

A Markov Decision Model for Cooperative Virtual Power Plants Market Participation

Crédo Paniah, Cédric Herpson, and Javier Gil-Quijano

Abstract—The advent of renewable energy induces new problematic electrical network configurations. Indeed, the renewable energy characteristics (intermittency, production uncertainty) combined with their inherent spatial distribution implies to aggregate them and to design specific control tools to connect them to the grid and allow them to participate to the market. Based on an industrial case study, we propose in this work to gather and coordinate numerous independent producers through a Cooperative Virtual Power Plant (CVPP). To perform an efficient market participation of such coalition of producers, we model its behavior as a Markov Decision Process. Our model takes into account renewable generation prior uncertainty, market constraints and optimizes sequentially the utilization of available resources. Experiments realized using realistic datasets show the efficacy of the proposed model and speak in favor of renewable and controllable producers' coalition in CVPP.

Index Terms—Cooperative virtual power plant, partially observable markov decision processes, renewable market participation, storage management.

I. INTRODUCTION

European countries adopted in 2008 the climate-energy package [1], setting as objectives to reach by 2020 a 20% decrease of greenhouse gas emissions, a 20% share of renewable energy in the total energy consumption, a 20% improvement of the energy efficiency. To achieve these objectives, European renewable installed capacities has known a remarkable expansion (Fig. 1). Yet the difficult integration of *Renewable Energy Sources* (RES) into the electrical grid (termed the Grid, hereafter) challenges the 2020's goals.

This difficulty is essentially due to renewable generation characteristics, which are hardly compatible with the current electrical system [2]. Indeed, renewable primary sources (wind, sun, etc.) are variable and intermittent, thus their behavior is difficult to forecast accurately. As electrical networks impose pull off and put in equilibrium, requiring a good knowledge of both demand and offer before operation, these characteristics strongly limit the integration of renewable generation in the Grid, as is.

One of the principal methods used to ensure the Grid equilibrium is to match offer and demands announcements on electricity exchange markets. They force participants to respect their production engagements or being subject to substantial penalties. A producer unable to ensure the

equality of generation announcements and realization thus jeopardizes:

- The stability of the physical network,
- Its competitiveness in the electricity market.

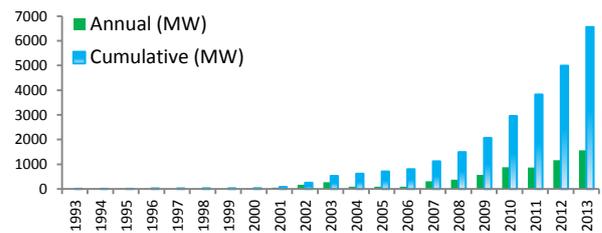


Fig. 1. European annual and cumulative offshore wind capacity installed – EWEA 2013 [3].

The case of off-shore wind power is representative of the challenges of RES. Indeed, their advantages (stronger and steadier generation profile, attractive selling prices [4]) make probable their increasing presence in the future. For the sustainability of the electrical system, such wind farms are interconnected by multi-terminal large offshore networks, connecting them as well to transmission networks [5]. Fig. 2 illustrates the classical configuration considered by stakeholders, to connect off-shore RES and storage units to the main grid.

The WINPOWER¹ project aims to design realistic, robust and stable control for off-shore networks connecting renewable sources and controllable ones to the main electricity system. Each actor must embed its own local physical control system, but global coordination mechanisms should be designed as well. Indeed, inherent RES characteristics and the addressed network configuration do not allow stakeholders to independently participate to the electricity market [6] as market rules constrain the different stakeholders on the network to appear as a unique (virtual) entity, from the point of view of the market, requiring management methods specifically suited to this task.

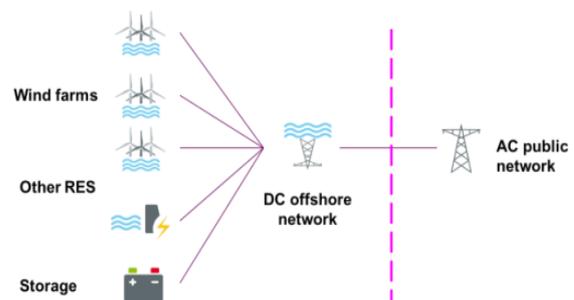


Fig. 2. Physical configuration.

Manuscript received April 28, 2014; revised July 3, 2014. This work was supported part by the French National Research Agency (ANR) in the frame of the WINPOWER project (WINPOWER-ANR-10-SEGI-016).

The authors are with the French Alternative Energies and Atomic Energy Commission, Saclay, 91191 France (e-mail: credo.paniah@cea.fr).

¹ Research project federating 4 companies and 5 research institutes, funded by the French National Research Agency.

Existing solutions for renewable market participation (short-term markets, *Virtual Power Plants* (VPP)) do not apply to these new configurations [7], mostly because they are not designed to handle multi-actors configurations – It is the case of the network addressed in our work (Fig. 2): as producers got to collaborate in order to participate as one to electricity markets, global control strategies are needed to manage them, addressing especially:

- The transmission and storage of supplemental energy while keeping the sustainability of the system,
- The optimization of all parties' financial equations (according to their own interests).

Such a solution falls into a paradigm recently introduced in the renewable management literature, *Cooperative Virtual Power Plants* (CVPP) aiming to model the collaboration between numerous decentralized, autonomous (renewable) producers. CVPP promote the coalition of RES to profitably sell their aggregated (i.e. more reliable and efficient) energy to the Grid. Our work lies within this context.

Within the project, our goal is to optimize day-ahead market participation of producers willing to collaborate and behave as an aggregated entity. We propose in this paper to model the CVPP behavior as a *Markov Decision Process* (MDP). This allows us to express the multi-actors configuration and to handle the uncertainty of RES in order to optimize day-ahead market participation of the CVPP.

In Section II, we first describe the physical network configuration considered in the WINPOWER project. We present the market context, and position our work into renewable management literature. In Section III, we formalize the problematic of multi-actors resources aggregation and control for market participation. Then we detail our aggregation mechanism proposal in Section IV. In Section V, we present the simulations performed and show our proposals utility to renewable producers gathered in a CVPP. Finally, Section VI discusses obtained results and present short term perspectives to this work.

