Comparison of different model formulations for modelling future power systems with high shares of renewables – The Dispa-SET Balkans model

Matija Pavičević[a]*, Konstantinos Kavvadias[b], Tomislav Pukšec[c], Sylvain Quoilin[a],

[a]Faculty of Engineering Technology, KU Leuven, 2440 Geel

[b]Joint Research Centre, European Commission, 1755 Petten

[c]Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, 10000 Zagreb

doi: https://doi.org/10.1016/j.apenergy.2019.113425

Please cite this article as: Matija Pavičević, Konstantinos Kavvadias, Tomislav Pukšec, Sylvain Quoilin, Comparison of different model formulations for modelling future power systems with high shares of renewables – The Dispa-SET Balkans model, Applied Energy, Volume 252, 2019, 113425, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2019.113425.

Comparison of different model formulations for modelling future power systems with high shares of renewables – The Dispa-SET Balkans model

Matija Pavičević[a]†, Konstantinos Kavvadias[b], Tomislav Pukšec[c], Sylvain Quoilin[a],

[a]*Faculty of Engineering Technology, KU Leuven, 2440 Geel*
[b]*Joint Research Centre, European Commission, 1755 Petten*
[c]*Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, 10000 Zagreb*

A B S T R A C T

Power system's operational flexibility represents its ability to respond to predicted or unexpected changes in generation and demand. Traditional policy and planning models usually do not consider the technical operating constraints directly responsible for its operational flexibility. Nevertheless, this capability becomes increasingly important with the integration of significant shares of renewables. Incorporating flexibility can significantly change optimal generation strategies, lower the total system costs and improve policy impact estimates. The goal of this research is to prove that, for computational efficiency reasons, it is useful to cluster some of the original units into larger ones. This process reduces the number of continuous and binary variables and can, in certain conditions, be performed without significant loss of accuracy. To this purpose the Dispa-SET unit commitment and power dispatch model which focuses on balancing and flexibility problems in the European grids has been applied to the Western Balkans power system. Various clustering methods are implemented and tested on the same dataset and validated against the "No clustering" formulation. "Per unit" aggregates very small or very flexible units into larger ones with averaged characteristics, "Per typical unit" considers one typical power plant per technology; and "Per technology" additionally simplifies the mathematical formulation by completely neglecting units flexibility capabilities.. The results have shown that the difference between disaggregated and clustered approaches remains acceptable and for certain accuracy metrics falls within a 2 % margin. This is especially true in case of highly interconnected regional systems with relatively high shares of hydro energy.

Keywords:
Dispa-SET; Optimal Dispatch; Unit Commitment; Model formulation; Energy system modelling; Clustering

1. Introduction

The IPCC's Fifth Assessment Report-AR5 [1] confirms unequivocally global warming and provides evidence of its substantial and wide-ranging consequences such as permafrost melting, heavy precipitations, floods, droughts wildfires etc. However, despite the global commitment achieved in the Paris Agreement during the 21st session of the Conference of the Parties to the United Nations Convention (COP21), countries' pledges [2] are still not sufficient to face the climate change challenge [3]. In fact, a comprehensive portfolio of climate change strategies must include both mitigation and ad hoc

* *Matija Pavičević.* Tel.: *+32 14 721 496*
E-mail address: matija.pavicevic@kuleuven.be

adaptation actions that allow achieving multiple goals in the sustainable development areas [4]. In order to tackle these issues, the European Union has set a target to collectively reach a share of at least 27% renewables in the final energy consumption by 2030 [5]. This energy-based goal could translate into 50% of total electricity production from renewable energy sources (RES-E) and reduction of greenhouse gas (GHG) emissions by 40% [6]. For the 2050 framework, these targets increase further, with a reduction of GHG emissions by 80% and an increase in the share of RES in electricity consumption of up to 97% [7]. A survey on climate change and adaptation policies in South East Europe [8] has shown that the Western Balkans region is an interesting geopolitical area. It consists of two EU member states (Croatia and Slovenia), four candidate countries (Albania, North Macedonia, Montenegro and Serbia) and two potential candidate countries (Bosnia and Herzegovina and Kosovo), all six members of the Energy Community [9], that have started to implement some the Unions 'acquis communautaire' and will eventually contribute to the common 2030 and 2050 climate targets.

European institutions, transmission system operators, scientific researchers and private companies have put a lot of effort to analyse the behaviour of future power systems. One of the main research fields, among others, is modelling of future power systems and flexibility requirements under high penetration of RES-E. Nowadays, there are numerous different models available, which need to find a compromise between highly-detailed operational power system and low time resolution long-term planning models [10].They can be classified into six main groups: generation expansion planning (GEP), production cost optimization (PCO), hydro-thermal coordination (HYTHCO), maintenance optimization (MO), unit-commitment (UC) and economic dispatch (ED) models. They are all based on the same physical and/or economic principles, but their formulations are quite different and depending on the size and complexity [11] can be binary for UC, mixed integer linear programming (MILP) for GEP, PCO, mixed integer nonlinear programming (MINLP) for GEP, HYTHCO or linear programming (LP) for GEP, ED and PCO. The GEP models focus to optimize investments to provide economically least expensive power systems while maintaining reliability and meeting environmental and other constraints such as emission targets, system flexibility etc. Due to many constraints and long time horizons, this can be computationally intensive if no modelling simplifications such as linearization [12] and relaxation [13] are introduced. As such, according to [14], the main challenge for 21st century energy system optimization models will be resolving the space and time dimensions. According to [15], GEP models that consider endogenous technology cost learning can take up to several days. When applied to a single zone such as the UK power system, simulations can range from 43 h upwards even if power plants are clustered (32 units in total) and a total of 296 GB RAM is used. Another example of such On the other hand, authors from [13] and [16] have proven that a long-term capacity expansion planning model for an electric power system integrating large-size renewable energy technologies can simplify the complexity of such models reducing the computation time down to several hours. It is even possible to reduce the simulation times down to minutes, as proposed by [12]. Common simplifications in long term GEP models usually include some sort of time sampling [17] or slicing such as representative days proposed by Koltsaklis et. al. [18]. Pineda et. al. [19] has proposed a more efficient chronological time period clustering for optimal GEP with storage units. Blanford et. al. [20] developed a method for selecting representative hours to preserve key distributional requirements for regional load, wind, and solar time series with a two-orders-of-magnitude reduction in dimensionality of the problem. Some GEP models, as proposed by Slednev at. al. [21] even go beyond existing models by considering the assignment of RES-E potentials to geographical grid nodes as a variable. Most of these GEP models focus on long time horizons and neglect the need for operational flexibility [22]. Although in some cases those flexibility requirements are considered to a certain extent, e.g. as proposed by Quoilin et. al. [23], the future energy and technology mixes from long term planning models are not validated by a detailed UC and/or ED model or formulation. To authors knowledge, such cross-model validation has only been proposed by several authors. Pietzcker et. al. [24] has analysed the system integration of wind and solar power in integrated system assessment models, Pavičević et. al. [25] flexibility of power demand and supply in the EU power system by soft-linking JRC-EU-TIMES and the open-source Dispa-SET model and Bistline et. al. [26] explores how enhanced flexibility through lower minimum load levels (also called "turndown" limits) for coal- and natural-gas-fired power plants can impact operations and profitability using soft-linked GEP and UC models. Palmintier et. Al. [27] emphasises how important it is to consider flexibility requirements when planning capacity expansion and dispatch strategies in future energy systems. Neglecting them can result in a different capacity and energy mix and emissions, that can lead to a system that is unable to simultaneously meet demand, carbon, and flexibility requirements. Strengths and weaknesses of soft-linking and direct linking and integrating GEP and UC and ED models have been summarized in [28]. The authors conclude that the inherent differences between the methodologies mean that each will integrate short term variations differently into the modelling process and assess the flexibility of the system differently. To this date those methodologies have been successfully applied in separate models and data sets, making the result comparison a difficult task.

