Electronic appendix of paper
Using Machine Learning to Enable Probabilistic Reliability Assessment in Operation Planning

Laurine Duchesne, Efthymios Karangelos, Louis Wehenkel

March 2018

1 Introduction

This document is the electronic appendix of [1]. It is organised as follows. Section 2 gives the detailed mathematical implementation of the day-ahead decision-making program simulating day-ahead operation planning and of the *Security Constrained Optimal Power Flow* (SCOPF) used to model the behaviour of the control-room operators in real-time. Section 3 briefly describes the data used in the case study of [1] and Section 4 introduces the supervised learning algorithms used in the paper and the machine learning setting.

2 Day-ahead and real-time operation models

This section details the implementation of the day-ahead and real-time operation simulators. It begins with introducing the notations used in the mathematical models, then it presents the day-ahead decision-making program and finally it describes the real-time SCOPF.

2.1 Notations

Indices

c	Index of contingencies
d	Index of demands
g	Index of generating units
k	Index of piece-wise linear dispatchable generation cost curve segments
l	Index of transmission elements (lines, cables and transformers)
n	Index of nodes
t	Index of hours in a day
w	Index of wind power generators

Sets

$\mathcal{C}$	Set of contingencies
$\mathcal{D}$	Set of demands
$\mathcal{D}_n$	Subset of demands connected at node n
$\mathcal{G}$	Set of dispatchable units

$\mathcal{K}$	Set of piece-wise linear dispatchable generation cost curve segments
$\mathcal{L}$	Set of transmission elements (lines, cables and transformers)
$\mathcal{N}$	Set of nodes
$\mathcal{W}$	Set of wind power generators
$\mathcal{W}_n$	Subset of wind power generators connected at node n

Parameters

$P_{d,t}^{forecast}$	Forecast of load active power of demand d at time t
$P_{w,t}^{forecast}$	Forecast of generation of wind power generator w at time t
$P_{d,t}^{RT}$	Realisation of load active power of demand d at time t
$P_{w,t}^{RT}$	Realisation of generation of wind power generator w at time t
on_g^{init}	Initial status of generating unit g at the beginning of the day-ahead decision-making (1 if started up, 0 otherwise)
$t_g^{up,init}$	Minimum number of time periods generating unit g must stay up at the beginning of the considered day
$t_g^{dn,init}$	Minimum number of time periods generating unit g must stay down at the beginning of the considered day
$t_g^{up,min}$	Minimum number of time periods generating unit g must stay up once started up
$t_g^{dn,min}$	Minimum number of time periods generating unit g must stay down once shut down
c_g	Redispatch marginal cost of generating unit g
c_g^0	Start-up cost of generating unit g
$c_{g,k}^{inc}$	Marginal running cost of generating unit g at the segment k of its piece-wise linear curve
P_g^{max}	Capacity of generating unit g
P_g^{min}	Minimum stable output of generating unit g
ΔP_g^-	Ramp-down limit of generating unit g (for 60min)
ΔP_g^+	Ramp-up limit of generating unit g (for 60min)
$\Delta P_g^{-,c}$	Ramp-down limit of generating unit g in case of corrective actions (for 20min)
$\Delta P_g^{+,c}$	Ramp-up limit of generating unit g in case of corrective actions (for 20min)
$P_{g,k}^{inc,max}$	Maximum power output of generating unit g at the segment k of its piece-wise linear curve
v_d	Voll of demand d in euro/MWh
p_w	Wind penalty for curtailment of wind power generator w in euro/MWh
R^+	Minimum up spinning reserve required per hour for one area
R^-	Minimum down spinning reserve required per hour for one area
f_l^{max}	Long-term thermal rating of transmission element l

r_l — Ratio of the short-term thermal rating to the long-term thermal rating of transmission element l ($r_l \geq 1$)

X_l — Reactance of transmission element l

$\beta_{n,l}$ — Element of the flow incidence matrix, taking a value of one if node n is the sending node of element l, a value of minus one if node n is the receiving node of element l, and a zero value otherwise.

$a_{l,c}$ — Binary parameter taking a zero value if element l is unavailable under contingency c.