II. WORKING CONTEXT

A. Market Context

In EU, the integrated and coordinated market has been initiated to establish competition in electricity exchanges. Its objective is to integrate a maximal part of electricity flows (in order to create a low electricity price, reflecting demand and offer signals on the market) and to reach an optimal mix (i.e. a mix with the highest social welfare²) using an auction mechanism. If other types of exchanges exist (bilateral contracts, options, etc.), they follow auction price signals. We will thus focus on the main auction market: the *Spot Day-ahead Market*.

In the day-ahead market, participants provide orders (offer/demand) for the following day [8]. They day is subdivided into 24 equal time-slots; for each time-slot, participants state electricity quantities and desirable prices at which they wish to buy/sell them. Exchange modalities (prices and volumes for each participant and time-slot) are defined by the market auction, the day ahead. On D-day, for

each time slot, participants have to respect contracts made the day ahead or pay penalties whose values are defined in function of the distance between the contract and its realization.

As introduced in the previous section, RES such as the wind farms of the WINPOWER project need to appear as one entity from the point of view of the market; it allows them as well to operate more reliably and to gain from this configuration. A framework to allow that is thus needed. We present, in the following section, existing approaches towards renewable optimal market participation and their respective limits.

B. Related Work

Current solutions used to manage renewable characteristics are of two types:

- Short-term markets participation,
- Virtual Power Plants (VPP).

Short-term markets aim at taking advantage of the higher accuracy of short-terms meteorological predictions [9]. It is more profitable for a renewable producer to participate in these markets, as their relatively short horizon permits to avoid imbalances and thus maximize market revenues [10]. They can also be used as a balancing solution to avoid long-term market contracts imbalances. Yet, the shorter a market horizon is, the relatively lower its volumes of exchanges are; thus, this solution cannot be considered as the only commercial outlet.

Another way to avoid renewable imbalances is to use controllable means to balance them; this solution is termed as VPP.

Indeed, the stochastic nature of renewable generation results in prediction errors; which need to be managed. VPP associate renewable to controllable devices in order to take advantage from the both. Virtual Power Plants have been defined as coalitions of multiple energy producers and, possibly, energy storage providers that come together to sell electricity as an aggregate [11]. Often, controllable sources in VPP must be scheduled, taking into account uncertainty information available prior to delivery time, to give an optimal utility [12].

To our knowledge, in most existing models, controllable resources utilization is planned only after market participation, to meet furniture contracts priority made [13], [14]; moreover, a number of them deal with the case where a unique stakeholder owns the whole VPP production means, and use controllable resources as backup means. Or, in a configuration with substantial storage amounts to control, controllable units cannot be considered only as adjustment variables; indeed, their respective utilities pass by their own market participation.

So, the VPP paradigm is not suited for a *collaborative* multi-actors configuration [15], *where producers coalesce willingly to sell their production (without relying on any external entity), as it is their best interest to do so* [7]. To overcome this limit, recent works proposed Cooperative VPP (CVPP) as a framework to aggregate small and distributed energy resources in the electricity supply network [16]. They propose payment schemes which incentive the actors of CVPP to provide truthful predictions of the future generation. However (and unlike the studies aforementioned), as controllable resources are not scheduled, given predictions

² defined as benefits of market participation for producers and consumers [8].

uncertainty, the resulted reactive operation is suboptimal as it does not consider the risks of difference between contracts and future realizations.

Our current work fits in this paradigm: In this paper, our goal is to develop an optimal control framework (allowing risk management and resources planning for market participation) for CVPP implementation. In contrast to studies previously cited, we thus consider controllable production means in the computation of the offer to submit to the day-ahead market. Associated with the consideration of RES production uncertainty, we believe that our proposal leads to significant improvements regarding both the network stability and the economic viability of RES stakeholders. In the following section, we first present the control architecture proposed to deal with off-shore CVPP before detailing the key points of renewable and controllable sources aggregation.

III. ORGANIZATIONAL MODEL AND FORMAL DEFINITIONS

The norm IEC-62325 [17], describes the different existing roles in the energy market and details the possible interactions between the actors. The organizational model that we propose for a CVPP relies on it. In our case, physical actors have a *producer* role and virtual actors have either an *aggregation* or a *market participation* role.

These three types of roles are distributed over three layers on our organization proposal (see Fig. 3):

- Production layer: Each *production actor* represents a wind farm or a storage device, owned by a stakeholder. It sends relevant data (i.e. prediction and uncertainty information associated for RES, storage characteristics storage devices owners) to the aggregation layer.
- Aggregation and Decision layer: The *aggregation actor* determines the optimal aggregated resources planning for market interaction.
- Market Interaction layer: The *market interaction actor*

uses the planning elaborated at the previous level to interact with the electricity market and contract generation amounts for the following day, on behalf of the whole coalition.

The key layer in this organization is the aggregation and decision one and it is the issue we address in our work.

Off-shore wind farms make prediction about their future production but there are errors in the predictions. Other producers like mass storage devices can take advantage of their controllability in such a physical configuration. The question this layer aims to answer is thus: *How to plan beforehand their utilization such as the market participation of the CVPP is optimal and that the different stakeholder's utility are maximized.*

A. Production Prediction and Uncertainty

The aggregation function takes as input individual production predictions and associated uncertainty. The output is a production offer for the coalition.

$$AggFct: (\mathbb{V} \times \mathbb{P}^p)^{n+m} \rightarrow \mathbb{V}^p$$

$$(pred_i, uncert_i)_{1 \leq i \leq n} \mapsto plan_{CVPP}$$

where:

- $n + m$ is the number of producers (n is the number of renewable producers and m is the number of controllable resources)
- p is the number of time slots
- \mathbb{V} is the set of possibly produced values
- \mathbb{P} is the set of uncertainty functions
- $pred_i \in \mathbb{V}^p$ is the p-values sequence of predicted energy amounts of producer i
- $uncert_i \in \mathbb{P}^p$ is the p-values sequence of uncertainty functions of producer i
- $(plan_{CVPP}) \in \mathbb{V}^p$ is the CVPP planning, defining its future generation amount.

