Active network management for electrical distribution systems: problem formulation and benchmark

Quentin Gemine, Damien Ernst, Bertrand Cornélusse

Montefiore Institute, Department of Electrical Engineering and Computer Science
University of Liège, 4000 Liège, Belgium
{qgemine,dernst,bertrand.cornelusse}@ulg.ac.be

Abstract

In order to operate an electrical distribution network in a secure and cost-efficient way, it is necessary, due to the rise of renewable energy-based distributed generation, to develop Active Network Management (ANM) strategies. These strategies rely on short-term policies that control the power injected by generators and/or taken off by loads in order to avoid congestion or voltage problems. While simple ANM strategies would curtail the production of generators, more advanced ones would move the consumption of loads to relevant time periods to maximize the potential of renewable energy sources. However, such advanced strategies imply solving large-scale optimal sequential decision-making problems under uncertainty, something that is understandably complicated. In order to promote the development of computational techniques for active network management, we detail a generic procedure for formulating ANM decision problems as Markov decision processes. We also specify it to a 75-bus distribution network. The resulting test instance is available at http://www.montefiore.ulg.ac.be/~anm/. It can be used as a test bed for comparing existing computational techniques, as well as for developing new ones. A solution technique that consists in an approximate multistage program is also illustrated on the test instance.

Index terms — Active network management, electric distribution network, flexibility services, renewable energy, optimal sequential decision-making under uncertainty, large system

1 Introduction

In Europe, the 20/20/20 objectives of the European Commission and the consequent financial incentives established by local governments are currently driving the growth of electricity generation from renewable energy sources [1]. A substantial part of the investments lies in the distribution networks (DNs) and consists of the installation of units that depend on wind or sun as a primary energy source. The significant increase of the number of these distributed generators (DGs) undermines the fit and forget doctrine, which has dominated the planning and the operation of DNs up to this point. This doctrine was developed when DNs had the sole mission of delivering the energy coming from the transmission network (TN) to the consumers. With this approach, adequate investments in network components (i.e., lines, cables, transformers, etc.) must constantly be made to avoid congestion and voltage problems, without requiring continuous monitoring and control of the power flows or voltages. To that end, network planning is done with respect to a set of critical scenarios consisting of production and demand levels, in
order to always ensure sufficient operational margins. Nevertheless, with the rapid growth of
DGs, the preservation of such conservative margins comes at continuously increasing network
reinforcement costs.

In order to avoid these potentially prohibitive costs, *active network management (ANM)*
strategies have recently been proposed as alternatives to the fit and forget approach. The
principle of ANM is to address congestion and voltage issues via short-term decision-making
policies \[2\]. Frequently, ANM schemes maintain the system within operational limits in quasi
real-time by relying on the curtailment of wind or solar generation \[3, 4, 5\]. Curtailment of
renewable energy may, however, be very controversial from an environmental point of view and
should probably be considered as a last resort. In that mindset, it would be worth investigating
ANM schemes that could also exploit the flexibility of the loads, so as to decrease the reliance
on generation curtailment. It is worth noting that exploiting flexible loads within an ANM
scheme comes with several challenges. One such challenge is that modulating a flexible load
at one instant is often going to influence its modulation range at subsequent instants. This is
because flexible loads are often constrained to consume a specific amount of energy over a certain
duration. In this context, it is therefore important for a distribution system operator (DSO) not
to take myopic decisions, but rather to make decisions by planning over a relevant time span
\[6\]. The uncertainty of future power injections from DGs relying on natural phenomena (wind,
sun, etc.), as well as the uncertainty of the power consumption of the loads, would necessarily
imply that the DSO should implement an ANM scheme able to plan adequate control actions
under uncertainty.

Many authors have already attempted to provide solutions to these operational planning
problems. Since they rely on different formalisms, it is difficult for one author to rebuild on top
of another’s work. However, a thorough reading of these works suggests that it is possible to have
a common decision paradigm. One reason for this is that these formalisms can be considered
as an extension of the *optimal power flow (OPF)* problem \[7\]. More specifically, they can be
assimilated to sequential decision-making problems where, at each time step, constraints that
are similar to those used for defining an OPF problem are met. In this article, we aim to
reduce the fragmentation of research efforts and facilitate the comparison of solution techniques
that have been developed. To that end, we first propose a generic formulation of ANM related
decision-making problems. More specifically, we detail a procedure to state these problems as
Markov Decision Processes (MDP), where the system dynamics describe the evolution of the
electrical network and devices, while the action space encompasses the control actions that are
available to the DSO. Afterwards, we instantiate this procedure to a 75-bus network and use
the elements of the resulting MDP to build a simulator of the system, which is available at
http://www.montefiore.ulg.ac.be/~anm/.

The rest of this paper is structured as follows. The ANM problem of a DN is described
in Section 2, where the electrical model and the network operation details are explained. In
Section 3, a formulation of the problem as a Markov decision process is introduced. A cast of
this formulation as a mixed-integer and non-linear mathematical program is then presented in
Section 4, along with a literature review of the recent research in this field and an identification
of potential solution techniques that could be applied to the ANM problem. The benchmark is
described in Section 5 and an illustrative solution technique is finally presented in Section 6.

## 2 Problem Description

We consider the problem faced by a DSO willing to plan the operation of its network over
time, while ensuring that operational constraints of its infrastructure are not violated. This
amounts to determining the optimal operation of a set \(\mathcal{D}\) of electrical devices over time. We
describe the evolution of the system by a discrete-time process having a time horizon $T$. This time horizon represents the number of periods used for the planning phase. The real time between two periods is chosen equal to 15 minutes, by analogy with the typical duration of a market period. We neglect the fast dynamics of the system and assume that the power injection and withdrawal levels are constant within a single period. The control actions in this section are aimed to directly impact these power levels. These actions can potentially introduce time-coupling effects.

2.1 Network infrastructure

The network infrastructure of the DSO is the set of electrical elements that compose its network: the nodes and the links (i.e., lines, cables, and transformers). Every node $n \in \mathcal{N}$ is characterized by a nominal voltage level $V_{n}^{(nom)}$ and every link $l \in \mathcal{L}$ connects a pair of nodes $n, n' \in \mathcal{N}$ and is modeled by its $\pi$-model, composed of three admittances $Y_{nn}, Y_{ne}, Y_{n'n'}$ (see Figure 1). We denote by $V_{n,t}$ the complex voltage at node $n \in \mathcal{N}$ and time $t \in \mathcal{T}$ and by $|V_{n,t}|, e_{n,t}$, and $f_{n,t}$ its magnitude, real part, and imaginary part, respectively.

In particular, we focus on the operation of the medium-voltage (MV) network of the DSO, which usually includes its whole infrastructure, except the elements that are closest to the residential consumers. These elements constitute the low-voltage (LV) network.

