

Unified Unit Commitment Formulation and Fast Multi-Service LP Model for Flexibility Evaluation in Sustainable Power Systems

Lingxi Zhang, *Student Member, IEEE*, Tomislav Capuder, *Member, IEEE*, and Pierluigi Mancarella, *Senior Member, IEEE*

Abstract — Classical unit commitment (UC) algorithms may be extremely time-consuming when applied to large systems and for long term simulations (for instance, a year) and may not consider all the features required for flexibility assessment, including analysis of different reserve types. In this light, this paper presents a novel flexibility-oriented unified formulation of a large-scale scheduling model considering multiple types of plants (including storage) and reserves, which can seamlessly model binary (BUC), mixed integer linear programming (MILP), and relaxed linear programming (LP) UC. Comparisons are carried out on several case studies for a reduced model of Great Britain, assessing loss of accuracy (as measured according to various metrics specifically introduced) against computational benefits in different renewables scenarios with more or less flexible systems. It is demonstrated how the computational time of the LP model is significantly less than the BUC and MILP approaches while capturing with relatively high precision all the relevant flexibility requirements and allocation of multiple types of reserves to different types of plants. The results indicate that the proposed fast LP model could be suitable for various computationally intensive flexibility studies (e.g., Monte Carlo simulations or planning), with significant reduction in simulation time and only minor errors relative to established MILP models.

Keywords: Flexibility, Linear Programming (LP), Mixed Integer Linear Programming (MILP), Renewable energy sources, Energy storage, Unit commitment.

ACRONYMS

BUC	binary unit commitment
GM	generation mix
CCGT	combine cycle gas turbines
LP	linear programming
MILP	mixed integer linear programming
MSG	minimum stable generation
OCGT	open cycle gas turbines
PFR	primary frequency response

The work is partly supported by the project FENISG - Flexible Energy Nodes in Low Carbon Smart Grid funded by Croatian Science Foundation under project grant No. 7766.

Lingxi Zhang and Pierluigi Mancarella are with the University of Manchester, School of Electrical and Electronic Engineering, M13 9PL, Manchester, UK (e-mail: lingxi.zhang@postgrad.manchester.ac.uk, p.mancarella@manchester.ac.uk).

Tomislav Capuder is with the University of Zagreb Faculty of Electrical Engineering and Computing, Zagreb, Croatia (email: tomislav.capuder@fer.hr).

PP	power provision
PSPP	pumped-hydro storage power plant
RES	renewable energy sources
SUR	secondary up reserve
TUR	tertiary up reserve

NOMENCLATURE

Indices

t, τ, T	time period (h)
i, I	generation unit/cluster, set of generation units/clusters
s, S	output power segment, set of segments
k, K	storage unit/cluster, set of storage units/clusters

Variables

C^{total}	total system operational cost (£)
$C_{i,t}^{flex}$	cost of flexibility provider (£)
$C_{i,t}^{inflex}$	cost of insufficient flexibility (£)
$U_{i,t}^{Gen_{SU/D}}$	generator start-up or shut-down index
$U_{i,t}^{Gen_{ON}}$	generator online/offline commitment state in BUC or the number of online units in MILP/LP
$U_{k,t}^{St_{ON}}$	storage generation mode index
e_t^{shed}	load shedding volume (MW)
$e_{i,t}^{over}$	over-generation volume (MW)
$FLEX_t$	sum of flexible components (MW)
$INFLEX_t$	sum of inflexible components (MW)
$FLEX_{k,t}^{St_{UP/DOWN}}$	storage discharging or charging power (MW)
D_t^{FLEX}, D_t^{INFLEX}	flexible, inflexible demand (MW)
$FLEX_{i,t}^{Gen}$	total power output of generator (MW)
$FLEX_{k,t}^{St}$	storage net power output (MW)
μ_t, Ω_t	curtailed wind, solar generation (MW)
$E_{k,t}^{St}$	storage energy content (MWh)
$P_{s,i,t}$	segment of generator output power (MW)
$f_{i,t}^{UP}$	primary frequency response contribution (MW)

$r_{i,t}^{Sec_spinUP/DOWN}$	secondary up or down spinning reserve contribution (MW)
$r_{k,t}^{Sec_stUP/DOWN}$	secondary up or down storage reserve contribution (MW)
$r_{i,t}^{Ter_spinUP/DOWN}$	tertiary up and down spinning reserve contribution (MW)
$r_{k,t}^{Ter_stUP/DOWN}$	tertiary up and down storage reserve contribution (MW)
$r_{i,t}^{Ter_standingUP}$	tertiary up standing reserve contribution (MW)
M^{cost}	operational cost metric (%)
M^{GM}	generation output metric (%)
M^{PP}	power deviation metric (%)
$M^{flexibility\ services}$	flexibility services deviation metric (%)
Parameters	
Δ_t	simulation time step (h)
$\Delta_{r_Sec}, \Delta_{r_Ter}$	secondary, tertiary reserve deployment timeframe (h)
T_{r_Sec}, T_{r_Ter}	maximum period for sustaining secondary, tertiary reserve (h)
W_t, PV_t	potential wind, solar production (MW)
D_t	total demand (MW)
C_i^{NL}	no-load cost of generation unit (£/h)
$C_i^{GenSU}, C_i^{GenSUE}$	generation unit start-up fuel, emission cost (£)
$FU_{s,i}^{Gen}, FU_{s,i}^{GenE}$	segment generation fuel, emission cost (£/MWh)
$C^{shed/over/curt}$	load shedding, over-generation or renewable generation curtailment penalties (£/MWh)
$P_i^{GenMAX/MIN}$	maximum or minimum output of generator (MW)
$F_k^{StMAX/MIN}$	maximum or minimum storage discharging power (MW)
$P_k^{StMAX/MIN}$	maximum or minimum storage charging power (MW)
$E_{k,t}^{StMAX/MIN}$	maximum or minimum storage energy level (MWh)
FC_i^{UP}	maximum frequency response capability (MW/unit)
γ_i^{UP}	frequency response function slope ratio
F_t^{UP}	primary frequency response requirement (MW)
PR_k^{StUP}	maximum percentage of storage capacity for each reserve provision (%)
G_i	maximum available number of units within the cluster
$V_i^{UP/DOWN}$	generation unit maximum ramping up or down rate (MW/h)

$T_i^{UP/DOWN}$	generation unit minimum on, off time (h)
$T_i^{stup/shdn}$	generation unit from off to on, on to off transition period(h)
$\eta_k^{c/dc}$	storage charging, discharging efficiency (%)
$R_i^{SecUP/DOWN}$	secondary up or down reserve requirement (MW)
$R_i^{TerUP/DOWN}$	tertiary up or down reserve requirement (MW)

I. INTRODUCTION

THE share of renewable energy sources (RES) in power systems around the world is rapidly increasing and this trend is expected to continue in the future as policy makers are setting more and more ambitious goals for emission reductions. The stochastic nature of RES has increased the uncertainty and variability historically present in power systems and, as the share of RES rises, so will the flexibility requirements, thus challenging the principles of today's power systems operation and planning. These arising flexibility requirements question the adequacy of existing market principles and services [1] and call for rethinking technical limitations of power system component (e.g., as to whether it is possible to lower the minimum generation level of power plants or increase their ramp rates [2]). The common final goal of flexibility requirement studies is to enhance systems capability to cope with the variability and uncertainty of renewable production, possibly also facilitated or supported by the integration of other low carbon technologies such as demand side response and storage [3], [4].

Assessment of the operational capability of the entire power system is a computationally intensive task, aimed at coordinating a large number of participants in order to maintain the supply-demand balance over different time horizons and services. Traditionally, unit commitment (UC) models have been developed to realistically represent day-ahead and intra-day system (and market) operation and multiple reserves scheduling, including reserves and in case frequency control requirements [5], [6]. The objective of these models is to schedule the resources available in the system in order to provide the desired service (in case through a relevant market) while minimising the operational cost and complying with a number of system and units' technical constraints such as minimum up and down times, ramp rates, minimum stable generation, and so forth. The complexity of these models increases with the integration of large share of RES, particularly if they mean to capture multiple time scales and services such as day-ahead and intra-day operation, primary and secondary frequency control, and tertiary reserves [7].

Based on the above, there is a continuous need for developing fast and reliable models capable of assessing power systems operational flexibility by capturing all relevant technical constraints and, by doing that, indirectly providing a backbone for planning future power systems. On the other hand, while the concept of UC is intrinsically associated to the need for scheduling individual power plants (and in case

energy storage), this level of detail may not be necessary for flexibility assessment and “generic” units could be modelled, provided that they are representative of the “flexibility issue” under investigation. In this light, this paper brings together the following contributions:

- A novel flexibility-oriented formulation of UC modelling approach using binary, integer and continuous variables (the latter two relaxing the problem and improving its solution’s computational performance) is presented, showing how power system operation can be seamlessly modelled in a unified way for the three classes of algorithms and specifically taking into consideration the technical constraints relevant to flexibility provision. In particular, the latter element is supported by a formulation that highlights the interaction between flexibility providers and insufficient flexibility indicators for increased flexibility requirements, particularly due to RES.
- Comparison of the scheduling and computational time performance of different classes of algorithms based on metrics specifically introduced and also considering the provision of different flexibility/reserve services from different plants (including energy storage). In this respect, clear demonstration is given of the merit of the proposed fast UC algorithm based on linear programming (LP) and benchmarked against classical mixed integer linear programming (MILP) approaches considering different generating unit clustering schemes as well as standard binary unit commitment (BUC), for assessing low-carbon systems’ flexibility.

