages and Power-to-gas. Regarding their restrictions, all those flexibility options are built up similarly, due to their underlying storage characteristic. However, to meet individual technical characteristics of each technology, modifications are implemented. In the following we introduce the main equations that all flexibility options of this type are subject to and outline the respective features when necessary.

Storage Level Equation:

$$V_{a,i,t}^{STORAGE} = \left(1 - \eta_{a,i}^{STOLOSS}\right) \cdot V_{a,i,t-1}^{STORAGE}$$

$$+ \eta_{a,i}^{LOADEFF} \cdot \left(W_{a,i,t}^{DA} + W_{a,i,t}^{+} - W_{a,i,t}^{-}\right) \cdot \Delta t - \left(P_{a,i,t}^{DA} + P_{a,i,t}^{DA} - P_{a,i,t}^{DA}\right) \cdot \Delta t$$
$$- \left(i_{a,i^{EV},t}^{LEAVE} \cdot v_{a,i^{EV}}^{STORAGE,MAX} \cdot SOC_{a,i^{EV},t}^{LEAVE}\right) + \left(i_{a,i^{EV},t}^{ARRIVE} \cdot v_{a,i^{EV}}^{STORAGE,MAX} \cdot SOC_{a,i^{EV},t}^{ARRIVE}\right)$$

$$,\forall a \in A, i \in I^{ESTORAGE\_DSM}, t \in T^{SPOT}$$

The storage level equation computes the storage content ($V_{a,i,t}^{STORAGE}$) at the end of the current time step and is built up of three main elements. The first one is the storage level of the previous time step ($V_{a,i,t-1}^{STORAGE}$) whereby also storage losses ($\eta_{a,i}^{STOLOSS}$) are considered. Secondly (dis-)charging activities ($W_{a,i,t}^{DA}$) of the calculation period are incorporated. These activities are subdivided into charging and discharging and include (un-)loading efficiency as well as up/down adjustments on the intraday market ($W_{a,i,t}^{+}$).

Work in OSMOSE has focused notably on a detailed and consistent modelling of electric vehicles (EVS). The third term in the preceding equation in this context considers the option that a storage facility is only available for a limited period of time. In the case of electric vehicles, this is reflected by capacity reductions or increases from leaving or arriving vehicles. For leaving EVs, this term is calculated by multiplying the share of leaving vehicles ($i_{a,i^{EV},t}^{LEAVE}$) with their maximum storage capacity ($v_{a,i^{EV}}^{STORAGE,MAX}$) and the expected state of charge when leaving ($SOC_{a,i^{EV},t}^{LEAVE}$). Arriving vehicles are accounted for accordingly. Alternatively, EVs can be modelled as DSM units, as suggested by WP1. In this case a percentage of the total installed capacity is included as a battery according to the assumed load shifting potential.

The capacity for the loading process of a storage unit is restricted by the following equation.

$$W_{a,i,t} + W_{a,i,t}^{+} - W_{a,i,t}^{-} + W_{a,i^{SPIN},t}^{ANC,-} + W_{a,i,t}^{NONSPIN,ANC,-} \leq \begin{cases} w_{a,i}^{MAX} & \forall i \notin I^{HEATPUMP}, \\ \dfrac{q_{a,i}^{MAX_{PROD}}}{\eta_{a,i}^{FULLLOAD}} & \forall i \in I^{HEATPUMP} \end{cases}$$

$$,\forall a \in A, i \in I^{ESTORAGE\_DSM}, t \in T^{SPOT}$$

It assures that the scheduled charging capacity on the day ahead market, corrected for the up or down adjustments on the intraday market as well as the blocked capacity ($W_{a,i^{SPIN},t}^{ANC,-}, W_{a,i,t}^{NONSPIN,ANC,-}$) to provide negative reserves (i.e. intake of oversupply) does not exceed the loading capacity of a storage ($w_{a,i}^{MAX}$) or the heat capacity of a heat pump ($q_{a,i}^{MAX_{PROD}}/\eta_{a,i}^{FULLLOAD}$), multiplied with an outage factor.

To ensure that the scheduled capacity for loading a storage unit at the day-ahead market is lower than or equal to the available capacity, the considered units are additionally subject to the following restriction.

$$W_{a,i,t} \le \begin{cases} w_{a,i}^{MAX} & \forall\, i \notin I^{HEATPUMP}, \\ q_{a,i}^{MAX_{PROD}} / \eta_{a,i}^{FULLLOAD} & \forall\, i \in I^{HEATPUMP} \end{cases} \quad , \forall\, a \in A, i \in I^{ESTORAGE\_DSM}, t \in T^{SPOT}$$

Further, the following restriction limits the maximal contribution of storage units in the provision of positive reserves.

$$W_{a,i,t} + W_{a,i,t}^{+} - W_{a,i,t}^{-} - W_{a,i,t}^{NONSPIN,ANC,+} - W_{a,i,t}^{ANC,+} \ge 0$$

$$, \forall a \in A, i \in I^{ESTORAGE\_DSM}, t \in T^{SPOT}$$

All storage units are also subject to a standard maximum and minimum content restriction. Regarding the lower bound, all FO are limited to positive storage contents except for Industrial DSM units. The content for those units may, depending to their technical characteristics, reach negative values for a certain duration – this corresponds to delayed electricity consumption.

Heat pumps have an analogous but somewhat simpler structure, since no reserve provision is foreseen. Other restrictions to ensure appropriate modelling like no simultaneous charging and discharging are also included in the JMM but not further described in this document.

3.6.1.1 Hydro Power

In line with the already described flexibility options, hydro reservoirs are also modelled as storage technologies. Due to their exogenous inflow, they are subject to different equations within the JMM framework. To calculate the current storage level (within the lower and upper bound), the storage level equation is used. It considers the content at the previous time step, the power production, the natural water inflow as well as possible spillage. It should also be noted that hydro reservoirs are modelled with the capability of pumping water, thus allowing them to reflect both characteristics, being considered a seasonal and a pumping storage. In this way our model can also consider pumps that are added in retrofit measures.

$$V_{a,t}^{HYDRO} = V_{a,t-1}^{HYDRO} + i_{a,t}^{INFLOW} \cdot \Delta t - \mathrm{p}_{a,t}^{SPILL} \cdot \Delta t$$

$$- \sum_{i \in I_a^{HYDRORES}} \left(P_{a,i,t}^{DA} + P_{a,i,t}^{+} - P_{a,i,t}^{-}\right) \cdot \Delta t + \eta_{a,i}^{LOADEFF} \cdot \left(W_{a,i,t}^{DA} + W_{a,i,t}^{+} - W_{a,i,t}^{-}\right) \cdot \Delta t$$

$$, \forall a \in A, i \in I^{HYDRO}, t \in T$$

Aside the storage equation, the hydro reservoir units are subject to the already described restrictions.

3.6.1.2 Power-to-Gas

Power-to-Gas (PtG) as a flexibility option of the future is accounted for as well. From a modelling perspective, it can be considered as a storage technology. As its functionality is considerably different from other storage technologies, it is yet described in a standalone section. A distinction between Power-to-Methane and Power-to-Hydrogen can be made through different input parameters, i.e. conversion rates and use values, yet the general implementation is similar and is described in the following. As the focus of our modelling approach is on the economic perspective and interactions between the electricity market, the CO_2 price and PtG. we refrain from a detailed technical modelling of the conversion process and a consideration of demand for methane or hydrogen. In the JMM framework, the PtG facility is utilized if the electricity price is lower than (or equal to) the value of the converted gas, i.e. the use value. This general calculus forms the basis for the implementation.

The use of power-to-gas leads to an additional electricity consumption in the specific region in the system. This electricity consumption is interpreted as filling of a storage. The filling rate is thereby restricted by a maximum filling rate per hour (MWh/h) and corresponds to the installed PtG capacity. The following equation describes the storage level ($V_{a,i,t}^{PtG}$) and is constructed similarly to the aforementioned storage equation.

$$V_{a,i,t}^{PtG} = V_{a,i,t-1}^{PtG} + \eta_{a,i}^{LOADEFF} \cdot \left(W_{a,i,t}^{DA} + W_{a,i,t}^{+} - W_{a,i,t}^{-}\right) \cdot \Delta t\,, \qquad \forall a \in A, i \in I^{PtG}, t \in T$$

The maximum storage capacity is chosen to reflect the maximum acceptable capacity of the gas infrastructure for synthetic gases. Because of the relatively large storage capacity in comparison to the installed PtG capacity, the limitation of the storage capacity has no effective impact. The electricity consumption of the electrolyser $W_{a,i,t}^{DA}$ is considered in the balance constraint and increases the electricity demand that has to be covered by generation and imports. Additionally, Power-to-gas facilities are able to participate in the reserve markets. In case of an electricity surplus, the utilization of a PtG facility and therefore the electricity consumption can be increased, whereas positive reserve can be used by decreasing the scheduled utilization.

The reconversion of synthetic gases produced by PtG facilities to electricity is not modelled in this case study. It is assumed that these gases are used in other sectors e.g. industry or mobility so that storage discharging does not contribute to the electricity generation of the system. To adequately reflect this cross-sector utilisation, PtG facilities receive a monetary compensation, namely the use value which represents a revenue for every MWh of consumed electricity as an incentive for their utilization. Consequently, PtG units are just used in time steps when the price for electricity is lower than the use value so that the use value represents an upper electricity price limit for PtG utilization in the modelling approach.

3.6.2 Reserve Modelling

Flexibility options in the JMM framework can, subject to their technical characteristics, provide balancing services. The JMM covers the Frequency Control Reserve (FCR), the automatic Frequency Restoration Reserve (aFRR) and the manual Frequency Restoration Reserve (mFRR). As each of these may include a positive and negative component, six restrictions control the provision of balancing services. The day-ahead markets for reserves (i.e. FCR) are described by demand restrictions for reserve provision in a so-called TSO region. These TSO regions may or may not coincide with the bidding regions for the day-ahead and intraday electricity market The exogenously given demand for the different balancing services for upward regulation (positive reserves) can be supplied either by increased power production of the power producing unit groups or by reduced loading of electricity storages and use of heat pumps. The exogenously given demand for downward regulation (negative reserves) can be met by decreasing the power production or by increasing the loading of electricity storages and use of heat pumps. Only spinning unit groups can provide FCR and aFRR. The eventual activation of balancing services during actual grid operation is not considered further. The provision of reserves has only an indirect impact on electricity market prices: The provision of reserves itself does come at no cost, yet it restricts the range of operation for units that provide balancing services and may thus induce price changes. In the process of the OSMOSE WP2 work, we added corresponding components to the JMM model framework.

First and foremost, we have extended the possibility of interzonal exchange of reserves from solely mFRR, to FCR and aFRR. In line with the outcomes from WP1, we include reserves for each country / market zone and this requires interzonal exchange of reserves since not all countries possess sufficient reserve capabilities to meet their demand on their own. Even though the activation of balancing power is not modelled, the corresponding capacities must be available, so the cross-border provision of reserves obviously affects transfer capacities. Additionally, we now include the possibility of providing reserves from wind power capacities by implementing an optional switch and adapting the restrictions. In order to provide positive reserve, wind generation must be reduced by the corresponding amount taking the expected wind infeed as reference. The negative reserve provision does not affect the wind power generation as we model only the capacity provision and not the actual activation.

4 Agent Based Modelling considering uncertainties with ATLAS

ATLAS (for *Agent-based short-Term eLectricity mArkets Simulation*) is an agent-based model designed to simulate a sequence of electricity markets. It was conceived and is developed by RTE within the PROMETHEUS cloud platform.

Its value lies in its ability to simulate profit-maximizing agents exposed to uncertain information of which the precision is gradually improved over time. As this information gets revealed to market participants, they adapt their planning and infer the orders they should submit to subsequent energy markets. This sequential modelling therefore enables the computation of uncertainty-dependent prices, as well as other market indicators which might be relevant to analyse a market design's performance.

The ATLAS model features a specific type of agent-based modelling designed to represent the real market process, with its market gate closure time and standard products. The approach is therefore different from classical and more theory-oriented models such as equilibrium problems with equilibrium constraints (EPEC) or mathematical problems with equilibrium constraints (MPEC) models, where the agent and market responses can often be expressed analytically[3]. Instead, the ATLAS model sequentially simulates real market orders with their volume, price, and link constraints, then runs a market clearing algorithm that replicates the real EUPHEMIA algorithm in a simplified fashion, and then simulates the response to that market

[3] Vendosa et al., Electricity market modelling trends, Energy Policy, 2005

from market participants. These different steps can be modified and parameterized, making it possible to test different player behaviours, different market rules or different energy mixes.

The main advantages of this approach are its modularity and the possibility to represent complex but realistic behaviours and constraints. A key contribution of this study will therefore be to evaluate the contribution of such a modelling approach to the analysis of future market operation. The development and interfacing of the required model components is however extremely complex; the Prometheus platform and the ATLAS model were based on results of the Optimate FP7 project[4].

As an agent-based model, ATLAS adopts the point of view of market players, who submit buy- or sell-orders on electricity markets, and program their dispatchable generation and flexibility assets to maximise their profit. In practice, for the study, two agents are taken into account for each country: a single consumer and a single generator.

The model is structured in modules that each achieve a specific task corresponding to a stage in the market sequence. Every module is divided into sub-modules that can be used to model the behaviour of a specific technologies.

The Antares-simulator[5], used for the WP1 studies, is part of these modules. Antares-Simulator is an Open-Source software which simulates the hourly dispatch of large-scale interconnected power systems over many weather years. In contrast to ATLAS, it makes a perfect foresight assumption, and adopts a centralised benevolent monopoly point of view. Within ATLAS, Antares is used to provide market players with Day-ahead price forecasts used for the formulation of market orders. Antares studies can also be used as an input data format: the studies performed in this work package uses the "2030 current goals achieved" study of WP1 as its input.

The input data and main modules are described in more detail in the following paragraphs. Special care was taken in specifying the differences in modelling principles and assumptions between all the modules.