2.2 Operational limits

The operational limits represent a set of voltage constraints at the nodes and of current constraints in the links. These constraints have to be respected to avoid compromising the operation of the network. The voltage constraints model that the voltage magnitude $|V_{n,t}|$ at node $n$ cannot deviate too much from its nominal value $V_{n}^{(nom)}$. If this occurred, it could, for example, damage the electrical devices that are connected to this node. So, we can write:

$$\forall (t, n) \in \mathcal{T} \times \mathcal{N} : V_{n} \leq |V_{n,t}| \leq \overline{V}_{n},$$

where $\underline{\cdot}$ and $\overline{\cdot}$ denote lower and upper bounds, respectively.

Concerning the current limits, or thermal limits, they model a congestion phenomenon that occurs in the links when the magnitude of the current that goes through them is too large. A protection system usually disconnects the affected links in the case of limit violations, which can potentially lead to the loss of some parts of the network. These constraints can be expressed as:

$$\forall (t, l) \in \mathcal{T} \times \mathcal{L} : |I_{l,t}| \leq \overline{I}_{l},$$

where $I_{l,t}$ is the current in link $l$ at time $t$ and $|I_{l,t}|$ is its magnitude.

Figure 1: $\pi$-model of a link that connects nodes $i$ and $j$. 
2.3 Electrical devices

The set $D$ of electrical devices is made of elements that are connected to nodes $n \in N$ of the network and that exchange electrical energy with it. They can be classified into two distinct subsets:

- the set $C \subset D$ of loads, which withdraw power from the network by consuming electrical energy;
- the set $G \subset D$ of generators, which inject power into the network by producing electrical energy.

Within each subset, we also distinguish two types of device models. The first ones represent individual injection and withdrawal points. They can model certain types of DGs or consumers that are directly connected to the MV grid (e.g., wind farms, some companies and factories, etc.). The others model an aggregate set of devices that are assimilated to a single connection point at the MV grid (e.g., residential consumers and solar panels). Correspondences between some physical elements and their device model are illustrated in Figure 2. At node 3, a set of residential loads and a set of distributed solar units have been aggregated into a single load model and a single generator model.

For each time period $t \in T$, an active power injection value $P_{d,t}$ and a reactive power injection value $Q_{d,t}$ are associated with every device $d \in D$. Even if only active power is useful to the devices, the capacitive and inductive properties of the electrical system inevitably induce reactive power flows. We assume that all the devices are operating at a constant power factor, so the ratio between reactive and active powers, denoted as $\tan \phi_d$, remains unchanged:

$$\forall (t,d) \in T \times D : \frac{Q_{d,t}}{P_{d,t}} = \tan \phi_d.$$ (3)

Conventionally, we consider that a power injection is positive if it supplies the network (generation) and negative if it takes energy from the network (consumption).
2.4 Electrical state

Determining the electrical state of the network is equivalent to computing the real and imaginary parts of the voltage at the nodes (or the magnitude and phase, depending on the coordinate system). The admittance values of the links and the power injection values of the devices, together with the power flow equations, allow for the computation of this state. We can express, \( \forall n \in \mathcal{N} \), the power flow equations as:

\[
\sum_{d \in \mathcal{D}(n)} P_{d,t} = \sum_{l \in \mathcal{L}} \left( (g_{nn}^{(l)} + g_{nn'}^{(l)}) (e_{n,t}^2 + f_{n,t}^2) - g_{nn'}^{(l)} (e_{n,t} e_{n',t} + f_{n,t} f_{n',t}) + b_{nn'}^{(l)} (e_{n,t} f_{n',t} - f_{n,t} e_{n',t}) \right),
\]

(4)

\[
\sum_{d \in \mathcal{D}(n)} Q_{d,t} = \sum_{l \in \mathcal{L}} \left( b_{nn}^{(l)} (e_{n,t} e_{n',t} + f_{n,t} f_{n',t}) + g_{nn'}^{(l)} (e_{n,t} f_{n',t} - f_{n,t} e_{n',t}) - (b_{nn}^{(l)} + b_{nn'}^{(l)}) (e_{n,t}^2 + f_{n,t}^2) \right),
\]

(5)

where \( g \) and \( b \) are the real and imaginary parts of the admittances \( Y \), respectively, and where \( \mathcal{D}(n) \) denotes the set of devices that are connected to node \( n \).

Once the voltages at the different nodes are known, the magnitude of the current \( I_{l,t} \) in link \( l \in \mathcal{L} \), which connects nodes \( n \) and \( n' \), is obtained using the following equation:

\[
I_{l,t}^2 = (g_{nn'} + b_{nn'})^2 ((e_n - e_{n'})^2 + (f_n - f_{n'})^2), \quad \forall l \in \mathcal{L}.
\]

(6)

From now on, we use the following shorthand notations:

\[
\forall n \in \mathcal{N} : g_n(e, f, P_{n,t}, Q_{n,t}) = 0,
\]

(7)

\[
\forall n \in \mathcal{N} : h_n(e, f, P_{n,t}, Q_{n,t}) = 0,
\]

(8)

\[
\forall l \in \mathcal{L} : i_l(e, f) = 0,
\]

(9)

to refer to equations (4), (5) and (6), respectively, where \( e \) and \( f \) are the vectors of the real and imaginary parts of the voltage at the buses and with

\[
P_{n,t} = \sum_{d \in \mathcal{D}(n)} P_{d,t},
\]

(10)

\[
Q_{n,t} = \sum_{d \in \mathcal{D}(n)} Q_{d,t}.
\]

(11)

2.5 Control actions

We now describe two control means of the system, the modulation of generation and the modulation of the demand, as well as one of the possible interaction schemes between the actors of this system. In practice, other approaches exist to control the system, such as modulating the tariff signal(s), acting on the topology of the network, or using distributed storage sources. For a single control method, there are also several ways to organize the interactions between the actors.

We consider that the DSO can use control actions that modify the operation of the devices that are connected to its network. For each device belonging to the set \( \mathcal{G} \subset \mathcal{D} \) of DGs, it can impose a curtailment instruction, i.e. an upper limit on the production level of the DG. The impact of such a control action on the power injection of a DG is illustrated in Figure 3. This request can be performed up to the time period immediately preceding the one concerned by the curtailment and it is acquired in exchange for a fee. This fee is used to compensate the producer for the energy that could not be sold during modulation periods. From the alternative
payment structures outlined in [8], we adopt a scheme where the compensation is defined by the day-ahead electricity market price for this period, for each MWh that could not be produced. If the curtailment decision is taken far in advance by the DSO, the market price might not yet be known because, in Europe, the market price for every hour of a day is known at noon the day before. For the sake of simplicity, we assume that this price is known.