The rest of this work is organized as follows: In Section II the current research on system operation modelling and flexibility assessment is investigated. Then, the flexibility-oriented unified formulation proposed for BUC, MILP and LP UC is described in Section III. Section IV introduces the metrics used for performance assessment of the three classes of algorithms. Section V contains the description of several case study applications for multiple RES and flexibility scenarios starting from a reduced system model of Great Britain. Section VI analyses the results to demonstrate the features of the proposed model, highlighting the pros and cons of an LP approach. Section VII finally concludes the paper.

II. CURRENT RESEARCH

A number of models have been proposed for fast assessment of power system operation, such as for instance the one in [8]; however, often these models do not sufficiently represent all the set of constraints and frequency control and reserve services needed to explicitly capture relevant system flexibility characteristics and metrics. With the advent of faster and faster optimization solvers commercially available, MILP methods have recently been developed in the direction of reducing computational time while maintaining reasonable accuracy in capturing all the relevant technical and economic aspects of system operation, particularly with inclusion of multiple services provided in different markets. In this respect, benefits of MILP modelling over more classical Lagrangian Relaxation techniques are well known [9]. However, the

complexity of MILP models is still relatively high when dealing with realistic large scale systems with hundreds of generators. In addition, in many cases when performing flexibility analysis there is a need to carry out studies over long time frames (for instance over a year in order to capture all the possible time varying correlation between RES supply and demand and therefore net load flexibility requirements), with fine resolution (30 minutes or below) and for multiple scenarios. Similarly, MILP models have proven to be computationally too expensive to be applied for planning purposes [10] thus requiring planners to make assumptions such as using representative week simulations instead of entire year analysis and by doing so, missing out on capturing all the seasonal characteristics and recognizing potential flexibility bottlenecks and requirements. Several techniques to increase the computational efficiency of MILP models for large scale systems have also been proposed, such as for instance [11], where an efficient formulation is proposed so as to require fewer binary variables and constraints. However, only one type of reserve is discussed there and the model is presented only for thermal units. The authors in [12], [13] use a clustering algorithm to simulate multiple power systems and interactions focusing on integration of electric vehicles in the presence of large share of renewable sources. However, they do not focus on assessing how accurate such an approach is relative to a full binary UC model or on the comparison between MILP and LP formulations. An effective approach has been proposed in [14], where a model based on integer variables rather than binary variables is used to represent clusters of thermal generators, thus increasing computational efficiency while also considering primary frequency regulation and tertiary reserve. On the same line, and with more focus on flexibility modelling, recent work [15] has demonstrated how clustering in MILP algorithms can significantly reduce the computational demand of UC studies with multiple services while maintaining very high level of accuracy. However, the authors do not elaborate as to how energy output and different services, such as frequency response and reserves, are allocated to specific clusters, nor are the results compared to those obtained when scheduling the same flexibility services using BUC. A similar operational modelling approach has been applied to power system generation planning [16]; however, again even with reduced requirements in computational time, analysing multiple services and multiple scenarios requires focusing on representative weeks only. Interesting work investigating which constraints can be relaxed for planning future low-carbon systems is presented in [17], but again not discussing how this relaxation might affect the allocation of specific services associated to flexibility provision (such as primary frequency response (PFR) and secondary and tertiary reserves) from a single or a cluster of units. In addition, energy storage providing multiple services is not included in the assessment. Hence, generally speaking no analysis could be found in the literature with respect to the use of LP approaches for UC, and especially for flexibility analysis purposes.

III. UNIFIED UNIT COMMITMENT MODEL FOR FLEXIBILITY ANALYSIS IN SUSTAINABLE POWER SYSTEMS

A. Objective function

While, as mentioned above, there are a number of UC formulations proposed in the literature, the one presented here presents a flexibility-oriented mathematical description distinguishing between costs associated to “flexible providers” and costs associated to presence of “insufficient flexibility” or “flexibility requirements” (which together eventually represent drivers for additional flexibility). The objective function is expressed as in (1), where it is emphasised how power system operation is driven by the minimisation of total operational cost (C^{total}), expressed as the sum of cost of flexibility providers (C^{flex}) and “penalty” costs for insufficient flexibility (C^{inflex}):

$$\text{minimise } C^{total} = \sum_{t \in T, i \in I} \left[C_{i,t}^{flex} + C_{i,t}^{inflex} \right] \quad (1)$$

In (1), C denotes operational cost, t the time periods during the considered interval T , and i refers to a specific unit (in the case of BUC) or cluster of units¹ (in the case of MILP or LP) in set I .

The first part of the objective function (C^{flex}), is capturing all flexible units that can adjust their output according to market or system requirements. Traditionally, these are flexible conventional generators and more recently storage (which are the focus of this paper), but in the context of future power systems new sources also arise, e.g., demand response and controllable wind power plants [17]. The operational cost of each generating unit or cluster i consuming fuel as input is modelled as a sum of: start-up cost (C_i^{GenSU})² times the number of start-ups ($U_{i,t}^{GenSU}$); operating cost, which includes the no load cost ($C_{i,t}^{NL}$) and the segment fuel cost ($FU_{s,i}^{Gen}$) (a concept used for piece-wise linearization of the generator cost curve); and carbon emission cost (C_i^{GenSUE} and $FU_{s,i}^{GenE}$), where Δ_t indicates the simulation time step. This is shown in Eq (2):

$$C_{i,t}^{flex} = \left[C_{i,t}^{NL} \cdot U_{i,t}^{GenON} + \sum_{s \in S} P_{s,i,t} \cdot (FU_{s,i}^{Gen} + FU_{s,i}^{GenE}) \right] \cdot \Delta_t + (C_i^{GenSU} + C_i^{GenSUE}) \cdot U_{i,t}^{GenSU} \quad (2)$$

On the other hand, the second part of the objective function (1) captures all the components characterizing “insufficient” flexibility in the system, namely, penalised by the cost of load shedding (C^{shed}), over-generation (C^{over}), and renewable curtailment (C^{curt}), as in Eq (3). The variables e_t^{shed} , $e_{i,t}^{over}$, μ_t and Ω_t quantify the energy of load shed, over-generation, and wind and photovoltaic (PV) curtailment, respectively.

¹ As discussed later, the equations below can be seamlessly applied to both individual units and cluster of units, and we will refer to either of them depending on the specific case.

² This term includes the costs potentially associated to additional fuel consumptions and emissions (in the case emissions are penalized).

$$C_{i,t}^{inflex} = \left[C^{shed} e_t^{shed} + C^{over} \sum_{i \in I} e_{i,t}^{over} + C^{curt} (\mu_t + \Omega_t) \right] \Delta_t \quad (3)$$

B. Energy balances

As shown in Eq (4), the “flexibility providers” are the controllable conventional generation units ($FLEX_{i,t}^{Gen}$) and storage ($FLEX_{k,t}^{st}$), along with, in general (although not discussed in detail here), flexible demand (D^{FLEX}). Some research papers also interpret curtailment of RES (μ_t and Ω_t) as an additional provider of flexibility for the purposes of economic system operation [17], [18]. However, in this paper we prefer to treat RES curtailment as an indicator of insufficient flexibility in the system, in line with current operational practices of system operators. Therefore, the non-flexible energy balance terms in Eq (5) include inflexible demand and components indicating insufficient capability of the system to avoid over-generation ($e_{i,t}^{over}$), load shedding (e_t^{shed}), and harness all the renewable energy that could be potentially produced (W_t and PV_t). The over-generation variables are used to ensure sufficient ability to provide reserve, specifically during periods of high renewable generation and low demand. In this sense the over-generation variable ($e_{i,t}^{over}$) indicates if the output level of the online units needs to be kept higher than demand in order to satisfy the downward reserve requirement. The definition of the generation-side flexibility providers is shown in Eq (6), and it is constrained by the maximum output power (P_i^{GenMAX}) and minimum stable generation (MSG) level (P_i^{GenMIN}) of each generator, as shown in (7). In (8), the system’s total load is modelled as the sum of flexible and inflexible shares. The operational constraints of storage, which is used in this work for energy arbitrage purposes, are modelled in Eqs. (9)–(13). More specifically: Eq. (9) illustrates the flexibility that storage can provide, with $FLEX_{i,t}^{stDOWN}$ representing downward flexibility provision while it is charging and $FLEX_{i,t}^{stUP}$ upward flexibility provision of discharging operation; the energy content of storage is constrained in (10) by energy inflows considering charging and discharging efficiencies (η_k^c, η_k^{dc}) as well as the energy content in the previous time step ($E_{k,t-1}^{st}$); Eq (11) describes that the energy content of storage is constrained by its minimum and maximum value ($E_{k,t}^{stMIN}, E_{k,t}^{stMAX}$); both charging and discharging operations are constrained by minimum and maximum powers, as in (12) and (13); finally, the two flexible and inflexible components ($FLEX$ and $INFLEX$) need to be kept in equilibrium, as from (14). It is also worth highlighting that, without loss of generality, in the case studies Pumped-Hydro Storage Power Plant (PSPP) has been taken as typical example of storage, so that charging and discharging can be associated to turbine and pumping operation respectively, which is also reflected in the commitment variables and constraints in (12) and (13).