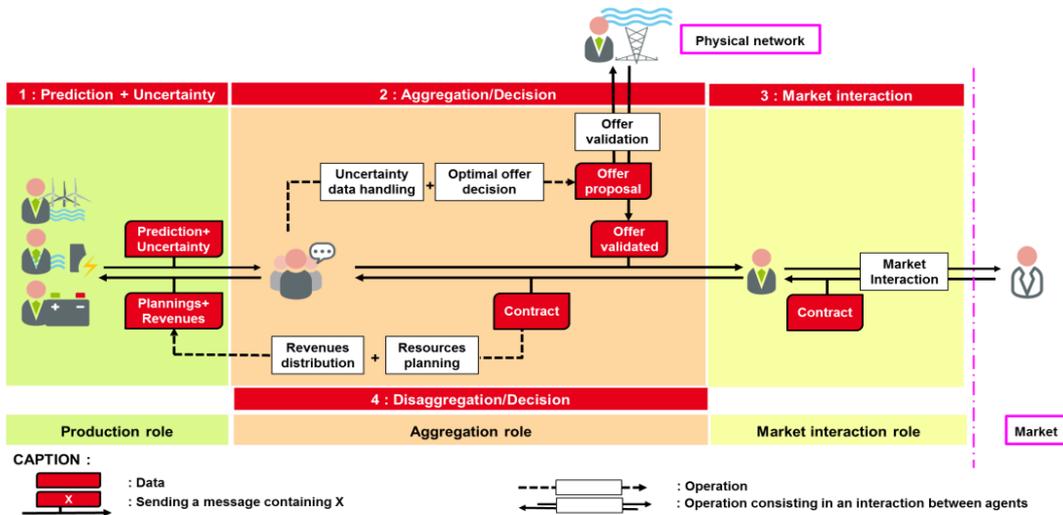


Fig. 3. Organization of the roles in the CVPP.

As introduced in Section II, on the day-ahead market, $p = 24$ and the utility of a planning mainly depends of its distance to its effective realization (as revenues and network safety directly depends of that distance). Thus, the optimal planning should maximize contracted revenues and minimize

the risk induced by the eventual distance with its realization.

$$AggFct: (\mathbb{V}^{24} \times \mathbb{P}^{24})^{m+n} \rightarrow \mathbb{V}^{24}$$

$$(pred_i, uncert_i)_{1 \leq i \leq n} \mapsto$$

$$argmax_{plan_{CVPP}} \Phi(rev_{CVPP}, \Delta(plan_{CVPP}, uncert_{CVPP}))$$

where:

- $(rev_{CVPP}) \in \mathbb{R}^{24}$ is the CVPP 24-values revenues vector
- The function Δ expresses the estimated gap between the computed planning and the future possible realization. This distance is predicted based on the producers' uncertainty functions
- The function Φ expresses that the utility of $plan_{CVPP}$ is function of the revenues rev_{CVPP} of the production plan if it were to be perfectly realized, and of the possible deviation between planning $plan_{CVPP}$ and effective realization, represented by $uncert_{CVPP}$.

Reference [18] defines the risk over a decision as a function of:

- The likelihood of the decision being a regrettable one,
- The amount by which the decision is regrettable.

The controllable means (i.e. those which offer variable generation volumes) makes it problematic to find the quantity that provides maximal contracted revenues and minimal penalty expectation, in each market time-slots. More, there exists a temporal dependence between market time-slots decisions; local optima computed for each time-slot can lead to global (over a day time) suboptimal decisions.

A solution is to keep a certain part of available capacity of charge (respectively discharge), in order to use it as backup, on the delivery day. Because of the possible future dynamic of the backup capacity utilization, it comes down to optimize the controllable means under uncertainty, with the information available day-ahead.

Several works studied the obtainment of such uncertainty information [19]; we will not focus on this aspect, but its utilization to obtain optimal controllable resources set points in a CVPP environment.

B. Background on Markovian Decision Processes

Markov Decision Processes are a mathematical control and decision paradigm which allows to model such a multi-producers configuration, controllable devices dynamics and to make decision over resources utilization, given possible outcomes and temporal dependencies over decisions.

The Markov property reflects the characteristic of a system to embed in its state at time t all the information necessary to predict its state at time $t + 1$; knowing previous states doesn't provide more useful information.

A MDP is defined by its set of states, actions and for each combination of state and action, the one step dynamic of the environment [20]. Formally, it is a tuple $M = \langle X, U, A, G, d \rangle$, with :

- X is a finite state of states
- U is a finite state of actions
- $A: X \times U \times X \rightarrow [0,1]$ is a transition function, associating to each state $x \in X$ combined to any action $u \in U$, the probability to go in a state $x' \in X$
- $G: X \times U \rightarrow \mathbb{R}$ is a reward function, associating to each state $x \in X$ combined to any action $u \in U$, an immediate expected reward
- $d \in [0,1]$ is a discount factor reducing future rewards value compared to current decision rewards. In our work, we will set it to 1, considering future rewards to

be as valuable as immediate ones.

A policy π associates to each state $x \in X$, the action $u \in U$ a decision maker should make; and the expected cumulated reward can be recursively expressed as the sum of immediate reward and weighted (by probabilities) expected rewards in future states (future times) :

$$V^\pi(x) = G(x, \pi(x)) + d \sum_{x' \in X} A(x, \pi(x), x') * V^\pi(x')$$

Then, the optimal policy belongs to the set of policies that respects the following optimality equation:

$$\pi^*(x) = \operatorname{argmax}_{u \in U} \left(G(x, u) + d \sum_{x' \in X} G(x, \pi(x), x') \times V^\pi(x') \right)$$

This formulation allows sacrificing short-term performance in order to optimize long-term one, using sequential recursive optimization. MDP have been an intensive research subject for years and several and efficient resolution algorithms have been proposed, experienced and confirmed; although their efficiency is limited by the uncertainty information accuracy, it constitutes a practical and easily implementable solution paradigm.

In the following section, we propose a formal definition of the MDP model expressing the behavior of the CVPP.

C. Notation

In the remainder, we use superscripts to denote particular subsystems (i.e. a participant of the coalition) and subscripts for the time index. Thus x_t^i denotes the state of the subsystem i at time t .

$\prod_{1 \leq i \leq n} (x_t^i) = x_t^1 \times x_t^2 \times \dots \times x_t^n$ is the Cartesian product (the product in case of cardinals) of variables $x_t^i, i \in \{1, \dots, n\}$.