We also consider that the DSO can modify the consumption of the flexible loads. These loads constitute a subset \( \mathcal{F} \) out of the whole set of the loads \( \mathcal{C} \subset \mathcal{D} \) of the network. An activation fee is associated with this control means and flexible loads can be notified of activation up to the time immediately preceding the start of the service. Once the activation is performed at time \( t_0 \), the consumption of the flexible load \( d \) is modified by a certain value during \( T_d \) periods. For each of these modulation periods \( t \in [t_0 + 1; t_0 + T_d] \), this value is defined by the modulation function \( \Delta P_d(t - t_0) \). An example of modulation function and its influence over the consumption curve are presented in Figure 4 and Figure 5, respectively.

Loads cannot be modulated in an arbitrary way. There are indeed constraints to be imposed on the modulation signal, which are inherited from the flexibility sources of the loads, such as an inner storage capacity (e.g., electric heater, refrigerator, water pump) or a process that can be scheduled with some flexibility (e.g., industrial production line, dishwasher, washing machine). In any case, we will always consider that the modulation signal \( \Delta P_d \) has to respect the following conditions:

- A downward modulation is followed by an increase of the consumption, and conversely.
The integral of the modulation signal is null in order to ensure that the consumption is only shifted, not modified.

The $\Delta P(t-t_0)$ curve of Figure 4 respects these conditions. The modulation sign for the second part of the time interval $[1;T_d]$ is opposite than it was in the first part and the integral is null ($\Delta E_- + \Delta E_+ = 0$).

3 Optimal sequential decision making formulation

In Section 2, we described the very basic elements of distribution network operation. We now formalize the problem of computing the right control actions as an optimal sequential decision-making problem. The sequentiality is induced by the modulation service that is provided by flexible loads. Indeed, if such a service is activated at time $t_0$ for a flexible load $d$, the action will influence the system for the set of periods $[t_0+1,t_0+T_d]$. In addition to being a sequential problem, it is also stochastic, because the evolution of the system and the outcome of control actions are affected by several uncertain factors. These factors are the wind speed, the level of solar irradiance, and the consumption level of the loads. In this section, we model this problem as a first-order Markov decision process with mixed integer sets of states and actions. We thus consider that the transition probabilities of the stochastic elements of the system from a period $t$ to a period $t+1$ only depend on the state at time $t$. Using a relevant model of the system state, we think that this stochastic process represents the uncertainty of the system with satisfactory accuracy. However, we stress that it would be interesting to perform a detailed study that would aim at quantifying the approximation error induced by the Markov hypothesis with respect to the most detailed models of the literature. Finally, the notion of optimality is defined using a reward function that associates an immediate reward (or score) to every transition of the system. The better the cumulated reward over a system trajectory, the better the sequence of control actions for this trajectory.

3.1 System state

Thanks to equations (4), (5) and (6), the value of the electrical quantities can be deduced from the power injections of the devices. If the representation of these injections requires the knowledge of the individual consumption of every load, it is possible to obtain the production
of DGs given the power level of their energy source (i.e. the wind speed or the level of solar irradiance). This brings us to defining a first state set $S^{(1)}$ such that the vectors $s^{(1)}_t \in S^{(1)}$ are defined by

$$s^{(1)}_t = (P_{1,t}, \ldots, P_{|G|,t}, i_{rt}, v_t)$$

and contain, at time $t \in T$, the minimal information that is required to know the power injections of the whole set $D$ of devices. The $i_{rt}$ and $v_t$ components represent the level of solar irradiance and the wind speed, respectively. If, for the sake of simplicity, the generators that we consider are only the solar and wind ones, others could easily be integrated by increasing the dimension of $S^{(1)}$. Similarly, the Markovian decision process could be extended to an order $n$ by integrating to this vector the history, over $n$ periods, of the power injections of the loads, the level of solar irradiance, and the wind speed. The model that we consider corresponds to a one-period-long history (first order).

Given that the DSO can influence the injection values of the devices, we introduce a second set $S^{(2)}$ that represents the modulation instructions. The vectors $s^{(2)}_t \in S^{(2)}$ are defined as

$$s^{(2)}_t = (P_{1,t}, \ldots, P_{|G|,t}, s^{(f)}_{1,t}, \ldots, s^{(f)}_{|F|,t})$$

and contain, at time $t \in T$, the upper limits in active power injection of the DGs $g \in G$, as authorized by the DSO, and the indicators $s^{(f)}_{d,t}$ of the flexibility service state of the loads $d \in F$, defined as

$$s^{(f)}_{d,t} = \begin{cases} 
\text{number of active periods left} & \text{if service is active} \\
0 & \text{if service is inactive.}
\end{cases}$$

Using the two sets $S^{(1)}$ and $S^{(2)}$, we can finally define the set of states $S$ of the system as

$$S = S^{(1)} \times S^{(2)} \times [1, 96],$$

where the component that belongs to the integer interval $[1, 96]$ identifies the quarter of an hour in the day to which the state corresponds. This component allows the definition of several transition functions that rely on this information. In the following, we denote this component by $q_t$.

### 3.2 Control actions

The control means that are available to the DSO to control the system are modeled by the set $A_s$ of control actions. This set depends on the state $s_t$ of the system because it is not possible to activate the flexibility service of a load if it is already active. The components of vectors $a_t \in A_s$ are defined by

$$a_t = (\tilde{p}_t, a^{(f)}_t),$$

with $\tilde{p}_t \in \mathbb{R}^{|G|}$ such that $\tilde{p}_{g,t}$ indicates the maximum level of active power injection for period $t+1$ and for each of the generators $g \in G$. On the other hand, the vector $a^{(f)}_t$ represents the activation indicators of the flexibility services of the loads $d \in F$, where each component $a^{(f)}_{d,t}$ belongs to $A^{(f)}_{d,s}$, which is defined as

$$A^{(f)}_{d,s} = \begin{cases} 
\{0, 1\} & \text{if } s^{(f)}_{d,t} = 0 \\
\{0\} & \text{if } s^{(f)}_{d,t} > 0,
\end{cases}$$

to ensure that a load which is already active is not activated.
By using such a representation of the control actions, we consider that a curtailment or flexibility activation action targeting a period \( t \) must always be performed at the period \( t - 1 \) that precedes it. As already described in Section 2, these control actions can indeed be notified by the DSO up to the period immediately preceding their implementation. Even if it were possible to consider a model of the action vectors that integrates components implying an effective control delayed of several periods, it would induce even larger time-coupling effects, while not improving the extent of control of the DSO. Indeed, for a period \( t \), the last recourse of the DSO is the period \( t - 1 \) and the interaction model considered is such that the cost associated with an action is independent of the delay between its notification and when it becomes effective.