[4] (PDF) OPTIMATE: An Open simulation Platform to Test Integration in MArkeT design of massive intermittent Energy (researchgate.net)
[5] Shedding light on the future of the energy system (antares-simulator.org)

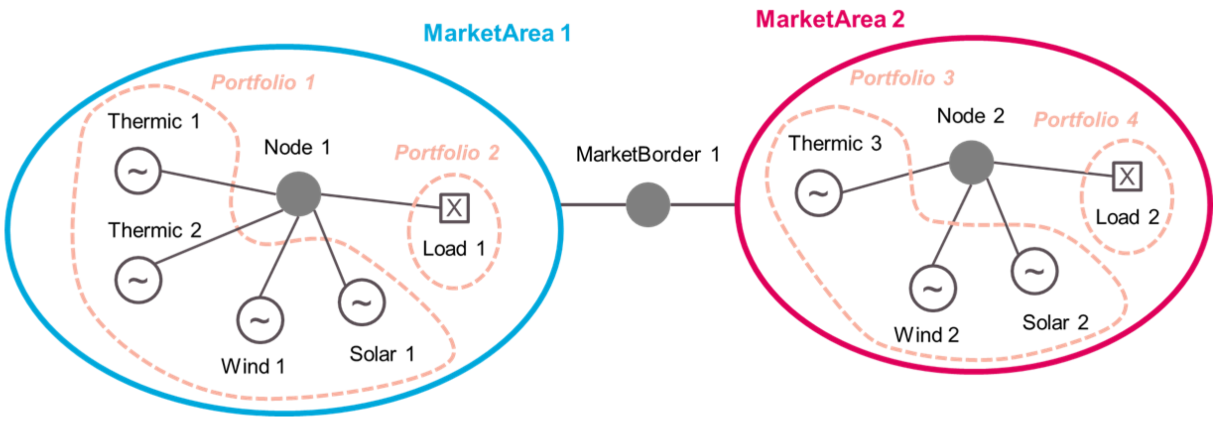


Figure 8: Example of a typical ATLAS environment

4.1 Overview of ATLAS methodology

4.1.1 ATLAS process input and data models

The main inputs to the process are:

- Antares studies, which contain wind and solar generation data, load data and a description of the thermal power plants and flexibility assets, with the costs and constraints associated with each of these.
- Additional csv files, to express technical and economic constraints not accounted for in Antares: operating ramps, starting and stopping ramps and minimal duration at a given value;
- Load, wind and solar production forecasts by timeframe, given by the forecasting error model described in deliverable D2.1.

Those inputs are converted to fit the ATLAS data format, which is consistent with that of ATLAS module outputs. It consists of five sets, each one containing one or several classes, as shown in Figure 9. Each class has a number of properties, which can take the form of values (binary or real), time-series or matrices.

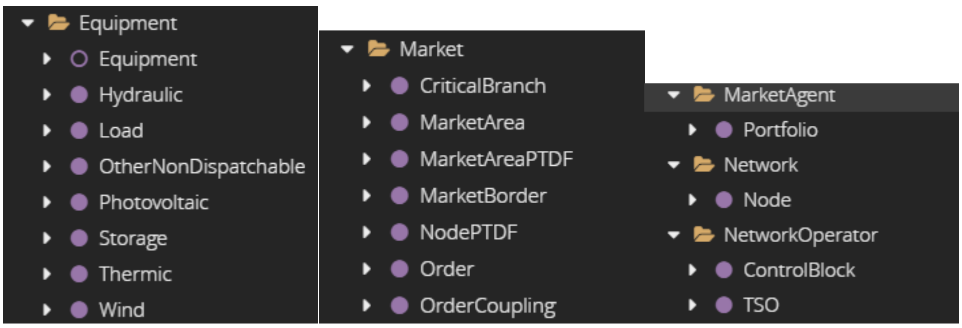


Figure 9: Sets and classes of ATLAS data model

All asset-related data is grouped together in the “equipment” set, which holds all parameters and time series: technical constraints, availability, costs, dispatch schedules, etc. Among these is the “power” matrix, which is filled in gradually by ATLAS modules, and contains each equipment’s programs defined at different lead-times. For example, the solar power matrix (see figure below) is filled in at the beginning of the simulation with solar generation forecasts for each lead-time, given by the forecasting error model.

	be_pv	be_pv	be_pv	be_pv	be_pv
Timestamp	C1 01/07/2028 01:00:00	C2 01/07/2028 02:00:00	C3 01/07/2028 03:00:00	C4 01/07/2028 04:00:00	C5 01/07/2028 05:00:00
2028-07-01 01:00:00	0	0	0	0	0
2028-07-01 02:00:00	0	0	0	0	0
2028-07-01 03:00:00	0	0	0	0	0
2028-07-01 04:00:00	0.591049203547594	0.566422153399777	0.5500041199679	0.426868869228818	0
2028-07-01 05:00:00	168.957982047452	164.672875321732	155.068325764084	128.060660768645	143.641374495497
2028-07-01 06:00:00	623.991987628656	623.868852377917	606.391855789683	515.189680075603	516.560585867165
2028-07-01 07:00:00	1347.72352835597	1347.71531933925	1310.43817443217	1222.07631850181	1173.74983709508
2028-07-01 08:00:00	2089.34251650731	2089.3343074906	2031.49357571009	1895.2731523258	1899.91945578702
2028-07-01 09:00:00	2710.01806138273	2708.94268019294	2643.46756286661	2642.95039481351	2575.84168316071
2028-07-01 10:00:00	3081.74696533059	3081.73875631387	3081.76338336402	3080.49919478976	2929.54358640036
2028-07-01 11:00:00	3338.9600860911	3338.95187707438	3338.27052868696	3336.87499584525	3103.76354816273
2028-07-01 12:00:00	3445.37356977981	3445.37356977981	3445.38177879653	3444.65117630881	3134.19437312872
2028-07-01 13:00:00	3412.22556028085	3412.22556028085	3412.23376929757	3411.51958484328	2944.57429600725
2028-07-01 14:00:00	3198.52843713153	3198.52843713153	3198.52843713153	3197.87171579426	2703.90234392935
2028-07-01 15:00:00	2771.18344493319	2771.18344493319	2771.18344493319	2771.18344493319	2296.47242628388
2028-07-01 16:00:00	2066.92369185608	2066.92369185608	2066.92369185608	2066.92369185608	1816.79495252143
2028-07-01 17:00:00	1211.42101480452	1211.42101480452	1211.42101480452	1211.42101480452	1063.98707458626
2028-07-01 18:00:00	485.284232179438	485.284232179438	485.284232179438	485.284232179438	445.503337173999
2028-07-01 19:00:00	71.229638044201	71.229638044201	71.229638044201	71.229638044201	68.1348387422921
2028-07-01 20:00:00	0	0	0	0	0
2028-07-01 21:00:00	0	0	0	0	0
2028-07-01 22:00:00	0	0	0	0	0
2028-07-01 23:00:00	0	0	0	0	0
2028-07-02 00:00:00	0	0	0	0	0
2028-07-02 01:00:00	0	0	0	0	0
2028-07-02 02:00:00	0	0	0	0	0
2028-07-02 03:00:00	0	0	0	0	0
2028-07-02 04:00:00	0	0	0	0	0.689557404138859

Figure 10: Power matrix for solar infeed in Belgium

Several dates are included in this matrix, as expressed in Figure 11:

- StartDate and EndDate define the forecasted period (often referred as T), StartDate and Enddate are different lines in the matrix;
- ExecutionDate is prior to this period, it is the date at which the forecast is made. If the ExecutionDate changes, the vision of the Forecasted Period will change. Execution dates correspond to columns in the matrix.

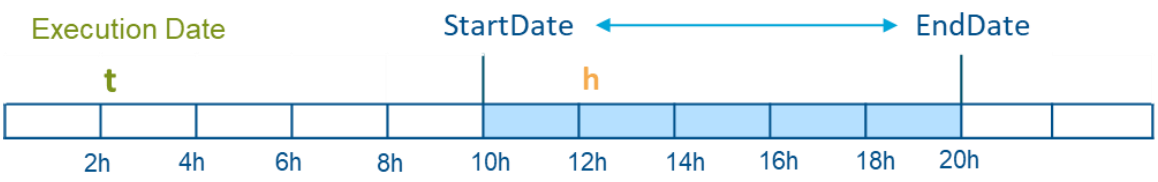
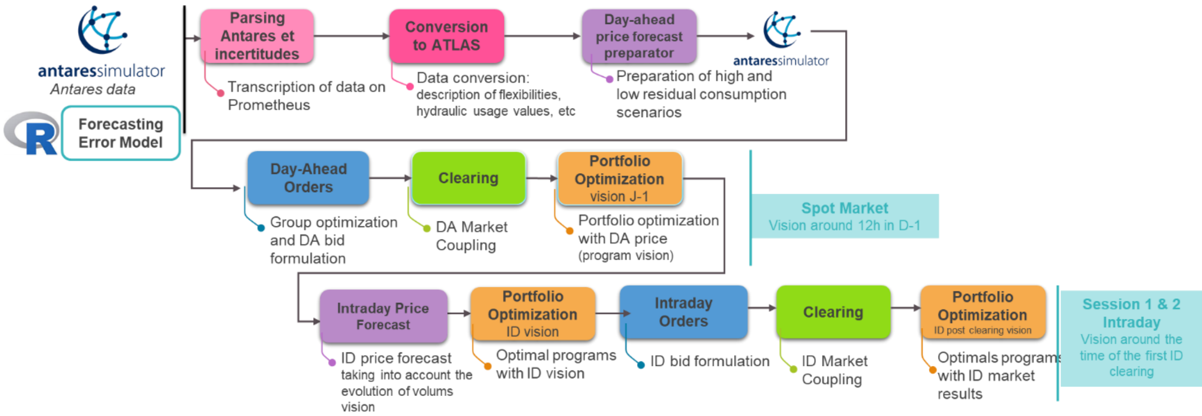


Figure 11: Dates definitions in ATLAS

4.1.2 Steps of the ATLAS modelling process

The ATLAS modelling process consists of a series of steps, involving different modules. We



will describe the global process before discussing the main modules (indicated by *) in more detail:

Figure 12: ATLAS Process

The ATLAS process is composed of 3 levels:

- The first converts Antares data to the ATLAS data format and generates the price forecasts used by market agents;
- The second simulates the stages of the day-ahead market (spot market), at the end of which a first generation plan of the power plants is derived;
- The third simulates the stages of the intraday market, which allows adjustments to be made to the generation plan set up on D-1 as the forecast uncertainties decrease.

At the first level, the main steps are:

- “Parsing” of data from Antares and from load, wind and solar forecasts;
- “Conversion to ATLAS”: this step converts Antares thermal generation and flexibility asset parameters into the ATLAS format, additional parameters such as minimal duration of power steps can be included at this stage. It also interpolates times series if the chosen time step is smaller than one hour (more details in section 4.4) and calculates Bellman value for hydro power (more details in section 4.5.4).
- “Day-ahead price forecast”: builds high, median and low net load (=load - RES generation) time series based on the uncertainty distributions per zone. These different scenarios are then used as inputs for a set of Antares simulations, which generates price forecasts consistent with the three net load time series. These price scenarios will serve as inputs for the bidding strategies of the market players in day ahead market.

The second level, the day ahead market, uses the following modules:

- "Day-ahead Orders"*: formulates the buy or sell orders submitted by the market players. These orders are submitted by equipment, in order to maximize the profit for each equipment assuming the agent is a price taker (i.e. it has a limited impact on prices).

- "Clearing"*: accepts and rejects the resulting buy- and sell-orders to match supply and demand while maximising social welfare and minimising exchanges between the different market areas. Builds the market clearing prices.
- “Portfolio optimization”*: derives the asset behaviours that maximise their agent’s profit, while considering 1) the orders accepted by the “clearing” module, 2) all technical constraints, 3) the expected price of imbalance penalties. This optimisation can lead to different results to those obtained by “day-ahead-orders”, since all sell-orders will likely not have been accepted, and the unit-commitment implied may not satisfy all technical constraints (the format of the sell-orders does not allow all the technical constraints to be taken into account).

Finally, for each session of intraday market, the main steps are:

- “Intraday Price Forecast”: updates the price forecasts using the updated load and variable renewable generation forecasts.
- ”Portfolio optimization”*: this idea is the same as the one used in the day-ahead level, but it is based on updated forecasts and the problem to be solved is slightly different.
    - Commitments have been made, which changes the optimization to be performed,
    - Real time is approaching and we are more concerned about imbalances, so we include them in the optimization objective.
- ”Intraday orders”* module: based on the optimal programs defined by portfolio optimization and the commitments, this module will create orders and requests on the intraday market;
- “Clearing”: determines the accepted buy- and sell-orders.
- ”Portfolio Optimization”: used in the same way as at the end of the day-ahead level, it determines final asset behaviours.

Additional intraday sessions can be run with the same modules using updated forecasts.

4.1.3 “Day-Ahead Orders”, “Portfolio optimization” and “Intraday Orders”

The “Day-Ahead Orders” and “Intraday Orders” modules simulate market participants’ order submission process respectively in the Day-Ahead market and intraday market. The order formulation strategy is based on an agent-focussed profit maximisation. The outcome of the module is composed of a list of market orders and a list of order coupling links. A market order takes a standard form, described below:

Table 1: Order structure and associated decision variables

Price information		
p_o	Energy price	(€/MWh)
Volume information		
q_{max_o}	Maximum power	(MW)
q_{min_o}	Minimum power	(MW)
Temporal information		
t_{start_o}	Start date of order	
t_{end_o}	End date of order	
t_{exec_o}	Submission date of order	
Binary information		
σ_o	Order sign: +1 for purchase, -1 for sale	
Decision variables		
x_o	Power quantity accepted	(MW)
δ_o	Binary variable: 1 if accepted, 0 else	

Each order concerns one time step only. The following orders couplings can be created between orders:

- “Same volume”: the linked orders must be activated for the same energy volume. This type of linking can for example be used for the start-up orders of thermal units and represents the binary aspect of a start-up decision and the related temporal coupling. The orders representing start-ups must all be activated at the same time for a given volume (the minimum stable power output).
- “Parent-child”: the child order can only be activated if the parent order is also activated. This link is used in particular for the flexible order of thermal power plants, which can only be accepted if the orders representing the start-up are also accepted.
- “Exclusivity”: two exclusive order cannot be accepted at the same time. It can for example be used to exclude orders made at the same hour but for different price scenarios.