We use canonical representations of Markov decision process, either by expressing the dynamic of the system $x_{t+1} = \Phi(x_t, u_t, w_t)$ or by writing transition (A) and reward (r) matrices of the MDP, such that:

A is a sequence $A = A_1, A_2, \dots$ with $A_1: X \rightarrow [0,1]$ such that

$$\begin{cases} A_1(x) \geq 0 \\ \sum_{x \in X} A_1(x) = 1, \forall x \in X \end{cases}$$

And

$\forall t > 1, A_t: X \times X \times U \rightarrow [0,1]$ such that

$$\begin{cases} A_t(x', x, u) \geq 0 \\ \sum_x A_t(x', x, u) = 1, \forall x, x' \in X \text{ and } u \in U \end{cases}$$

r is a sequence r_1, r_2, \dots with $r_t: X \times U \rightarrow \mathbb{R}, \forall t \in \{1, \dots, p\}$

D. Model Inputs

Let's consider the participation of the WINPOWER

coalition into the day-ahead market, for a given day. The first task of the aggregator/CVPP is to collect all the relevant information available. The second is to compute an optimal offer which:

- Integrates as much as possible CVPP generation into the market,
- Minimize the risk of seeing the coalition effective generation be far from the offer made on the market.

It reduces to a maximization of revenues collected from market participation of the coalition (it corresponds to producers' revenue maximization, if a relevant distribution mechanism is available, but this task is out of the limits of our current work so we will assume it available).

As we said before, the market is divided into $p = 24$ operation time-slots a resources optimization need to be performed to make an optimal offer in each of them. Available information is provided to the central controller (the aggregator) by each actor in the coalition and for each time-slot:

*RES actors provide forecast information about their expected generation at day+1 associated to uncertainty information*³.

Formally:

At time $t \in \{1, \dots, p\}$, and for the actor $i \in \{1, \dots, n\}$, the information provided is the tuple $X_t^i = (f_t^i, E_t^i, P_t^i)$ where :

- f_t^i is the forecast point prediction
- E_t^i is a finite set of possible errors over the forecast value f
- $P_t^i: E_t^i \rightarrow [0,1]$ is a function such that $\forall e_t^i \in E_t^i$ (e_t^i is a possible error), $P_t^i(e_t^i) \geq 0$ and $\sum_{e_t^i} P_t^i(e_t^i) = 1$. It is a function giving the probability of each possible error to appear, the following day.

This tuple allows to express that the possible amounts of energy generated by the actor i at time t g_t^i is :

$$g_t^i = f_t^i + e_t^i$$

with $-f_t^i < e_t^i < Cap_{max}^i - f_t^i$

where Cap_{max}^i corresponds to the maximum production capacity of i .

When the forecast is perfectly accurate, the effective generated amount of energy equals the forecast point prediction i.e. $e_t^i = g_t^i - f_t^i = 0$.

*The optimal state of charge*⁴ of a storage device j is computed in each time slot, from its characteristics (minimal state of charge (SoC_{min}^j), maximal state of charge (SoC_{max}^j), charge (R_{charge}^j) and discharge rates ($R_{discharge}^j$)⁵, initial state of charge (SoC_0^j) and if applicable, previous decisions made.

Formally:

At time t , the state of the actor $j \in \{n + 1, \dots, n + m\}$ can be computed from the following initial informations and the set points decided at the previous times :

- The tuple $(SoC_{min}^j, SoC_{max}^j, R_{charge}^j, R_{discharge}^j)$ is provided with SoC_0^j that denotes the state of charge of the device before the first computing time.
- At time t , the set T_t^j such that $SoC_t^j \in T_t^j$ is the set of possible set points of the controlled device j . It is a function of the state at the previous time, the dynamics and the characteristics of the device :

$$T_t^j = \varphi(SoC_{t-1}^j, (SoC_{min}^j, SoC_{max}^j, R_{charge}^j, R_{discharge}^j))$$

To perform its decision, the CVPP needs a view of the future state of the market. Market clearing prices predictions are computed for each market time slot; each market point forecast can be associated with some uncertainty information.

We'll consider in our work that we only have a point prediction of the MCP either perfect or imperfect, and possibly along with an associated uncertainty.

The penalty risked by the CVPP in each market time-slot, is a function γ of the future MCP and a penalty factor k predefined in market rules such as the penalty cost is $\gamma(MCP_t, k) = MCP_t + k * MCP_t$

IV. MDP AGGREGATION MODEL

Recent theoretical works [21] showed that networked systems coupled via communication links, each one evolving as a Markov process, and controlled by a centralized actor, taking optimal sequential decisions over a finite state space and in partially observable environments, can be represented as an unique *Partially Observable MDP* (POMDP) which itself can be converted into an *Information State MDP* (IS-MDP), and solved using classic dynamic programming algorithms, if a relevant information state can be built.

We draw on this work to propose in the following section a representation of the CVPP under the form of a MDP, allowing the aggregation agent to perform optimal decision about the control of available resources.

A. A First Model Proposal

To model our CVPP, a first approach is to consider it as a whole discrete time dynamic system, where the state of the system at time $t > 0$ is denoted x_t . The system dynamics are :

$$x_{t+1} = \Phi(x_t, u_t, w_t) \quad (1)$$

With $x_t \in X_t$ is the state of the system at time t such that $X_t = \prod_{1 \leq i \leq n} (X_t^i) \times \prod_{1 \leq i \leq m} (T_t^j)$

From this definition,

- x_t is the combination of local states of the CVPP actors at time t . The state of an RES i at time t is the error $e_t^i \in E_t^i$ associated to the prediction f_t^i and the state of controllable devices j is the set point SoC_t^j decided as action at the previous time $t - 1$.

- $u_t \in U_t$ is the control action or the decision taken at time t such that $U_t = \prod_{1 \leq i \leq m} (T_t^j)$. From this definition, u_t is the combination of local decisions made on any controllable device being a part of the CVPP, at time t . We define a local action $u_t^j \in T_t^j$ as the choice of a set

³This latter can take, for example, the form of a discrete probability distribution over possible errors of the forecast tool.

⁴For the sake of simplicity, we assume in this paper that the SOC determination error is negligible.

⁵We assume that these rates are constant regarding computation times and states of charges.

point SoC_t^j .

- w_t is a random variable which influences the transition of the system between two computation periods. Here, $w_t \in W_t$ is defined by the combination of random processes $W_t = \prod_{1 \leq i \leq n} P_t^i : E_t^i \rightarrow [0,1]$ of errors over the generation prediction or RES actors.

Equation (1), expressing the dynamic of the system, from a given time to the following one, means that the state of the system is defined as the combination of its components state and that it changes between two computation periods, as a function of decisions taken and wind realization.