### 3.3 Transition function

The system evolution from a state \( s_t \) to a state \( s_{t+1} \) is described by the transition function \( f \). The new state \( s_{t+1} \) depends, in addition to the preceding state, on the control actions \( a_t \) of the DSO and of the realization of the stochastic processes, modeled as Markov processes. More specifically, we have

\[
f : S \times A_s \times W \to S,
\]

where \( W \) is the set of possible realizations of a random process. The general evolution of the system is thus governed by relation

\[
s_{t+1} = f(s_t, a_t, w_t),
\]

(12)

where \( w_t \in W \) and such that it follows a conditional probability law \( p_W(\cdot|s_t) \). In order to define this function in more detail, we now describe the various elements that constitute it.

#### 3.3.1 Load consumption

The uncertainty about the behavior of consumers inevitably leads to uncertainty about the power level they draw from the network. However, over a one-day horizon, some trends can be observed. For example, consumption peaks arise in the early morning and in the evening for residential consumers, but at levels that fluctuate from one day to another and among consumers. We model the evolution of the consumption of each load \( d \in C \) by

\[
P_{d,t+1} = f_d(P_{d,t}, q_t, w_{d,t}),
\]

(13)

where \( w_{d,t} \) is a component of \( w_t \sim p_W(\cdot|s_t) \). The dependency of functions \( f_d \) to the quarter of an hour in the day \( q_t \) allows for capturing the daily trends of the process. Given the hypothesis of a constant power factor for the devices, the reactive power consumption can directly be deduced from \( P_{d,t+1} \):

\[
Q_{d,t+1} = \tan \phi_d \cdot P_{d,t+1}.
\]

In Section 5, we describe a possible procedure to model the evolution of the consumption of an aggregated set of residential consumers using relation (13).

#### 3.3.2 Wind speed and power level of wind generators

The uncertainty about the production level of wind turbines is inherited from the uncertainty about the wind speed. The Markov process that we consider governs the wind speed, which is assumed to be uniform across the network. The production level of the wind generators is then
obtained by using a deterministic function that depends on the wind speed realization, this function is the power curve of the considered generator. We can formalize this phenomenon as:

$$\begin{align*}
v_{t+1} &= f_v(u_t, q_t, w_t^{(v)}), \\
P_{g,t+1} &= \eta_g(v_{t+1}), \forall g \in \text{wind generators} \subset G,
\end{align*}$$

such that \( w_t^{(v)} \) is a component of \( w_t \sim p_W(\cdot | s_t) \) and where \( \eta_g \) is the power curve of generator \( g \). A typical example of power curve \( \eta_g(v) \) is illustrated in Figure 6. Like loads, the production of reactive power is obtained using:

$$Q_{g,t+1} = \tan \phi_g \cdot P_{g,t+1}.$$  

A possible approach to determine \( f_v \) from a set of measurements is described in Section 5.

### 3.3.3 Solar irradiance and photovoltaic production

Like wind generators, the photovoltaic generators inherit their uncertainty in production level from the uncertainty associated with their energy source. This source is represented by the level of solar irradiance, which is the power level of the incident solar energy per \( m^2 \). The irradiance level is the stochastic process that we model, while the production level is obtained by a deterministic function of the irradiance and of the surface of photovoltaic panels. This function is simpler than the power curve of wind generators and is defined as

$$P_{g,t} = \eta_g \cdot \text{surf}_g \cdot i_{r,t},$$

where \( \eta_g \) is the efficiency factor of the panels, assumed constant and with a typical value around 15%, while \( \text{surf}_g \) is the surface of the panels in \( m^2 \) and is specific to each photovoltaic generator. The irradiance level is denoted by \( i_{r,t} \) and the whole phenomenon is modeled by the following Markov process:

$$\begin{align*}
i_{r,t+1} &= f_{ir}(i_{r,t}, q_t, w_t^{(ir)}), \\
P_{g,t+1} &= \eta_g \cdot \text{surf}_g \cdot i_{r,t+1}, \forall g \in \text{solar generators} \subset G,
\end{align*}$$

such that \( w_t^{(v)} \) is a component of \( w_t \sim p_W(\cdot | s_t) \). The technique used in Section 5 to build \( f_{ir} \) from a dataset is similar to the one for the wind speed case.
3.3.4 Impact of control actions

The stochastic processes that we described govern the evolution of the state $s_t^{(1)} \in S^{(1)}$ of the consumption of loads (flexibility services excluded) and of the power level of energy sources of DGs. The following transition laws define the evolution of the components of the state of modulation instructions $s_t^{(2)} \in S^{(2)}$ by integrating the control actions of the DSO:

\[
\forall g \in G : \quad \overline{P}_{g,t+1} = P_{g,t}, \quad (18)
\]
\[
\forall d \in F : \quad s_{d,t+1}^{(f)} = \max(s_{d,t}^{(f)} - 1; 0) + a_{d,t}^{(f)} T_d, \quad (19)
\]
\[
\forall d \in F : \quad \Delta P_{d,t+1} = \begin{cases} \Delta P_{d,T_d} - s_{d,t+1}^{(f)} + 1 & \text{if } s_{d,t+1}^{(f)} > 0 \\ 0 & \text{if } s_{d,t+1}^{(f)} = 0 \end{cases}, \quad (20)
\]

From vectors $s_t^{(1)}$ and $s_t^{(2)}$, we can determine, for each node $n \in \mathcal{N}$, the active and reactive power injections and thus obtain the value of the electrical quantities of the network:

\[
P_{\sum,n,t} = \sum_{g \in G(n)} \min(\overline{P}_{g,t}; P_{g,t}) + \sum_{d \in C(n) \setminus F(n)} (P_{d,t} + \Delta P_{d,t}), \quad (21)
\]
\[
Q_{\sum,n,t} = \sum_{g \in G(n)} \min(\tan \phi_g \overline{P}_{g,t}; Q_{g,t}) + \sum_{d \in C(n) \setminus F(n)} (Q_{d,t} + \tan \phi_d \Delta P_{d,t}), \quad (22)
\]
\[
0 = g_n(e, f, P_{\sum,n,t}^{\sum}, Q_{\sum,n,t}^{\sum}), \quad (23)
\]
\[
0 = h_n(e, f, P_{\sum,n,t}^{\sum}, Q_{\sum,n,t}^{\sum}), \quad (24)
\]
and, in each link $l \in \mathcal{L}$, we have:

\[
i_l(e, f) = 0. \quad (25)
\]