$$FLEX_t = \sum_{i \in I} FLEX_{i,t}^{Gen} + \sum_{k \in K} FLEX_{k,t}^{st} - D^{FLEX}_t \quad (4)$$

$$INFLEX_t = D^{INFLEX}_t + \sum_{i \in I} e_{i,t}^{over} - e_t^{shed} - (W_t - \mu_t) - (PV_t - \Omega_t) \quad (5)$$

$$FLEX_{i,t}^{Gen} = \sum_{s \in S} P_{s,i,t} \quad (6)$$

$$P_i^{Gen_{MIN}} \cdot U_{i,t}^{Gen_{ON}} \leq FLEX_{i,t}^{Gen} \leq P_i^{Gen_{MAX}} \cdot U_{i,t}^{Gen_{ON}} \quad (7)$$

$$D^{FLEX}_t + D^{INFLEX}_t = D_t \quad (8)$$

$$FLEX_{k,t}^{st} = FLEX_{k,t}^{st_{UP}} - FLEX_{k,t}^{st_{DOWN}} \quad (9)$$

$$E_{k,t}^{st} = \left(\frac{1}{\eta_k^c} FLEX_{k,t}^{st_{DOWN}} - \eta_k^{dc} FLEX_{k,t}^{st_{UP}} \right) \Delta_t + E_{k,t-1}^{st} \quad (10)$$

$$E_k^{st_{MIN}} \leq E_{k,t}^{st} \leq E_k^{st_{MAX}} \quad (11)$$

$$F_k^{st_{MIN}} \cdot U_{k,t}^{st_{ON}} \leq FLEX_{k,t}^{st_{UP}} \leq F_k^{st_{MAX}} \cdot U_{k,t}^{st_{ON}} \quad (12)$$

$$P_k^{st_{MIN}} \cdot (1 - U_{k,t}^{st_{ON}}) \leq FLEX_{k,t}^{st_{DOWN}} \leq P_k^{st_{MAX}} \cdot (1 - U_{k,t}^{st_{ON}}) \quad (13)$$

$$FLEX_t = INFLEX_t \quad (14)$$

C. Reserves

All the flexible components in (4) may in principle contribute to provision of reserves. In this paper, the focus is put on generation units as reserve providers while more specific modelling and analysis of reserves from storage and flexible demand are left to future extension. The proposed model captures three types of reserve, namely, primary frequency response, secondary spinning reserve and tertiary spinning and standing reserves [19].

1) Primary frequency response

The total primary frequency response (F_t^{UP}) that needs to be available in case of contingency is typically a constant value, for instance defined in the Great Britain based on the largest generating unit. This needs to be provided by online generators ($f_{i,t}^{UP}$), as in (15), depending on their characteristics. Typically, thermal units can provide around 10% of their nominal capacity (FC_i^{UP}) within the time frame of primary reserve.

$$\sum_{i \in I} f_{i,t}^{UP} = F_t^{UP} \quad (15)$$

Provision of primary frequency response is constrained by the spare capacity of the generator excluding secondary and tertiary spinning reserve provision, the upper limit of the response contribution of individual generators, and the slope function (γ_i^{UP}) when the output power is close to maximum/minimum value, as shown in Fig. 1 and formulated in (16) and (17). Typical technical constraints are listed in Table 1 of the case study using similar values as in [20].

$$f_{i,t}^{UP} \leq U_{i,t}^{Gen_{ON}} \cdot FC_i^{UP} \quad (16)$$

$$\frac{f_{i,t}^{UP}}{\gamma_i^{UP}} + r_{i,t}^{Sec_spin_{UP}} + r_{i,t}^{Ter_spin_{UP}} \leq P_i^{Gen_{MAX}} \cdot U_{i,t}^{Gen_{ON}} - FLEX_{i,t}^{Gen} \quad (17)$$

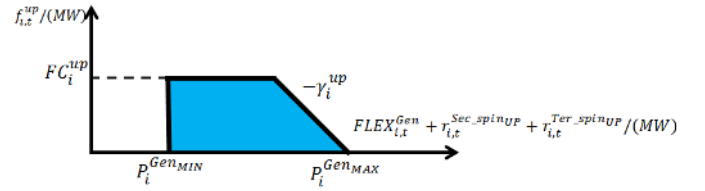


Fig. 1. Limitations of the generator response contributions.

2) Secondary reserve

Secondary up and down reserves (SUR and SDR) are to be provided by spinning generation (online) units ($r_{i,t}^{Sec_spin}$) and pumped-hydro storage units ($r_{i,t}^{Sec_st}$), as shown in Eq. (18)-(19). The secondary reserve requirement is calculated here based on the maximum existing generation plant capacity in the system and the potential volatility of the load and RES, which are normally a result of uncertainty and variability [10], [21]. An interesting work elaborating on variability and uncertainty time frames of wind generation can be found in [22].

The spinning reserve contribution of each generating unit is limited by its ramping ability associated with the required timeframe of reserve deployment (Δ_{r_Sec}), as defined in (20) and (21). In addition, the generator's spare capacity also needs to be considered. A spare capacity constraint for upward reserve is already considered in the primary up frequency response in Eq. (17). Eq. (22) adds spare capacity constraint for downward spinning reserves too. The reserve contribution of storage units is constrained by the minimum and maximum energy content and the maximum period required for sustaining the secondary and tertiary reserves (T_{r_Sec} , T_{r_Ter}), as shown in (23) and (24). In addition, the maximum percentage of storage capacity used for each reserve provision ($PR_k^{st_{UP}}$) is also limited in Eqs. (25) and (26). Further, the amount of reserve provision from storage is calculated based on the operating mode (turbine or pumping). More specifically, as shown in Eq (27), the maximum amount of up reserve provision is calculated either with the spare capacity in turbine mode or storage input power in pumping mode. A similar principle is used to constrain the downward reserve provision by using spare capacity in pumping mode or generation output in turbine mode, as defined in (28). The storage model used here is based on the work carried out in [23].

$$\sum_{i \in I} r_{i,t}^{Sec_spin_{UP}} + \sum_{k \in K} r_{k,t}^{Sec_st_{UP}} = R_t^{Sec_{UP}} \quad (18)$$

$$\sum_{i \in I} r_{i,t}^{Sec_spin_{DOWN}} + \sum_{k \in K} r_{k,t}^{Sec_st_{DOWN}} = R_t^{Sec_{DOWN}} \quad (19)$$

$$r_{i,t}^{Sec_spin_{UP}} \leq V_i^{UP} \cdot U_{i,t}^{Gen_{ON}} \cdot \Delta_{r_Sec} \quad (20)$$

$$r_{i,t}^{Sec_spin_{DOWN}} \leq V_i^{DOWN} \cdot U_{i,t}^{Gen_{ON}} \cdot \Delta_{r_Sec} \quad (21)$$

$$r_{i,t}^{Sec_spin_{DOWN}} + r_{i,t}^{Ter_spin_{DOWN}} \leq FLEX_{i,t}^{Gen} - P_i^{Gen_{MIN}} \cdot U_{i,t}^{Gen_{ON}} \quad (22)$$

$$r_{k,t}^{Sec_st_{UP}} \cdot T_{r_Sec} + r_{k,t}^{Ter_st_{UP}} \cdot T_{r_Ter} \leq E_k^{st} - E_k^{st_{MIN}} \quad (23)$$

$$r_{k,t}^{Sec_st_{DOWN}} \cdot T_{r_Sec} + r_{k,t}^{Ter_st_{DOWN}} \cdot T_{r_Ter} \leq E_k^{st_{MAX}} - E_k^{st} \quad (24)$$

$$\{r_{k,t}^{Sec_stUP}, r_{k,t}^{Ter_stUP}\} \leq PR_k^{stUP} \cdot F_k^{stMAX} \quad (25)$$

$$\{r_{k,t}^{Sec_stDOWN}, r_{k,t}^{Ter_stDOWN}\} \leq PR_k^{stDOWN} \cdot P_k^{stMAX} \quad (26)$$

$$r_{k,t}^{Sec_stUP} + r_{k,t}^{Ter_stUP} \leq F_k^{stMAX} \cdot U_{k,t}^{stON} - FLEX_{k,t}^{stUP} + FLEX_{k,t}^{stDOWN} \quad (27)$$

$$r_{k,t}^{Sec_stDOWN} + r_{k,t}^{Ter_stDOWN} \leq P_k^{stMAX} \cdot (1 - U_{k,t}^{stON}) - FLEX_{k,t}^{stDOWN} + FLEX_{k,t}^{stUP} \quad (28)$$

3) Tertiary reserve

Similar logic as for secondary reserve can be applied for tertiary up and down reserves (TUR and TDR), with the variability and uncertainty of load and RES generation depending on the relevant deployment time of tertiary reserve (Δ_{r_Ter}). Again, all participating units, namely spinning ($r_{i,t}^{Tert_spin}$) and standing ($r_{i,t}^{Tert_standing}$) and pumped-hydro storage ($r_{i,t}^{Tert_st}$) can be scheduled so as to leave ‘‘headroom’’ for provision of tertiary reserve, as in (29) and (30).

$$\sum_{i \in I} r_{i,t}^{Ter_spinUP} + \sum_{i \in I} r_{i,t}^{Ter_standingUP} + \sum_{k \in K} r_{k,t}^{Ter_stUP} = R_i^{TerUP} \quad (29)$$

$$\sum_{i \in I} r_{i,t}^{Ter_spinDOWN} + \sum_{k \in K} r_{k,t}^{Ter_stDOWN} = R_i^{TerDOWN} \quad (30)$$

$$r_{i,t}^{Ter_spinUP} \leq V_i^{UP} \cdot U_{i,t}^{GenON} \cdot \Delta_{r_Ter} \quad (31)$$

$$r_{i,t}^{Ter_spinDOWN} \leq V_i^{DOWN} \cdot U_{i,t}^{GenON} \cdot \Delta_{r_Ter} \quad (32)$$