As the state of the system and its possible transitions are dependent of the time, we affine the expression of the matrices as following :

$$\begin{cases} A_1(x_1) \geq 0, \forall t > 1 \\ \sum_{x_1 \in X_1} A_1(x_1) = 1 \\ A_1(x_t) = 0 \text{ if } t \neq 1 \end{cases}$$

$\forall t > 1, A_t : X_{t+1} \times X_t \times U_t \rightarrow [0,1]$ such that

$$\begin{cases} A_t(x_{t+1}, x_t, u_t) \geq 0, \forall x_t \in X_t, x_{t+1} \in X_{t+1} \text{ and } u_t \in U_t \\ \sum_{x_t} A_t(x_{t+1}, x_t, u_t) = 1, \forall x_t \in X_t \text{ and } u_t \in U_t \end{cases}$$

$$g_t : X_t \times U_t \rightarrow \mathbb{R}, \forall t \in \{1, \dots, p\}$$

Here :

A is the transition matrix that define for each time $t \in \{1, \dots, p\}$ the transition probabilities from each state $x_t \in X_t$ to each state $x_{t+1} \in X_{t+1}$ given any possible action at the time t $u_t \in U_t$.

Given the explanation we gave about the dynamic of the system, we can write for all $x_t = e_t^1 \times \dots \times e_t^n \times SoC_t^1 \times \dots \times SoC_t^m$ the following :

$$\begin{aligned} A_1(x_1) &= 0 \text{ if } \forall j \in \{1, \dots, m\}, SoC_1^j \neq SoC_0^j \\ A_1(x_1) &= \prod_{i \in \{1, \dots, n\}} P_t^i(e_t^i) \text{ else.} \\ A_t(x_{t+1}, x_t, u_t) &= 0 \text{ if } u_t \neq SoC_t^1 \times \dots \times SoC_t^m \\ A_t(x_{t+1}, x_t, u_t) &= \prod_{i \in \{1, \dots, n\}} P_t^i(e_t^i) \text{ else.} \end{aligned}$$

We define the reward of the transition to a state x_t as :

$$g_t(x_t) = \widehat{MCP}_t * (PA + SAS) - \gamma(\widehat{MCP}_t, k) * REA - \alpha(PA) - \beta(SAT)$$

where

- α est le cout de la production renouvelable is obtained from the cost provided by RES actors,
- β is the cost of using the storage,
- $PA = \sum_{i \in \{1, \dots, n\}} f_t^i$ is the total amount of generation predicted by RES actors,
- $EA = \sum_{i \in \{1, \dots, n\}} e_t^i$ is the total amount of errors in the particular transition scenario considered,
- $SAT = \sum_{j \in \{1, \dots, m\}} (SoC_t^j - SoC_{t-1}^j)$ is the storage amount used by the transition process,
- $SAR = EA - REA$ is the amount reserved by the

storage for the following day,

- $\begin{cases} REA = \max(0, EA - SAT) \text{ if } EA \geq 0 \\ REA = \min(0, EA - SAT) \text{ if } EA < 0 \end{cases}$ is the remaining error amount the CVPP has to pay penalties for,
- $SAS = SAT - SAR$ is the storage amount sold specifically in the day-ahead market.

This means that at each time, the set of possible states in which the MDP can be is driven by the probability distribution of errors over predictions and by the action made to handle it ; and that the cost of each action is given by how it handles the current error cost as well as how it allows controllable means to participate into the market to acquire additional revenues.

The model allows integrating the uncertainty information associated to MCP_t as well, either by computing a mean value of its future possible values, or by computing expected rewards of decisions, given MCP_t possible values.

However, this architecture presents two major drawbacks :

- It is very difficult to compute the solution of such a MDP because of its complexity and the size of the solutions set. As an illustration, the number of possible states is

$$Z = \prod_{i \in \{1, \dots, n\}} (card(E_t^i)) * \prod_{j \in \{1, \dots, m\}} (card(T_t^j))$$

where $card$ is the cardinality function.

And the size of the transition matrix (respectively the reward matrix) is

$$W = Z * Z * Q$$

where $Q = \prod_{j \in \{1, \dots, m\}} card(T_t^j)$ is the number of actions.

It requires substantial computation time and memory to compute and store the analytic form of the MDP, prior to an eventual solution computing. To overcome this latter drawback, we took advantage of matrices sparse representations [22] to reduce memory costs by storing only non-zeros value of the MDP matrices (i.e. transitions and rewards from states x_t to states x_{t+1}).

- This model is suited for the real time control of the CVPP as it assumes that the current state is actually known at the computation time. Yet, this is not our actual case as the idea is to set the needs for the following day, the errors values for each hour are not known. We are in the case of a Partially Observable MDP where the current state is not known. So, to obtain a relevant MDP expression, a solution is to extract this belief over the distribution of states at each computation time from the POMDP.

B. Information State Theory

The probability distribution over the possible states in which the system can be at each time constitutes a belief state that can be used as information state for the POMDP described above. The information state of a POMDP consists of either a complete history of actions and observations or their corresponding sufficient statistics [21].

Extract this information from the POMDP allows to reduce the state space of the system to the set of controllable devices

set points ; its dynamics are deterministic and are driven by the actions taken at every time ; the reward function over each state and action is dependent of the possible errors distribution.

We obtain an IS-MDP whose expression is :

$$x_{t+1} = \Phi(x_t, u_t, w_t) \quad (2)$$

With $x_t \in X_t$ is the state of the system at time t such that $X_t = \prod_{1 \leq i \leq n} (T_t^i)$. From this definition :

- x_t is the combination of local states of the controllable actors at time t .
- $u_t \in U_t$ is the *decision* taken at time t such that $U_t = \prod_{1 \leq i \leq m} (T_t^i)$. It is the combination of local decisions made on any controllable device being a part of the CVPP, at time t .
- As the random variable of errors over the predictions is extracted in the *information* state, the MDP is a deterministic one where the transition is driven only by the action (the state at time t is always the one chosen at the previous time).
- The belief state impacts the reward over each action; at time t , it is the distribution of

$$e_t = e_t^1 \times \dots \times e_t^n \in \prod_{1 \leq i \leq n} E_t^i$$

Such as $P(e_t) = \prod (P_t^i(e_t^i))$ and it impacts the reward of action decided at time t .