3.4 Reward function and goal

In order to evaluate the performance of a policy, we first specify the reward function $r : S \times A_s \times S \to \mathbb{R}$, which associates an instantaneous reward for each transition of the system from a period $t$ to a period $t + 1$:

\[
r(s_t, a_t, s_{t+1}) = -\sum_{g \in G} \max\{0, \frac{P_{g,t+1} - \overline{P}_{g,t+1}}{4}\} C_{\text{curt}}^{\text{d}}(q_{t+1}) - \sum_{d \in F} a_{d,t}^{(f)} C_{d}^{\text{flex}} - \Phi(s_{t+1}), \quad (26)
\]

where $C_{\text{curt}}^{\text{d}}(q_{t+1})$ is the day-ahead market price for the quarter of an hour $q_{t+1}$ in the day and $C_{d}^{\text{flex}}$ is the activation cost of the flexible loads, specific to each of them. The function $\Phi$ is a barrier function that allows for penalizing a policy that leads the system into a state that violates the operational limits. It is defined as

\[
\Phi(s_{t+1}) = \sum_{n \in \mathcal{N}} \left[ \chi(c_{n,t+1}^2 + f_{n,t+1}^2 - \overline{V}_n^2) + \chi(V_n^2 - e_{n,t+1}^2 - f_{n,t+1}^2) \right] + \sum_{l \in \mathcal{L}} \chi(|I_{l,t+1} - I_l|), \quad (27)
\]
where \( e_{n,t+1}, f_{n,t+1} \) (\( n \in \mathcal{N} \)) and \( I_{l,t+1} \) (\( l \in \mathcal{L} \)) are determined from \( s_{t+1} \) using equations (21)-(25) and
\[
\chi(x) = \begin{cases} 10^5 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}.
\] (28)
The higher the operational costs and the larger the number of violated operational limits, the more negative the reward function.

We can now define the return over \( T \) periods, denoted \( R_T \), as the weighted sum of the rewards that are observed over a system trajectory of \( T \) periods
\[
R_T = \sum_{t=0}^{T-1} \gamma^t r(s_t, a_t, s_{t+1}),
\] (29)
where \( \gamma \in [0; 1[ \) is the discount factor. Given that \( \gamma^t < 1 \) for \( t > 0 \), the further in time the transition from period \( t = 0 \), the less importance is given to the associated reward. Because the operation of a DN must always be ensured, it does not seem relevant to consider returns over a finite number of periods and we introduce the return \( R \) as
\[
R = R_\infty = \lim_{T \to \infty} \sum_{t=0}^{T-1} \gamma^t r(s_t, a_t, s_{t+1}),
\] (30)
that corresponds to the weighted sum of the rewards observed over an infinite trajectory of the system. Given that the costs and penalties have finite values and that the reward function \( r \) is the sum of an infinite number of these costs and penalties, a constant \( C \) exists such that, \( \forall (s_t, a_t, s_{t+1}) \in \mathcal{S} \times \mathcal{A} \times \mathcal{S}, \) we have \( |r(s_t, a_t, s_{t+1})| < C \) and thus
\[
|R| < \lim_{T \to \infty} C \sum_{t=0}^{T-1} \gamma^t = \frac{C}{1 - \gamma}.
\] (31)
It means that even if the return \( R \) is defined as an infinite sum, it converges to a finite value. One can also observe that, because \( s_{t+1} = f(s_t, a_t, w_t) \), a function \( \rho : \mathcal{S} \times \mathcal{A} \times \mathcal{W} \to \mathbb{R} \) exists that aggregates functions \( f \) and \( r \) such that
\[
r(s_t, a_t, s_{t+1}) = \rho(s_t, a_t, w_t),
\] (32)
with \( w_t \sim p_{W,Y}(\cdot|s_t) \). Let \( \pi : \mathcal{S} \to \mathcal{A}_s \) be a policy that associates a control action to each state of the system. We can define, starting from an initial state \( s_0 = s \), the expected return \( R \) of the policy \( \pi \) by
\[
J^\pi(s) = \lim_{T \to \infty} \mathbb{E}_{w_t \sim p_{W,Y}(\cdot|s)} \left\{ \sum_{t=0}^{T-1} \gamma^t \rho(s_t, \pi(s_t), w_t) \right\}.
\] (33)
We denote by \( \Pi \) the space of all the policies \( \pi \). For a DSO, addressing the operational planning problem described in Section 2 is equivalent to determine an optimal policy \( \pi^* \) among all the elements of \( \Pi \), i.e. a policy that satisfies the following condition
\[
J^{\pi^*}(s) \geq J^{\pi}(s), \forall s \in \mathcal{S}, \forall \pi \in \Pi.
\] (34)
It is well known that such a policy satisfies the Bellman equation [9], which can be written
\[
J^{\pi^*}(s) = \max_{a \in \mathcal{A}_s} \mathbb{E}_{w \sim p_{W,Y}(\cdot|s)} \left\{ \rho(s, a, w) + \gamma J^{\pi^*}(f(s, a, w)) \right\}, \forall s \in \mathcal{S}.
\] (35)
If we only take into account the space of stationary policies (i.e. that selects an action independently of time \( t \)), it is without loss of generality comparing to the space of policies \( \Pi' : \mathcal{S} \times \mathcal{T} \to \mathcal{A} \) because the return to be maximized corresponds to an infinite trajectory of the system [10].
4 Solution techniques

In this section, we identify three classes of solution techniques that could be applied to the operational planning problem. The first one is mathematical optimization, a technique for which we also provide a review of the literature concerning the research about solving multi-period OPF. The second approach that we consider is constituted by techniques relying on the dynamic programming framework, such as approximate dynamic programming and reinforcement learning. Finally, simulation-based optimization techniques are discussed. We also refer the reader to [11] for a more complete description of the set of techniques that could be used for solving this problem.

4.1 Mathematical programming

It is possible to address this operational planning problem by solving, for each period $t$, a mathematical program $M_t$ that accounts for the system evolution over a finite time horizon $T_t$, i.e. where $[t, t + T - 1]$ constitutes the set $T_t$ of the considered periods. The control action $a_t$ that is applied at time $t$ is then the action computed by the mathematical program $M_t$ and associated with the first period of $T_t$. The resulting approximate optimal policy $\hat{\pi}^*_{M_t}$ can be formalized as

$$\hat{\pi}^*_{M_t}(s) = \arg \max_{a \in \mathcal{A}_s} \mathbb{E} \left[ \sum_{t'=t}^{t+T-1} \gamma^{t'-t} \rho(s_{t'}, a_{t'}, w_{t'}) \right]$$

subject to

$$s_t = s$$

$$a_t = a$$

$$s_{t'} = f(s_{t'-1}, a_{t'-1}, w_{t'-1}), \quad \forall t' \in T_t \setminus \{t\}$$

$$a_{t'} \in \mathcal{A}_{s_{t'}}, \quad \forall t' \in T_t \setminus \{t + T - 1\},$$

where the shorter the horizon $T$, the higher the approximation error. However, it is not the only source of approximation. First, there is no exact numerical method to solve (36)-(40) without requiring a discrete approximation of the continuous stochastic processes [12]. Then, because of the non-linearity of power-flow equations on the one hand, and the integer variables that model the activation of flexibility services on the other hand, the resulting mathematical problem is very complex to solve. For this reason, it is often required to resort to local optimization techniques and to heuristics, which do not provide any guarantee about the optimality of the solutions.