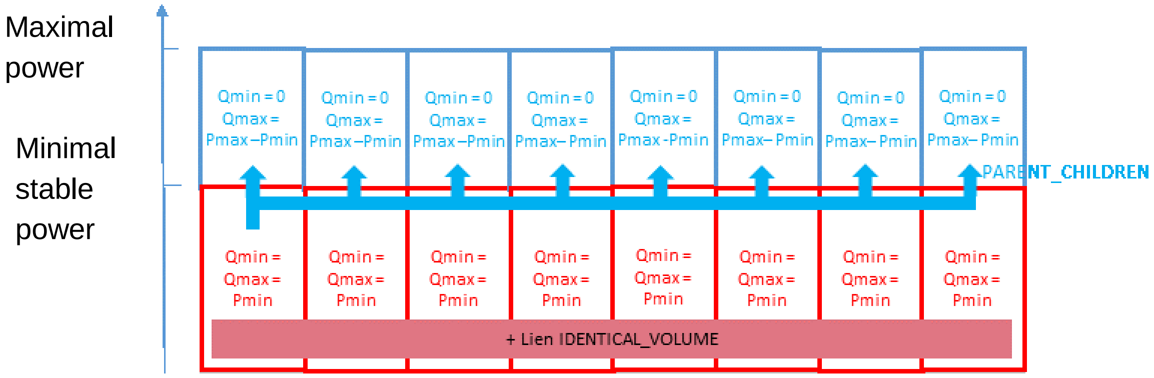


Figure 13: Orders and links between orders for thermal group in a simple case (no reserve, no start-up/shut-down ramp, only one price scenario)

"Day-Ahead Orders" and "Intraday Orders" modules are based on the same principles that will be first described, before explaining the main differences.

The order strategy is based on two steps: optimization and formulation of orders. "Day-ahead orders" has two sub-modules: optimization and formulation of orders, whereas those two functions are in two different modules for intraday ("portfolio optimization" for optimization and "intraday orders" for orders formulation).

- **In both cases, optimization is a profit maximization given a forecasted price scenario**, with technologic specificities, distinguishing with: load, thermic power stations, hydraulic power stations, storage assets, and non-dispatchable sources of energy. Each technology is represented with detailed representations. For example, thermal optimization considers the classical constraints such as maximal power, minimal stable power, minimal time on/off but also ramp, minimal duration at a given value and start/stop ramps. In the case of day-ahead orders, several optimizations are performed, corresponding to several price scenarios. The optimization for intraday process is made by portfolio optimization module.

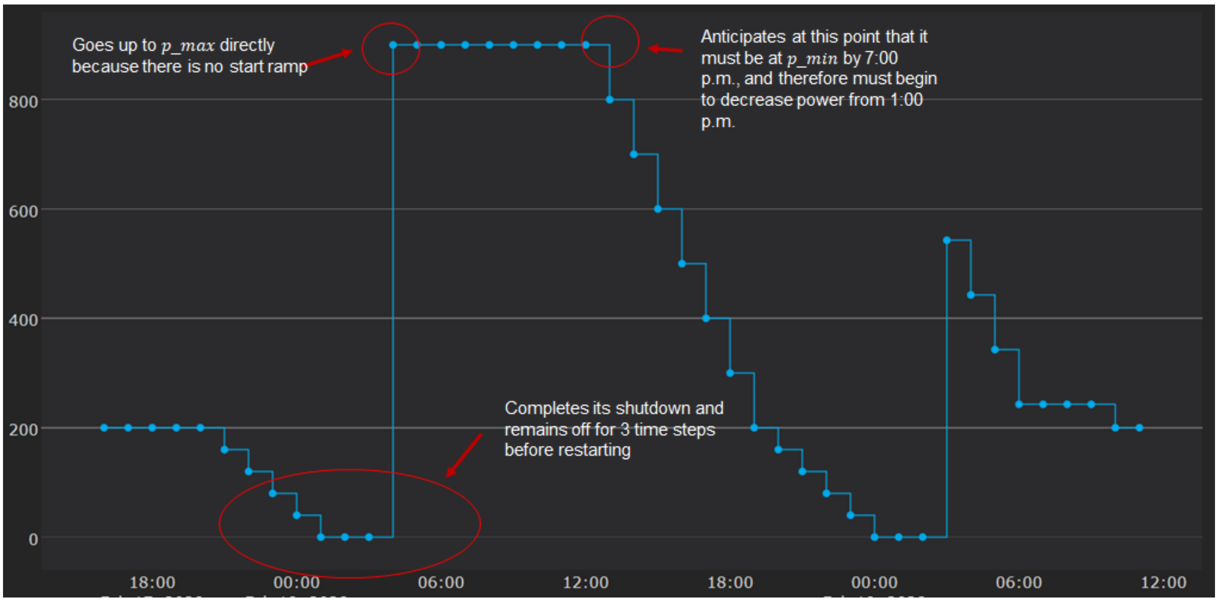


Figure 14: Example of optimization result in Prometheus

Formulation of orders: strategy of formulation depends on assets. It is described for flexibilities in section 4.5. For thermal units, the formulation is based on:

- Start-up orders (resp. shutdown orders), modelled by identical volume bids of quantities equal to the minimum stable power of the plant and costs equal to the start-up costs (resp. the benefit due to avoided start-up and the additional fuel cost).
- Start-up ramping orders (resp. shutdown ramping orders), modelled by parent-child orders and quantities adapted to represent the start-up ramp.
- Flexible orders, at marginal cost and a volume equal to Maximal power – Minimal power. Those orders have a parent-child link with start-up orders, i.e., start-up orders must be accepted to accept flexible orders.

Day-ahead and intraday orders process calculation are mainly differentiated by:

- The formulation of orders: in intraday, orders are formulated considering the difference between optimal program and previous engagements, whereas in day-ahead, there is no previous engagements so the formulation is simplified.
- The perimeter of optimization (unit-based vs portfolio-based): on the one hand, Day-Ahead optimization is made with a unit-based vision as portfolio-based optimization has little added value and is a lot more complex. On the other hand, intraday optimization is made with a whole portfolio vision because intraday is used to balance perimeters of balancing responsible party, which needs a portfolio vision.

4.1.4 Market Clearing

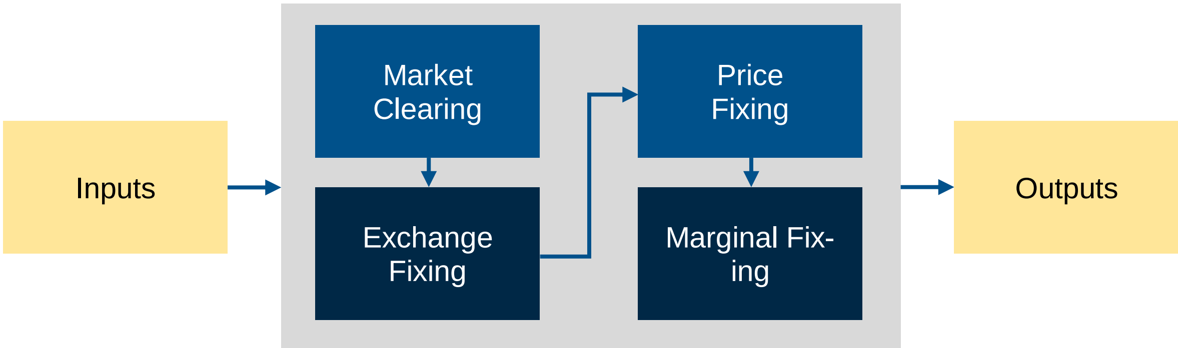


Figure 15: Structure of the market clearing module

The module performs a market coupling, taking as an input the market orders from each market area, as well as the market border constraints, whether this is the available transfer capacity (ATC) method or flow-based method. From these inputs, the algorithm matches buy and sell orders to maximize the overall social welfare. It is similar[6] to Euphemia, the algorithm that solves the problem associated with the coupling of the day-ahead power markets in Europe. It

[6] The four sub-modules have the same objective, but Euphemia can take into account more types of orders and link between orders. Euphemia also has a complex resolution process, using heuristics to converge more rapidly, whereas the market clearing module try to solve the optimization problems without heuristics.

furthermore calculates the market prices and other values and market indicators, including the cross-border power flows.

More precisely, the market clearing module is broken down into four main sub-modules (see Figure 15):

1. The clearing phase. This sub-module chooses a first version of the accepted and refused orders, by maximizing the social welfare on all the time steps, taking into account the various links between orders and the exchanges limits.
2. The exchange fixing phase. This sub-module calculates importation/exportation compatible with clearing phase results,
3. The Price Fixing phase. This sub-module selects prices for each zones and each time step. Prices must respect several properties: equality of prices between zones in the absence of constraints, absence of paradoxically accepted orders, and minimisation of price in case of equivalent solutions.
4. The marginal fixing phase. This sub-module marginally improves the solution calculated in the clearing phase by accepting the maximal quantity possible without modifying the social welfare

In this module the agents take no action.

Market clearing outputs are:

- For each order, quantities accepted;
- For each area and each time step: price (DA-price or intraday price depending on the execution date), and balance.

4.2 Downscaling forecast data at nodal level

The forecast process described in D 2.1 does not allow to consider geographical correlation between error forecasts. It is not a big issue for zonal studies, as this correlation is weak. However, for nodal studies, this correlation is much more important and it has been decided to use available historical data. Actually, the nodal description has been limited to 62 adjacent 400-225 kV substations in France, the rest of France (26 areas) and neighbouring countries are modelled with a zonal representation.

The data concerned are load, wind and solar generation forecasts for the different zones/substations of the study, with two forecasts in Day-ahead (11am and 7pm) and forecasts every three hours in intraday.

4.2.1 Extraction of the historical data

The data come from the "Daily Network Situations" calculated at RTE. Data concerning a week of November 2019 was extracted. It includes all the consumption and production forecast data, substation by substation for the two considered areas

These network files are based on the RTE transmission network, i.e., all RTE voltage levels are represented. As the study is carried out on the 225 and 400 kV levels, it was necessary to scale up each file. This scaling up allows to obtain equivalent productions and consumptions on the desired voltage levels: the total volumes of consumption and production are thus kept during the scaling up. These data indicate productions at certain substations, this does not mean that the productions are actually connected to these substations but that the evacuation of productions from lower levels to higher levels is equivalent to a set of fictitious groups connected to these substations.

For foreign countries, the forecast data are taken from the Transparency platform. As for France, load factors are calculated and then multiplied by the forecasted power in 2035 for the different stations. For consumption, a coefficient is taken into account to reflect changes in annual consumption volumes.

Possible improvements in the forecast of renewables generation by 2035 are not taken into account. In other words, for both the forecast and realized chronicles, historical load factors are used. This can be considered as conservative.

4.2.2 Changes to 2035

For wind and solar, the local load factors are calculated, and using scenario of installed capacity per substation by 2035, it is possible to have the generation forecast and realisation chronicles for these two sectors.

For substations with no connected RES production in historical data, but where development is anticipated between now and 2035, load factors were interpolated in order to obtain production forecast and actual data. The interpolation was performed using the 3 closest substations with load factor data for the technology considered. The interpolated load factor is calculated by averaging the load factors weighted by the distance of the substations. This method has the advantage of retaining the overflow effects induced by the high voltage level.

For consumption, a corrective coefficient is used to consider changes in consumption volume between 2019 and 2035. However, the profile remains identical.

4.3 Network description

4.3.1 Network description in for the zonal study

In ATLAS zonal study, the grid is simply modelled by NTC given as input to the “Clearing” module. The nodal study has a more complex representation, with a multi-level network modelling.

4.3.2 A multi-level network modelling for the nodal study

Western Europe is represented with: Austria, Belgium, Czech Republic, Denmark, France, Germany, Ireland, Italy, Netherlands, Poland, Portugal, Spain, Switzerland, United Kingdom. France is parcelled into 26 areas and other countries into several areas (see Figure 16). This modelling is taken from RTE internal studies where areas have been designed so that the future constraints are mainly located at the interface between the zones. Zones are connected by links, whose capacities and impedances have been calculated to be representatives of the connection between zones.

Two of these 26 areas (named 08_fr and 14_fr), are again divided into 34 nodes, which coincide with existing extra-high voltage stations of the French network. The 34 nodes thus created will be the study area for the nodal market sequence. These two areas have been selected for two reasons:

- First, they are positioned at the centre of France, thus eliminating the risk of taking into account any edge effects during the study.
- Second, these areas have thermal and nuclear assets connected to their inner network, as well as renewable production capacities.
- 26 stations out of 62 have a wind history: this will allow us to use the history to obtain production forecasts at different times.

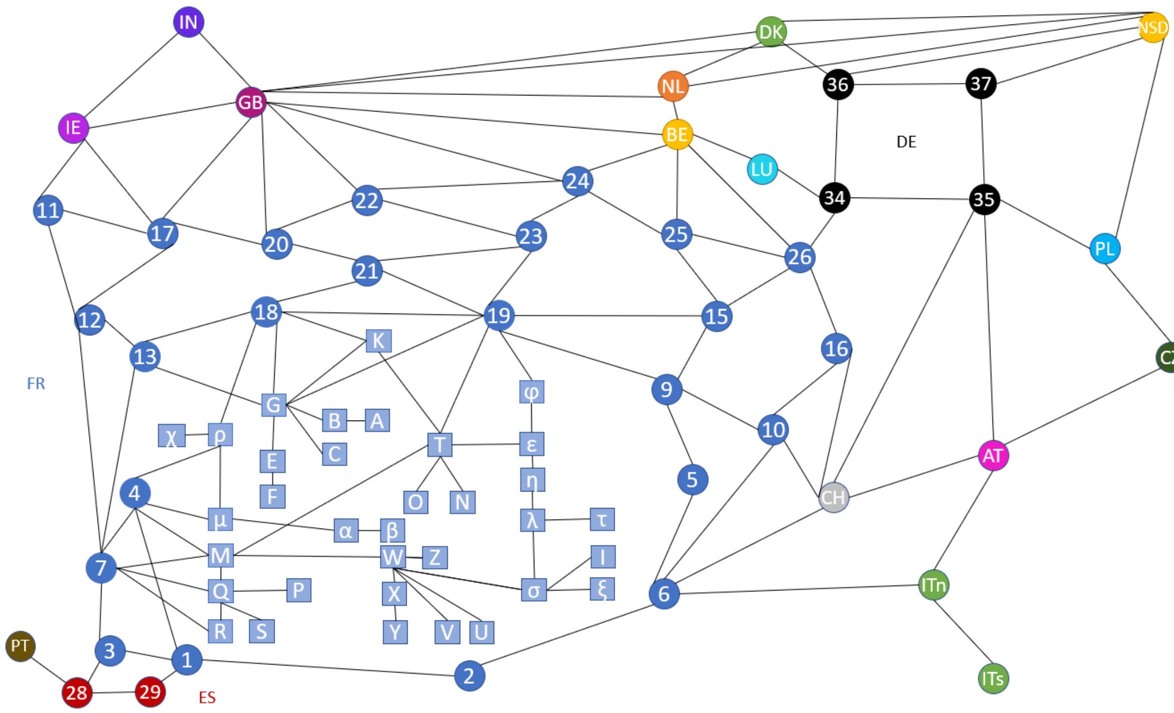


Figure 16: Visualization of the equivalent network in Antares and ATLAS integrating the nodal study area

Concerning date, the equivalent network used for the nodal market study is acquired by merging the information for 2030 from the TYNDP, for the European network, and a more detailed description of the French network from RTE database.
The network modelling approaches are equivalent yet different within Antares and ATLAS. In the former, the network is modelled using equivalent impedance and loop flows, while in the latter, PTDFs are used.