For shorter time frames that range from weeks, months, seasons or even up to a year [29], modeling enters into the realm of PCO, midterm considerations of MO, and HYTHCO must be synchronized with the shorter term UC and ED models. In those modeling frameworks, objective is usually related to cost optimization of operating the power system for an extended period of time. There are numerous ways to find an optimal strategy that would minimize the energy production costs of a power system that is for example linked to the energy storage system [30] or combined heat and power (CHP) generation system [31][32], or minimize the production costs of a standalone or microgrid system [33]. HYTHCO problems are usually not limited by the peak power capacity of the generators but by water availability in the accumulation reservoirs throughout the year [29]. The resulting water usage is allocated over the year to ensure enough capacity during dry seasons. This topic where optimization methods are applied for solving the short-term HYTHCO problem [34] is extensively covered by many publications with traditional UC formulations [35][36] or development of new algorithms such as quasi-opposition-based learning and self-learning mutation, an improved multi-objective "teaching" and "learning" optimization algorithm [37]. MO algorithms are part of mid-term flexibility requirements [29] and take into the account the scheduled maintenance operations of individual powerplants [38] and unforeseen outages [39]. MO are usually proposed in advance and are modelled as outage factors where the nominal capacity is reduced by a certain amount. When represented in form of a timeseries they are passed as input data for UC and ED models. Time horizons in UC models are usually in the range of hours and days and determine which generators should be turned on and have available power production [29]. This is computationally most demanding optimization task as such formulations are tightly constrained by the operational parameters that can only be formulated with binary and/or integer variables A review article of UC problems [40] describes how several mixed-integer linear programming formulations of the transmission-constrained unit commitment problem impact the computation time and end results. Authors from [41] provide the MIP formulation of convex hull descriptions of thermal units and cover all the basic operating constraints in UC models such as generation limits, start-up and shut-down capabilities, and minimum up and down times. It has also been proven that the inclusion of startup and shutdown trajectories [42] often yielded the largest improvements in schedule performance of conventional powerplants in systems with high penetration of RES-E. Complex nature of UC problems has inspired many researchers to find alternative and similarly accurate formulations such as mixed integer quadratic programming (MIQP) projected two-binary variable formulation [43], binary artificial sheep formulation

[44] and parallel-series hybrid meta-heuristic method [45]. Beside those novel methods security constrained unit-commitment (SCUC) problem is a common and well established MILP method for hydro and thermal scheduling [46], UC scheduling including transmission constraints [47] as well as simplified flexibility evaluation using UC formulation [35].

A literature review has concluded that UC problems are always formulated and then solved using one of the following three methods: binary, clustered MILP or simplified LP. This paper conducts a comprehensive comparison analysis of these three most commonly used and well-established formulations by analyzing several key metrics. It should be noted that there are already several formulations available in the literature. One such article compares two mathematical formulations of SCUC [48] and demonstrates their potential applicability to medium-scale and large-scale power systems. Meus et al [49] has proven that simplified formulations provide identical results to a traditional binary unit commitment formulation but this assumption only holds for portfolio not restricted by start-up and shut-down limitations. Palmintier et al [50] proposed heterogeneous unit clustering for efficient operational flexibility modelling. Furthermore, he demonstrated that for a 205-unit bus system, clustering introduces errors of 0.05%-0.9% across several metrics while providing several orders of magnitude faster solution times (400x) and (x2000) if further clustering is allowed in time horizons of one week. Impact of clustering groups of generators into categories has been further investigated by [51]. Authors have shown that such formulation reduces computation time by x5000, but operating costs and carbon emissions produce high errors of up to 39% when compared to traditional binary UC formulation. The previous publications mostly focus on model comparisons for isolated cases and short time horizons (one week or less). In this paper, we aim at proposing a similar comparison, but on a large interconnected system including several countries and using historical input data. This work also considers detailed technology-specific comparison in terms of load duration curves, capacity factors, probability distributions, hourly costs of running the system, RES-E curtailment and differences in cross-border energy flows. Analysis has been carried out in Dispa-SET, an open source UC and ED model focused on the balancing and flexibility problems in European grids. Dispa-SET is mainly developed within the Joint Research Centre of the EU Commission, in close collaboration with the University of Liège and the KU Leuven (Belgium).

The main contributions of this work are related to both the development of an open power system model for the Balkans region and the use of this model to compare various model formulations. In particular they include:

- analyzing the short-term-based strategic dispatch decisions under different clustering methodologies and the related model formulations
- proposing practical guidance to energy modelers by providing comparisons metrics to four different model formulations (No clustering, Per unit, Per typical unit and Per technology),
- providing a detailed model for the Western Balkans power sector, which can be further re-used and/or adapted by other researchers)
- analyzing the strengths of the Western Balkans power grid which gives us an idea about bottlenecks and potential reinforcement fixes.

This paper is divided into five sections. the first section covers the literature overview and motivation behind the research. The second section describes the model and different formulation as well as key metrics used for the comparison. the third section describes the scenarios in which the model was applied to the Western Balkans region. The fourth section is dedicated to the results analysis of different powerplant clustering formulations. The fifth and final section concludes the analysis of this research.

Nomenclature

Abbreviations

CHP	Combined heat and power
CO_2	Carbon dioxide
DH	District heating
ED	Economic dispatch
GEP	Generation expansion planning
GHG	Greenhouse gas
HYTHCO	Hydro-thermal coordination
LP	Linear programming
MILP	Mixed integer linear programming
MIQP	Mixed integer quadratic programming
MO	Maintenance optimization
NTC	Net transfer capacity
PCO	Production cost optimization
RES-E	Electricity production from renewable energy sources
SCUC	Security constrained unit-commitment
UC	Unit commitment
UC	Unit commitment
VRES	Variable renewable energy sources

Zones:

AL	Albania
BA	Bosnia and Herzegovina
HR	Croatia
ME	Montenegro
MK	North Macedonia
RS	Serbia
SI	Slovenia
XK	Kosovo

Technologies:

BEVS	Battery-powered electric vehicles
COMC	Combined cycle
GTUR	Gss turbine

HDAM	Conventional hydro dam
HPHS	Pumped hydro storage
HROR	Hydro run-of-river
ICEN	Internal combustion engine
PHOT	Solar photovoltaic
STUR	Steam turbine
THMS	Thermal storage
WTOF	Offshore wind turbine
WTON	Onshore wind turbine
Fuels:	
BIO	Bagasse, Biodiesel, Gas from biomass, Gasification, Biomass, Briquettes, Cattle Residues, Rice hulls or husk, Straw, Wood gas (from wood gasification), Wood waste liquids (excluding black lignite and including red liquor, Sludge, Wood spent sulphite liquor and other liquids, Wood and wood waste
GAS	Blast furnace gas, Boiler natural gas, Butane, Coal bed methane, Coke oven gas, Flare gas, Gas (generic), Methane, Mine gas, Natural gas, Propane, Refinery gas, Sour gas, Synthetic natural gas, Top gas, Volatile organic compounds gas & vapor, Waste gas, Wellhead gas
HRD	Anthracite, Other anthracite, Bituminous coal, Coker by-product, Coal gas (from coal gasification), Coke, Coal (generic), Coal-oil mixture, Other coal, Coal and pet coke mi, Coal tar oil, Anthracite coal waste, Coal-water mixture, Gob, Hard coal / anthracite, Imported coal, Other solids, Soft coal, Anthracite silt, Steam coal, Subbituminous, Pelletized synthetic fuel from coal, Bituminous coal waste)
LIG	Lignite black, Lignite brown, lignite
NUC	Uranium, Plutonium
OIL	Crude oil, Distillate oil, Diesel fuel, No. 1 fuel oil, No. 2 fuel oil, No. 3 fuel oil, No. 4 fuel oil, No. 5 fuel oil, No. 6 fuel oil, Furnace fuel, Gas oil, Gasoline, Heavy oil mixture, Jet fuel, Kerosene, Light fuel oil, Liquefied propane gas, Methanol, Naphtha, ,Gas from fuel oil gasification, Fuel oil, Other liquid, Orimulsion, Petroleum coke, Petroleum coke synthetic gas, Black liquor, Residual oils, Re-refined motor oil, Oil shale, Tar, Topped crude oil, Waste oil
SUN	Solar energy
WAT	Hydro energy
WIN	Wind energy
WST	Digester gas (sewage sludge gas), Gas from refuse gasification, Hazardous waste, Industrial waste, Landfill gas, Poultry litter, Manure, Medical waste, Refused derived fuel, Refuse, Waste paper and waste plastic, Refinery waste, Tires, Agricultural waste, Waste coal, Waste water sludge, Waste
OTH	Other fuel types and energy carriers
Sets	
f	Fuel types
l	Transmission lines between zones
n	Zones

p	Pollutants	
s	Storage units, including hydro reservoirs	
t	Hours	
u	Units	
j	Aggregated units	