As mathematical programs can explicitly take the constraints over the optimization variables into account, it is possible to modify the mathematical program to remove penalties from the objective function and to integrate operational limits as constraints. If we also consider a
discretization of the random process as a scenario tree [13], we obtain a policy \( \hat{\pi}^{*}_{M_t} \) defined as

\[
\hat{\pi}^{*}_{M_t}(s) = \arg \min_{a \in A(s)} \min_{\forall k \in K_t \setminus \{0\}} \sum_{k=0}^{[K_t]-1} \left[ P_k \gamma^{D_k} \text{cost}(s_k, a_k, w_k) \right] \\
\text{s.t. } s_0 = s \quad (42) \\
a_0 = a \quad (43) \\
s_k = f(s_{A_k}, a_{A_k}, w_{A_k}), \quad \forall k \in K_t \setminus \{0\} \quad (44) \\
a_{A_k} \in A_{s_k}, \quad \forall k \in K_t \setminus \{0\} \quad (45) \\
s_k \in S^{(ok)}, \quad \forall k \in K_t \setminus \{0\}, \quad (46)
\]

where the following notation is used:

- cost : \( S \times A_s \times W \) defines the modulation and flexibility costs associated with a transition of the system;
- \( K_t \) is the set of nodes of the scenario tree (0 \( \in K_t \) corresponds to the root node);
- \( P_k \) is the probability attached to node \( k \in K_t \);
- \( A_k \) is the closest parent of node \( k \) in the tree;
- \( D_k \) is the depth of node \( k \) in the tree and defined as \( D_k = D_{A_k} + 1 \) with \( D_0 = 0 \);
- \( S^{(ok)} \) represents the subset of the system states that respect operational limits.

4.1.1 Relaxations and recent solving attempts

Optimal power flow problems, although non-convex, have been solved for a long time using local non-linear optimization methods. Interior-point methods are probably the most widespread class of methods dedicated to this problem [14]. If the solution they provide has no guarantee to be globally optimal, then they have been made popular by their convergence speed and their ability to solve problems of large dimensions fairly efficiently. Recently, semidefinite programming (SDP) was successfully applied as a convex relaxation to the OPF problem [15]. The authors report no duality gap on some standard meshed test systems and randomized versions of these test systems. The zero duality gap property was thus observed experimentally on standard test systems, and further research resulted in sufficient conditions. This is the case, for example, if the objective function is convex and monotonically increasing with the active power generation, and if the network has a radial topology [16, 17]. Another approach aiming at global optimality relies on Lagrangian relaxation (LR) [18]. The author also describes a spatial branch and bound (B&B) algorithm to close the gap, should one exist. The ability of both SDP and LR to decrease the optimality gap within a B&B framework was evaluated in [19]. Although SDP appeared to be computationally more attractive, it showed that it could be very challenging to reach a significant gap reduction within reasonable time limits, even for small test systems. Multi-period applications related to energy storage are investigated in [20], where the SDP relaxation of [15] is successfully applied, as their particular application met the conditions of having no duality gap. The authors of [21] argue that extending [19] to a multi-period setting yields an SDP too large for current solvers to solve efficiently and suggest relaxing the time-coupling constraints using LR. However, it ended up being computationally too expensive to make the B&B approach worthwhile. Many papers consider the unit commitment problem
over an AC network, which is an instance of a multi-period OPF with discrete variables. For instance in [22], a generalized Benders decomposition divides the problem into a linear master problem with discrete variables and non-linear multi-period sub-problems. Benders cuts are generated from the sub-problems to tighten the MIP master problem. Finally, [23] focused on trying to solve a problem that is mathematically close to the one we consider and provides more information on related research.

4.2 Dynamic programming

Because the methods derived directly from mathematical programming have computational issues, it is relevant to consider algorithms that rely on the principles of dynamic programming (DP). In particular, approximate dynamic programming [24, 25] targets computing approximate optimal policies for large decision-making problems. These algorithms often rely on iterative schemes where, at each iteration $k$, they learn a function $J_k$ that is supposed to come closer to the function $J^*$ with the number of iterations. When the algorithm is done, the last function learned, which we call $\hat{J}^*$, can then be used to define an approximate optimal policy as follows:

$$\hat{\pi}^*(s) = \arg \max_{a \in A_s} \mathbb{E}_{w \sim p_W(\cdot | s)} \left\{ \rho(s, a, w) + \gamma \hat{J}^*(s') | s' = f(s, a, w) \right\}. \quad (47)$$

4.3 Simulation methods

Direct policy search [26] constitutes an important subset of simulation-based solution techniques. This approach first requires the definition of a space $\Pi_\theta \subset \Pi$ of parametric policies. At each iteration $k$, parameter values $\theta_i (i \in I^{(k)})$ are evaluated by simulating trajectories of the system that are associated with the policies $\pi_{\theta_i}$. The result of these simulations allows the selection of new parameter values $\theta_i (i \in I^{(k+1)})$ for the next iteration. The goal of such an algorithm is to converge as fast as possible towards a parameter value $\hat{\theta}^*$ that defines a good approximate optimal policy $\pi_{\hat{\theta}^*}$.

Another subset of simulation-based methods is the Monte-Carlo tree search technique [27, 28]. At each time-step, this class of algorithms usually relies on the simulation of system trajectories to build, incrementally, a scenario tree that does not have a uniform depth. The previous simulations are exploited to select the nodes of the scenario tree that have to be developed. When the construction of the tree is done, the action that is deemed optimal for the root node of the tree is applied to the system.

5 Test instance

In this section, we describe a test instance of the considered problem. The set of models and parameters that are specific to this instance, as well as documentation for their usage, are accessible at http://www.montefiore.ulg.ac.be/~anm/ as Matlab® code. It has been developed to provide a black-box-type simulator that is quick to set up. The DN on which this instance is based is a generic DN of 75 buses [29] that has a radial topology, which is presented in Figure 7. We bound various electrical devices to the network in such a way that it is possible to gather the nodes of this network into four distinct categories:

- each **residential** node is the connection point of a load that represents a set of residential consumers and of a generator that models the photovoltaic installations that are on the roofs on the houses;
- each **production** nodes allow the connection of a wind generator that has a capacity of 6 MW;
• *topological* nodes do not correspond to devices, but serve the topology by creating ramifications in the network;

• the *slack* node models the TN as seen from the DN; it guarantees the balance between the consumed and produced powers within the network and a constant voltage magnitude is associated to this node.