4.3.2.1 Network description in Antares for nodal study

For Antares description, an equivalent network has been calculated starting from a detailed description of the 2030 European network. On a lot of generation/load situations, impedances and loop-flow between zones have been chosen to minimize the quadratic difference on inter-zonal flows between the estimated model and the results of a complete DC load flow[7]. Phase

[7] For more details, see *Zonal Reduction of Large Power Systems: Assessment of an Optimal Grid Model, Accounting for Loop Flows*. IEEE Transactions on Power Systems, Vol.30 Issue 1, 2015 (M. Doquet)

shifting transformers and HVDC are also taken into account by calculating their influence on each interzonal flows in each situation. Interzonal Capacities were calculated on all the situations, taking into account the capacities and flows in N and N-1 on each line influenced by interzonal exchange.

Then, the equivalent network has been integrated into Antares:

- The substations or areas are modelled as nodes in the network. They are considered as market areas where there can be consumption and different types of generation.
- These different substations are connected by lines representing the exchange capacities between these different nodes. In Antares, HVDC lines interconnecting different areas, such as France and Spain for instance, are added in parallel to AC lines.
- In order to represent the effects of different impedances on the exchange capacities and the different behaviours of interconnecting lines – AC or HVDC, constraints (Kirchhoff's current law) are introduced in the form of binding constraints.

4.3.2.2 Network description in ATLAS:

Into ATLAS, the different nodes, zones or countries each represent a market area without internal congestion. These are all interconnected using the equivalent network presented earlier. It is therefore necessary to choose a market coupling method between these zones that will calculate and allocate interconnection capacity.

For this purpose, the Flow-Based coupling method was chosen, as it allows the association of commercial exchanges between market areas and the physical flows caused by such an exchange. This method is now used for the D-1 spot market in the CWE zone, which includes France, Germany, Austria, Belgium and the Netherlands, since 2015.

This method relies on the use of:

Market zones: they are the same as the zones described in Antares, with a description of several zones per country in Western Europe and a part of France with zones representing 400 KV nodes.

Critical branches, i.e. the link that can limit the exchanges. We use all the lines of the equivalent network, distinguishing two critical branches for each link to take into account the direction[8]. For each of these branches, it is then necessary to determine all of the variables that are useful in the calculation of Flow-Based exchange capacity:

- The maximum capacities: in a Flow-Based coupling method, it is difficult to incorporate HVDC lines in the critical branches' description. Thus, new capacities have been experimentally computed from the ones used in the Antares study to take into account the binding constraints added in Antares to simulate the behaviour of HVDC lines and their impact on the other lines.
- The reference flow Fref: are the loop flows used in Antares study.

[8] Indeed, a critical branch generates two different constraints, depending on the direction of the flow: flow(A->B) < link capacity and flow(B->A) < link capacity.

- The FRM margin: this is not necessarily filled in because it only quantifies the uncertainties to be taken into account, particularly in terms of forecast errors. It could be filled in following the first results of this study, which would allow us to determine what safety margin to use.

Power Transfer Distribution Factors (PTDF), i.e., coefficients for distributing an exchange over the lines of a network. They are deduced from the equivalent network described above. In order to avoid biases related to the location of the slack node, it has been assumed to be distributed on the different nodes of the network.

4.4 Time downscaling: Increased Time Granularity in ATLAS

Time granularity plays an import role in market design for exchanging electricity products, especially in the presence of a flexibility-driven energy mix. Therefore, the ATLAS model features an interpolation module that converts hourly time series into sub-hourly ones (e.g., on 15 min or 5 min time steps), designed for price, load, wind generation, and solar generation forecasts. The interpolation method varies slightly from one time series to another due to the particular form, characteristics, and trends of each time series. However, the main principle remains the same. The methodology is designed for country or zonal data up to a minute time scale since they show a rather smooth profile.

The interpolation method has been implemented using mathematical optimization instead of pre-made algorithms to enable easy access to its kernel and allow for potential customization. The goal of this optimization-based algorithm is to find the “best” function to interpolate between the given measured points – “best” meaning different things depending on the time series’ nature, as detailed later in this section.

To perform an interpolation over time series that can potentially be several years long, the latter are divided into overlapping segments over which the interpolation is locally and sequentially performed. A delimiting window, called the optimization window, scrolls through all of these segments to perform the subsequent interpolation tasks, as illustrated in the figure below where orange points represent initial hourly data:

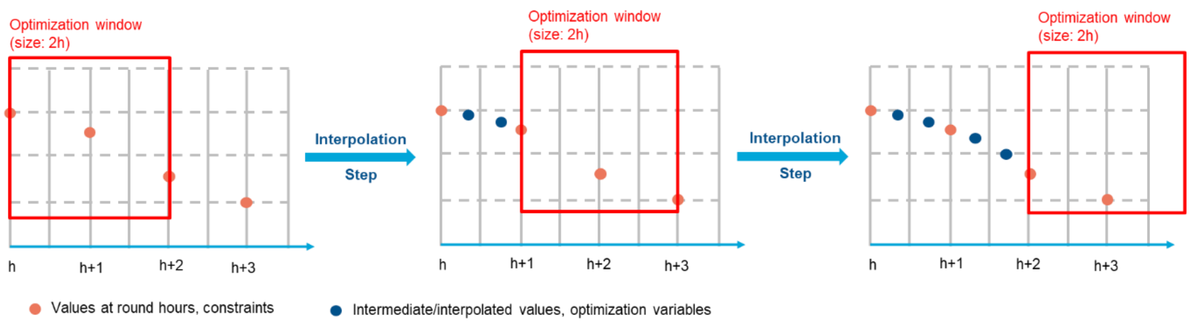


Figure 17: Interpolation with an optimization window

The length of this optimization window is constant over the whole time frame but it can be customized. In practice, it usually lasts two to three hours.

The computation performed at each step within the optimization window always consists in solving an optimization problem, but depending on the nature of the data to be interpolated, the objective function will differ, and some additional post-treatment may be applied.

The differences in objective functions are due to the form and signification of the initial data, and to the fact that some system variables fluctuate more than others. At a country level, price forecasts and wind generation can vary rapidly and significantly whereas solar generation forecast curves are smoother and bell-shaped. The latter therefore has a different objective function from the others, as detailed in the appendix 0.

The main post-treatment addresses a specificity encountered with two types of time series, namely price and load forecasts, which is the relative importance of their extreme values. Extreme values for prices and load have technical and economical meanings that will strongly impact the results of any quantitative study (negative prices, need for flexibility, etc.). Those data are therefore more sensitive to a change in extreme values compared to others, and must therefore be interpolated carefully. For both price and load forecasts, a post-treatment is applied after the optimization problem is solved to restrain the excursion of the interpolated values when necessary. This post-treatment is applied differently for both types of time series, as explained later in this section, but the goal is the same: avoiding deforming too much the extreme values whenever relevant. For example, on the figure below where blue points represent the interpolated values for a particular load forecast computed from the initial orange points, the two blue points between h+2 and h+3 may be slightly scaled down to avoid outstripping too much the peak value measured at h+3. The magnitude of this scaling is customizable, and can even be nulled to avoid the post-treatment.

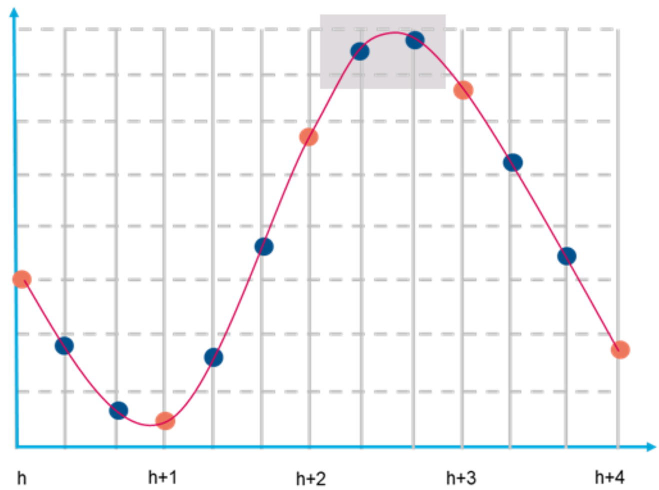


Figure 18: Controlling the range of interpolated values

Another post-treatment is applied on interpolated forecasts of wind generation, to ensure the interpolated values are all non-negative.

Below are given example of results for wind and solar time downscaling, further details on the methodology for each type of data is given in Appendix 1.

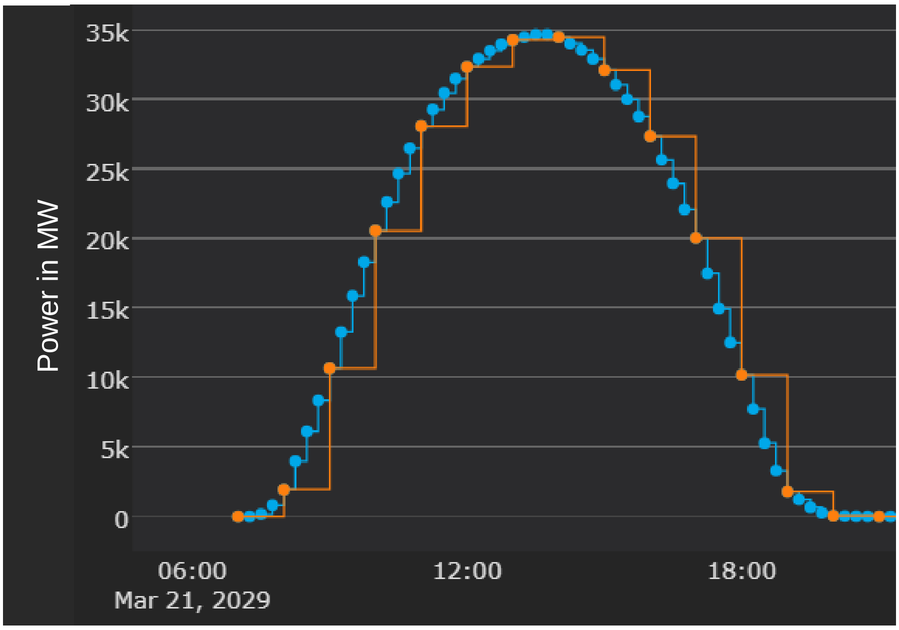


Figure 19: Example of interpolation of solar generation (T = 3, n = 4)

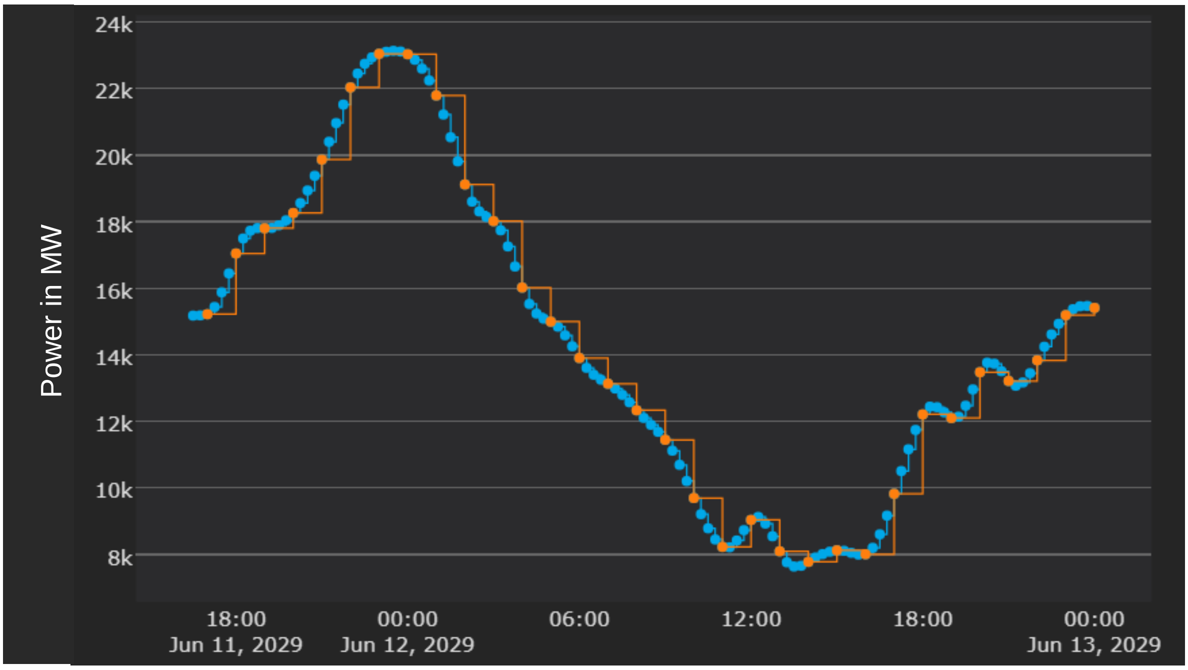


Figure 20: Example of interpolation of wind generation (T = 2, n = 4)

4.5 Flexibility solution modelling in Antares and Prometheus/ATLAS

Technologies modelled as storage technologies in Prometheus-ATLAS studies are: Stationary Batteries, Electric Vehicles, Hydro Storages and Power-to-gas.