Parameters

$\Delta S_u^{down,max}$	MW	Ramp down limit at the start up
$\Delta S_u^{up,max}$	MW	Ramp up limit at the start-up
$\Delta P_u^{down,max}$	MW	Ramp down limit
$\Delta P_u^{up,max}$	MW	Ramp up limit
$E_l^{trans,max}$	-	Maximal transmission capacity
$E_l^{trans,min}$	-	Minimal transmission capacity
$E_n^{shed,max}$	MW	Maximal load shedding
P_u^{max}	MW	Maximal load
P_u^{min}	MW	Minimal stable load
$P_u^{quickstart}$	MW	Quickstart power
$a_u^{min\,down}$	-	Minimum down time
$a_u^{min\,up}$	-	Minimum up time
$a_u^{quickstart}$	-	Number of quickstart units in off state
c_l^{trans}	€/MW	Fixed costs of using the line
$c_{n,f,t}^{fuel}$	€/MWh	Fuels costs in any given time period
$c_{n,t}^{pwr}$	€/MW	Costs of lost load due to max and min power
$c_{n,t}^{rsrv}$	€/MWh	Costs of lost load due to up and down reserves
c_n^{shed}	€/MW	Fixed costs for load shedding
c_p^{polut}	€/t	Fixed price of individual pollutants
$c_{u,t}^{ramp}$	€/MWh	Costs of lost load due to ramping rates
c_u^{down}	€/MW	Ramping down costs
c_u^{fix}	€	Fixed costs
c_u^{fuel}	€/MWh	Additional fixed costs per start
c_u^{up}	€/MW	Ramping up costs
f_u^{start}	MWh/start	Fuel use per start-up
$l_{l,n}^{node}$	-	Location of transmission nodes between two zones
$l_{s,n}$	-	Location of storage units
$l_{u,n}$	-	Location {binary: 1 u located in n}
$l_{u,n}$	-	Location of the unit

η_u	-	Electrical efficiency of any given unit
Variables		
$C_{chp,t}^{CHP}$	€	Costs associated to CHP units
$C_{n,t,u}^{VOLL}$	€	Costs of lost load
$C_{n,t}^{shed}$	€	Load shedding costs
$C_{t,l}^{trans}$	€	Transmission costs
C^{tot}	€	Total system costs
$C_{\mathbf{u},t}^{ramp}$	€	Ramping costs
$C_{\mathbf{u},t}^{start}$	€	Start-up costs
$C_{\mathbf{u},t}^{var}$	€	Variable costs
$C_{\mathbf{u}}^{fix}$	€	Fixed costs of running the unit
$D_{n,h}$	MW	Demand
$EM_{u,p}$	t	Emission rates as a function of pollutants
$E_{l,t}^{trans}$	MW	Load in the transmission lines
$E_{n,t}^{shed}$	MW	Amount of power being shedd
$F_{u,f}$	MWh	Fuel usage
$LL_{n,t}^{\substack{max\\pwr}}$	MW	Deficit in terms of maximum power
$LL_{n,t}^{\substack{min\\pwr}}$	MW	Suficit when power is exceeding the demand
$LL_{n,t}^{2D}$	MWh	Deficit in reserve down
$LL_{n,t}^{2U}$	MWh	Deficit in reserve up
$LL_{n,t}^{3U}$	MWh	Deficit in non-spinning reserve up
$LL_{u,t}^{\substack{ramp\\down}}$	MW	Deficit in terms of ramping down
$LL_{u,t}^{\substack{ramp\\up}}$	MW	Deficit in terms of ramping up
$P_{u,t}$	MW	Power output
$Stor_{s,h}$	MW	Storage inflow
$S_{u,t}/D_{u,t}$	-	Start-up and shot-down events
$U_{u,t}$	-	Number of start-up events

2. Methods

2.1. Model formulations

The section describes different formulations and clustering methods implemented into the Dispa-SET model. Note that, since model formulation and clustering methods are tightly interlinked, both expressions will be used indifferently in the rest of this paper.

Four model formulations and clustering methods are compared in this work: “No clustering”, “Per unit”, “Per typical unit” and “Per technology”. It is worthwhile to note that each modelling formulation or clustering method can be applied to the same input dataset [52]. This allows an easy comparison of different methods in terms of computational efficiency and accuracy. Moreover, importance of such standardized and flexible input datasets is an ongoing discussion among energy modeling experts within the Open Energy Modelling Initiative (OpenMod) [53].

2.1.1. No clustering

A binary formulation [54] in which each power plant in the system is considered individually is the core formulation of the model. This formulation is highly detailed and most accurate but computationally most difficult to solve. It allows constraints such as minimum up/down times, minimum load, ramping limits, etc. for each individual power plant (there is therefore one binary variable per unit).

When compared to other formulations the “No Clustering” case can be considered a reference point (i.e. results from other formulations can be compared to this baseline through the various metrics described later on in the paper). All model equations are available in the annex A of this paper, in the Dispa-SET online documentation [55] or directly in the source code open repository [56].

2.1.2. Per unit clustering

Per unit clustering is a MILP based formulation of the unit commitment problem. It is also the standard formulation of the Dispa-SET tool based on a hypothesis which assumes that, for computational efficiency reasons, it is useful to merge some of the units that share the same characteristics into larger units [55]. This, depending on the size of the problem, can significantly reduce its space-time dimension. This reduction of binary and continuous variables can in some cases be performed without a significant loss of simulation accuracy. In this formulation, units with small installed capacity and/or units that are highly flexible are merged into larger units. When two units are merged the minimum $\mathrm{P_u^{min^*}}$ and maximum $\mathrm{P_u^{max^*}}$capacities (MW) of newly aggregated units are formulated as follows:

$$P_u^{min^*} = min\left(P_j^{min}\right) \quad , \quad P_u^{max^*} = \sum_{j\in \boldsymbol{J}} \left(P_j^{max}\right) \tag{1}$$

where $\mathrm{P_j^{min}}$ and $\mathrm{P_j^{max}}$ are minimum and maximum power capacities (MW) of original units j. Variable costs, $\mathrm{C_{u,t}^{var^*}}$, (€/MW) of these newly created units u in each time interval t, are formulated as follows:

$$C_{u,t}^{var^*} = \frac{\sum_{j\in \boldsymbol{J}}\left(P_j^{max} \cdot C_{j,t}^{var}\right)}{P_u^{max^*}} \tag{2}$$

where $\mathrm{C_{u,t}^{var}}$ are original variable costs (€/MW) of units j in time interval t. Due to aggregation the start-up and shut-down costs are transformed into ramping costs as given by:

$$C_{u,t}^{ramp} = \frac{\sum_{j\in \boldsymbol{J}}\left(P_j^{max} \cdot C_{j,t}^{ramp}\right)}{P_u^{max^*}} + \frac{\sum_{j\in \boldsymbol{J}}\left(C_{j,t}^{start}\right)}{P_u^{max^*}} \tag{3}$$

where $\mathrm{C_{j,t}^{ramp}}$ are ramping costs (€/MW) of initial units j, and $\mathrm{C_{j,t}^{start}}$ are start-up costs (€/MW) of initial units j in each time interval t, Plant efficiency, minimum up and down times and carbon emissions are computed as weighted average of all aggregated units:

$$\eta_u{}^* = \frac{\sum_{j\in \boldsymbol{J}}\left(P_j^{max} \cdot \eta_j\right)}{P_u^{max^*}} \tag{4}$$