D. Unified UC formulation

Traditional UC models [24] use binary variables for modelling whether a unit is coming online, whether it is online and whether it is turning offline at a given time; this is modelled by $U_{i,t}^{GenSU}$, $U_{i,t}^{GenON}$ and $U_{i,t}^{GenSD}$. In [25], the author provides a detailed analysis of UC and compares the approach using binary and integer variables in a clustered UC model; excellent correlation of results of the models is shown. In the UC formulation proposed here, $U_{i,t}^{GenSU}$, $U_{i,t}^{GenON}$ and $U_{i,t}^{GenSD}$ can seamlessly be binary, integer or continuous variables, and this indeed allows writing a compact and unified formulation for the BUC, MILP and LP algorithms, respectively. In particular, and focusing our description on the clustered cases, the number of the online units $U_{i,t}^{GenON}$ is defined as in (33). Further, the number of online units at time t should be less than the total available number of units within the cluster (G_i) minus the offline units, as from (34). More specifically, the constraints on the number of online units in (34) can be explained by considering that the online units at time t need to be fewer than the units in the cluster minus the number of units which are shut down within $T_i^{down} + T_i^{stup} + T_i^{shdn} - 1$ before t . Similarly, the online units need to be more than the units that start up within $T_i^{up} - 1$ before t (Eq. (35)).

$$U_{i,t}^{GenON} - U_{i,t-1}^{GenON} = U_{i,t}^{GenSU} - U_{i,t}^{GenSD} \quad (33)$$

$$U_{i,t}^{GenON} \leq G_i - \sum_{\tau \in [1, t - (T_i^{shdn} + T_i^{DOWN} + T_i^{stup} - 1)]} U_{i,\tau}^{GenSD} \quad (34)$$

$$U_{i,t}^{GenON} \geq \sum_{\tau \in [1, t - (T_i^{UP} - 1)]} U_{i,\tau}^{GenSU} \quad (35)$$

In (34) and (35), T_i^{DOWN} represents the minimum down-time, T_i^{UP} the minimum up-time, and T_i^{stup} and T_i^{shdn} the transition times from and to off-status, respectively. These transition times are calculated from the unit’s ramping up (V_i^{UP}) or down (V_i^{DOWN}) rates, as shown in (36) and (37), respectively:

$$T_i^{stup} = \text{ceil}(P_i^{GenMIN} / V_i^{UP}) \quad (36)$$

$$T_i^{shdn} = \text{ceil}(P_i^{GenMIN} / V_i^{DOWN}) \quad (37)$$

The reason for explicitly considering these transition times is that when a unit switches from online to offline state, the unit takes time to ramp down to its final off-status. Similarly, when a unit is switched on it takes a certain time before reaching its MSG level and therefore its online state. Hence, it is necessary to include ramping times too, and not minimum up and down times only, to fully account for the actual constraints relevant to transitions to and from the off-status.

The ability of a single unit as well as unit cluster to change operational point between two time steps is bound by its ramping constraint, which is in turn related to the number of the existing committed units and the new units coming online. In particular, the increase in cluster’s output is limited by the ramping limits of the units already online at the previous time step plus the maximum output power of the new online units. This is defined in (38):

$$FLEX_{i,t}^{Gen} - FLEX_{i,t-1}^{Gen} \leq U_{i,t-1}^{GenON} \cdot V_i^{UP} \cdot \Delta_t + U_{i,t}^{GenSU} \cdot V_i^{UP} \cdot T_i^{stup} \quad (38)$$

In addition, the power increase that can be achieved is also limited by the maximum power of the units already online, as from (39):

$$FLEX_{i,t}^{Gen} - FLEX_{i,t-1}^{Gen} \leq U_{i,t-1}^{GenON} \cdot P_i^{GenMAX} - FLEX_{i,t-1}^{Gen} + U_{i,t}^{GenSU} \cdot V_i^{UP} \cdot T_i^{stup} \quad (39)$$

Similarly, the decrease of the power output, formulated in (33), is constrained by the down-ramping limits of the units already online at the previous time step plus the output power of the units that have been shut down at the given time step t .

$$FLEX_{i,t}^{Gen} - FLEX_{i,t-1}^{Gen} \leq U_{i,t-1}^{GenON} \cdot V_i^{DOWN} \cdot \Delta_t + U_{i,t}^{GenSD} \cdot V_i^{DOWN} \cdot T_i^{shdn} \quad (40)$$

The mathematical model in equations (33) – (40) describes the operation of the flexible generating units, with, as aforementioned, $U_{i,t}^{GenON}$, $U_{i,t}^{GenSU}$ and $U_{i,t}^{GenSD}$ potentially being binary, integer or continuous variables, depending on the algorithm used (BUC, MILP or LP, respectively). This is again the core of the unified formulation that we propose here, which can seamlessly model the three classes of algorithm.

The following sections present and compare the results of all three approaches, highlighting how for LP relevant metrics (introduced next) are within limited margins of error but while running potentially several thousand times faster.

TABLE I
CONVENTIONAL GENERATOR CHARACTERISTIC IN CLUSTERING

Technology	No. generators (units)	P^{GenMAX}/P^{GenMIN} (MW/unit)	C^{NL} (£/(unit.h))	FJ^{Gen} (£/MWh)	C^{Gensu} (£/unit)	T_{UP}/T_{DOWN} (hour(s))	V_{UP}/V_{DOWN} (MW/h)	γ^{UP}	FC^{UP} (MW/unit)
Coal	13	1585 / 634	5869	30.6	47538	4 / 4	634 / 634	0.3	158.5
Nuclear	1	9373 / 4687	5680	7.1	-	-	170 / 120	-	-
CCGT	39	770 / 308	7554	40.3	13231	6 / 4	462 / 462	0.4	77

IV. METRICS FOR LP, MILP AND BUC SCHEDULING ALGORITHM COMPARISON

In order to assess the performance of the LP, MILP and BUC methods, specific metrics have been identified with respect to system operational cost and accuracy of generation mix (GM) (as in [15]), as well as accuracy of flexibility services (e.g., frequency response and secondary and tertiary up reserves) allocation (as a novel contribution of this work). These metrics have been selected owing to their ability to assess how two algorithms, for example MILP and BUC, can schedule energy and different types of reserves and allocate them to different plants. The differences between the two algorithms should thus emerge clearly when considering the metrics proposed, namely, differences in total cost, in energy generation share per plant type, in type of generation providing energy and each reserve at every time step, and in simulation time. All three methods are compared with each other to obtain information in relative terms.

A. Operational cost

This metric (M^{cost}) is used to determine the relative difference of the total system operational cost between two methods, as in (41):

$$M_{Method A/B}^{cost} = \frac{C_{Method A}^{total} - C_{Method B}^{total}}{C_{Method B}^{total}} \quad (41)$$

B. Generation mix

The following metric (M^{GM}) is used to capture the average difference of energy shares in between algorithms. The energy share supplied by each type of conventional generator is calculated; then the average absolute difference of these shares between two algorithms is determined, as from (42).

$$M_{Method A/B}^{GM} = Mean_{i \in I} \left| \frac{\left(\sum_{t \in T, s \in S} P_{s,i,t,A} \right)}{\left(\sum_{t \in T, s \in S, i \in I} P_{s,i,t,A} \right)} - \frac{\left(\sum_{t \in T, s \in S} P_{s,i,t,B} \right)}{\left(\sum_{t \in T, s \in S, i \in I} P_{s,i,t,B} \right)} \right| \quad (42)$$

C. Power and flexibility services provision

In order to compare the performance of different simulation methods in terms of generation and flexibility service provisions, the mean deviation is calculated for power provision (PP) of all generation ($M_{Method A/B}^{PP}$) and flexibility services ($M_{Method A/B}^{flexibility\ services}$) for all generator types and at each time steps. For example, Eq. (43) is used to calculate the relative total power output deviation:

$$M_{Method A/B}^{PP} = Mean_{t \in T, i \in I} \left| \frac{\sum_{s \in S} P_{s,i,t,A} - \sum_{s \in S} P_{s,i,t,B}}{\sum_{s \in S, i \in I} P_{s,i,t,B}} \right| \quad (43)$$

For the flexibility services, the calculation method is the same, but just replacing the power output with the contribution of each cluster on primary frequency response ($f_{i,t}^{UP}$), secondary reserve ($r_{i,t}^{Sec_spin}$) or tertiary reserve ($r_{i,t}^{Ter_spin}$).

D. Simulation time

The simulation time represents the time taken by the optimization solver, excluding input data loading and output data writing processes.

TABLE II
CLUSTERS WITH DIFFERENT SEGMENT SIZES

Segment size (MW)	Number of clusters per each type of plant			
	Total	Coal plant	Nuclear plant	CCGT
all	3	1	1	1
2000	5	2	1	2
1000	8	4	1	3
500	12	6	1	5
250	19	10	1	8
100	28	10	1	17

V. CASE STUDY DESCRIPTION

A. Base system description

In order to provide a comparison of the performance of LP, MILP and BUC algorithms, a reduced version of the current Great Britain generation portfolio is modelled considering historical demand data [26] and wind generation data [27] with half-hour resolution. Most of the Great Britain power plant capacity information is taken from [28]. However, only 30GW out of actual 35GW Combined Cycle Gas Turbine (CCGT) plant, 25 GW out of actual 30GW coal plant, and 9.4GW nuclear plant capacity values are presented there, making up 62 conventional plant units. Other types of plants such as biomass, oil and diesel plants are omitted due to their relatively small capacity (2.6 GW in total) and minor generation contribution to the system. The Open Cycle Gas Turbine (OCGT) plants are considered here as a standing reserve provider for the tertiary up reserve and are not included in the simulation of the generation portfolio. The original demand profile has a 56.6 GW peak, but considering the smaller generation capacity in the model, it has been scaled down to a profile with a 49.7 GW peak. This reduced plant model has been simulated according to a classical BUC. The conventional power plant characteristics for different clusters by fuel type are shown in Table II. The nine nuclear power plants are clustered into one equivalent unit as it is assumed that they will operate as “must-run”, therefore without start-up or shut down actions. The different power plants characteristics such as start-up cost, no load cost, ramping rates, etc. are scaled up/down based on their capacities. The frequency response requirement of the system

TABLE III
PROBLEM SIZE FOR ONE AND OBJECTIVE VALUE FOR THE DIFFERENT ALGORITHMS, ONE WEEK SIMULATION

Model	No. constraints	No. variables	No. non-zeros	No. global entities	Optimal objective value (£M)
BUC	384393	337344	1625789	98784	156.320
MILP, full size, 3 clusters	28216	23520	203333	4032	156.736
MILP, 2000MW, 5 clusters	40980	31584	256853	6048	156.711
MILP, 1000MW, 8 clusters	60126	43680	336480	9072	156.522
MILP, 500MW, 12 clusters	85654	59808	443520	13104	156.365
MILP, 250MW, 19 clusters	130328	88032	630187	20160	156.336
MILP, 100MW, 28 clusters	187766	124329	876904	29232	156.327
LP, full size, 3 clusters	28216	23520	203333	0	156.144

is fixed to 1900 MW based on National Grid regulation [29] and the wind and solar forecast error used for reserve requirement determination is obtained from [21]. The standing reserve provided by OCGT is fixed to 1524 GW based on National Grid information [30].