Variables

$P_{g,t}^{DA}$ — Dispatch of generating unit g at time t as per the day-ahead decision-making

$on_{g,t}^{DA}$ — Binary variable representing the status of generating unit g as per the day-ahead decision-making (1 if started up, 0 otherwise)

$st_{g,t}^{up}$ — Binary variable indicating when generating unit g is started-up (value 1 when started up, 0 otherwise)

$st_{g,t}^{dn}$ — Binary variable indicating when generating unit g is shut down (value 1 when shut down, 0 otherwise)

$WC_{w,t}^{DA}$ — Provisional curtailment of wind power generator w at hour t in day-ahead

$R_{g,t}^{+}$ — Upward redispatch flexibility provided by generating unit g at time t in day-ahead

$R_{g,t}^{-}$ — Downward redispatch flexibility provided by generating unit g at time t in day-ahead

$f_{l,t}^{DA}$ — Power flowing through transmission element l at time t under the pre-contingency state in day-ahead

$f_{l,t,c}^{DAST}$ — Power flowing through transmission element l at time t following contingency c in day-ahead

$\theta_{n,t}^{DA}$ — Voltage angle at node n under the pre-contingency state in day-ahead

$\theta_{l,t,c}^{DAST}$ — Voltage angle at node n following contingency c in day-ahead.

${}^{+}P_{g,t}^{RTp}$ — Preventive ramp-up of generator g in real-time at hour t

${}^{-}P_{g,t}^{RTp}$ — Preventive ramp-down of generator g in real-time at hour t

$LS_{d,t}^{RTp}$ — Preventive load shedding of demand d in real-time at hour t

$WC_{w,t}^{RT^p}$ — Preventive wind curtailment of wind power generator w in real-time at hour t

${}^{+}P_{g,t}^{RTc}$ — Corrective ramp-up of generator g in real-time at hour t

${}^{-}P_{g,t}^{RTc}$ — Corrective ramp-down of generator g in real-time at hour t

$LS_{d,t,c}^{RTc}$ — Corrective load shedding of demand d in real-time at hour t

$WC_{w,t,c}^{RTc}$ — Corrective wind curtailment of wind power generator w in real-time at hour t

$f_{l,t}^{p}$ — Power flowing through transmission element l under the pre-contingency state

$f^{ST}_{l,t,c}$ Power flowing through transmission element l following contingency c and prior to the application of corrective control.

$f^{c}_{l,t,c}$ Power flowing through transmission element l following contingency c and the successful application of corrective control.

$\theta^{p}_{n,t}$ Voltage angle at node n under the pre-contingency state

$\theta^{ST}_{l,t,c}$ Voltage angle at node n following contingency c and prior to the application of corrective control.

$\theta^{c}_{n,t,c}$ Voltage angle at node n following contingency c and the successful application of corrective control.

All the variables are continuous, except for $on_{g,t}$, $st^{dn}_{g,t}$ and $st^{up}_{g,t}$ which are binary variables. Powers flowing through transmission elements and voltage angles are continuous in $\mathbb{R}$ and the remaining variables are positive.

2.2 Day-ahead decision-making

We simulate day-ahead decision-making with a multi-period SCOPF in order to commit and dispatch the generating units of the system and also determine the provisional wind curtailment. We use the DC approximation [2] and consider as reliability criterion the N-1 criterion for transmission elements only.

The objective function minimizes generation cost as well as provisional wind curtailment:

$$\text{minimize} \quad \sum_{t=1}^{24} \left(\sum_{g \in \mathcal{G}} \left(c^0_g * st^{up}_{g,t} + \sum_{k \in \mathcal{K}} c^{inc}_{g,k} * P^{inc}_{g,k,t} \right) + \sum_{w \in \mathcal{W}} p_w * WC^{DA}_{w,t} \right). \tag{1}$$

The first set of constraints (2-12) of the day-ahead program concerns the minimum time a generating unit must stay up or down, either at the beginning of the day or during the day.