Figure 1 is a graphical representation of the conditions underlying the decision to cluster a specific unit or not. Starting from the left, units can be clustered only if they are of the same type (Gas, Coal, Hydro etc.), have the same values (Ramping rates, Start-up costs, Efficiency etc.) and their minimum power output is close to zero. The second possibility is to cluster highly flexible units of the same type whose start-up time is less than 1h. The third possibility groups into a single cluster small units whose maximum power output is less than 30 MW. As an example, three-zones (A, B and C) power system can be considered. In zone A there are 3 10 MW GAS fired GTUR units, in zone B there are 2 identical 500 MW GAS fired COMC units and one 200 MW GAS fired COMC unit, in zone C there are 3 (10, 25 and 50 MW) WAT powered HROR units. In zone A, Per unit clustering groups all three units into one single unit since their fuel and technology are identical and $\mathrm{P_j^{max}} < 30$ MW. In zone B, the units are grouped into one unit only if their flexibility is high (i.e. they can start/stop or ramp to full load in less than one hour). In zone C, all three HROR units are merged into a single one since their characteristics are similar and their minimum power is close to zero (the start-up and shut-down of these units is therefore barely noticeable).

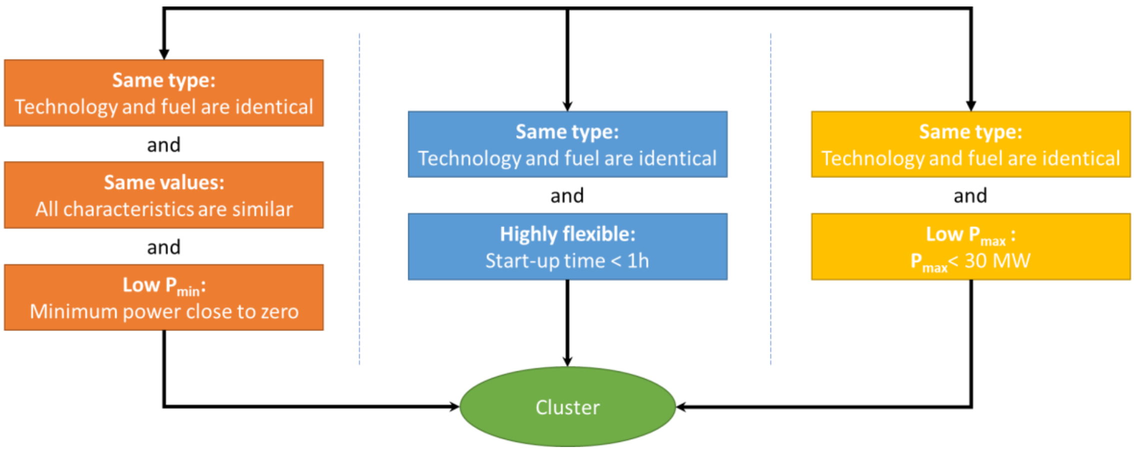

Figure 1 Graphical representation of different Per unit clustering decisions built in the Dispa-SET model [55].

2.1.3. Per typical unit clustering

In this integer formulation, a typical unit is considered for each technology, fuel and zone where it is located, and multiplied N times. Such unit is a typical or representative unit and is defined by averaging the characteristics of all units that belong to the cluster. Integer or binary commitment decisions are somewhat distinguishable from each other as with clustering the integer commitment state varies from zero to the number of units in the cluster. Such formulation still enables capturing of commitment decisions and to them associated relations for each unit. On the other hand, binary commitment decisions can only be related to the whole clusters and to them associated relations such as power output level, reserves contribution, etc. Integer formulation conserves the total number of units allowing a proper representation of constraints such as start-up costs, minimum up and down times and minimum stable load values. Figure 2 graphically represents the difference between binary and integer formulations. From there it is clear that in binary formulation each unit has its separate binary variable representing either on or off state of a unit while in clustering formation there is only one binary state dedicated to the commitment of the whole cluster and one integer variable dedicated to the number of units from that particular cluster being on-line. Example of such clustering has already been described in more detail in the No clustering paragraph.

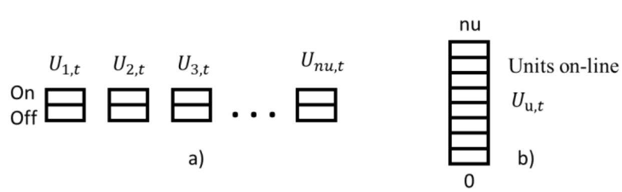

Figure 2 Comparison between binary (a) and integer (b) unit commitment formulation for a single type of unit in a single time period. Drawing inspired by [29].

Mathematically, clustering involves a modification of the original formulation of the unit commitment problem. All relations remain the identical to those in the binary formulation except ramping limits and minimum up and down times. They are now given by following inequalities:

$$P_{u,t-1} - P_{u,t} \le (U_{u,t} - S_{u,t}) \cdot \Delta P_u^{down,max} - P_u^{min} \cdot S_{u,t} + max(P_u^{min}, \Delta P_u^{down,max}) \cdot D_{u,t}$$
$$P_{u,t} - P_{u,t-1} \le (U_{u,t} - S_{u,t}) \cdot \Delta P_u^{up,max} + max(P_u^{min}, \Delta P_u^{up,max}, P_u^{quickstart}) \cdot S_{u,t} - P_u^{min} \cdot S_{u,t} \quad (5)$$

where $P_u^{quickstart} = a_u^{quickstart} \cdot P_u^{min}$. Formulation of the minimum down time now takes in to the account the number of units currently in off state. This leads to the following alteration of the equation (28) (provided in the Annex A of this paper):

$$n_g - U_{g,t} \ge \sum_{\tau = t - a_g^{mindown}}^{t} D_{g,\tau} \quad (6)$$

2.1.4. Per technology clustering

The Per technology clustering is a LP formulation of the power dispatch model, where all units are clustered by technology. In this formulation, all integer-based constraints are removed. Therefore, the Per technology clustering does not include constraints such as minimum up and down times, start-up costs and minimum stable load. Furthermore, since the start-up of individual units is not considered anymore, disaggregation isn't useful, thus all units that share the same technology, fuel and zone are merged into a single unit as proposed earlier. As an example, if one zone comprises three different GAS fired units: 10 MW GTUR, 500 MW COMC and 300 MW GTUR unit. Per technology clustering formulation leaves the COMC unit as is and groups the two GTUR units into a single one with no flexibility constraints what so ever. This is a commonly used approach in long term expansion planning models.

2.2. Performance metrics

Some key metrics such as computation time, total costs, energy mix and power dispatch have been proposed as good validation parameters by Palmintier et al. [50]. This is the foundation of this analysis and has been further expanded with more useful metrics such as analysis of power dispatch and load duration curves, congestion on interconnection lines, load shedding and curtailment. [57]

2.2.1. Computation Time

Computation time is reported as total CPLEX run time including GAMS model building and Python data preprocessing.

2.2.2. Total costs

Total costs of running the system are the values obtained by the optimization objective function and include all fixed, variable and operation costs. For comparison reasons, the percent difference is computed as follows:

$$\Delta C^{tot} = \frac{C^{tot} - C^{tot}_{base}}{C^{tot}_{base}} \tag{7}$$

where C^{tot} are total system costs (€) computed by alternative formulations and C^{tot}_{base} are total system costs (€) computed by No clustering formulation.

2.2.3. Energy mix

Energy Mix is based on total annual production aggregated by fuel type and divided in the same way as for clustering. All energy mixes ER_u are computed by summing the product of power outputs in all time intervals and divided by the total energy production of the system and is given by the following equation:

$$ER_u = \frac{\sum_{t \in \boldsymbol{T}} P_{\boldsymbol{u},t}}{\sum_{h \in H} D_{n,h}} \tag{8}$$

Difference between computed results and base results is given as a mean absolute difference and is formulated as follows:

$$\Delta EM_u = |ER_u - ER_u^{base}| \tag{9}$$

2.2.4. Power dispatch and Load duration curves

Statistical correlation between computed power outputs and load duration curves and base results is computed as the Spearman correlation index [58]. This method has been chosen since the expected data are not bivariate normal, making the other statistical indexes such as Pearson index inapplicable.