The MDP associated to such a dynamic is the tuple (A, g) defined for every $x_t = SoC_t^1 \times \dots \times SoC_t^m$ as:

- 1) $A_1(x_1) = 0$ if $\forall j \in \{1, \dots, m\}, SoC_1^j \neq SoC_0^j$
 $A_1(x_1) = 1$ if $\forall j \in \{1, \dots, m\}, SoC_1^j = SoC_0^j$
 $A_t(x_{t+1}, x_t, u_t) = 0$ if $u_t \neq SoC_t^1 \times \dots \times SoC_t^m$
 $A_t(x_{t+1}, x_t, u_t) = 1$ if $u_t = SoC_t^1 \times \dots \times SoC_t^m$
- 2) And the reward is computed in function of the belief state as :

$$g_t(x_t) = \widehat{MCP}_t * (PA + SAS) - \beta(SAT) - \alpha(PA) - \sum_{e_t} P(e_t) (\gamma(\widehat{MCP}_t, k) * REA)$$

This aggregation models allows the collaboration between renewable actors. Indeed, if a new RES actor is integrated in the coalition, the belief it has about its states is integrated in the belief state of the coalition and the optimal planning is computed as well with the updated reward function. If a new controllable device is integrated in the coalition, the state of the MDP is updated with this new information and the reward function remains the same.

V. SIMULATION AND RESULTS

We implemented the proposed model on a network test-case based on the physical architecture presented in Fig. 2. We simulate a cooperative operation (CVPP) and compare the results to a non-cooperative configuration (i.e. when every producer acts on his own). Then, we show that a day-ahead optimization of CVPP resources as we propose gives better results than using controllable resources in the

CVPP as a backup solution, as proposed in the reference work of [13]. We present hereafter our simulations modalities and objectives, and we end by an analysis of the simulations' results.

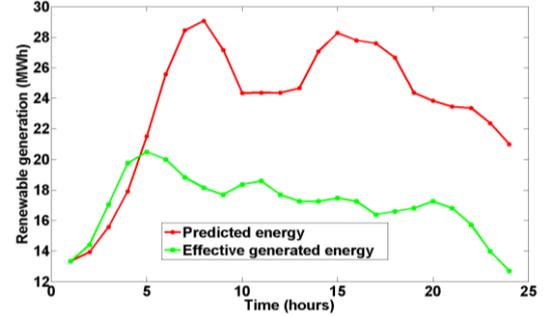


Fig. 4. Example of renewable day-ahead predictions and effective generation.

A. Simulation's Description

1) Wind farms

We consider two wind farms with a nominal capacity of 100 MW and two sets of data over their operation between February 1st and March 31st 2008:

- Day-ahead prediction data is provided for the 24 time-slots of the following day.
- Effective production data of each hour of the following day is provided as well.

Data is simulated on the basis of a statistical model calculated over several wind farms, as described in [23]. Fig. 4 shows an example of prediction and production values, for the 24 hours of a day.

2) Controllable device

The network that we consider associates the two wind farms to a storage device: it is an on-shore STEP storage device (a well-known technology, for which utilization costs are available), with a nominal capacity of 200 MWh. We assume charge and discharge rate to be 200 MWh. Variable cost of storing (respectively retrieving) one MWh into the storage device (respectively from the device) are approximated to a constant value of 10 €. It is a likely value obtained through off-line optimization of storage operation. Network links are designed to allow the storage in (respectively the retrieval from) the device of the whole nominal capacity. These realistic hypotheses have been validated by our industrial partners in the context of the simulation of our proposals.

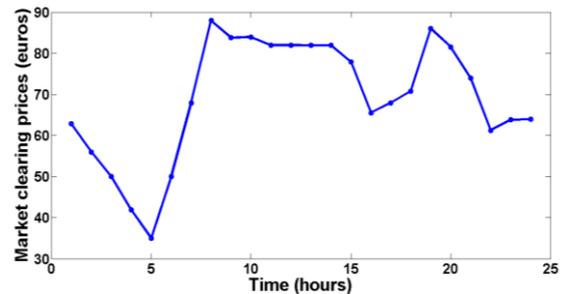


Fig. 5. An example of market prices evolution within a day operation. Prices for February the 1st 2008 [24].

3) Market prices and penalties

We assume that the coalition's generation is negligible on

the day-ahead spot market i.e. it does not considerably influence the market auction outputs. Given this assumption, we use relevant historical market data to simulate a market environment. We use data (available on the EPEX website [24]) of February and March 2008. We consider a perfect prediction of future Market Clearing Prices. Fig. 5 illustrates prices variability within a same day for February the 1st 2008. The CVPP is a price-taking entity into the market, so the whole amount proposed is always accepted on the market. Indeed, on the time period considered, renewable generation marginal costs are quite low compared to MCP values. We assume that any difference between the contracted and realized generation is penalized, and that no feed-in tariffs are considered for renewable producer (i.e. the CVPP acts as a conventional producer). The market penalties for a given hour slot t is calculated using the following equation:

$$Penalties_t^{CVPP} = MCP_t \times (1 + k) \times \Delta_t^{CVPP}$$

where Δ_t^{CVPP} is either the deviation in MW from the contracted amount by the CVPP at time t (given by the optimization of controllable resources in the CVPP), or the sum of wind farms predictions absolute errors values (when their independent market participation is assumed).

Finally, we use the penalization factor $k = 0.05$ applied in 2008 by RTE, the French electricity transport network manager [25].

4) Management strategies

In order to evaluate the performance of our proposal we compare it to different management strategies. The considered strategies are:

a) Perfect market participation

It is the reference case, when producers forecast perfectly their future generation and contract these amounts with the

market. The storage device optimizes its independent participation to the market, given market clearing prices evolution within the day.

b) Independent producers' basic strategy

It is an independent operation of producers, with a penalization of prediction errors. Every producer acts on its own, there is no coordination between actors and the risk is not managed (i.e. penalization amounts are maximal).

c) CVPP reactive backup strategy

The storage is used on the delivery day to meet contracts established with the market the day-ahead. It is a reactive resource used exclusively to backup predictions errors; it does not take into account the storage dynamics or market penalties evolution.

d) CVPP optimized backup:

This strategy implements the coordination solution as proposed in [13]. The storage is optimized the delivery day as a backup solution, to minimize market penalties. Nevertheless, as we stated in our work, it corresponds to an optimization case where no uncertainty information is used to allow a day-ahead participation of controllable devices. It constitutes also the lower bound of efficiency of our method proposal.

e) CVPP MDP strategy

This strategy implements the Markov decision process described in section IV to coordinate the operation of the actors in the CVPP in the day-ahead market. We simulate an operation case where the aggregator knows prediction and used them to optimally schedule the storage operation for error management and in the day-ahead market for additional revenues. It constitutes an upper bound of efficiency of our method proposal.