The studies conducted use three models for each of the flexibilities:

- One for the Antares study serving as input to the ATLAS process: for the zonal study, it is the WP1 study to 2030; for the nodal study, a new Antares study has been created but it uses the same flexibilities modelling as the WP1 study;
- One for program optimizations given a price forecast: this type of optimization is used before the markets to feed the order strategy (both in "Day Ahead Orders" and in intra-day in "portfolio optimization"), but also after the markets for the programming ("portfolio optimization");
- One for the bidding strategy based on the schedule: this involves translating the optimal flexibility schedule into buy- and sell-orders on the market. More precisely, order formulation strategy consists in specifying the minimum and maximum hourly quantities proposed, their price and any dependencies between these blocks (identical volumes, parent child or exclusivity).

4.5.1 Electric vehicles

ATLAS allows several modelling for EVs. We will describe the modelling chosen to fit with WP1.2 Antares modelling (see deliverable 1.3), whose main characteristics are:

- No possibility to resell energy, EVs can only displace their charging;
- A daily energy to be consumed

4.5.1.1 Antares modelling for EV

As previously explained, Antares modelling of most of flexibilities is the same as in WP1.2: it is achieved via virtual nodes, by enforcing constraints on flows on the link between the virtual node and the real area to which it is connected. This virtual node's modelling strategy has been used repeatedly by many Antares users for several years: its versatility allows the consideration of new power system assets without having to resort to cumbersome and time-consuming software development.

In Antares, one 00_EV_STO virtual node is created for the whole nodal study area, and is then linked to each of the 34 nodes.

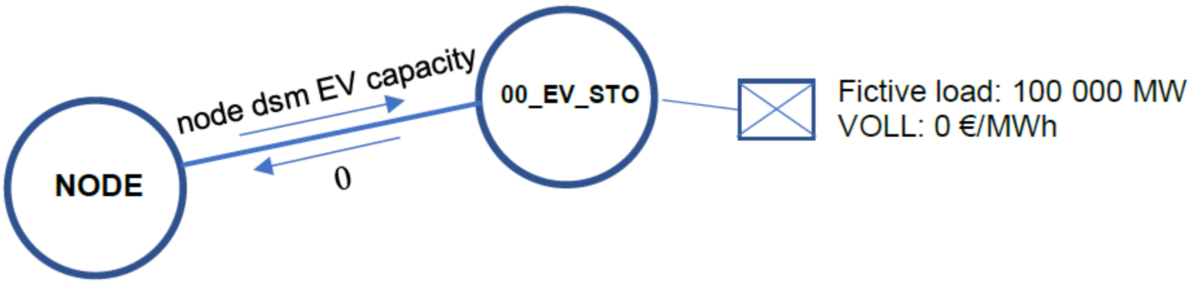


Figure 21: Model representation for electric vehicles through a virtual node

The configuration specific to each node will be set thanks to the following parameters:

- The 'node dsm EV capacity', in MW: set at the maximal charging power observed during the day. To take into account the fact that less electric vehicles may be connected to a charging station from 9am to 5pm, only a certain percentage of this capacity may be enabled during these hours.
- A transit capacity from 00_EV_STO to the node that can be non-zero if a modelling of the vehicle-to-grid needs to be added
- The daily binding constraint that ensures the pilotable electric vehicle load is actually consumed writes:

$$\sum_{day} flow(NODE \rightarrow 00\_EV\_STOR) = p \cdot \sum_{day} flexible\ electric\ vehicle\ load$$

where p equals the percentage of flexible electric vehicles in the fleet.

The rest of the electric vehicle load is directly added to the load of the node concerned.

4.5.1.2 Program optimizations given a price forecast

In ATLAS, the modelling of this flexibility is inspired by the latter but needs to be adapted to the ATLAS format. Instead of adding a virtual node to the network on ATLAS, the fleet of flexible electric vehicle is modelled by a storage equipment for each node. As the virtual load no longer exists, the flexible part of the electric vehicle load is added as well to the node's load.

This time, the configuration specific to each fleet will be set through the storage equipment's following properties:

- The maximum power, in MW: it is set similarly to the 'node dsm EV capacity' in Antares.
- The maximum energy, in MWh: it represents the maximum value of energy which can be stored by the equipment during the optimization period, here a day. It thus needs to be at least equal to the pilotable electric vehicle load which needs to be met daily.
- The displacement energy, in MWh: it represents the quantity of energy to be optimized during the given period. In the case of an electric vehicle, it gives the amount of energy consumed by the EV since its last reload.

The optimization of the demand placement takes into account the price forecasts and these constraints, over a configurable period of time. We choose to optimize over two days (the day we are working on and the next day): anticipating the next day allows us to have a more optimal behaviour.

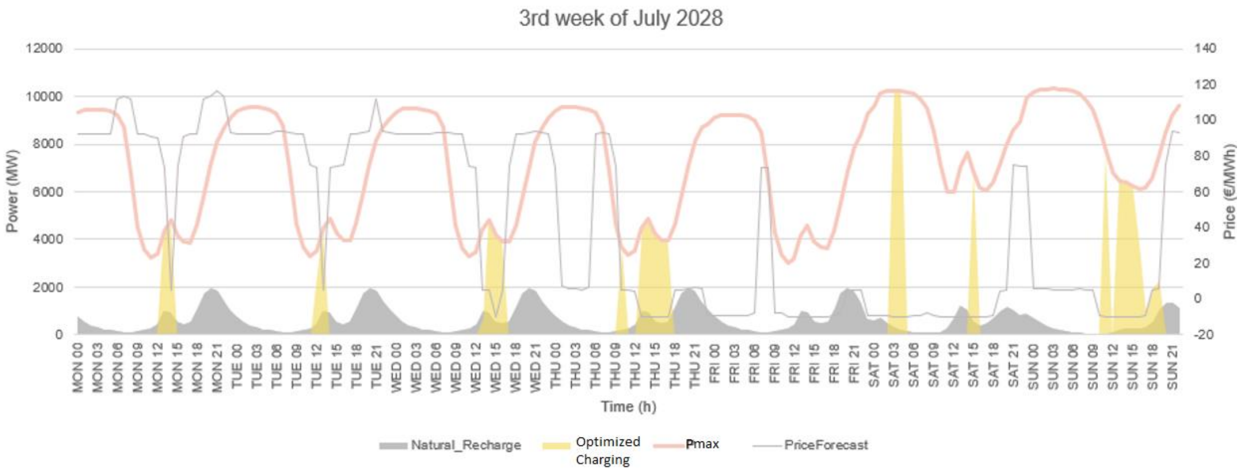


Figure 22: Optimization of EV charging compared with charging as soon as connected ("Natural Recharge")

In order to avoid unrealistic power demand peaks, we decided to divide the maximum power into several fragments, and add a slight overhead on the recharge/injection for each fragment. Indeed, in case of equivalent solutions, i.e., identical prices on several time steps, we will have a smoother curve, which seems more realistic.

4.5.1.3 Order formulation strategy

The buy-order strategy in this case is quite simple, because there is no possible sale for electric EVs although the model offers the possibility:

- The demands are perfectly flexible demands,
- The quantities are those obtained by the optimization previously described
- The price depends on the market:
  - For Day-ahead market, it is the maximum price anticipated on the hours on which the demand has been placed. Acceptance of those orders is not certain but the missing energy can be bought intraday market.
  - For intraday market, the order price is the maximum market price to avoid imbalances.

4.5.2 Industrial DSM and Power-to-gas

- In the WP1 Antares study used as input, DSM and Power-to-Gas are simply represented as a load that can be cancelled at a given price. In practice, this cancellation is modelled through a virtual thermal unit[9].
- As there is no complex technical constraint to consider, bids strategy do not need a previous optimization. It consists in taking into account the maximal load for industrial and power-to-gas into account in load and offering the load cancellation at the price defined.

4.5.3 Stationary battery and pumped hydro

Stationary battery and pumping hydraulic have similar modelling with difference for some parameters: efficiency and energy constraints.

In this section, we will not give the modelling details of all storage technologies because this detail is already given by WP1 (see Deliverable 1.3). The modelling uses the same type of tricks that the one for EVs but with an injection.

The optimization problem used is a very classical formulation of storage revenue maximization given the constraints:

- Minimum and maximum energy level in the storage,
- Maximum injection and withdrawal power.

Taking into account the yield and the past evolution of the energy level in the storage (to have a coherence between the sequentially simulated days).
As for the electric vehicle, the duration of the optimization is configurable. We choose a duration of two days for the batteries (working day plus one additional day) and one week for the hydraulic pumping (the latter often having a weekly behaviour).

[9] Indeed, the cancellation of a load at a given price is the same as the use of a thermal group à this price.

The strategy is based on previous optimization. The buy and sell quantities are those obtained by the optimization previously described. They are considered as perfectly flexible (no links between orders, no minimal quantities).

There are two prices used for the day: one for sell and the other for buy. They are calculated from the DA price expectations respectively at the times when storage is selling and buying. The prices are modified to respect:

$$Purchase\ price = storage\ yield \cdot sale\ price$$

Indeed, this relationship ensures that the buy-sell is profitable, while maximizing the chances of being accepted. Assuming that the storage holder does not have no market power, it is a priori not detrimental to revenue, since buy- and sell-orders are remunerated at the clearing price (similar to the marginal cost sell-order for thermal units).

4.5.4 Hydro power

The fine representation of the behaviour of hydraulic structures is a complex subject which requires at the same time a detailed knowledge of the topologies of the valleys (contributions, travel time of water, head, solid transport), of the local constraints (reserved flows, tourist coasts, agricultural withdrawals) and of the equipment (capacity to take part in the system services, maintenance periods). As it is not possible to have an exhaustive knowledge of these phenomena, hydraulic modelling requires simplification choices.

In Antares and Prometheus/ATLAS, production is separated into three categories:

- run-of-river, as a non-dispatchable generation, is inflexible;
- pumped hydro stations which pump water back to an upper reservoir to be stored, as described previously;
- lake (all the reservoirs whose stock allows generation to be deferred), is described below.

4.5.4.1 Antares

Antares has a zonal management of the hydraulic stock:

- In the medium term: dividing the annual energy into daily energy, according to a net consumption criterion and taking into account reservoir level constraints. This criterion and these reservoir constraints have been calculated by WP1 in order to reflect a realistic behaviour;
- In the short term: placing this energy in an economically optimal way, but by constraining the average power produced during the day to be not too low compared to the peak production. This also makes it possible to obtain a more realistic zonal hydraulic production.

4.5.4.2 Optimization in ATLAS

ATLAS does not use the same modelling as Antares but is based on a classical strategy for water reservoirs.

The lakes are managed in a seasonal way, it happens that the actor chooses not to turbinate at all in summer, to keep his stock limited to winter, because energy is more expensive in winter. This long management horizon makes the use of a deterministic optimization method unsuitable. On the one hand, the optimization over several months with hourly time steps is quite expensive in terms of computation time. On the other hand, the large uncertainty on the prices prevents us from using a deterministic optimization problem.

Thus, stochastic dynamic programming has been used to simulate the management of these peak reserves. This method, based on the recursive application of Bellman's optimality principle, brings out water use values. It then becomes possible to assign this fictitious cost to the hydraulic stocks, and use it to make production decisions.

The use of peak reserves, such as lake hydraulics, always corresponds to a trade-off between its direct use (an immediate gain), or its conservation (and the preservation of an expectation of future gain).

Let's take the example of a hydroelectric dam manager. Faced with a higher-than-normal demand in November, he will not necessarily use his water immediately, to make money at the market price, if he thinks that this water will be more useful in the middle of winter, in January-February. The decision to use the stored energy is a trade-off between the immediate gain and the lost expectation of gain that has been lost by discarding this level of storage. The lost expectation of gain is the estimated value for a lake level at that time of the year, and for that same level reduced by the turbined volume. These are the Bellman values. The hydraulic trade-off can thus be summarized in Figure 23.

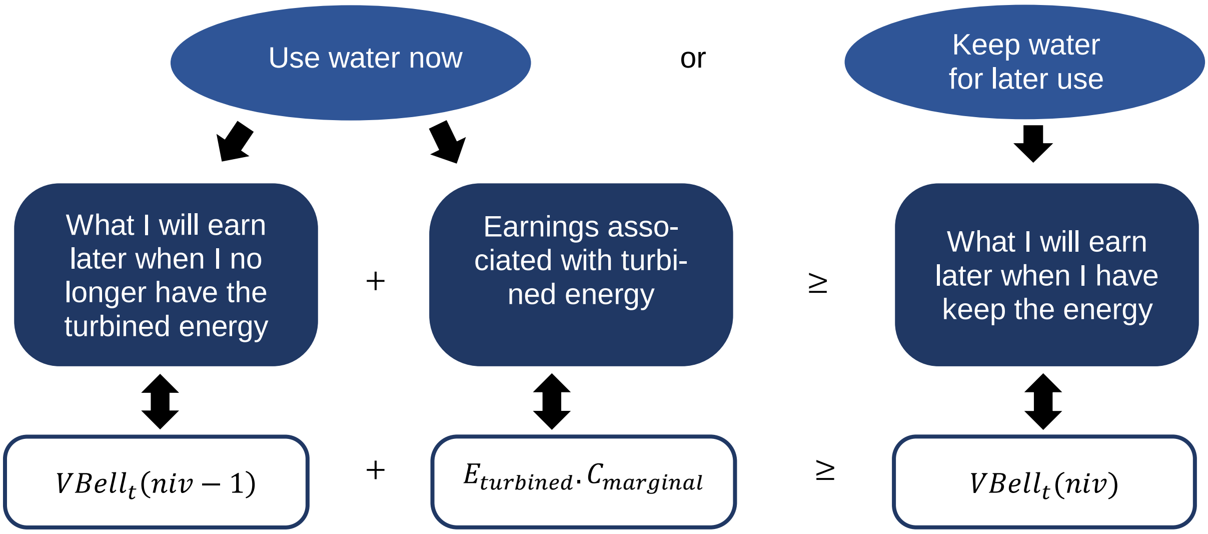


Figure 23: Water use strategy of a hydraulic producer using Bellman values

Stochastic dynamic programming is very complex and time consuming. In practice, those values are abacuses: values depend on hour of the year and reservoir level. In ATLAS model, it has been chosen to calculate water use value to an aggregate stock of all dispatchable units of the areas represented. We will call them “national water use value”, even if areas definition doesn't always correspond to countries.