2.2.5. Merit order analysis

Comparison of power dispatch curves sorted by merit order is analyzed similarly to the energy mix, but instead of annual analysis, the difference between the base and alternative clustering is computed on an hourly basis:

$$PR_u = \frac{P_{\boldsymbol{u},t}}{D_{n,h}} \tag{10}$$

Difference between computed results and base results is given as a mean absolute difference and is formulated as follows:

$$\Delta MO_u = |PR_u - PR_u^{base}| \tag{11}$$

2.2.6. Startups

A number of startup events is an important metric that can only be applied to the No Clustering, Per unit and Per typical unit formulations. It is computed as the mean absolute difference of a number of commitment events for each time period and is normalized based on the total number of units online in that time period when compared to the baseline one.

2.2.7. Congestion

Congestion is another important metrics for analyzing the flows in the cross-border interconnection lines. It is computed as the difference between the number of congestion hours from the simulation and the number of congested hours from the base, normalized to the base case.

3. Scenario analysis

The proposed methods have been applied on three different scenarios. The first scenario is a replica of the Balkans energy system from the year 2015 and serves as a baseline for all further analysis. For that purpose, a complete input dataset (production, demand, prices, power plants, storage units etc.) has been gathered and is released as open data [52]. A preliminary version of this data was already proposed by the authors in [59]. The geographical area covers eight interlinked countries: Albania, Bosnia and Herzegovina, Croatia, Kosovo, North Macedonia, Montenegro, Slovenia and Serbia. The second and the third scenarios are a projection of the same energy system but for the years 2030 and 2050. They correspond to a transition of all individual countries in the region to a low carbon society by the integration of large amounts of RES-E. There are numerous GEP studies already published on this topic alone but a most recent one is the South East Europe electricity roadmap published by Szabó et al. [60]. Future generation and technology mixes from this publication have been taken as input data for this analysis with the addition of battery electric vehicles (BEVS) and CHP power plants. BEVS power and storage capacity and charging/discharging behaviour have been modelled as described by Beltramo et al. [61]. Number of electric vehicles in the region has been set to 25% in 2030 and 70% in 2050. The total number of vehicles has been estimated by the logistic growth function with an annual growth of 5% and upper bound of 800 that correlates the number of cars per 1000 inhabitants in currently most developed countries. Future fuel prices and carbon tax have been estimated based on the 2% annual inflation. Costs and operational parameters of conventional powerplants such as efficiency, minimum up and down times, ramping rates, ramping costs, start-up costs, no-load costs, minimum partial load, start-up time and CO_2 intensity have been taken from the Vilavenico et al. [62]. It is important to note that the same unit parameters have been conserved for the 2030 and 2050 scenarios because of the lack of data regarding future power plant characteristics. This hypothesis could be refined in future works, as it could possibly impact the results for future scenarios. CHP related parameters such as the power to heat and power loss factors have been obtained as described by Jiménez Navarro [32]. Hydro dam (HDAM) and hydro pumped storage (HPHS) self-discharge rates have been described by [39]. Self-discharge rates of thermal storage coupled by the steam turbine (STUR) or combined cycle (COMC) units by [59] and BEVS by [61]. Regional scope and nodal approximation of the analyzed region is presented in Figure 3.

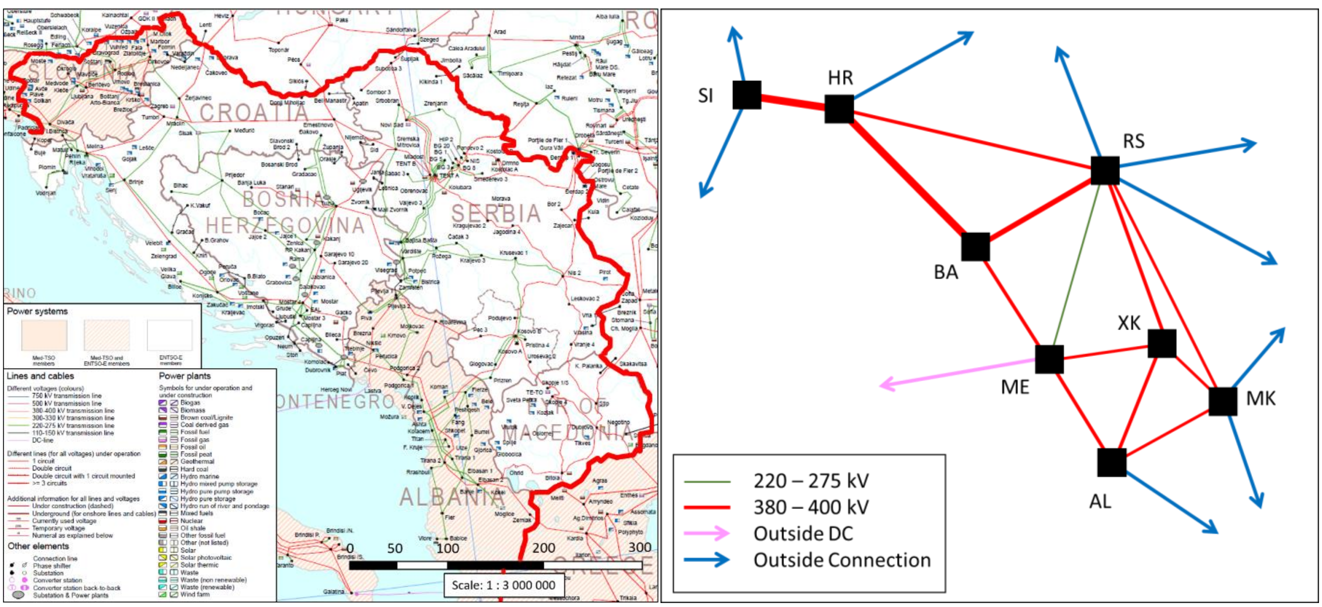

Figure 3 Geographical region covered in the scenarios. On the left is a cut-out piece from the interconnected network of Med-TSO 2018 [63]. The lines represented, constitute the relevant electricity grid, i.e. the portion of the grid that can affect the interconnections between the power systems. On the right is a nodal approximation used by the Dispa-SET model. Thickness of the interconnection lines highlights the total NTC capacities (as of 2018) and arrows highlight outside interconnections.

Total installed capacities aggregated by fuel type and location for all three scenarios are presented in Figure 4. From there it is clear that the reference scenario is dominated by fossil fuels such as lignite, coal, oil and gas. There are also substantial amounts of hydro capacities available in the region (45.3%), especially in Albania where the total hydro amounts to 98% of the peaking load. Other countries are dominated by lignite, the only two exceptions beeing Croatia and Slovenia who besides conventional fossil based technologies also have nuclear and renewables in form of sun and wind. Total RES-E capacity is 3% without and 23.8% with hydro run of river. It is important to note that Second scenario introduces new technologies such as BEVS (represented as OTH in the capacity plots) and RES-E across all eight countries. The share of renewable capacities is 24.5% without and 50.9% with the hydro run of river. The total share of hydro capacities amounts to 42.3%. Projections suggest that total installed capacity of BEVS will amount to 3.7% of the total installed capacities in the region. It is important to note that 40% of conventional powerplants will be decommissioned and out of operation. Third scenario is almost entirely based on RES-E technologies. There is only 1.9% of gas and 1% of lignite and coal combined. Most DH systems are replaced by CHP powerplants that run on biofuels such as biomass, biogas or other bio derivatives. Their total capacity amounts to 4.2%. The share of RES-E amounts to 52.1% without and 72.4% with the hydro run of river. The total share of BEVS is 8.2% without and 12.9% with pumped hydro storage.