For the metric comparison, a week simulation (starting from 2013.04.14) with half-hour time steps is considered. The demand pattern of this week has been chosen as having the least deviation from the average demand pattern of 13 weeks in the spring season of the year.

B. Case studies

1) Size of clusters

To assess the impact of clustering different plants of each type when running MILP and LP algorithms, different numbers of clusters have been considered based on fuel type and “segment size”. The total number of clusters thus range from 3 (units clustered based on fuel type only, as in Table I) to 28. In the latter case, the units have been clustered based on fuel type and 100MW segments; this means that plants with same fuel type and size in the range of 0÷100 MW have been clustered into one group, plants with same fuel type and capacity in the range 100÷200MW into another group, and so forth. The details of the different clusters with different “segment size” are shown in Table II.

2) Renewable energy integration scenarios

With higher RES penetration level, flexibility requirements (for reserves, in particular) will change; this, along with the different levels of energy produced by RES, will in turn affect the operations and utilisation levels of different generators. Therefore, it is important to investigate the impact of different RES scenarios on simulation result accuracy, particularly in the comparison between MILP and LP, which is the core of this work. Four scenarios have thus been identified (adapted from National Grid’s Gone Green scenario [31]):

- *Traditional system*, with no RES;
- *Current system*, with 11.5GW wind (8.86% energy penetration level) and 5GW PV (1.26% energy level);
- *2020 system*, with 21.7GW wind (16.32% energy level) and 6.7GW PV (1.69% energy level);
- *2024 system*, with 30.9GW wind (23.24% energy level) and 8.2GW PV (2.07% energy level).

3) Scenarios with different levels of system flexibility

To further demonstrate comparisons of MILP and LP models, the four RES scenarios introduced above have also been analysed considering three different system characteristics in terms of flexibility, as from below:

- The first (“*Normal*”) flexibility scenario is the base case with characteristics given in Table I.
- In the second (“*Flexible*”) flexibility scenario, the system’s flexibility has been changed by altering some key technical constraints, namely, system’s MSG, ramp rates, and minimum down times. In this case, following [32], the system’s MSG is reduced by reducing the MSG of coal and CCGT generators from 40% to 30%; the ramp rate of coal and CCGT units is increased by 66%; the minimum down time of coal units is reduced from 4 to 2 hours;
- The third (“*Flexible with storage*”) flexibility scenario is the same as the second one with the addition of energy storage for energy balancing and secondary and tertiary reserve provision, which are the most likely services that pumped-hydro plants would provide [32]. This also means, in practice, that the system’s MSG is further reduced and its ramp rate further increased. As aforementioned, a PSPP has been selected as storage, with the current Great Britain characteristics [34], namely, 2.8 GW of installed capacity, and 27.6 GWh of energy volume [35], and assuming 87% storage discharging (turbine) efficiency and 87% storage charging (pumping) efficiency. The maximum percentage for reserve provision (PR_k^{StUP}) is set to 50%.

C. Simulation setup and problem size

The optimization solver applied here is FICO XPRESS 7.6 mixed-integer solver [36]. The input, output and solver setting are configured through MATLAB 2013a. The Mixed-Integer Program (MIP) Gap is set to 0.1%. The time step of simulation profile is half hour and the data is processed on a weekly basis (so as to also capture typical arbitrage operations of storage, in case). This results in 336 time steps in one optimization program. The duration of one simulation is limited to 10 hours. The computer used in the simulation has an 8 core CPU with 3.7 GHz clock speed (Intel i7-3770) and 16 GB RAM.

The number of constraints, variables, non-zeros and global entities (i.e., binary or integer variables) for different UC models before pre-solving, as well as their achieved objective values, are shown in Table III for a one-week simulation. It

can be seen that clustering similar units and using integer variables for the clusters of units with similar technical constraints (and thus similar behaviour) is an effective way to reduce the problem size, passing from binary to integer variables. This should also significantly reduce the solution time, as will be shown later.

VI. ANALYSIS OF THE RESULTS

A. Base case: clustering by capacity size

The performance of the three algorithms for the base case (taken here as “current system” with “normal flexibility”) are compared for the different cluster sizes of Table II in Fig. 2 (showing the computational speed, in logarithmic scale, of the different algorithms vs different cluster sizes), while Fig. 3 shows the performance comparison according to the different metrics introduced in Section IV per cluster group. The values in the brackets in the legend in Fig. 2 and Fig.3 represent the total number of clusters.

From Fig. 2, it can be noticed that the BUC simulation hits the predefined time constraint which is 10 hours. The simulation time of MILP exponentially increases with increasing number of clusters and reaches the simulation time limit when the group number of clusters increases to 8. However, for the first two clustering levels the MILP algorithm runs 35800 and 4300 times faster than the BUC (taking as reference the 10h of time limit that the BUC hits, without converging to the pre-set 0.1% gap – see also Section VI.D for further discussion). This confirms that clustering similar units and using integer variables for the clusters of units with similar technical constraints (and thus similar behaviour), besides being an effective way to reduce the problem size, also significantly reduces the solution time. Furthermore, the LP algorithm can be 3 times (for the 3-clusters case) to 16500 times (8-clusters case) faster than MILP, thus showing the computational benefits of further relaxing integer into continuous variables.

It can be seen from Fig. 3³ that all the metrics considered for different cluster groups feature errors smaller than 1.6% and the largest error occurs when comparing secondary up reserve between LP and MILP algorithms. The secondary up requirement has the largest value among all three flexibility services in this case. Also, the tertiary up reserve requirement is always higher than the requirement of secondary up reserve, but the former can be partially provided by standing reserve; that is why the spinning reserve of tertiary up reserve is smaller than the secondary up reserve. For the MILP, its performance relative to BUC generally improves with more clusters. In addition, it can be noticed that when the number of clusters increases from 19 to 28, the MILP performance doesn’t improve noticeably. Moreover, it is interesting to notice that, based on the result of the cost metric, the MILP gives higher total costs (by 0.004% to 0.27%) relative to

BUC⁴. This is because the BUC can commit units of smaller size and with wider choice relative to the units in MILP, which are characterised by their average capacity; hence, in spite of being more constrained, BUC results into cheaper commitment, as no-load and start-up costs can be lower. With regards to LP, since the online number of units is relaxed to a continuous value and the algorithms is less constrained, as it could be expected it is characterised by lower commitment cost relative to both MILP and BUC. In addition, the LP simulation maintains similar performance across different clustering cases when using the BUC performance as a benchmark (in the order of -0.15% for the cost metric, 0.5% for primary frequency response, 0.8% for secondary up reserve, and 0.35% for tertiary up reserve). In fact, the LP algorithm allows generators to be “partially online”, so the change in cluster size will not have a significant impact on the simulation result. Furthermore, the difference between the performances of MILP and LP decreases with increasing cluster number, as the maximum value of any metric reduces from 1.5% to 0.8% in this case.

The above findings indicate that relaxing the MILP discrete variables to continuous ones (making the problem an LP one), as proposed here, can result into substantial computational efficiency gain without losing significant accuracy in the results relative to a full BUC. These benefits may be of high value in many cases, particularly when running strategic flexibility studies for large scale systems and long temporal horizons. Hence, in order to explore more the robustness of the MILP-to-LP relaxation, more studies are discussed below.

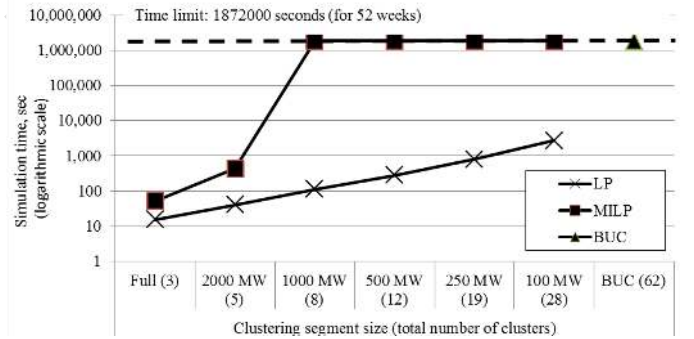


Fig. 2. Computation time of LP and MILP with different clusters and BUC: annual simulation on a weekly basis, “current system” with “normal flexibility”.

B. Algorithm comparison for different RES scenarios

As aforementioned, increasing the number of clusters reduces the differences between MILP and LP. Therefore, in order to be conservative in the comparison between a relaxed LP and a “non-relaxed” MILP, only the “worst case” with three clusters (by fuel type) is considered below⁵. On the other hand, the 3-cluster LP model also proves to be very close to a full BUC in terms of performance, within 0.8% in any metric value, as from Fig. 3.