When sizing the devices, we considered that the network supplies twenty thousand dwellings with a mean number of 2.4 occupants per dwelling. These dwellings are allocated among 53 loads, as indicated in Figure 8. In order to determine the capacity of the generators at residential nodes, we assumed a mean surface of 1.3 $m^2$ of solar panels per person. These values have been used to build the stochastic models of injections and withdrawals. They are illustrated in Figure 9, where realizations of the power withdrawn from the network over a one-day horizon are presented. The procedure used to build these stochastic processes is described at the end of this section.

In addition to the curtailment actions that can be applied to both solar and wind generators, at a cost that is defined in Section 3.4 and using the market price illustrated in Figure 10, we consider that all the loads offer a flexibility service. For this instance, we have, $\forall d \in \mathcal{F} = \mathcal{C}$:

$$ t'_d = T_d - s_{d,t}^{(f)} + 1 $$

$$ \Delta P_d(t'_d) = \begin{cases} 
\pm \Delta P_d^{nom} \sin(1.8\pi \frac{t'_d - 0.5(T_d + 1)}{T_d - 1}) & \text{if } t'_d \in [1, T_d] \\
0 & \text{otherwise}
\end{cases} $$

(49)
where the sign of $\Delta P_d(t')$ depends on the type of modulation that is offered by the load $d$ (i.e. upward or downward). About half of the loads offer a downward modulation, followed by an upward rebound effect, and inversely for the other half. These flexibility services are illustrated in Figure 11. Parameters $\Delta P_d^{nom}$ and $T_d$ are specific to each load and we consider that the activation costs are proportional to the magnitude of the modulation signals ($C_d^{flex} \propto \Delta P_d^{nom}$).

The approach used to build the transition functions of the stochastic quantities (i.e. the consumption of the loads, the wind speed, and the level of solar irradiance), is to learn, from a dataset, the functions $\mu_i(s_{i,t}, q_t)$ and $\sigma_i^2(q_t)$, which predict the mean and the variance, respectively, for period $t+1$ of each of these quantities $i$. The input values used for this purpose are the realization $s_{i,t}$ of the quantity $i$ at period $t$ and the quarter of an hour in the day $q_t$ to which it corresponds. The following procedure has been used to build the approximate functions $\hat{\mu}_i$ and $\hat{\sigma}_i^2$:

1. formatting of the dataset into $K$ tuples $(s_{i,t}^{(k)}, q_t^{(k)}, s_{i,t+1}^{(k)})$;
2. let $\hat{\mu}_\theta$ be a neural network and $\theta$ its learning parameters, $\hat{\mu}_i$ is determined by solving

$$\theta^*_i = \arg \min_\theta \sum_{k=1}^K [s^{(k)}_{i,t+1} - \hat{\mu}_\theta(s^{(k)}_{i,t}, q^{(k)}_t)]^2$$

with $\hat{\mu}_i = \hat{\mu}_{\theta^*_i}$;

3. let $\hat{\sigma}^2_\eta$ be a neural network, $\eta$ its learning parameters and $\mathcal{J}_k$ a set defined as $\mathcal{J}_k = \{ j \mid q^{(j)}_t = q^{(k)}_t \}$, $\hat{\sigma}^2_\eta$ is determined by solving

$$\eta^*_i = \arg \min_\eta \sum_{k=1}^K \left[ \hat{\sigma}^2_\eta (q^{(k)}_t) - \frac{1}{|\mathcal{J}_k|} \sum_{j \in \mathcal{J}_k} [s^{(j)}_{i,t+1} - \hat{\mu}_i(s^{(j)}_{i,t}, q^{(j)}_t)]^2 \right]^2$$

with $\hat{\sigma}^2_\eta = \hat{\sigma}^2_{\eta^*_i}$.

Then, by choosing $p_{\mathcal{W}}(\cdot|s_t)$, such that each component $w_{i,t}$ of vector $\mathbf{w}_t \sim p_{\mathcal{W}}(\cdot|s_t)$ follows a standard normal law $\mathcal{N}(0, 1)$, the transition functions of the stochastic quantities can be defined
as

\[ P_{d,t+1} = f_d(P_{d,t}, q_t) = \hat{\mu}_d(P_{d,t}, q_t) + w_{d,t} \sqrt{\hat{\sigma}_d^2(q_t)}, \quad \forall d \in \mathcal{C} \]  

(50)

\[ v_{t+1} = f_v(v_t, q_t) = \hat{\mu}_v(v_t, q_t) + w_{v,t} \sqrt{\hat{\sigma}_v^2(q_t)} \]  

(51)

\[ ir_{t+1} = f_{ir}(ir_t, q_t) = \hat{\mu}_{ir}(ir_t, q_t) + w_{ir,t} \sqrt{\hat{\sigma}_{ir}^2(q_t)} \]  

(52)

The datasets that we used to learn the functions are real measurements of the wind speed\(^1\) and of the solar irradiance\(^2\), while, for residential consumption, they correspond to the aggregation of scenarios generated from a detailed stochastic model [30].

6 Example of policy

The goal of this section is twofold: to illustrate the operational planning problem and the test instance of Section 5 and to provide a reference to assess the performance of the more advanced solution techniques that will be developed. To that end, we present a solution technique that belongs to the class of mathematical programming approaches and, in particular, that is a simplified version of Problem (41)-(46).

At each time step \( t \) of the decision process, a scenario tree must first be built, as it defines the set of nodes \( K_t \) used by the mathematical program. We are looking for a decent approximation of the uncertainty while keeping the resulting mathematical program computationally affordable. A widespread procedure is to generate many trajectories of the system over a time horizon \( T \), i.e. sequences \( w_t, w_{t+1}, \ldots, w_{t+T-1} \) of realizations of the random process, and then to cluster them into a scenario tree [31]. It is possible to obtain trajectories using sampling, by running simulations with the Matlab\(^{\circledR} \) code associated to the test instance (cf. Section 5). We chose to aggregate these \( N_{\text{trajs}} \) trajectories into a scenario tree that has a branching factor of \( N_{\text{scens}} \) at its root node and of 1 everywhere else. A k-mean clustering algorithm is used for this purpose [32], where the distance between two trajectories \( i \) and \( j \) is the Euclidean distance between the resulting time series of active power injections of the devices:

\[
\text{dist}(\text{traj}^{(i)}, \text{traj}^{(j)}) = \sqrt{\sum_{k=1}^{T} \sum_{d \in \mathcal{D}} \left( P_{d,k}^{(i)} - P_{d,k}^{(j)} \right)^2},
\]

(53)

where, for each trajectory \( i \), the power injections \( P_{d,k}^{(i)}, \forall (d,k) \in \mathcal{D} \times \mathbb{Z}_{[1,T]} \), are fully defined by the initial state \( s_t \) and by the sequence \( w_t^{(i)}, w_{t+1}^{(i)}, \ldots, w_{t+T-1}^{(i)} \) associated to the trajectory. Such a tree is illustrated in Figure 12, where we highlighted how it models the decision process by labeling the nodes and edges of the tree with the corresponding states and actions.