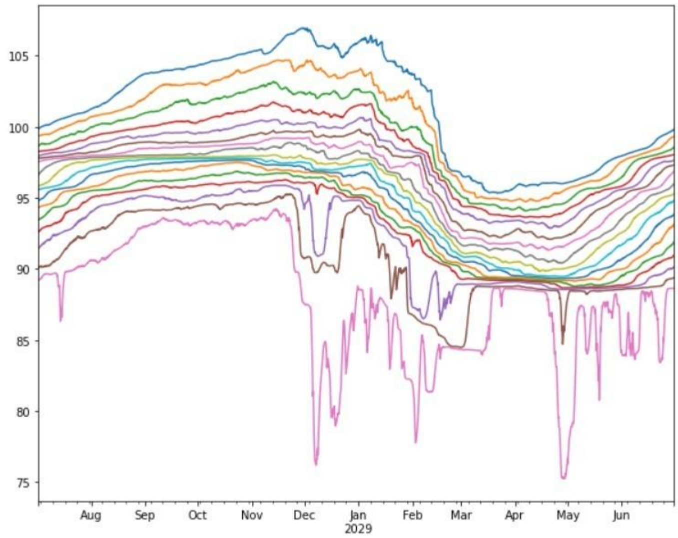


Figure 24 : Water use value for France. The different lines represent different reservoirs levels.

Hydraulic stock is in practice made up of numerous power stations that do not have the same characteristics, in particular, more or less storable energy compared to their power. All the stocks of the country do not have the same values of use. Thus, a power plant which will have on average inflows allowing to turbinate one hour per day will have higher usage values than a power plant allowing to turbinate five hours per day. In the first case, the usage value will reflect the average of the maximum daily prices, whereas in the second case, the usage value will reflect the average of the 5 hours of highest prices.

This price diversity was reflected by separating the hydraulic power into five to seven different fragments (depending on the zone) to which was associated a different price deviation from the “national water use value”. More explicitly:

- Hydraulic assets are represented by five to seven units, with different maximal power. The sum of the maximal power of those entities are the maximal dispatchable hydraulic power for the area.
- Five to seven water use values are be defined by:

$$Water\ use\ value\ for\ unit\ 1 = \text{“}national\ water\ use\ value\text{”} + delta1$$
$$Water\ use\ value\ for\ unit\ 7 = \text{“}national\ water\ use\ value\text{”} + delta7$$

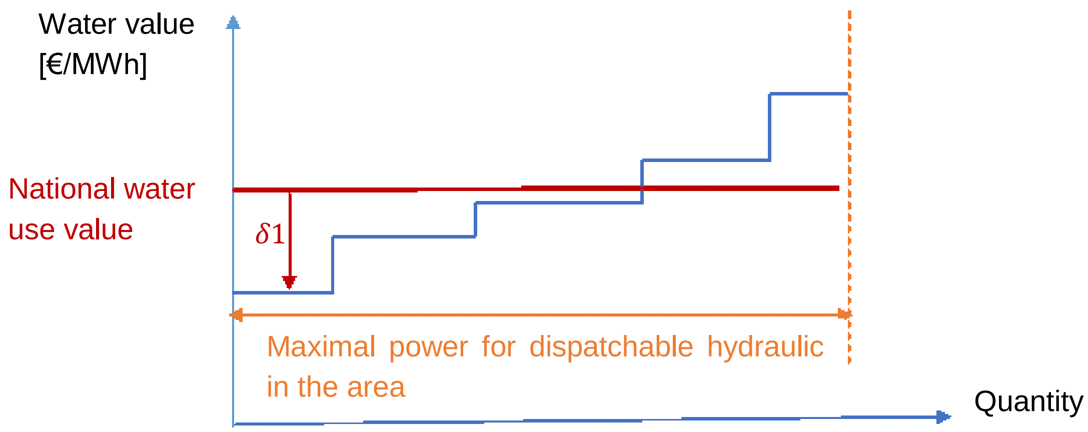


Figure 25: Illustration of the available hydraulic power for an area (x-axis) and different water use values (y-axis)

The division of maximal power into units and the $\delta 1 \ldots \delta 7$ have been calculated in relation to the orders of the French hydraulic power plants of the year 2019 on the French balancing, after having checked that they reflected the values of use on the French hydraulic. As a consequence, the heterogeneity of water value is supposed to be similar in all the modelled areas.

4.5.4.3 Sell-order strategy

The usual sell-order strategy is to submit orders at the water use value. Indeed, those orders allow to sell energy when the water use value is lower than the market price and to keep energy otherwise.

In ATLAS case, one order is made by time step and by units (5 to 7 units):

- those orders are flexible,
- with maximal quantities equal to maximal power of the unit,
- with price equal to the water use value for the unit.

So, for each time step, a supply curve similar to Figure 25 is submitted.

5 Modelling of Distribution Grid

Since the flexibility, needed by both the Transmission System Operator (TSO) and Distribution System Operators (DSOs) to cope with grid operation challenges, is mainly found at the distribution system level, modelling the distribution grid is fundamental for spatially downscaling the whole model for market mechanisms' simulation.

The methodology proposed by ENSIEL, briefly described in this report, aims at assessing how the use of flexibility by the TSO can impact the DSO activities and at quantifying the expected costs. For these purposes, local distribution market models where the Distribution Energy Resources (DERs) offer flexibility to the TSO and DSO have been hypothesized. The final goal is quantifying the flexibility that the TSO and market players can procure from the distribution system without harmful impact on the distribution network operation.

Two complementary and interconnected tasks are performed:

1. The first task builds the distribution network model in terms of lines, topology, conductors, demand and production by using open data only. The results are useful for the TSO and for the stakeholders that do not know the distribution grid in detail.
2. The second task aims at quantifying the availability of flexibility products and the relevant costs by using local market models that optimize the DERs dispatching.

5.1 Distribution network modelling using open data

The object of the model is the network downstream to a distribution transformer (HV/MV transformer into a primary substation - PS). This transformer represents the point of common coupling (PCC) where the MV distribution system (DSO managed) is connected to the HV transmission system and, thus, it represents the TSO/DSO interface.

Publicly available open data on the energy consumption and production of a region or wider area, opportunely processed, allow to obtain realistic load and generation profiles for each distribution network in the examined region/area. Furthermore, by combining these profiles with geo-spatial and socio-economic data, accurate synthetic models of the distribution grid can be built with all data necessary for power flow analysis (e.g., number of lines, type of conductors, loads, and generators). Thus, voltage profiles and power congestions can be obtained as well as the level of flexibility that DERs in a given area can offer to the TSO without harmful effects on the distribution system.

The procedure to produce synthetic networks, pictorially depicted in Figure 6, can be split into the following steps:

1. Choice of the geographic area to examine (the clearer the boundary of the area, the more accurate the results). The choice of such area should be driven by the available data on the total energy consumption and production publicly provided by regulatory bodies or power system operators (i.e., TSO or DSOs). A typical example of an area is an Italian region or a French region or department.
2. Geo-spatial study of the examined area to locate the PS (points of coupling between HV and MV systems). Among PS in the area, a classification is made to distinguish stations used to connect relatively big customers or producers (User Primary Station, UPS) from the primary stations whose distribution transformers are dedicated to feed distribution networks (Distribution Primary Station, DPS). The classification is necessary since there is no interest to model private distribution networks beyond a UPS.
3. Geo-spatial analysis of the territory in the examined area to study the social-economical texture. The region is divided into small portions, gradually smaller (e.g., provinces and municipalities) to gather detailed georeferenced data on the territory, the population, the use of land, the economy, the cities' planning, etc. Geographical information systems (GIS) are used to analyse maps and layers with related information. The goal of this point is to give to each territorial portion a share of energy consumption and production of the region and estimate the demand and production of the DPSs.
4. Building an incidence matrix to associate the territorial portions to one or more DPSs.
5. Assessing the total energy consumption and generation for each DPS.
6. Define the time series of active power for each DPS.
7. Associate the geographical information with each DPS, for the grid modelling.

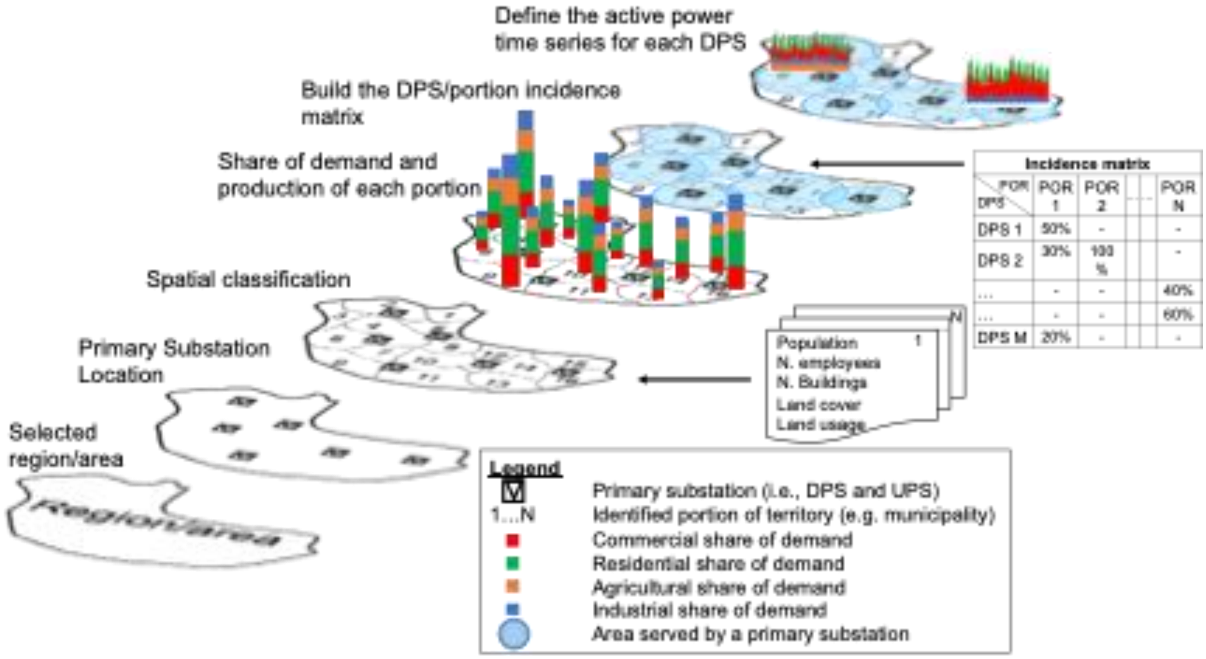


Figure 26: Modelling the distribution system using open data only: exchange profiles

The estimation of distribution grid flexibility must consider the grid technical constraints that can be calculated only with a detailed knowledge of the distribution grids. This information is never available for different reasons (e.g., for security reasons) and confined into the DSO databases. The main goal of the activity is the definition of synthetic networks capable to accurately represent the distribution networks fed by each DPS and to offer TSO an estimate of flexibility in terms of quantities, prices and temporal locations. The method is useful even in the cases where the information on distribution grid is open, since the distribution networks are continuously reconfigured, and it is much more efficient for high level operational planning studies to refer to a good representation of the distribution system instead of trying to follow the reconfigurations.

Thus, once the exchange profiles of each DPS are estimated, the network model in terms of lines and position of loads and generators is built to by using representative network portions derived from a clustering process that identifies the most typical network arrangements. Figure 7 depicts the GIS-based procedure to produce synthetic networks. The layers that include information about buildings in municipalities, unelectrified areas (i.e., lakes, ponds, forests, etc.), land usage, etc. are superimposed on to each other for building synthetic networks that result from a combination of elementary network portions located in an integrated final layer. Such an approach can be summarized as follows:

1. Association of the geographical/socio-economical information to the territory supplied by a given DPS, by exploiting GIS tools and applications (shape files and intersections).
2. Assessment of the shares of rural, industrial, urban areas according to the geo-spatial study of the territory.

3. Combination of elementary portions of representative networks for building the synthetic network that models the grid fed by the HV/MV transformer of the examined DPS.

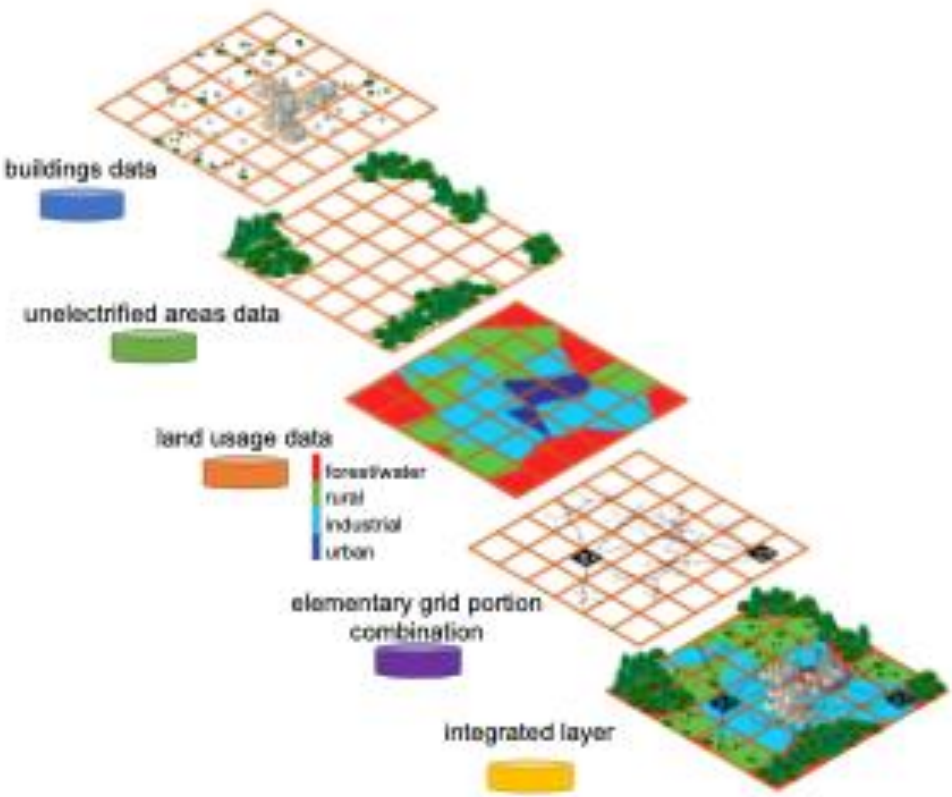


Figure 27: Modelling the distribution system using open data only: synthetic network building

Once the distribution network is modelled, several scenarios can be applied to model the flexibility provisions (typical scenarios are those with the existing or the expected level of Distributed Generation (DG), Demand Response, and Electrical Vehicles). For instance, a DG scenario includes different combinations of non-programmable RES and programmable energy sources (i.e., CHP or biomass generators) and reproduces the realistic DG scenario derived from the previous step 5 (i.e., the one used for defining the power profiles at the TSO/DSO interface). Starting from the known or estimated production of a distribution network, generators are allocated in the representative feeders that compose the passive model of the network until it achieves the supposed real DG penetration.