For $t = 1$, $\forall g \in \mathcal{G}$:

$$st^{up}_{g,t} - st^{dn}_{g,t} = on^{DA}_{g,t} - on^{init}_{g} \tag{2}$$

$$st^{up}_{g,t} + st^{dn}_{g,t} \leq 1 \tag{3}$$

$\forall t = 2, ..., 24$, $\forall g \in \mathcal{G}$:

$$st^{up}_{g,t} - st^{dn}_{g,t} = on^{DA}_{g,t} - on^{DA}_{g,t-1} \tag{4}$$

$$st^{up}_{g,t} + st^{dn}_{g,t} \leq 1 \tag{5}$$

$$\tag{6}$$

$\forall g \in \mathcal{G}$:

$$\sum_{t'=1}^{t^{up,init}_g} \left(1 - on^{DA}_{g,t'}\right) = 0 \tag{7}$$

$$\sum_{t'=1}^{t^{dn,init}_g} on^{DA}_{g,t'} = 0 \tag{8}$$

$$\tag{9}$$

$\forall g \in \mathcal{G}$, $\forall t = 1, ..., \left(24 - t_g^{up,min}\right)$:

$$\sum_{t'=t}^{t+t_g^{up,min}} on_{g,t'}^{DA} \geq st_{g,t}^{up} \cdot t_g^{up,min} \tag{10}$$

$$\tag{11}$$

$\forall g \in \mathcal{G}$, $\forall t = 1, ..., \left(24 - t_g^{dn,min}\right)$:

$$\sum_{t'=t}^{t+t_g^{dn,min}} \left(1 - on_{g,t'}^{DA}\right) \geq st_{g,t}^{dn} \cdot t_g^{dn,min} \tag{12}$$

$$\tag{13}$$

The following set of constraints limits the power output of each generating unit between its minimum stable output and its maximum capacity and also imposes ramping constraints to go from one committed dispatch to the one of the next period in one hour:

$$-P_{g,t}^{DA} + R_{g,t}^{-} \leq -P_g^{min} * on_{g,t}^{DA} \tag{14}$$

$$P_{g,t}^{DA} + R_{g,t}^{+} \leq P_g^{max} * on_{g,t}^{DA} \tag{15}$$

$$P_{g,t+1}^{DA} - P_{g,t}^{DA} \leq \Delta P_g^{+} * on_{g,t}^{DA} + P_g^{max} * (1 - on_{g,t}^{DA}) \tag{16}$$

$$-\left(P_{g,t+1}^{DA} - P_{g,t}^{DA}\right) \leq \Delta P_g^{-} * on_{g,t}^{DA} + P_g^{max} * (1 - on_{g,t+1}^{DA}) \tag{17}$$

We assume a piece-wise linear cost function of $|\mathcal{K}|$ segments for the marginal running cost of a generating unit g, which gives eq. (18) and (20).
$\forall g \in \mathcal{G}$, $\forall k \in \mathcal{K}$, $\forall t = 1, ..., 24$:

$$P_{g,k,t}^{inc} \leq on_{g,t}^{DA} \cdot \ P_{g,k}^{inc,max} \tag{18}$$

$$\tag{19}$$

$\forall g \in \mathcal{G}$, $\forall t = 1, ..., 24$:

$$P_{g,t}^{DA} = \sum_{k=1}^{K} P_{g,k,t}^{inc} \tag{20}$$

Equation (21) represents the balancing of the system and equations (23)-(25) the transmission constraints in case of the DC approximation.

$\forall t = 1, ..., 24$, $\forall n \in \mathcal{N}$:

$$\sum_{w \in \mathcal{W}_n} (P_{w,t}^{forecast} - WC_{w,t}^{DA}) + \sum_{g \in \mathcal{G}_n} P_{g,t}^{DA} - \sum_{l \in \mathcal{L}} \beta_{n,l} * f_{l,t}^{DA} = \sum_{d \in \mathcal{D}_n} P_{d,t}^{forecast} \tag{21}$$

$\forall t = 1, ..., 24$, $\forall w \in \mathcal{W}$:

$$0 \leq WC_{w,t}^{DA} \leq P_{w,t}^{forecast} \tag{22}$$

$\forall t = 1, ..., 24$, $\forall l \in \mathcal{L}$:

$$f_{l,t}^{DA} - \frac{1}{X_l} \sum_{n \in \mathcal{N}} \beta_{n,l} * \theta_{l,t}^{DA} = 0 \quad (23)$$

$$f_{l,t}^{DA} \leq f_l^{max} \quad (24)$$

$$-f_{l,t}^{DA} \leq f_l^{max} \quad (25)$$

Equations (26)-(29) force the system to still be secure in the case of the loss of one transmission element.