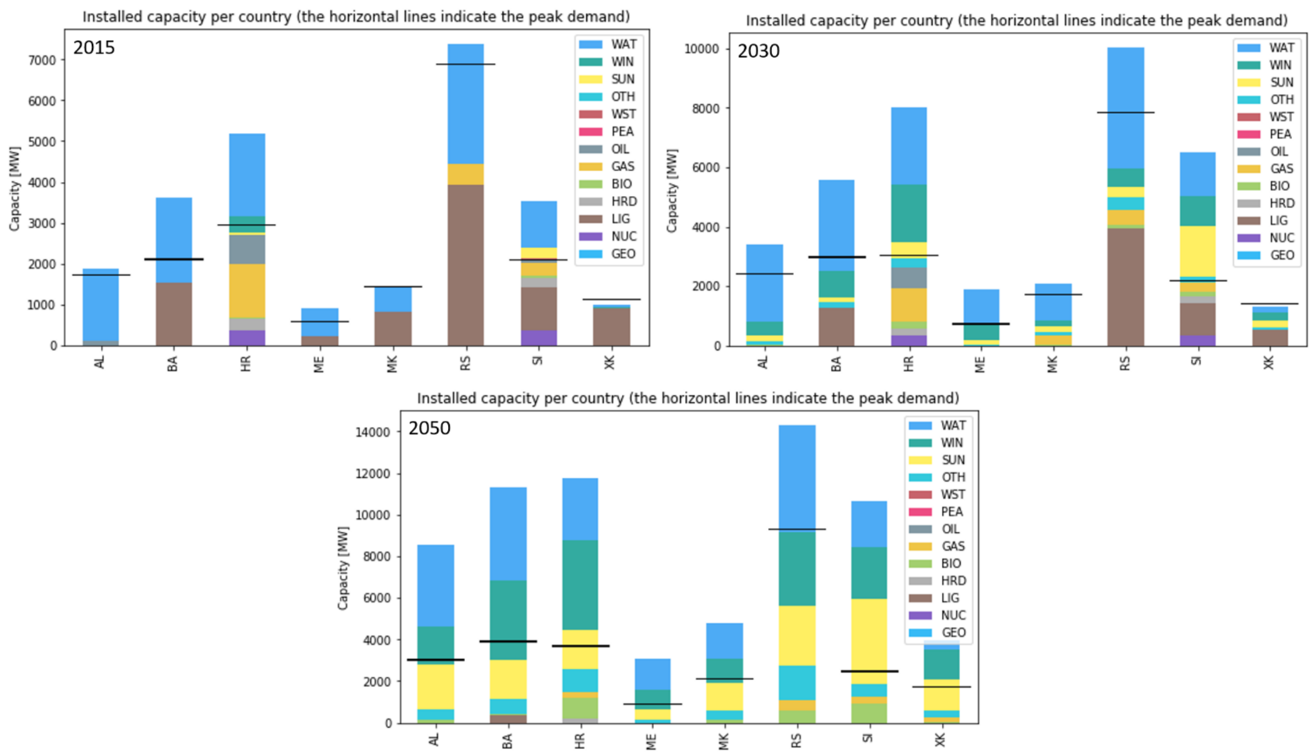

Figure 4 Graphical representation of installed capacities in all analysed countries. 2015 scenario (top), 2030 scenario (middle) and 2050 scenario (bottom)

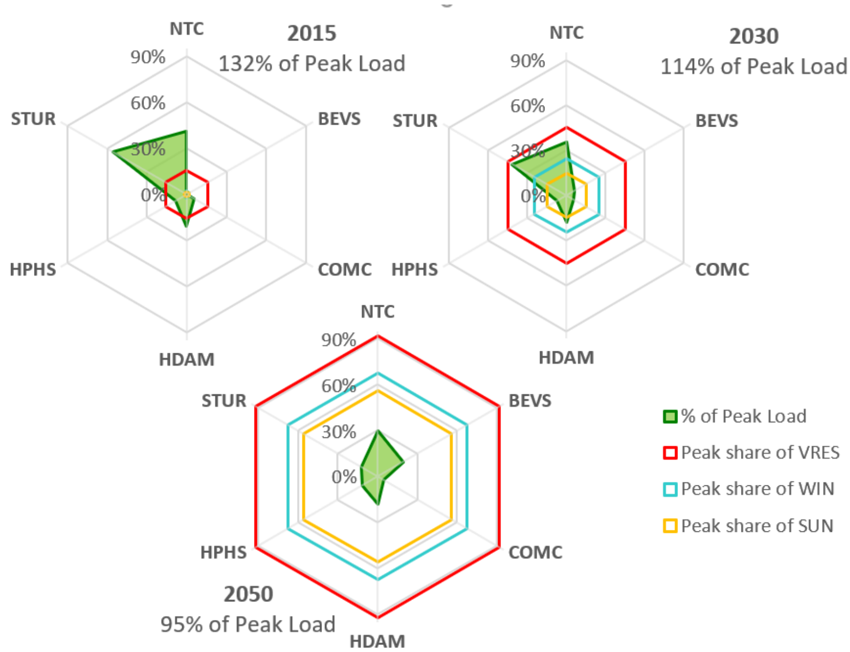

Figure 5 Flexibility charts providing an overview of generation-based flexibility in the region as well as the shares of wind and solar power and the share of all VRES combined.

In order to ensure that enough flexibility is available at all time, peeking demand of variable RES-E (VRE) is set to always be lower than the total capacity of technologies that can provide flexibility to the system. In reference scenario, 132% of the peak load can be covered by flexible technologies, 114% in the second scenario and 95% in the third scenario. For flexibility comparison purposes three flexibility charts, inspired by Yasuda et al. [64], have been created and presented in Figure 5. They provide a brief snapshot overview of the whole region.

Available NTC capacities between the zones are presented in Figure 6. Expansion of the NTC capacities in the second scenario has been modeled according to the planned projects from the ENTSOE TYNDP [65]. NTC capacities in the third scenario have been expanded in such a way that the whole region is well interconnected and has always enough capacity to shift RES-E between the countries.

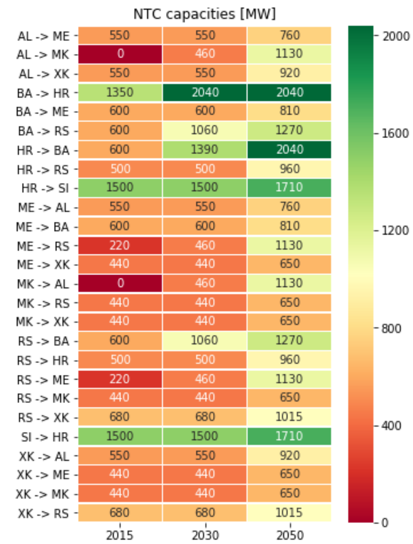

Figure 6 Available NTC capacities in the region. The red-yellow-green color scheme is used for easier identification of potential bottlenecks (dark red) and well interconnected zones (dark green) in the system.

Net residual load duration curves with timestep equaling one hour are plotted for all three scenarios as shown in Figure 7. They indicate the system's reliability to handle three different shares of VRES. It is clear that despite increased load shifting potential due to new BEVS and HPHS capacities, part of the RES-E generation has to be curtailed. RES-E time series downloaded from [66][67] is fixed for all three scenarios. Since the prediction of future hydrological flows is not in the scope of this paper, hydro inflows are also fixed for all three scenarios.