The results of the whole procedure are (i) the exchange power profiles at the TSO/DSO interface that characterize an equivalent generator at the PCC with a four-quadrants capability curve (i.e., in import or export mode, from/to the bulk grid), and (ii) for each DPS a synthetic network that can be used for all the studies that need to check the grid limitations.

5.2 Quantitative assessment of the market potential of the distribution grid

DERs connected to the distribution systems, as renewable energy sources (RES) based generators, fuel generators (e.g., combined heat and power, CHP), energy storage systems, as well as active users, can offer flexibility products basically represented by a variation of the scheduled/expected active and reactive powers. By considering the transmission grid point of view these variations of the scheduled/expected working points of DERs result in bids of flexibility services that, if awarded in the service market, produce a variation of the power profile at the PCC with the transmission grid (i.e., the TSO/DSO interface).

To estimate such a market potential of DERs connected to a distribution network, the proposed approach starts from the defined profiles at the PCC and hypothesizes the participation profiles of the DERs involved in offering flexibility products.

The DERs behaviour in the market is described in terms of pairs price/quantity of flexibility, and the bids are differentiated between upward and downward. It is assumed that the new market players behave rationally in the market (i.e., by trying to maximize their profits), and the bid prices are defined according to the ATLAS model.

For a given distribution network, the quantitative assessment of its market potential is obtained by performing the following steps for each considered time interval:

1. Calculate the maximum variations in upwards and downwards of the expected working point at the TSO/DSO interface by considering the hypothesized participation profile of each DER (i.e., maximum/minimum local generation and the minimum/maximum demand). Thus, the most extensive range of potential bids of the virtual power plant at the TSO/DSO interface is defined.
2. Represent with a fair number of points the range of upward and downward offers.
3. Perform power flow (PF) calculations to verify the compliance of the distribution grid with the technical constraints by applying the generation/load conditions corresponding to the points obtained in step 2. The PF calculation is performed on the synthetic model of the real distribution network.
    a. If no violations are found, the relevant flexibility can be used by TSO with no adjunctive cost than the one correspondent to the price of the bids times the quantity purchased.
    b. If operational issues are found (e.g., voltage regulation and power congestions), an Optimal Power Flow (OPF) calculation is performed for identifying the optimal set points of local resources required for fixing such distribution issues. The OPF aims at minimising the distribution system operation costs that have to be added to the price of the bids.
    c. It may happen that, in particularly critical distribution networks, some operational issues cannot be solved by resorting the local resources. In such cases, the flexibility that can be offered to the TSO is limited by the distribution constraints. The difference between the potential bids calculated in step 1 and the one resulting by eliminating the critical working points represent the measure of the *non-feasible* bids. DERs can bid such offers to the service market, but, if awarded, they will be blocked by the DSO (i.e., they will receive a red traffic light).

The result of this methodology is the price/quantity curves for upward and downward offers from a distribution network, for each time interval considered. The extra costs possibly sustained for the flexibility products or the blocks imposed by the DSO for avoiding harmful impacts on the distribution network operation are also outcomes of the methodology.

In Figure 28 it is shown an example of price/quantity curves assessed for a model of distribution network derived by a real case in Italy, at noon of two working days (wd) in winter (WIN) and summer (SUM). The grey zone highlights the blocked potential.

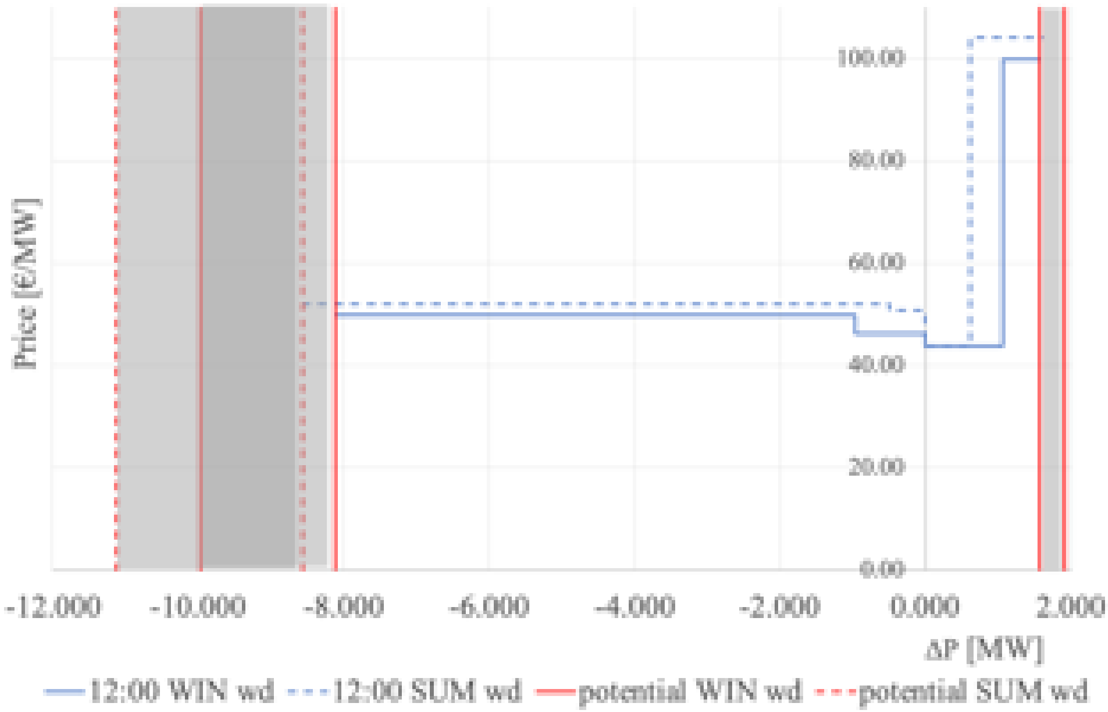


Figure 28: Price/quantity curves (noon of WIN and SUM wd)

The methodology has been applied to six Italian regions and to a selected list of French distribution primary substations spread in the central regions of France.

5.3 Integration of the models

For integrating the outcomes of this task into the whole market model for market mechanisms simulation described in this report, it is important to accurately check the boundaries of interest. In particular, it is necessary to consider the voltage levels of each model.

For modelling the distribution grid, the TSO/DSO interface point (HV side) is represented by an HV node of the sub-transmission network (namely, 33 kV, 63 kV, 70 kV, and 150 kV). The interfaces between TSO and major DSOs are at this level of the system, and such grids need to be modelled with the procedure proposed in the previous paragraphs. The MV side of the TSO/DSO interface supplies the distribution networks with nominal voltage ranging from 20 kV down to 6 kV depending on the DSO (the most common voltage level is 20 kV).

The reason for the choice of modelling the interface with the sub transmission network is that below the sub transmission voltage level the TSO has limited observability of the system and, even though the RFG (UE 2016/631) is determining an improvement on the knowledge of power production in the distribution system, the information on distribution network characteristics and constraints are still little at the TSO point of view with reference to the procurement of flexibility services.

A good model that represents the distribution system behind the TSO/DSO interface is an equivalent power plant that operates on four quadrants by including both positive and negative values of active and reactive powers, with a circular or rectangular capability curve, if active power P and reactive power Q are totally decoupled.

Therefore, for integrating such models of the distribution grid into the model of the region and the areas used for the nodal market modelling, one or more virtual power plants that reproduce the distribution system behaviour are aggregated to the nodes of the extra-high voltage grid, considered in the paragraph 5.1.

6 Conclusion

The complementary methodology previously described (cf. Figure 29) is used to perform the market studies. As can be seen from the detailed description of the different modelling approaches, they complement each other so that we are able to achieve a high level of accuracy in each dimension (functionalities or use cases).

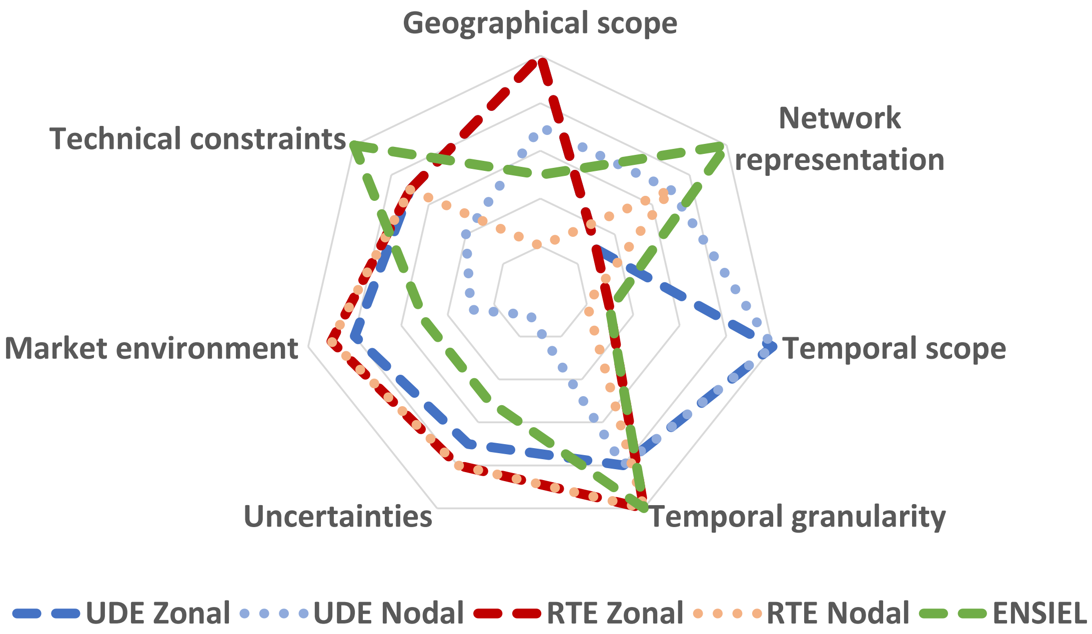


Figure 29: Comparison of different complementary model frameworks of UDE, RTE and ENSIEL

The outcomes of the market studies will be presented and compared in the Deliverables D2.4 and D2.5.

7 Appendix: Detailed methodology for time downscaling in ATLAS per data type

7.1 Solar generation forecasts

The typical solar production profile is bell-shaped with null values during the night. For this reason, a method of interpolation that minimizes deformations and based on the curve's second derivative is used.

The optimization problem used for solar generation time series is defined as follows:

- The decision variables are the interpolated values $x_{h,i}$ of solar generation at the i-th point of measurement of hour h, as illustrated below where n is the number of interpolated values per hour:

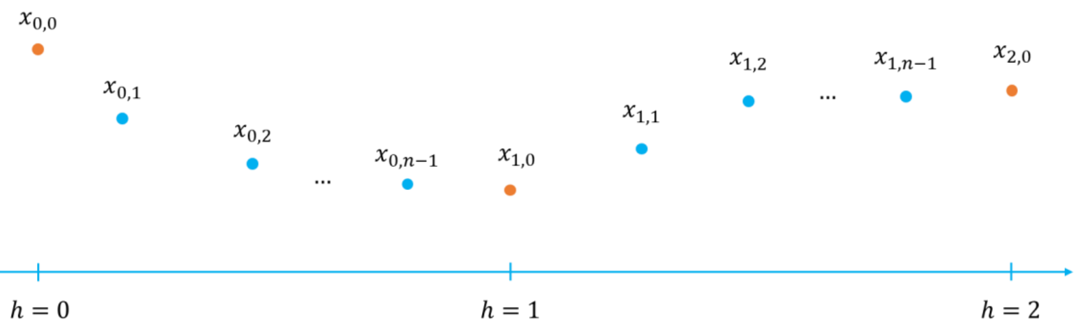


Figure 30: Decision variables for solar time series' interpolation

- The objective function is a linear combination of both L^1 and L^∞- norms of the discrete second derivative α_k along the curved drawn by all $x_{h,i}$ points. It can be written as:

$$\theta\left(\sum_{k=0}^{nT-3} |\alpha_k|\right) + (1-\theta)\left(\max_{0\leq k\leq nT-3} \alpha_k\right)$$

Where θ is a customizable weigh parameter, set to 0.75 by default, and T is the length of the optimization window, expressed in hours.

The second order derivative takes into account previously computed values if available, and starts from $x_{0,0}$ otherwise, when initializing. Provided all indices are valid, the formula giving α_k at the k-th point along the curve, starting at $k = 0$ (so that α_0 is measured at point $x_{0,0}$, α_{n-1} at point $x_{0,n-1}$, and α_n at point $x_{1,0}$ for instance) is the following:

$$\alpha_k = x_{\left\lfloor \frac{k}{n} \right\rfloor,\, k\,[n]} - 2x_{\left\lfloor \frac{k+1}{n} \right\rfloor,\, k+1\,[n]} + x_{\left\lfloor \frac{k+2}{n} \right\rfloor,\, k+2\,[n]}$$

With $\lfloor x \rfloor$ denoting the integral part of real number x, and $k\ [n]$ the remainder of quotient $\frac{k}{n}$.

The only constraints are that all $x_{h,i}$ must be non-negative, and that all $x_{h,0}$ must equal the hourly measurement of hour h. A mean constraint can be implemented instead of the latter, when hourly values do not represent measurements but average values resulting from an aggregation process.