³ It has to be noted that, since during the simulation it was found that the BUC could hardly converge to the desired 0.1% MIP gap due to the size of the problem, in order to allow a like-for-like performance comparison and enable the results to converge to a uniform MIP gap the three algorithms have been tested on a daily simulation instead of a weekly one. More discussions on this aspect are also provided in Section VI.D.

⁴ However, it has to be noticed that since a 0.1% MIP gap has been used, this is the highest accuracy that is meaningful in the comparisons carried out by means of the cost metric.

⁵ However, although the details are not shown here for brevity, our further studies indicate that other cluster sizes would bring similar results and trends as in Fig. 3.

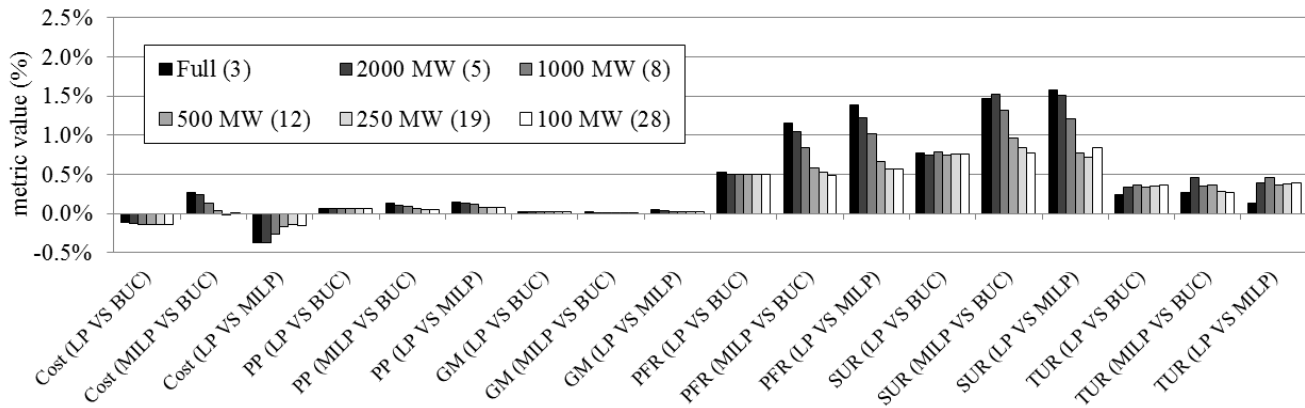


Fig. 3. Value of metrics in the simulations of LP and MILP with different clusters and BUC: week simulation on a daily basis, “current system” with “normal flexibility”.

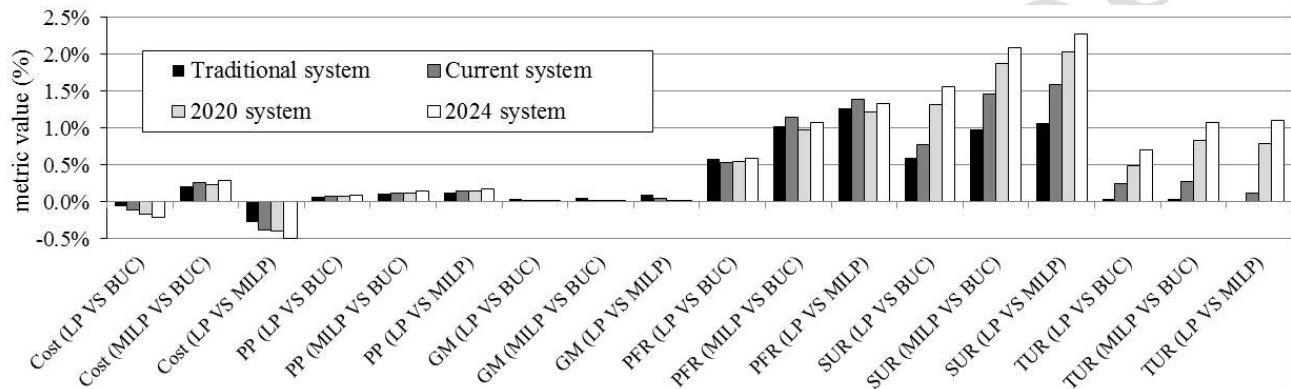


Fig. 4. Value of metrics in the simulations of LP and MILP with three generation clusters and BUC: weekly simulation on a daily basis, at different RES penetration levels with “normal flexibility”.

The values of the considered metrics for different RES penetration scenarios are shown in Fig. 4. The results generally indicate that the error increases with the RES penetration. This increase is particularly evident in the secondary and tertiary reserve metrics, since their requirements are also increasing with the renewable penetration level. Again, the highest error is incurred for secondary up reserve at the 30.9 GW wind capacity case, which gives around 2.2% error between LP and MILP. On the other hand, increasing RES penetration only causes minor increase on the output power deviation and generation output assessment, with these metrics staying within 0.15%.

As a key finding, it can be noticed that the LP gives a better or similar performance relative to MILP across all the metrics at different RES penetration levels, if using the BUC as benchmark. Therefore, again these studies suggest that there is only a minor loss of accuracy in adopting the proposed relaxed LP as opposed to a commonly used clustered MILP for the purposes of assessing flexible system operation, while creating noticeable benefits by reducing the required computational time.

C. Analysis of different scenarios of system’s flexibility

Further studies have been run to specifically compare the performance of LP and MILP approaches, again with three clusters. As a result, the cost, primary frequency response and secondary up reserve metrics are shown in Figs. 5, 6 and 7, respectively, for all combinations of RES and flexibility scenarios. These three metrics have been selected since the

cost metrics is the most indicative for system operation, while the previous studies have shown that the other two metrics typically had the largest value. The simulation timeframe has been set to one week, which is also the typical maximum cycling period for the PSPP storage size considered here.

The results illustrate how the values of all three metrics generally decrease with increasing flexibility of the system (meaning that MILP and LP get closer to each other in terms of performance). In particular, in Fig. 7 the worst case of the cost metric occurs for the 2024 scenario, with the metric’s absolute value decreasing from 0.46% to 0.28% when the plants become more flexible and then further to 0.2% for the case with storage. For the primary frequency response, as it can be seen in Fig. 8, the metric value does not vary noticeably when increasing the plants’ flexibility, and stays in the order of 1.3%÷1.4% across different RES scenarios. However, when including storage, the value of the metric is slightly reduced to 1.07%÷1.16%. A similar trend can be seen in Fig. 9 in terms of secondary up reserve. For example, the metric in the 2024 system decreases from 2.25% to 1.89%. It is also of interest to notice that, for each metric, the differences across RES scenarios for the “Flexible plants with storage” case tends to be lower than in the other two flexibility scenarios. In fact, storage increases system flexibility, which in turn supports RES integration, and as a result the two algorithms get closer to each other. This is also consistent with the results of Fig. 3, whereby with an increase in flexibility (which in Fig. 3 can be associated to an increase

in number of MILP clusters) the LP and MILP algorithms perform closer to each other.

It is also interesting to highlight that while the system becomes more flexible, relevant savings in operational costs may be associated to the “value of flexibility”. In this respect, when using MILP, the total system costs between “normal plants” and “flexible plants with storage” decrease by 1.9% (in the “traditional system”) to 3% (in the “2024 system”). On the other hand, the results in Fig. 5 indicate that MILP and LP perform extremely close to each other (with a difference in assessing the value of flexibility in the order of 2÷3%). Hence, considering that MILP algorithms are widespread for flexibility studies, it can be said once again the proposed relaxed LP algorithm (which is the focus of our work) can also be very effective to quantify the economic value of flexibility, especially in scenario and planning studies.

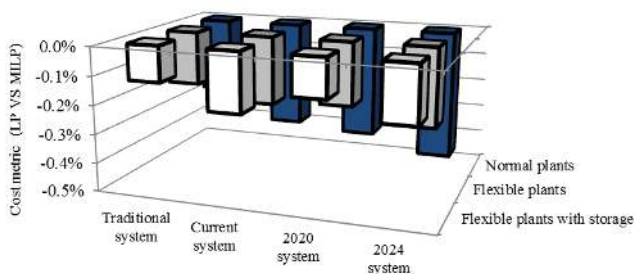


Fig. 5. Value of cost metric in the simulations of LP and MILP with different RES scenarios and system’s flexibility characteristics, weekly simulation.

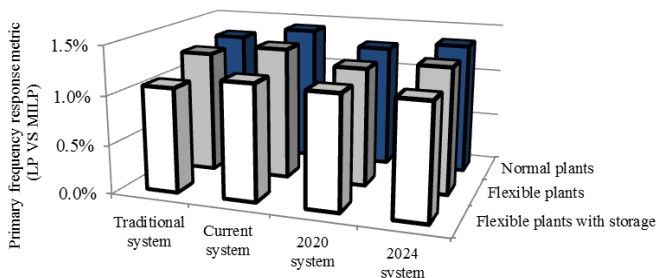


Fig. 6. Value of primary frequency response metric in the simulations of LP and MILP with different RES scenarios and system’s flexibility characteristics, weekly simulation.

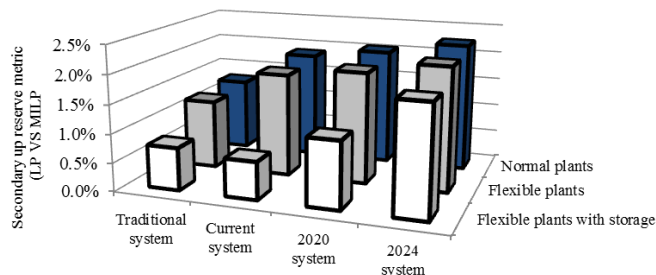


Fig. 7. Value of secondary up reserve metric in the simulations of LP and MILP with different RES scenarios and system’s flexibility characteristics, weekly simulation.