$\forall t = 1, ..., 24$, $\forall c \in \mathcal{C}, \forall n \in \mathcal{N}$:

$$\sum_{w \in \mathcal{W}_n} (P_{w,t}^{forecast} - WC_{w,t}^{DA}) + \sum_{g \in \mathcal{G}_n} P_{g,t}^{DA} - \sum_{l \in \mathcal{L}} \beta_{n,l} * f_{l,t,c}^{DAST} = \sum_{d \in \mathcal{D}_n} P_{d,t}^{forecast} \quad (26)$$

$\forall t = 1, ..., 24$, $\forall c \in \mathcal{C}, \forall l \in \mathcal{L}$:

$$f_{l,t,c}^{DAST} - a_{l,c} * \frac{1}{X_l} \sum_{n \in \mathcal{N}} \beta_{n,l} * \theta_{l,t,c}^{DAST} = 0 \quad (27)$$

$$f_{l,t,c}^{DAST} \leq a_{l,c} * f_l^{max} \quad (28)$$

$$-f_{l,t,c}^{DAST} \leq a_{l,c} * f_l^{max} \quad (29)$$

Equations (30)-(35) determine the minimum size of the up and down spinning reserves per area in the system (in this work, we have three areas and the same spinning reserve requirements per area).

$\forall t = 1, ..., 24$, $\forall g =\in \mathcal{G}$

$$\sum_{g \in \mathcal{G}_{area1}} R_{g,t}^{+} \geq R^{+} \quad (30)$$

$$\sum_{g \in \mathcal{G}_{area1}} R_{g,t}^{-} \geq R^{-} \quad (31)$$

$$\sum_{g \in \mathcal{G}_{area2}} R_{g,t}^{+} \geq R^{+} \quad (32)$$

$$\sum_{g \in \mathcal{G}_{area2}} R_{g,t}^{-} \geq R^{-} \quad (33)$$

$$\sum_{g \in \mathcal{G}_{area3}} R_{g,t}^{+} \geq R^{+} \quad (34)$$

$$\sum_{g \in \mathcal{G}_{area3}} R_{g,t}^{-} \geq R^{-} \quad (35)$$

2.3 Real-time operation

In order to simulate real-time operation along a system trajectory, we solve sequentially the 24 hourly steps of the trajectory. That is we solve 24 single period problems corresponding to the 24 hours of one day.

We model real-time operation with a SCOPF problem with the N-1 reliability criterion, again considering only transmission elements. We consider preventive (pre-contingency) as well as corrective actions (post-contingency) and we do not forget the intermediate state after the occurrence of a contingency but before any corrective action can be applied, that we call short-term post-contingency state.

Note that continuous variables from the day-ahead decision-making program are parameters for this problem.

2.3.1 Objective function

The objective function (36) minimizes the redispatch cost (upward and downward) as well as load shedding and wind curtailment, both in preventive and corrective modes. The value of lost load and wind penalty should be such that load shedding and wind curtailment are used only where no other solution exists. Note that, in order to favour corrective actions over preventive ones, we multiply the total preventive cost by a large factor M.

$$\begin{aligned} \text{minimize} \quad & M * \left(\sum_{g \in \mathcal{G}} c_g \left({}^{+}P^{RTp}_{g,t} +^{-} P^{RTp}_{g,t} \right) + \sum_{d \in \mathcal{D}} v_d * LS^{RTp}_{d,t} + \sum_{w \in \mathcal{W}} p_w * WC^{RTp}_{w,t} \right) \\ & + \sum_{c \in \mathcal{C}} \left(\sum_{d \in \mathcal{D}} v_d * LS^{RTc}_{d,t,c} + \sum_{w \in \mathcal{W}} p_w * WC^{RTc}_{w,t,c} + \sum_{g \in \mathcal{G}} c_g \left({}^{+}P^{RTc}_{g,t,c} +^{-} P^{RTc}_{g,t,c} \right) \right) \end{aligned} \tag{36}$$