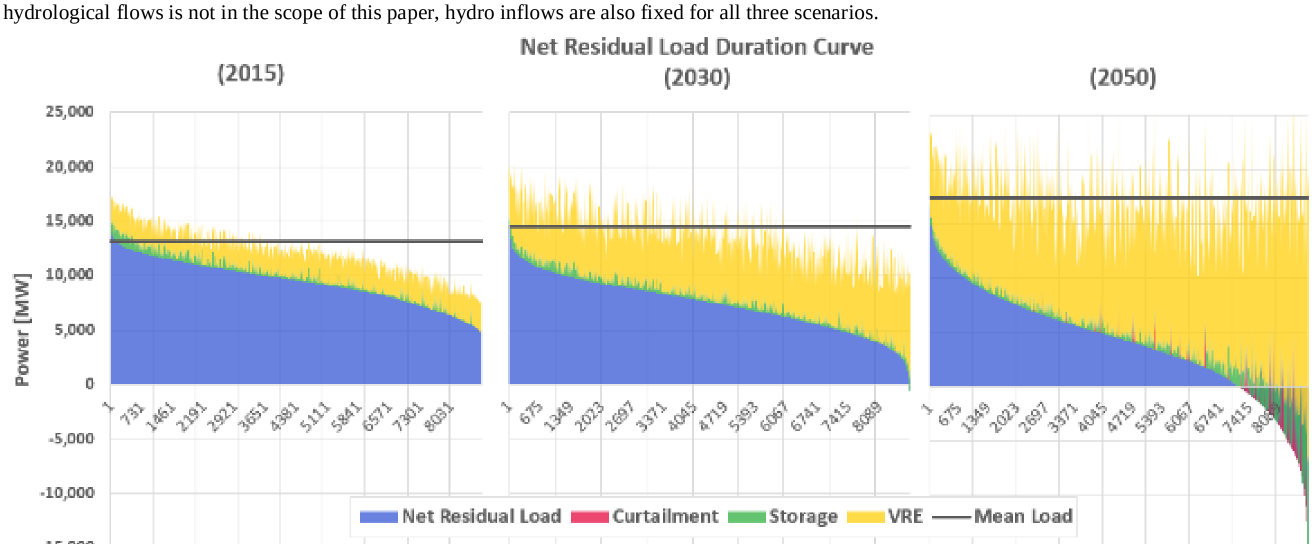

Figure 7 Net residual load duration curves (timestep = 1h) from all three scenarios. Negative values indicate V-RES overproduction that, in order to be utilized, needs to be shifted to time periods where V-RES production is lower than the actual demand (positive side of the y-axis).

4. Results

This section provides a detailed analysis of the results obtained from all twelve conducted case studies, four from each of the three scenarios. The main focus is the comparison of different formulations and clustering methods where key metrics such as speed and accuracy of the obtained results are analyzed. Besides, some other previously unpublished results from the Dispa-SET Balkans model is presented and discussed.

4.1. Computational efficiency

One of the key comparison metrics is the computation speed of different formulations subjected to different input conditions. Table 1 summarizes the computation time from all sixteen runs. All simulations have been run on a IntelI CoreI i7-5960 @ 3.00GHz (16 cores), 16 GB RAM shared server machine. Results show that for this particular system configuration, the difference between computation time can be in the range from 23.8%, for the Per unit formulation in 2015 scenario, up to 95.2%, in case of Per technology formulation in 2030 scenario. On average, Per unit formulation took x1.4, Per typical unit formulation x2.1 and Per technology formulation x19 less time than the base No clustering one.

Table 1 Computation time from all analyzed scenarios and cases

Time (s)	No clustering	Per unit	Per typical unit	Per technology
2015	5,765.31	4,392.51	2,804.01	351.05
2030	11,167.02	7,329.70	5,025.11	537.44
2050	9,248.48	6,478.15	4,703.67	464.79

The difference in memory usage is less drastic. The main reason is the long time horizons and many different technologies spread out through eight zones which limits the clustering possibilities significantly. Table 2 summarizes the memory usage from all sixteen runs. From there it is clear that memory usage can on average be reduced by x1.2 for Per unit, x1.5 for Per typical unit and x1.6 for Per technology clustering. Table 3 is a summary table of total units present in all three alternative formulations.

Table 2 Memory usage from all analyzed scenarios and cases

Memory (MB)	No Clustering	Per unit	Per typical unit	Per technology
2015	1,664	1,376	1,000	920
2030	1,984	1,672	1,368	1,254
2050	1,765	1,490	1,276	1,168

Table 3 Number of units present in all analyzed scenarios and cases

Units (-)	No Clustering	Per unit	Per typical unit	Per technology
2015	82	68	48	48
2030	101	86	70	70
2050	87	74	64	64

4.2. Accuracy

Key metrics for measuring the accuracy of different formulations have been discussed in more detail earlier. The main results are presented as follows: Costs, Energy mix, dispatch, merit order, load duration curves, startups and congestion

4.2.1. Costs

Average annual generation cost is another important accuracy metric. It is worth mentioning that all methods compute similar generation costs as presented in Table 4. The highest deviation from base case is 2.7% in case of Per typical unit clustering formulation in the third scenario. There is a trend of higher deviations at higher RES-E penetrations, but the main reason for that are generally much lower overall prices that are x1.6 lower in 2030 and x3.1 in 2050. Alternative clustering formulations tend to underestimate the electricity price in the range from 0.01 to 0.16 €/MWh. This can be associated with the startup and ramping events that are to some extent simplified in alternative formulations when compared to the No clustering one.

Table 4 Average annual generation cost from all analyzed scenarios and cases

Average annual generation cost (€/MWh)	No clustering	Per unit	Per typical unit	Per technology
2015	18.21	18.13	18.16	18.08
2030	11.54	11.53	11.59	11.52
2050	5.95	5.88	5.79	5.83

Total costs of running the system are presented in Table 5. This cost represent the computed objective function values. It is clear that higher shares of RES-E in alternative scenarios offset the increased demand, as costs are lower than in low RES 2015 scenario. Total costs are consistent in all four formulations. Depending on the amount of RES-E total costs tend to be higher in Per unit and Per typical unit formulations. Total costs of running the system in Per technology formulation are always lower than the ones from the No clustering formulation.

Table 5 Total costs of running the system from all analyzed scenarios and cases

Price (million €)	No clustering	Per unit	Per typical unit	Per technology
2015	1,790.055	1,791.257	1,794.835	1,786.881
2030	1,283.183	1,282.146	1,288.601	1,281.619
2050	814.546	804.933	792.258	797.747

4.2.2. Energy mix

Annual energy mix is an important measure and usual output from long term GEP models. Thus, it is important accuracy metrics that can be used as a judgment tool of overall energy planning. Results have shown that the absolute annual energy mix error varies in a range from 0.217% whit higher shares of RES-E in 2050 scenario up to 1.756% in the reference scenario. There is an obvious downward trend from 2015 scenario where higher amounts of fossil-based power plants are a norm to high RES-E 2050 scenario where most of the conventional powerplants have been decommissioned and replaced by new and more flexible technologies.

Table 6 Absolute annual energy mix error

	Per unit	Per typical unit	Per technology
2015	0.654%	0.347%	1.756%
2030	0.301%	0.385%	1.496%
2050	0.217%	0.603%	0.818%

Energy mix in all eight countries from the region, computed by the No clustering formulation is presented in Figure 8. Countries that were completely relying on the lignite production such as Serbia will have to face serious challenges after 2030 when most of the conventional power fleet will have to be decommissioned due to the old age of the units. Serbia, now almost entirely self-sufficient could become a major electricity importer by 2050. Diversification of energy production units could greatly benefit smaller nations with high RES-E potential such as Slovenia, Montenegro and Albania who could become major exporters in the region.

Although energy mix is an important factor from a long term planning perspective, when analyzed on its own it doesn't provide any indication for the flexibility requirements of the system. As shown later in the analysis, a small deviation in energy mix can significantly impact the merit order and dispatch curves.