D. Analysis and discussion of the computational time requirements

The computational time requirements of the LP model (with three clusters) for the simulation of the Great Britain power system’s annual operation keep in the order of 15

seconds across the multiple scenarios with different RES penetration and flexibility cases. The simulation duration increases slightly (in the order of a few seconds) when storage is also considered, as this adds more constraints to system operation.

In the case of the MILP model (with three clusters), the computational time in the “normal flexibility” case increases from 45 seconds in the traditional system to 835 seconds in the 2024 system scenario. In the scenario when plants become more flexible the computational time reduces, by in the order of 10% for traditional system to 70% for 2024 system. However adding storage again increases the computational time, similarly to LP, due to a larger number of constraints in the problem. In general, the ratio of computational time between MILP and LP across multiple scenarios ranges between about 3 times (for the case of “traditional system” with “normal plants”) and about 50 times (for the case of “2024 system” with “normal plants”) in favour of LP.

As a further point to highlight, it was shown in Fig. 3 how the computational time ratio between LP and MILP increases with larger number of clusters. This result is consistent with the fact that with larger number of clusters MILP becomes more and more similar to BUC, with an increasing number of discrete variables. On the other hand, the number of constraints and variables in the LP problem increase too, but the problem keeps being fully linear and therefore relatively easier to solve. Although the computational comparison across scenarios is based on the 3-cluster model, as the one with the largest error between LP and MILP, our further studies confirm that simulating more clusters would favour LP even more in terms of computational time required, in line with the general trend shown in Fig. 3 and Fig. 4.

In order to further investigate computational time issues and the ability of MILP and BUC to converge to the pre-set MIP gap, additional simulations for all three algorithms and clustering approaches have been conducted. In particular, Fig. 8 illustrates the simulation time for a weekly simulation in the base scenario, for a number of LP and MILP clusters as well as for BUC, along with the remaining MIP gap in those cases when time limit was reached for the MILP and BUC algorithms. Furthermore, the figure now also demonstrates the performance of the different algorithms and clustering approaches with respect to a further dimension of problem complexity, namely, the number of reserves used in the simulations. In fact, considering multiple reserves is critical to properly addressing flexibility issues, but may become computationally onerous for larger clusters. To confirm this, different cases named “single reserve”, “two reserves” and “full reserves” have been run: “single reserve” means that only one reserve type is considered, which can be either primary, secondary or tertiary reserve; in the “two reserve” cases, three combinations are analysed combining the reserves in groups of two, *i.e.*, primary and secondary, primary and tertiary, and secondary and tertiary⁶; finally, the “full reserves” cases mean

⁶ For simplicity, in order to show the results in a compact form and due to space reasons, the values reported in Fig. 8 represent the average computational time of the three reserves (in the “single reserve” cases) or of

that all three reserve types are included in the algorithm, which are also the general formulation and case studies considered above in the paper⁷.

The results in Fig.8 indicate that increasing the number of reserves leads to longer average computational time. For example, in the “8 clusters” scenario the MILP algorithm takes 16600, 25000 and 36000 seconds for the “single”, “two” and “full” reserve cases, respectively, with 36000 seconds (=10h) hitting the time limit with a remaining MIP gap equal to 0.17%. Likewise, the BUC algorithm converges in about 26600 seconds on average in the “single reserve” case, but hits the pre-set simulation time limit of 10h with a remaining MIP gap of 0.19% and 0.49% in the “two reserves” and “full reserves” cases, respectively⁸. Hence, the results from Fig. 8 clearly show how considering multiple reserves can substantially increase the problem complexity and capability and speed of the solver to converge to a solution within the pre-set gap, while many of the UC algorithms that are used for different studies and particularly flexibility ones often only use one or at most two reserves. On the other hand, as demonstrated above in the paper, the proposed relaxed LP algorithms manages to perform very close to and (much) faster than MILP and BUC when providing “full reserves”, which further strengthens the contribution of this work.

Since the BUC algorithm and MILP algorithms with high number of clusters may not converge within the given time cap when adopting two or three reserves, additional analyses for a smaller problem (in terms of fewer simulation time steps), namely, simulation of one day, are shown in Fig. 9 for LP and MILP for different number of clusters as well as BUC. The results once again support the previous findings of the paper. For a single day analyses all three models converge to a uniform MIP gap of 0.1% within the computational time limit. The performance of LP compared to MILP in terms of computational time is in line with similar results on a weekly basis (Fig. 2 and Fig. 3). The simulation times of LP algorithms are between 0.19 and 2.32 seconds across different clustering methods, and then increase to in a range of 0.25 to 362 seconds in the MILP cases and further to 3090 seconds in BUC case. The computational time difference between MILP and BUC (besides LP) algorithms can thus now also be fully appreciated.

the three reserve combinations (in the “two reserve” cases). Generally speaking, the different algorithms are slower when using primary reserves than when using secondary and in turn than when using tertiary, and this trend becomes more evident with increasing number of clusters.

⁷ Regarding the modification of relevant constraints in the problem formulation when considering different number of reserves, if a specific reserve type is not considered then its relevant variable is to be set to zero for all considered simulation time steps. More specifically, Eq. (16) is neglected when primary frequency response is not considered. Ramping constraints (20) and (21) are discarded in the cases without secondary reserve. Similarly, Eqs (31) and (32) are excluded when tertiary reserve is not considered. In addition, f_{it}^{up} is eliminated from (17) if the primary frequency response is not considered; $r_{i,t}^{Sec,spin}$ is taken out from (17) and (22), for the cases without secondary reserve; and $r_{i,t}^{Ter,spin}$ is neglected in Eqs (17) and (22) for the cases without tertiary reserve.

⁸ In this respect, again it can be seen and somehow quantified how BUC “takes longer” than MILP, as even when neither algorithm converges the gap for BUC is still larger than for MILP.

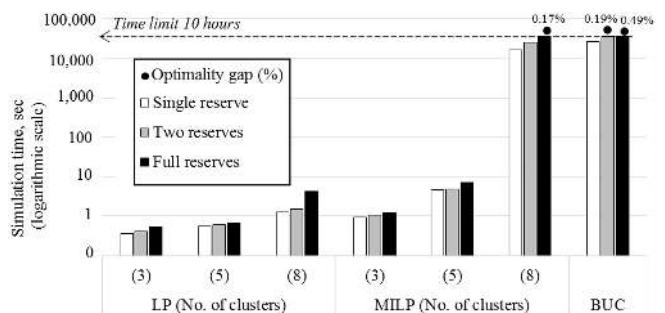


Fig. 8. Computational time and optimality gap of LP and MILP with 3, 5 and 8 clusters and of BUC: weekly simulation, “current system” with “normal flexibility” and different reserve numbers.

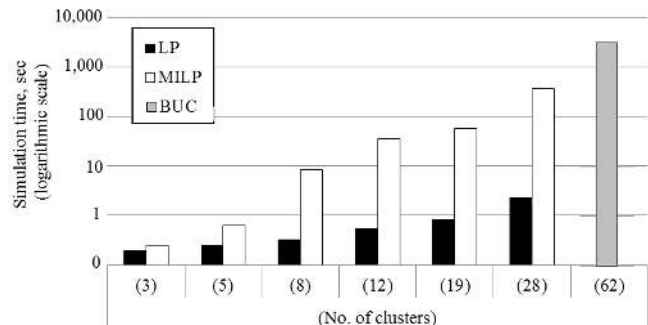


Fig. 9. Computational time of LP and MILP with different clusters and BUC: daily simulation, “current system” with “normal flexibility” and full reserves.

VII. CONCLUSIONS

In this paper, we have introduced a flexibility-oriented unified formulation that can seamlessly model binary (BUC), MILP and LP unit commitment approaches by representing the generating plants’ commitment variables as binary, integer or continuous values, respectively. In particular, it has been shown how an LP algorithm can be modelled within the proposed formulation by simply relaxing the variable representing the online units from integer (MILP) to continuous for different generators’ clustering levels of the MILP representation. Frequency response and reserve services and all relevant constraints have been explicitly considered, and the flexible and inflexible “components” of the UC problem formulation have been highlighted.

The performance of the three methods (BUC as well as MILP and LP with different clustering levels) have been compared through different metrics specifically introduced, namely, performance metrics such as relative differences in system’s operational cost, power output composition per cluster, and allocation of flexibility services (primary frequency response, secondary and tertiary up reserves) per cluster, as well as computational efficiency. To demonstrate the proposed approach, several case studies have been run considering a 62-unit reduced version of the Great Britain power system. In the considered base case, the MILP UC can perform 4300 to 35800 faster than BUC when using the smaller number of clusters, but reaches the pre-set time constraint when the number of clusters increases to 8. On the other hand, the LP UC can perform tens to thousands times faster than MILP across the 3 to 28 cluster cases. All performance metrics stay within around 1.5%, and increasing the number of clusters does not generally give much

improvement in terms of results for LP cases (relative to BUC) but improves the MILP performance. Furthermore, for different RES penetration scenarios (run for a conservative 3-cluster case), the maximum difference in results between the LP and MILP algorithms is confined to 2.2% for any performance metric. This difference also decreases when the system becomes more flexible, making the generating units more flexible and then further adding storage to the system. On the other hand, the LP always brings substantial computational gains, running about 3 to 50 times faster than a MILP, depending on the scenario. The complexity of the different models when considering different types of reserve services has also been investigated in detail, also quantifying the optimality gap of MILP and BUC algorithms when the pre-set gap could not be reached within the defined time limit was also investigated in detail. In addition, the computational performance of LP, MILP and BUC for a smaller case study (one-day simulation) has also been analysed, confirming the trends found for larger scale cases.