2.3.2 Pre-contingency state

The following equations determine the preventive actions.The possible redispatch of generating units is limited by maximum and minimum output power of generating units as well as by ramping constraints of one hour. Equations (41) and (42) also impose that with the re-dispatch of a unit g, it is still possible to go in one hour to the dispatch of the generating unit g at time $t+1$ as per the day-ahead decision-making.

$\forall g \in \mathcal{G}$:

$$P^{DA}_{g,t} + ({}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t}) \geq P^{min}_{g} * on^{DA}_{g,t} \tag{37}$$

$$P^{DA}_{g,t} + ({}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t}) \leq P^{max}_{g} * on^{DA}_{g,t} \tag{38}$$

$${}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t} \leq \Delta P^{+}_{g} \tag{39}$$

$$-({}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t}) \leq \Delta P^{-}_{g} \tag{40}$$

$$P^{DA}_{g,t+1} - \left(P^{DA}_{g,t} + ({}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t} \right) \leq \Delta P^{+}_{g} \tag{41}$$

$$-\left(P^{DA}_{g,t+1} - \left(P^{DA}_{g,t} +^{+} P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t} \right) \right) \leq \Delta P^{-}_{g} \tag{42}$$

The next constraints correspond to the classical DC approximation and limits on maximum load shedding and wind curtailment.

$\forall n \in \mathcal{N}$:

$$\sum_{w \in \mathcal{W}_n} (P^{RT}_{w,t} - WC^{DA}_{w,t} - WC^{RT^p}_{w,t}) + \sum_{g \in \mathcal{G}_n} \left(P^{DA}_{g,t} + ({}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t}) \right) - \sum_{l \in \mathcal{L}} \beta_{n,l} * f^{p}_{l,t} = \sum_{d \in \mathcal{D}_n} (P^{RT}_{d,t} - LS^{RT^p}_{d,t}) \tag{43}$$

$\forall l \in \mathcal{L}$:

$$f^p_{l,t} - \frac{1}{X_l} \sum_{n \in \mathcal{N}_n} \beta_{n,l} * \theta^p_{l,t} = 0 \tag{44}$$

$$f^p_{l,t} \leq f^{max}_l \tag{45}$$

$$-f^p_{l,t} \leq f^{max}_l \tag{46}$$

Finally, we ensure that we do not shed more load and wind generation than what is possible: $\forall d \in \mathcal{D}$:

$$0 \leq LS^{RT^p}_{d,t} \leq P^{RT}_{d,t} \tag{47}$$

$\forall w \in \mathcal{W}$:

$$0 \leq WC^{DA}_{w,t} + WC^{RT^p}_{w,t} \leq P^{RT}_{w,t} \tag{48}$$

2.3.3 Short-term post-contingency state

In this stage, a contingency occurred but the operator did not react yet. Since we are in emergency state, the line thermal ratings correspond to the short-term ones.

$\forall c \in \mathcal{C}, \forall n \in \mathcal{N}$:

$$\sum_{w \in \mathcal{W}_n} (P^{RT}_{w,t} - WC^{DA}_{w,t} - WC^{RT^p}_{w,t}) + \sum_{g \in \mathcal{G}_n} \left(P^{DA}_{g,t} + ({}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t}) \right) - \sum_{l \in \mathcal{L}} \beta_{n,l} * f^{ST}_{l,t,c} = \sum_{d \in \mathcal{D}_n} (P^{RT}_{d,t} - LS^{RT^p}_{d,t}) \tag{49}$$

$\forall c \in \mathcal{C}, \forall l \in \mathcal{L}$:

$$f^{ST}_{l,t,c} - a_{l,c} * \frac{1}{X_l} \sum_{n \in \mathcal{N}_n} \beta_{n,l} * \theta^{ST}_{l,t,c} = 0 \tag{50}$$

$$f^{ST}_{l,t,c} \leq a_{l,c} * r_l * f^{max}_l \tag{51}$$

$$-f^{ST}_{l,t,c} \leq a_{l,c} * r_l * f^{max}_l \tag{52}$$

2.3.4 Corrective control

Finally, corrective actions can be applied to keep the system secure.