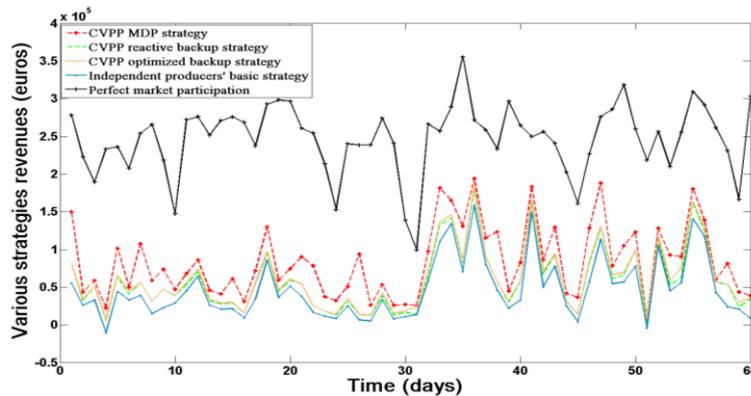


Fig. 6. Various independent and CVPP strategies comparison.

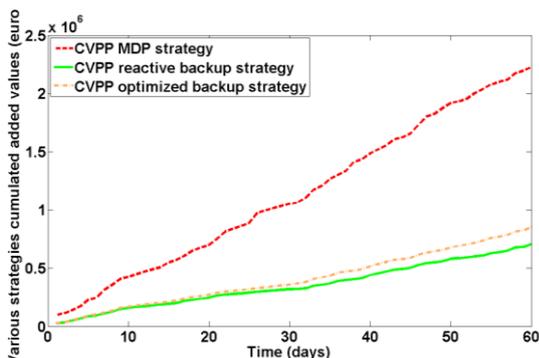


Fig. 7. CVPP strategies comparison on cumulated revenues.

B. Analysis of Results

Fig. 6 depicts the results of executing the different management strategies in terms of revenues. The black curve represents the first case, when producers in the CVPP perfectly predict their generation amounts; it represents the maximal amount of revenues the CVPP can acquire on the market, if perfect predictions tools were used for the day-ahead market participation. When no error management is performed, as producers participate independently to the market (case 2 illustrated by the blue curve), the revenues decrease significantly – inducing the need of a coordination solution. In a CVPP configuration with the storage used as a

backup reactive solution (case 3 illustrated by the green dashed line), the cooperative operation results in an immediate revenues increase. *The optimization of this backup strategy the delivery day*, as proposed in the literature gives better results (depicted by the orange line) than the reactive strategy (even if they are not significant, because of the substantial amount of storage capacities available to be used in a reactive way). *The CVPP MDP strategy* results are bounded by these latter results (our method proposal always outperforms the state of the art solution) and the red line (this upper red line depicts the best results that can be obtained with our method, whatever uncertainty information is available to the storage device for its operation optimization). Fig. 7 shows the cumulated revenues of the three CVPP management strategies over the computation time. Our strategy outperforms the other two CVPP strategies. As shown in Fig. 7, by the end of the simulation period the MDP CVPP strategy can increase the revenues of around 38% (increase of ~1.4Meuros) compared to the optimized backup strategy.

It has to be noted that although it outperforms the optimized backup strategy, the performance of the CVPP MDP strategy we propose is highly dependent on the quality of the generation errors distribution. The quality of this information is a challenging issue to obtain optimal market participation results.

VI. CONCLUSION AND PERSPECTIVES

Simulations results presented in the previous section showed the performance of our proposal for CVPP participation to the day-ahead market. We presented at the beginning of this paper the general context of our work and the physical network it addresses. We highlighted that the literature in renewable market participation management lacks solutions for the particular issue addressed, a CVPP (with substantial controllable resources available) participation into the day-ahead market. We formalize this issue and show how Markov decision processes constitutes a relevant solution approach to resolve the cooperative market participation of a coalition of RES actors and controllable ones.

We proposed a classical MDP model and showed that our problem lies in the particular context of partially observable MDP. We took advantage of information state theory to recast the POMDP into a classical MDP, solvable using classical resolution algorithms developed in the literature.

Then we simulated its results through different strategies comparisons and showed how it outperforms the literature management solutions (up to 38% better, depending of the uncertainty information available to the CVPP).

Nevertheless, this work can be considerably improved by an addition of the intra-day market participation to the day-ahead participation of the CVPP. Indeed, to optimize the near real-time operation of a CVPP allows increasing revenues as the information available is more accurate. This real-time optimization makes possible to balance residual errors (thus market penalties) from deviations between contracts made and effective generation amounts. Moreover, several market participation aspects which are not the core of our work are not considered in the study, for instance,

obtaining the uncertainty information associated to day-ahead predictions, the prediction of market clearing prices, etc. Nevertheless, these subjects have been widely studied in the literature.

The main pitfall of our proposal is that producers are constrained to share sensitive information with the aggregator, i.e. the distribution of generation's errors and storage market participation strategies. Besides the producers do not intervene in the decision making process, it is totally centralized by the aggregator. In order to increase the acceptability of our solution, those issues must be solved. The most promising research direction that we are currently exploring is to model the decision making mechanism in a distributed manner. The key idea is to maintain the confidentiality of the sensitive information and give the control of the decision making process to the producers. This involves some adaptations of our work and a new formalization as a distributed coordination problem. We are currently modeling and implementing this new approach by using the multi-agent systems paradigm and the recent research results on distributed Markov decision process.

ACKNOWLEDGMENT

This work was done under the partial funding of French National Research agency (ANR) in the frame of the WINPOWER project (WINPOWER-ANR-10-SEGI-016). Authors would like to thank Robin Girard and ARMINES for providing the wind power data used for the experiments.