In order to limit the complexity of the resulting mathematical program and, especially, to restrain the number of integer variables, the possibility to activate flexible loads is only considered for the first action \( a_t \) while the recourse actions at the following steps are limited to curtailment instructions. Using notations of Section 4, the policy \( \hat{\pi}^{*} \) that is derived from the
The presented solution technique can be written as:

\[
\hat{\pi}^* (s) = \arg \min_{a \in A_k(s)} \min_{k \in K_t} \left[ \sum_{\gamma \in \mathcal{D}_k} \frac{\Delta P_{g,k}}{4} \sum_{g \in \mathcal{G}} \left( C_{\text{curt}}(q_k) + \epsilon_2 \Delta P_{g,k}^2 \right) - \epsilon_1 \Delta M_{g,k} + \epsilon_2 \Delta M_{g,k}^2 \right] + \sum_{d \in \mathcal{F}} C_{\text{flex}}(f) (54)
\]

subject to:

\[
s_0 = s (55)
\]

\[
a_0 = a (56)
\]

\[
s_k = f(s_{h_k}, a_{h_k}, w_{h_k}), \quad \forall k \in K_t \setminus \{0\} (57)
\]

\[
a_{h_k} \in A_{h_k}, \quad \forall k \in K_t \setminus \{0\} (58)
\]

\[
a_{h_k}(f) = 0, \quad \forall k \in \{ k \in K_t \mid \mathcal{D}_k > 1 \} (59)
\]

\[
\mathcal{S}_{ok}(s_k) = \{ s \in S \mid \sum_{g \in \mathcal{G}} P_g + \sum_{d \in \mathcal{C}} (P_d + \Delta P_d) < \overline{C} \}, (63)
\]

where Equation (59) enforces that the activation of flexible loads is not accounted as a recourse action. The set \( \hat{\mathcal{S}}_{ok} \) is an approximation of the set \( \mathcal{S}_{ok} \) of the system states that respect operational limits. For the test instance presented in this paper, this set is defined using a linear constraint over the upper limits of active production levels and over the active consumption of loads:

\[
\hat{\mathcal{S}}_{ok}(s_k) = \{ s \in S \mid \sum_{g \in \mathcal{G}} P_g + \sum_{d \in \mathcal{C}} (P_d + \Delta P_d) < \overline{C} \}, (63)
\]

where \( \overline{C} \) is a constant that can be estimated by trial and error and with \( \Delta P_{d,k} \) defined as in Equation (20). The physical motivation behind this constraint is that issues usually occur when a high level of distributed production and a low consumption level take place simultaneously.

In order to get control actions that are somehow robust to the evolutions of the system that would not be well accounted in the scenario tree, the objective function of Problem (54)-(62) includes, for each node \( k \) but the root node, the following terms:

\[
P_k^\gamma \sum_{g \in \mathcal{G}} \left( \epsilon_2 \Delta P_{g,k}^2 - \epsilon_1 \Delta M_{g,k} + \epsilon_2 \Delta M_{g,k}^2 \right), (64)
\]

where \( \epsilon_1 \) and \( \epsilon_2 \) are small positive parameters. The goal of these terms is to drive the solution towards curtailment instructions that are as equally shared as possible among generators and towards margins between the production upper limits and the forecasted production levels that are as large as possible while being also equally shared among generators. By doing so, the

This formulation is used for the sake of understanding. It differs from the exact implementation but defines an equivalent mathematical program.
outcome of the actions is less dependent on the realization of the power injections of a particular generator.

The implementation has been done using the Matlab® code mentioned in Section 5 to simulate the system, YALMIP [33] to build the mathematical programs and Gurobi® [34] to solve them. The values of the parameters that have been used are a time horizon $T$ of 15 periods, a discount factor $\gamma$ of 0.99, and 250 sampled trajectories ($N_{\text{trajs}}$) at each time step, which are then clustered into 5 scenarios ($N_{\text{scens}}$). The parameters $\epsilon_1$ and $\epsilon_2$ were set to $10^{-2}$ and $10^{-3}$, respectively. In order to evaluate the performance of this policy, it was simulated over 300 runs of 2 days (i.e. of 192 transitions). The results are shown in Figure 13. The empirical estimations of the expected return reported on this Figure are computed as:

$$\mathbb{E}_{s \sim p_0(\cdot)} \left\{ J^{\hat{\pi}^*}(s) \right\} \approx \frac{1}{N_{\text{runs}}} \sum_{n=1}^{N_{\text{runs}}} \sum_{t=0}^{191} \gamma^t r_t^{(n)},$$

where $\left(r_t^{(n)}\right)_{t=0}^{191}$ corresponds to the sequence of instantaneous rewards observed during the $n^{th}$ simulation run and using policy $\hat{\pi}^*$. For each of the simulation runs, the initial state of the system was drawn randomly by the simulator, following a probability distribution that we denote by $p_0(\cdot)$. Running on a 2.6 GHz machine and limited to a single core, the described solution technique took in average 23.13 s by time step to compute the control actions.

Finally, Figure 14 presents a simulation run of the system, subject to the control actions of the policy. At around time step 140, we observe that both production curtailment and load modulation decisions are taken, the latter relieving the needs for curtailment, which is a more expensive control mean.

7 Conclusions

In this chapter, we have proposed a generic formulation for the ANM problem of an electrical distribution system that takes the form of a Markov decision process. This formulation highlighted the sequential and uncertain nature of the problem and allowed the identification of potential solution techniques. In order to promote the development of computational techniques in this field, the formulation has also been specified to a particular distribution system,
Figure 14: Example of a simulation run of the system controlled by policy (54)-(62), over two
days.

which is supported by Matlab\textsuperscript{®} code that implements a simulator of this test instance (cf.
http://www.montefiore.ulg.ac.be/~anm/). In addition, an example of solution technique
was presented and its performance was reported.

8 Acknowledgment

This research is supported by the public service of Wallonia - Department of Energy and Sustain-
able Building within the framework of the GREDOR project. The authors give their thanks
for the financial support of the Belgian Network DYSCO, an Inter-university Attraction Poles
Program initiated by the Belgian State, Science Policy Office.

The authors would also like to thank Raphaël Fonteneau for his precious advices and com-
ments.

References


tributed generation into electric power systems: A review of drivers, challenges and oppor-


Variable DG and active networks. IEEE Transactions on Power Systems, 25(1):87–95,
2010.

tion power flow management utilizing an online optimal power flow technique. IEEE