$\forall c \in \mathcal{C}, \forall g \in \mathcal{G}$:

$$P^{DA}_{g,t} + ({}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t}) + ({}^{+}P^{RTc}_{g,t,c} -^{-} P^{RTc}_{g,t,c}) \geq P^{min}_g * on^{DA}_{g,t} \tag{53}$$

$$P^{DA}_{g,t} + ({}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t}) + ({}^{+}P^{RTc}_{g,t,c} -^{-} P^{RTc}_{g,t,c}) \leq P^{max}_g * on^{DA}_{g,t} \tag{54}$$

$${}^{+}P^{RTc}_{g,t,c} -^{-} P^{RTc}_{g,t,c} \leq \Delta P^{+,c}_g \tag{55}$$

$$-({}^{+}P^{RTc}_{g,t,c} -^{-} P^{RTc}_{g,t,c}) \leq \Delta P^{-,c}_g \tag{56}$$

$\forall c \in \mathcal{C}, \forall n \in \mathcal{N}$:

$$\sum_{w \in \mathcal{W}_n} (P^{RT}_{w,t} - WC^{DA}_{w,t} - WC^{RT^p}_{w,t} - WC^{RT^c}_{w,t,c}) + \sum_{g \in \mathcal{G}_n} \left(P^{DA}_{g,t} + ({}^{+}P^{RTp}_{g,t} -^{-} P^{RTp}_{g,t}) + ({}^{+}P^{RTc}_{g,t,c} -^{-} P^{RTc}_{g,t,c}) \right) - \sum_{l \in \mathcal{L}} \beta_{n,l} * f^{c}_{l,t,c} = \sum_{d \in \mathcal{D}_n} (P^{RT}_{d,t} - LS^{RT^p}_{d,t} - LS^{RT^c}_{d,t,c}) \tag{57}$$

$\forall c \in \mathcal{C}, \forall d \in \mathcal{D}$:

$$0 \leq LS_{d,t}^{RT^p} + LS_{d,t,c}^{RT^c} \leq P_{d,t}^{RT} \tag{58}$$

$\forall c \in \mathcal{C}, \forall w \in \mathcal{W}$:

$$0 \leq WC_{w,t}^{DA} + WC_{w,t}^{RT^p} + WC_{w,t,c}^{RT^c} \leq P_{w,t}^{RT} \tag{59}$$

$\forall c \in \mathcal{C}, \forall l \in \mathcal{L}$:

$$f_{l,t,c}^{c} - a_{l,c} * \frac{1}{X_l} \sum_{n \in \mathcal{N}_n} \beta_{n,l} * \theta_{l,t}^{ST} = 0 \tag{60}$$

$$f_{l,t,c}^{c} \leq a_{l,c} * f_l^{max} \tag{61}$$

$$-f_{l,t,c}^{c} \leq a_{l,c} * f_l^{max} \tag{62}$$

3 Case study: the IEEE-RTS96 - short description of the data

In [1], we test our methodology on the IEEE-RTS96 benchmark [3], where 19 windfarms have been added as in [4].

We consider in this case study the first day of the year, with a peak demand per area of 3135MW. Demand, generating units and line ratings data come from [3], while the forecast wind generation ('favorable') as well as the initial states of the generating units are borrowed from [4]. Note that line ratings have been reduced by a factor of 20%.

Concerning the reliability criterion, we use the N-1 criterion for transmission elements only (transmission lines, cables and transformers) and thus we have 120 contingencies. It is important to note that for the real-time simulator we do not consider contingencies of lines 49 and 82 in order to avoid islanding of nodes 207 and 307. Therefore, we have only 118 contingencies in the real-time problem.