REFERENCES

- [1] European Commission, "Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions - 20 20 by 2020 - Europe's climate change opportunity," Tech. Rep., 2012.
- [2] R. Piwko, D. Osborn, R. Gramlich, G. Jordan, D. Hawkins, and K. Porter, "Wind energy delivery issues [transmission planning and competitive electricity market operation]," *IEEE Power and Energy Magazine*, vol. 3, no. 6, pp. 47-56, 2005.
- [3] E. W. E. Association, "The European offshore wind industry - Key trends and statistics 2013," European Wind Energy Association, 2013.
- [4] Bloomberg. (May 5, 2014). U.S. wind power blows new records. Again And Again. [Online]. Available: <http://www.bloomberg.com/news/2014-04-07/u-s-wind-power-blows-new-records-again-and-again.html>.
- [5] G. Damm and R. Girard, "Anglais WINPOWER Project - Wind energy integration by DC network," presented at Anglais EWEA OFFSHORE 2011, Amsterdam, Pays-Bas, 2011.
- [6] T. Hammons, "Integrating renewable energy sources into European grids," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 8, pp. 462-475, 2008.
- [7] G. Chalkiadakis, V. Robu, R. Kota, A. Rogers, and N. R. Jennings, "Cooperatives of distributed energy resources for efficient virtual power plants," in *Proc. the 10th Int. Conf. on Autonomous Agents and Multiagent Systems, International Foundation for Autonomous Agents and Multiagent Systems*, 2011, vol. 2, pp. 787-794.
- [8] R. Djabali, J. Hoeksema, and Y. Langer, "Cosmos description cwe market coupling algorithm," Public Report, Cosmos Development Group, 2010.
- [9] M. Moreno, M. Bueno, and J. Usaola, "Evaluating risk-constrained bidding strategies in adjustment spot markets for wind power producers," *International Journal of Electrical Power & Energy Systems*, vol. 43, no. 1, pp. 703-711, 2012.
- [10] G. Bathurst and G. Strbac, "Value of combining energy storage and wind in short-term energy and balancing markets," *Electric Power Systems Research*, vol. 67, no. 1, pp. 1-8, 2003.
- [11] D. Pudjianto, C. Ramsay, and G. Strbac, "Virtual power plant and system integration of distributed energy resources," *Renewable Power Generation, IET*, vol. 1, no. 1, pp. 10-16, 2007.

- [12] J. Usaola and J. Angarita, "Bidding wind energy under uncertainty," in *Proc. IEEE International Conference on Clean Electrical Power*, 2007, pp. 754-759.
- [13] F. Bourry, "Management of uncertainties related to renewable generation participating in electricity markets," Ph.D. dissertation, MINES ParisTech - École nationale supérieure des mines de Paris, 2009.
- [14] M. He, S. Murugesan, and J. Zhang, "Multiple timescale dispatch and scheduling for stochastic reliability in smart grids with wind generation integration," in *Proc. 2011 IEEE INFOCOM*, 2011, pp. 461-465.
- [15] M. Vasirani, R. Kota, R. Cavalcante, S. Ossowski, and N. Jennings, "An agent-based approach to virtual power plants of wind power generators and electric vehicles," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1314-1322, 2013.
- [16] V. Robu, R. Kota, G. Chalkiadakis, A. Rogers, and N. Jennings, "Cooperative virtual power plant formation using scoring rules," in *Proc. the 11th International Conference on Autonomous Agents and Multiagent Systems*, 2012, vol. 3, pp. 1165-1166.
- [17] S. Rohjans, M. Uslar, R. Bleiker, J. González, M. Specht, T. Suding, and T. Weidelt, "Survey of smart grid standardization studies and recommendations," in *Proc. 2010 First IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2010, pp. 583-588.
- [18] L. Costa, F. Bourry, and G. Kariniotakis *et al.*, "Stochastic optimization techniques for the optimal combination of wind power generation and energy storage in a market environment," in *Proc. European Wind Energy Conference & Exhibition EWEC 2008*, 2008.
- [19] P. Pinson, "Estimation of the uncertainty in wind power forecasting," Ph.D. dissertation, Mines, 2006.
- [20] R. S. Sutton and A. G. Barto, *Introduction to Reinforcement Learning*, MIT Press, 1998.
- [21] S. Adlakha, S. Lall, and A. Goldsmith, "Networked markov decision processes with delays," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 1013-1018, 2012.
- [22] J. R. Gilbert, C. Moler, and R. Schreiber, "Sparse matrices in matlab: Design and implementation," *SIAM Journal on Matrix Analysis and Applications*, vol. 13, no. 1, pp. 333-356, 1992.
- [23] P. Pinson, R. Girard, and P. McSharry, "Stochastic power generation," Technical University of Denmark, *Tech. Rep.*, 2013.
- [24] EPEX SPOT SE. (May 5, 2014). Market data auction. [Online]. Available: https://www.epexspot.com/en/market_data/auction.
- [25] RTE. Prix du règlement des écarts [Online]. Available: https://clients.rte-france.com/lang/fr/clients_consommateurs/vie/mecanisme/jour/prix.jsp.



Créo Paniah is a Ph.D. student at CEA-LIST Laboratory, Saclay, France since 2011.

He received a MSc degree in artificial intelligence from Paris-Dauphine University, Paris, France in 2011. His research interests are in distributed artificial intelligence and decision-making under uncertainty. He currently works on applying those solutions for the control of networked renewable energy resources and storage

devices and their participation to electricity markets.

Cécile Herpson received a Ph.D. degree in computer science from the University Pierre and Marie Curie, France, 2012.

He works since then as a researcher at CEA-LIST Laboratory, Saclay, France. He joined the CEA to pursue his work in the field of distributed energy management and optimization for smart grids. His research interests are in the area of artificial intelligence and include agents' autonomy, especially decision-making under uncertainty, multi-agent coordination and heuristic search under real-life hypotheses.



Javier Gil-Quijano was born on July 29th, 1976 in Bucaramanga, Colombia.

He is a PhD holder in computer sciences of University Pierre et Marie Curie, Paris, France since 2007 and received a MSc degree in artificial intelligence of University Los Andes, Bogota, Colombia in 2002. He works since 2010 as a research engineer at CEA-LIST Laboratory, Saclay, France. He works on multi-agent modelling and control of technical complex systems. Throughout its different work experiences he has turned his interests into modeling of biological complex systems and modeling of human behaviors for energy usages. Currently at CEA LIST, he is a technical-leader on several projects involving the use of multi-agent systems for distributed monitoring of complex systems (buildings, distribution networks) and the distributed energy management and optimization at different scales (district management systems, virtual centrals for wind power generation).