A result example is shown below, with the orange curve representing prospective hourly measurements of solar generation in France, and the blue curve the corresponding interpolated values at a 15 minute time step:

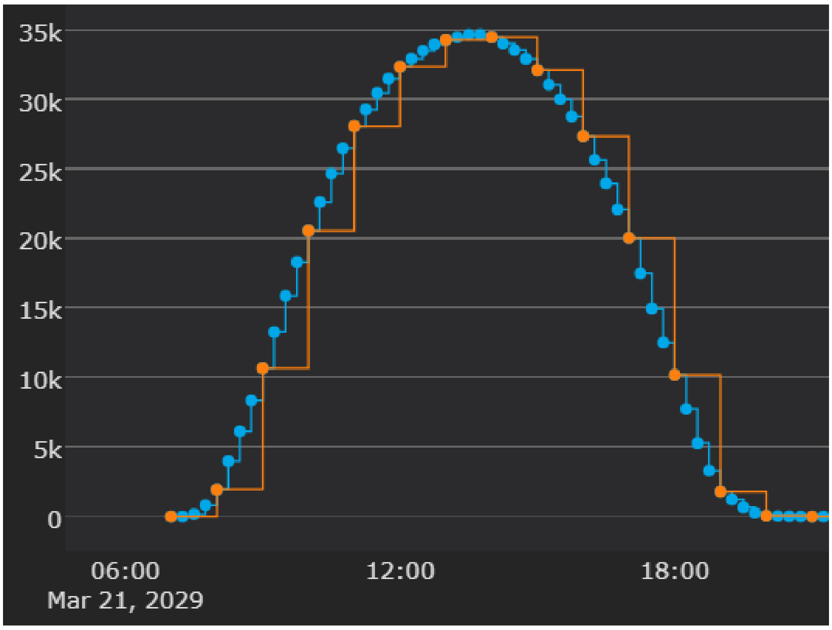


Figure 31: Example of interpolation of solar generation. Unit is MW (T = 3, n = 4)

7.2 Wind generation forecasts

Wind generation fluctuates a lot more than solar generation or consumption, even during short periods of time. Since norm-based interpolation methods did not prove effective, a spline-based interpolation was implemented instead.

The optimization problem used for wind generation time series is defined as follows:

- The decision variables over one optimization window covering T hours are the coefficients a_i of a polynomial P of degree d defined over that optimization window:

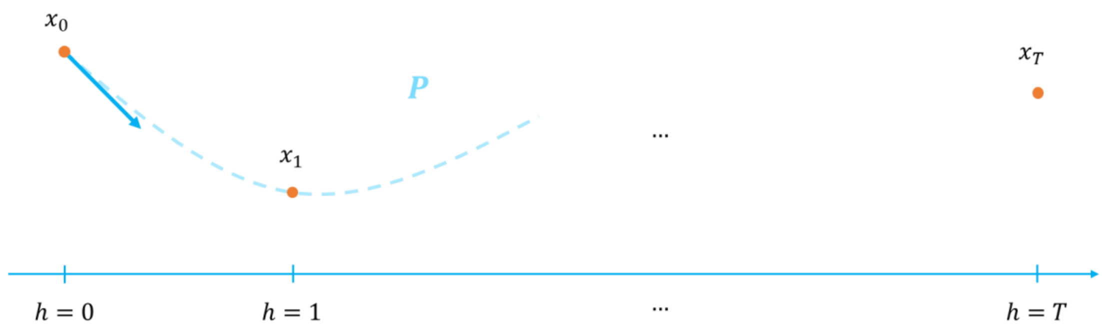


Figure 32: Visualizing decision variables for wind time series' interpolation

Depending on whether or not the optimization window is at the very beginning of the time series to interpolate, d is either set to $T + 1$ or $T + 2$. When the first optimization is performed at the beginning of a wind forecast time series, d is set to $T + 1$. Afterwards, d is set to $T + 2$ to compensate for an additional constraint on the first derivative of polynomial P, as described below (adding one constraint requires us to add one degree of freedom).

The objective function is the L^1-norm of the vector whose h-th coefficient is $P(h) - x_h$, with x_h denoting the wind power predicted for hour h. It can be written as:

$$\sum_{h=0}^{T} \left| \sum_{i=0}^{d} a_i h^i - x_h \right|$$

The only constraint that applies appears exclusively if the optimization window is not at the very beginning of the time series. It requires that the first order derivative of polynomial P at hour 0 (which is simply a_1) is equal to the first order derivative of the polynomial computed for the preceding window position at hour 1. This constraint ensures the final spline is always of smoothness at least C^1, which avoids undesirable breaks and guarantees a smooth final curve.

Unlike for solar generation time series, the technique used to interpolate wind generation forecasts does not provide any control on the sign of the interpolated values, and does no guarantee that in-between values will not be negative (on the contrary, the use of polynomials exposes the methodology to Runge's phenomenon of oscillating polynomials, although the purpose of splines is to limit that effect using piecewise polynomials with lower degrees). Therefore, a post-processing step has been implemented to ensure all values are non-negative: if one interpolated point is negative within a particular hour, then all interpolated values in that hour undergo an affine transformation that ensures the minimum value becomes zero, while leaving untouched the lowest hourly measurement bounding that hour. In practice, this rarely happens, and the securing post-processing procedure is virtually never called.

A result example is shown below, with the orange curve representing prospective hourly measurements of wind generation in France, and the blue curve the corresponding interpolated values at a 15 minute time step:

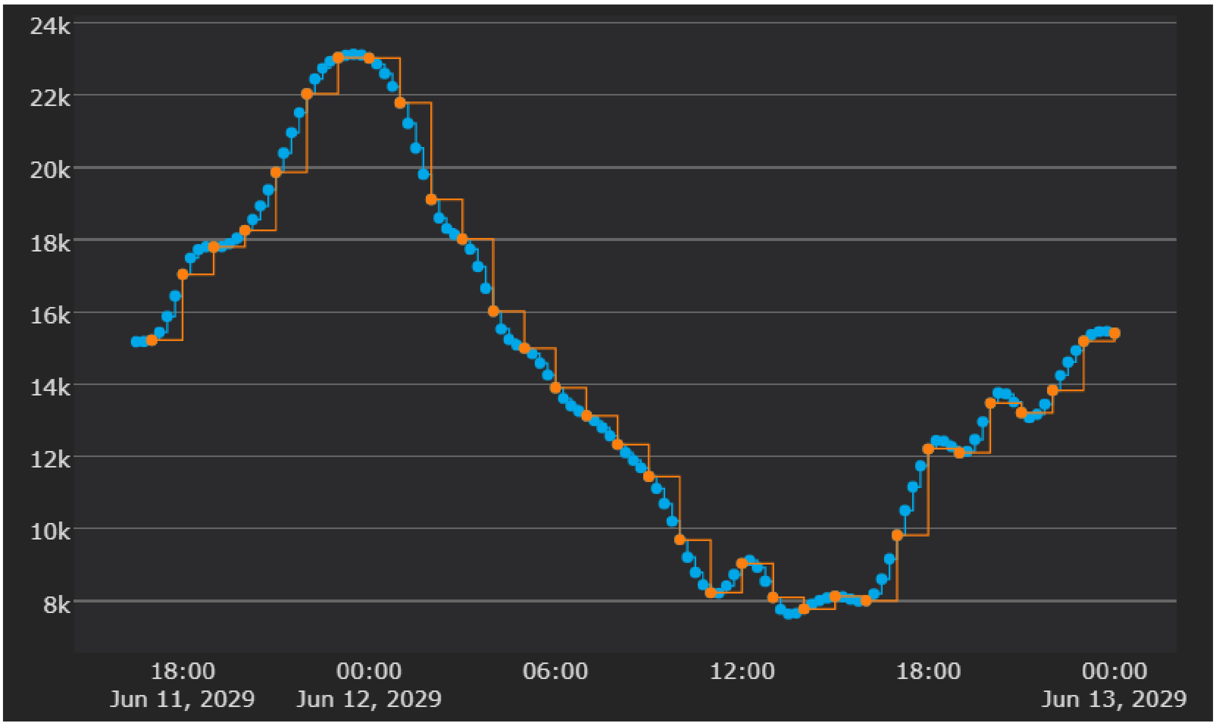


Figure 33: Example of interpolation of wind generation. Unit is MW (T = 2, n = 4)

7.3 Load forecasts

Load forecast interpolation is also based on splines, and the optimization problem used to compute interpolated points is mathematically identical to that of wind forecast generation. The difference between the two interpolation methods lies in the post-processing step.

Because load and wind forecasts share the same optimization problem form, the optimization problem that performs the interpolation of load forecasts will not be described in this section. Only the post-processing treatment will be detailed. Readers interested in the optimization problem for load forecasts are invited to consult that of wind forecasts in the previous section.

Just like wind forecasts, interpolating load forecasts with polynomials exposes the result to potential polynomial oscillations, which could, in theory, generate negative or absurdly high values in the interpolated time series. This phenomenon is partially controlled by the use of splines, but a post-processing step remains necessary.

Since usual values of load forecasts are numerically high when expressed in MW, it is almost impossible to observe negative values after an interpolation step. However, interpolated values sometimes exceed hourly measurements and thus become new peak loads, or on the contrary, become new consumption dips. The goal of the post-processing step is precisely to control these variations and ensure they remain consistent with the real load shape observed in practice.

The post-processing step relies on a user-defined parameter called *LoadMaximumOverload*, expressed in MW, which corresponds to the largest margin that interpolated values can take outside the range of hourly measurements. This concept is illustrated on the following figure:

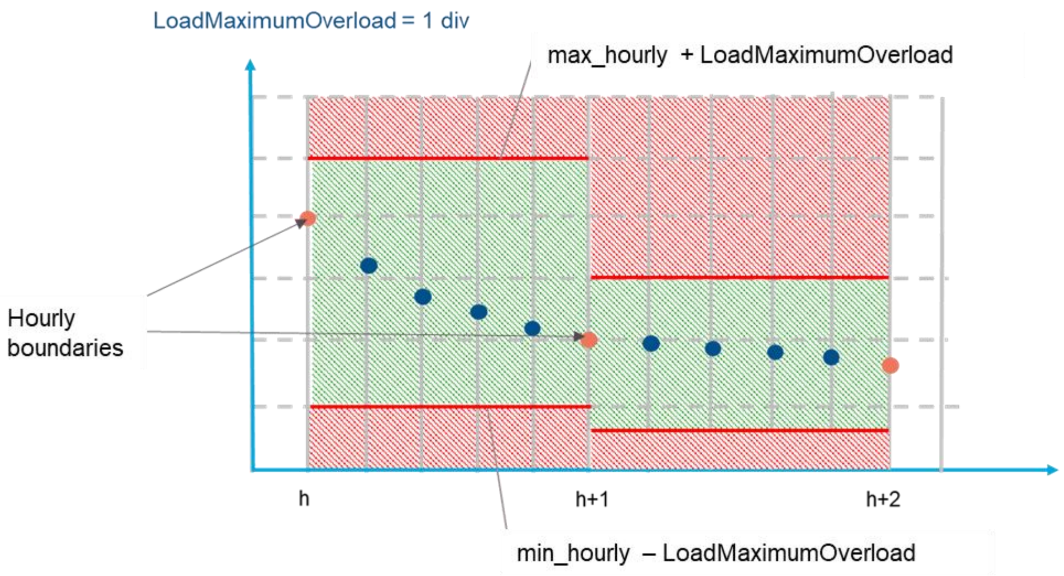


Figure 34: Controlling load excursions with the LoadMaximumOverload parameter. min_hourly and max_hourly are respectively the minimum and maximum values of the two hourly measurements bounding hour h.

If any interpolated value exceeds the range allowed by *LoadMaximumOverload* during a particular hour, then all interpolated values within that hour will undergo an affine transformation that narrow the excursion to the limit defined by *LoadMaximumOverload*, without changing the hourly measurements bounding that hour. In practice, this post-processing step takes effect only a few hours per year, and on small variation excesses.

A result example is shown below, with the orange curve representing prospective hourly measurements of load in France, and the blue curve the corresponding interpolated values at a 15 minute time step:

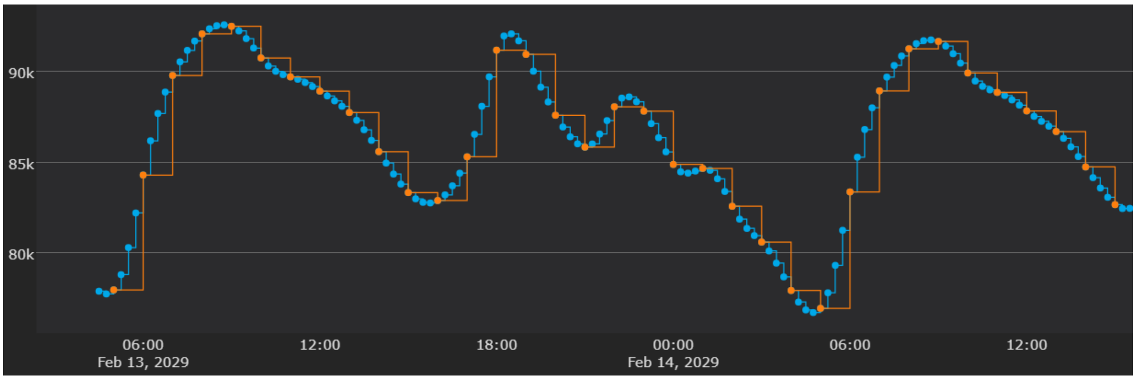


Figure 35: Example of interpolation of load. Unit is MW (T = 2, n = 4, LoadMaximumOverload = 1000 MW)

7.4 Price forecasts

Price variations can be much less smooth than all other physically-induced time series. However, in the absence of historical data close enough to the European market structure to calibrate a specialized interpolation model, and in an effort of simplification, the implemented interpolation methodology mirrors that of other time series.

Thus, the algorithm is very similar to load interpolation, since the objective function and the control of the excursions are both the same.

It is worth noting that this will not prevent the appearance of abrupt price changes in the simulations, as it is only the input price forecasts delivered to market players at the beginning of the simulation that will undergo this interpolation process. The subsequent market simulation steps will determine the output price values, and although they may be expected to be close the forecast, they will be able to vary much more freely.