We choose as minimum up and down spinning reserve per area 300 MW. The wind penalty is 300€/MWh and the voll is an average of the coefficients from [5] converted in e and is thus equal to 4018.2€/MWh.

Concerning the objective function of the real-time simulator, we set M=150($> |\mathcal{C}|$) in order to be sure that corrective actions will always be favoured over preventive ones.

4 Machine learning settings and predictors

In this section, we begin by introducing briefly the learning algorithms used in the paper, then we describe the procedure used to train and test the models and we list the values of the meta-parameters tested to improve the performance of our models.

4.1 Two regression algorithms: extremely randomized trees and neural networks

We tested two types of predictors : extremely randomized trees (ET) [6] and artificial neural network (NN)[7].

The ET algorithm is a Random Forest algorithm [8]. It is an ensemble of regression trees where each tree is built with some randomness. The final prediction is the average of the predictions of each tree in the forest. This method as three meta-parameters: the number of trees in the forest, the number k of features selected randomly at each split and the minimum number of samples required to split a node n_{min}.

The artificial neural network we studied is a multi-layer perceptron with a ReLU activation function. We tuned the number of hidden layers and the number of neurons per layer.

4.2 Description of the learning procedure

In order to avoid overfitting, we divide the dataset randomly into two sets: a learning set and a test set. Each model is learnt with the learning set and its performances are assessed on the test set. It allows to see how well each model generalises on unseen data. The measure used to compare each model is the R^2-score (coefficient of determination), which is computed on the basis of N cases by [10]:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2},$$

where y_i is the true output of case i, $\hat{y}_i$ is the predicted output, and $\bar{y}$ is the mean of the N true values. The best possible score is 1 and corresponds to a model that perfectly predicts all the target output values of the dataset used to estimate its value.

As said in the previous section, each model has some meta-parameters that we can tune in order to improve their performance. In order to select the best meta-parameters, we use a 5-fold cross validation. For the ET algorithm, we tested the following parameters: $k = 1, p/3, p/2, p$, where p is the total number of features and $n_{min} = 2, 4, 6, 8, 10, 20$. The number of trees was set to 1000, which is good trade-off between performance and time needed to train and predict. For the NN algorithm, we tried the following configurations: two or three hidden layers with 10, 50 or 100 neurons per layer. The best meta-parameters vary in function of the output we want to predict or the setting used.

References

[1] L. Duchesne, E. Karangelos, and L. Wehenkel, "Using machine learning to enable probabilistic reliability assessment in operation planning," *Submitted for presentation at Power Systems Computation Conference (PSCC), Dublin, Ireland, June 2018*, October 2017.

[2] A. J. Wood and B. F. Wollenberg, *Power generation, operation, and control.* John Wiley & Sons, 2012.

[3] P. Wong, P. Albrecht, R. Allan, R. Billinton, Q. Chen, C. Fong, S. Haddad, W. Li, R. Mukerji, D. Patton *et al.*, "The ieee reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee," *Power Systems, IEEE Transactions on*, vol. 14, no. 3, pp. 1010–1020, 1999.

[4] H. Pandzic, Y. Dvorkin, T. Qiu, Y. Wang, and D. Kirschen, "Unit Commitment under Uncertainty - GAMS Models," Library of the Renewable Energy Analysis Lab (REAL), University of Washington, Seattle, USA, [Online]. Available at: http://www.ee.washington.edu/research/real/gams_code.html.

[5] M. Fotuhi-Firuzabad and R. Billinton, "Impact of load management on composite system reliability evaluation short-term operating benefits," *IEEE Transactions on power systems*, vol. 15, no. 2, pp. 858–864, 2000.

[6] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, 2006.

[7] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer Series in Statistics, 2009.

[8] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[9] L. Duchesne, E. Karangelos, and L. Wehenkel, "Machine learning of real-time power systems reliability management response," *PowerTech Manchester 2017 Proceedings*, 2017.

[10] Scikit-learn developers, "Model evaluation: quantifying the quality of predictions," [Online], Available at http://scikit-learn.org/stable/modules/model_evaluation.html, accessed on 2016-05-11.