The results of this work generally indicate that the proposed relaxed LP UC approach, in case with low level of clustering, can be used instead of popular MILP UC approaches for strategic flexibility analysis for large systems with long time horizons and simulations, with minor (and in many cases negligible) loss of accuracy in the different metrics considered, while gaining noticeable computational speed. Hence, LP UC algorithms could be particularly suitable for time-consuming flexibility studies that might for instance entail scenario analysis or Monte Carlo/stochastic simulations, as well as for planning purposes. On the other hand, given the complexity of the issues discussed, the specific results might also be affected by the solver choice, solver settings and details of the problem formulation, and more research is needed to comprehensively address these points.

Work in progress aims at incorporating demand side resources in the analysis in a more comprehensive and systematic way, in particular for provision of reserves.

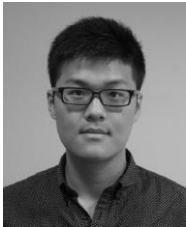
ACKNOWLEDGEMENTS

The authors would like to thank the anonymous Reviewers for their tremendous work that helped improve the paper significantly.

REFERENCE

- [1] Y. G. Rebours, D. S. Kirschen, M. Trotignon, and S. Rossignol, "A survey of frequency and voltage control ancillary services - Part I: Technical features," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 350–357, 2007.
- [2] B. Kirby and M. Milligan, "A Method and Case Study for Estimating the Ramping Capability of a Control Area or Balancing Authority and Implications for Moderate or High Wind Penetration," in *WINDPOWER*, 2005.
- [3] SuChua-Liang and D. Kirschen, "Quantifying the Effect of Demand Response on Electricity Markets," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1199–1207, Aug. 2009.
- [4] D. Pudjianto, M. Aunedi, P. Djapic, and G. Strbac, "Whole-Systems Assessment of the Value of Energy Storage in Low-Carbon Electricity Systems," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 1098–1109, Mar. 2014.
- [5] E. Ela, V. Gevorgian, A. Tuohy, B. Kirby, M. Milligan, and M. O'Malley, "Market designs for the primary frequency response ancillary service-Part II: Case studies," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 432–440, 2014.
- [6] E. Ela, V. Gevorgian, A. Tuohy, B. Kirby, S. Member, M. Milligan, and M. O. Malley, "Market Designs for the Primary Frequency Response Ancillary Service — Part I: Motivation and Design," *Power Syst. IEEE Trans.*, vol. 29, no. 1, pp. 421–431, 2014.
- [7] F. D. Galiana, F. Bouffard, J. M. Arroyo, and J. F. Restrepo, "Scheduling and Pricing of Coupled Energy and Primary, Secondary, and Tertiary Reserves," *Proc. IEEE*, vol. 93, no. 11, pp. 1970–1983, 2005.
- [8] E. Delarue, D. Cattrysse, and W. D'haeseleer, "Enhanced priority list unit commitment method for power systems with a high share of renewables," *Electr. Power Syst. Res.*, vol. 105, pp. 115–123, Dec. 2013.
- [9] T. Li and M. Shahidehpour, "Price-Based Unit Commitment: A Case of Lagrangian Relaxation Versus Mixed Integer Programming," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 2015–2025, 2005.
- [10] J. Ma, V. Silva, R. Belhomme, D. S. Kirschen, and L. F. Ochoa, "Evaluating and planning flexibility in sustainable power systems," *IEEE Trans. Sustain. Energy*, vol. 4, no. 1, pp. 200–209, 2013.
- [11] M. Carrion and J. M. Arroyo, "A Computationally Efficient Mixed-Integer Linear Formulation for the Thermal Unit Commitment Problem," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1371–1378, Aug. 2006.
- [12] R. Barth, H. Brand, P. Meibom, and C. Weber, "A stochastic unit-commitment model for the evaluation of the impacts of integration of large amounts of intermittent wind power," 2006 9th Int. Conf. Probabilistic Methods Appl. to Power Syst. PMAPS, 2006.
- [13] J. Kiviluoma and P. Meibom, "Methodology for modelling plug-in electric vehicles in the power system and cost estimates for a system with either smart or dumb electric vehicles," *Energy*, vol. 36, no. 3, pp. 1758–1767, 2011.
- [14] M. Aunedi, P. Kountouriotis, J. E. O. Calderon, D. Angeli, G. Strbac, and A. Parameters, "Economic and Environmental Benefits of Dynamic Demand in Providing Frequency Regulation," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 2036–2048, 2013.
- [15] B. S. Palmintier and M. D. Webster, "Heterogeneous Unit Clustering for Efficient Operational Flexibility Modeling," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1089–1098, May 2014.
- [16] A. Flores and R. Moreno, "A Column Generation Approach for Generation Expansion Planning with High Renewable Penetration," in *INFORMS 2014*, 2014, pp. 1–6.
- [17] M. A. Ortega-Vazquez, D. S. Kirschen, and Y. Dvorkin, "Wind generation as a reserve provider," *IET Gener. Transm. Distrib.*, vol. 9, no. 8, pp. 779–787, May 2015.
- [18] B. C. Ummels, M. Gibescu, E. Pelgrum, W. L. Kling, and A. J. Brand, "Impacts of Wind Power on Thermal Generation Unit Commitment and Dispatch," *IEEE Trans. Energy Convers.*, vol. 22, no. 1, pp. 44–51, Mar. 2007.
- [19] ENTSO-E, "Continental Europe Operation handbook Policy 1," 2013.
- [20] F. Teng, V. Trovato, and G. Strbac, "Stochastic Scheduling With Inertia-Dependent Fast Frequency Response Requirements," *IEEE Trans. Power Syst.*, pp. 1–10, 2015.
- [21] A. Van Stiphout, K. Poncelet, K. De Vos, and G. Deconinck, "The impact of operating reserves in generation expansion planning with high shares of renewable energy sources," in *IAEE European Energy Conference, Sustainable Energy Policy and Strategies for Europe*, 2014, pp. 1–15.
- [22] F. Bouffard and M. Ortega-Vazquez, "The value of operational flexibility in power systems with significant wind power generation," in 2011 IEEE Power and Energy Society General Meeting, 2011, pp. 1–5.
- [23] R. Jiang, J. Wang, and Y. Guan, "Robust Unit Commitment With Wind Power and Pumped Storage Hydro," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 800–810, May 2012.
- [24] H. Pandzic, Ting Qiu, and D. S. Kirschen, "Comparison of state-of-the-art transmission constrained unit commitment formulations," in 2013 IEEE Power & Energy Society General Meeting, 2013, pp. 1–5.
- [25] B. S. Palmintier, "Incorporating operational flexibility into electric generation planning Impacts and methods for system design and policy analysis," Massachusetts Institute of Technology, 2013.
- [26] National Grid, "Generation by fuel type," 2013. [Online]. Available: <https://www.elexonportal.co.uk/article/view/216?cachebust=k9nxjlaqx6>.
- [27] Y. Zhou, P. Mancarella, and J. Mutale, "Generation adequacy in wind rich power systems: Comparison of analytical and simulation approaches," in 2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2014, pp. 1–6.

- [28] I. Macleay, K. Harris, and A. Annut, "DIGEST OF UNITED KINGDOM ENERGY STATISTICS 2014," 2014.
- [29] National Grid, "Winter Outlook 2013/14," 2013.
- [30] National Grid, "Short Term Operating Reserve Fuel type Analysis Season 8.5."
- [31] National Grid, "UK Future Energy Scenarios," 2014.
- [32] F. Teng, D. Pudjianto, G. Strbac, F. Ferretti, and R. Bove, "Assessment of the value of plant flexibility in low carbon energy system," in 3rd Renewable Power Generation Conference (RPG 2014), 2014, pp. 7.1.4–7.1.4.
- [33] D. Lumb and N. T. Hawkins, "Provision of power reserve from pumped storage hydro plant," in IEE Colloquium on Economic Provision of a Frequency Responsive Power Reserve Service, 1998, vol. 1998, pp. 3–3.
- [34] Parliamentary Office of Science and Technology, "Electricity Storage," no. 306. 2008.
- [35] Energy research Partnership, "The future role for energy storage in the UK," 2011.
- [36] FICO, "FICO XPRESS 7.6." FICO, 2015.



Lingxi Zhang (S'13) received the B.Eng (Hons) degrees in Electrical and Electronic Engineering from the University of Manchester, UK, and North China Electric Power University, China, in 2013. He is currently pursuing a PhD degree in Electrical Energy and Power systems group at the University of Manchester. His research interests include multi-energy system modelling, power system flexibility and energy system operation and planning.



Tomislav Capuder (S'08–M'14) received the Ph.D. degree in power systems from the Faculty of Electrical Engineering and Computing, Zagreb, Croatia, in 2014. He is currently a Post-Doctoral Researcher with the Faculty of Electrical Engineering and Computing. His current research interests include multienergy generation systems, district energy systems, power and energy system flexibility, and operation and planning of energy systems.



Pierluigi Mancarella (M'08, SM'14) received the Ph.D. degree in Electrical Energy Systems from the Politecnico di Torino, Italy, in 2006. After being a Research Associate at Imperial College London, UK, in 2011 Pierluigi joined the University of Manchester, UK, where he is currently a Reader in Future Energy Networks.

Pierluigi's research interests include multi-energy systems modelling, power system integration of low carbon technologies, network investment under uncertainty, and risk and resilience of smart grids. He is author of several books and book chapters and of over 200 research papers on those topics. Pierluigi is an Associate Editor of the IEEE Systems Journal, and the Chair of the Energy working group of the IEEE European Public Policy Initiative.