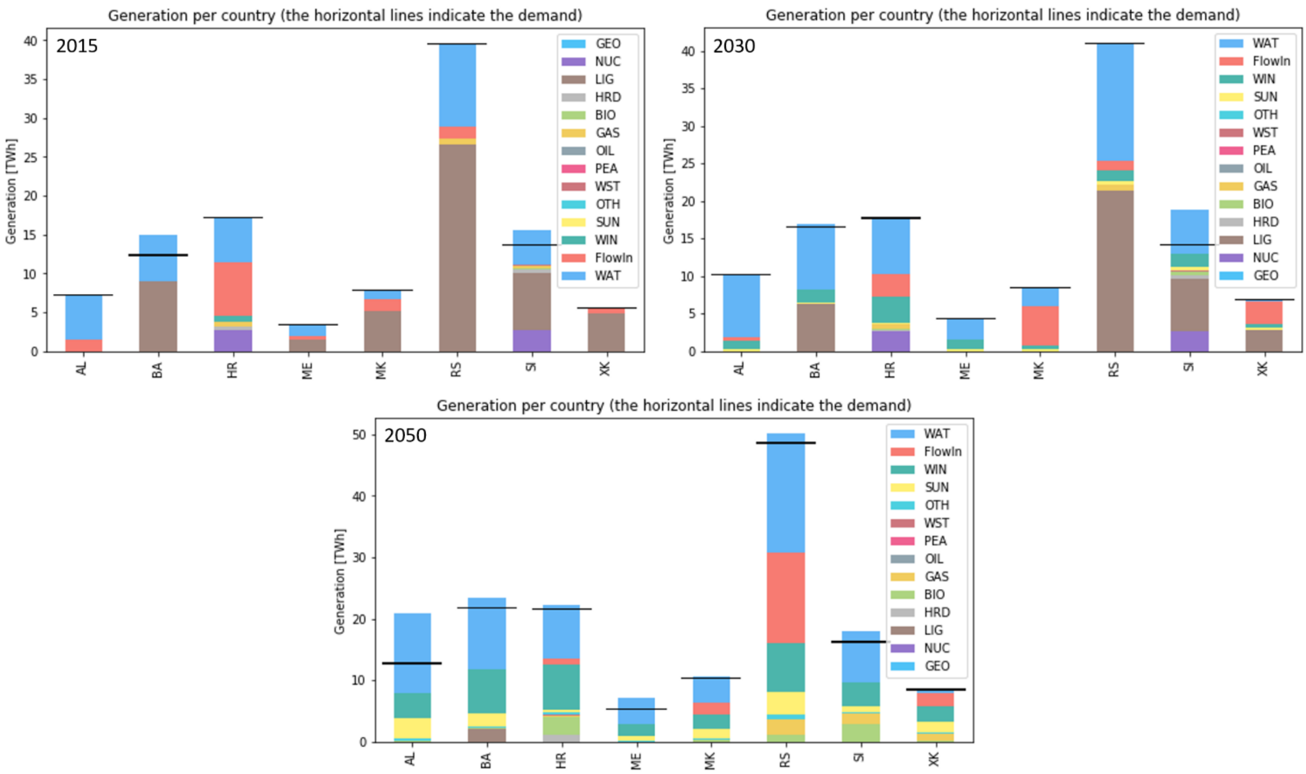

Figure 8 Energy mix in all three scenarios computed by No clustering formulation

4.2.3. Load duration curves

The load duration curves highlight the operation of the unit throughout the year. They can, for example, be used to check if the actual units are over or undersized. Load duration curves aggregated by fuel type for all three scenarios and all four formulations are presented in Figure 9. It can clearly be seen that most formulations have similar load duration curves with slight deviations from the No clustering one. The highest deviation can be observed for the Per technology formulation, which tends to overestimate the production from coal-fired units. This is explained by the lack of binary and integer variables, which allows unlimited start-ups and shut-downs. This causes the Per technology formulation to utilize the coal-fired power plants for frequency regulation.

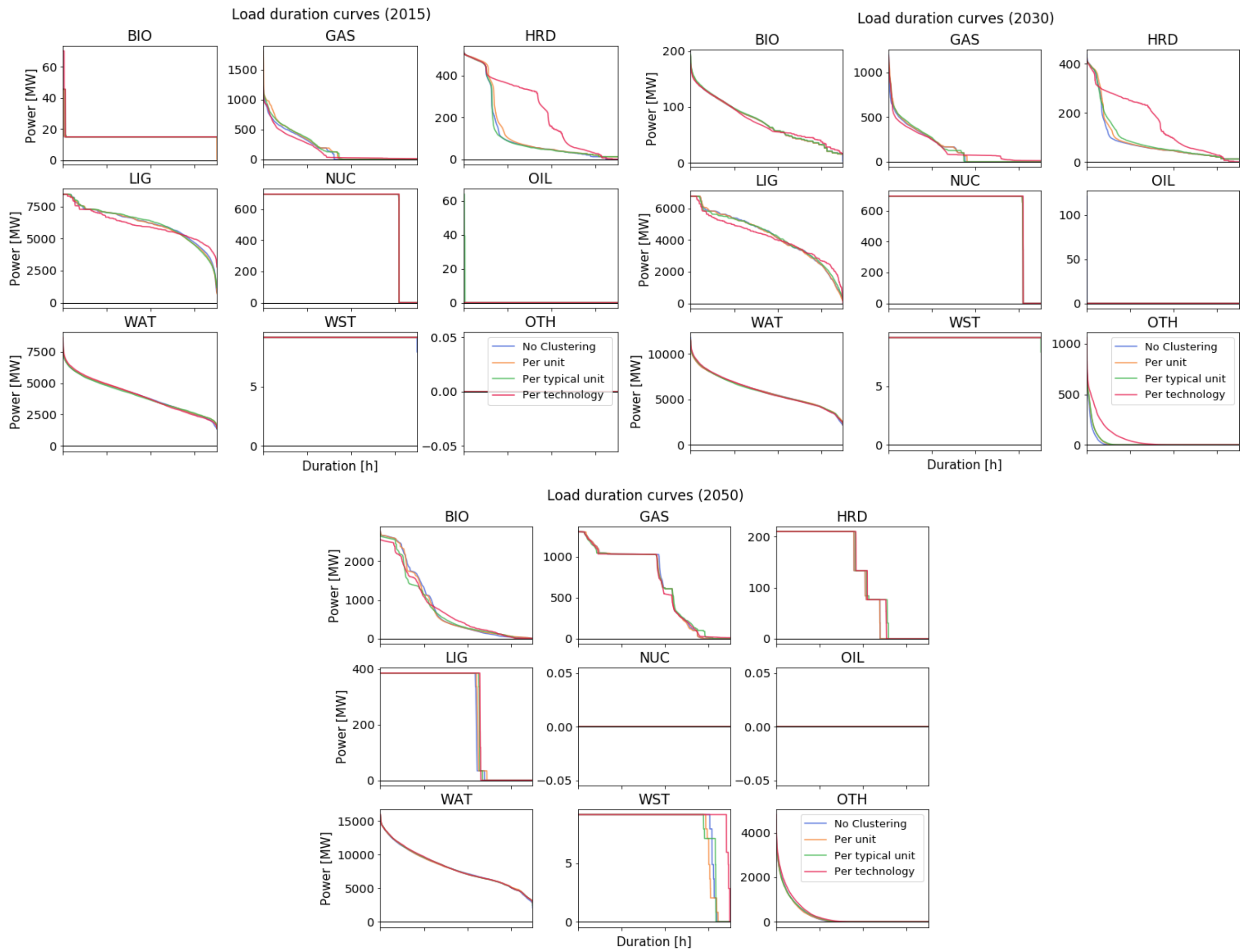

Figure 9 Load duration curves for all four formulations in all three scenarios. Per technology clustering (red) deviates the most from the No clustering formulation (blue) in all three scenarios. Per unit (orange) and Per typical unit (green) follow the load duration curves from the No clustering more closely.

4.2.4. Power dispatch

Table 7 provides the Spearman correlation coefficients computed for all alternative formulations when compared to the No clustering one. From the table, it is clear that Per unit clustering is in general most accurate, while the Per technology formulation is least accurate. On average, the Per technology clustering correlates only around 75% with the generation dispatch from the baseline. This difference becomes less significant in 2030 when it averages 78% and 89% in the 2050 scenarios with high penetrations of VRES. It is important to mention that cheapest energy sources such as nuclear seem to have a high correlation of over 98% to the base case since it is rarely used for flexibility.

Table 7 Spearman correlation coefficients between alternative formulations and binary formulation.

	2015			2030			2050		
	Per unit	**Per typical unit**	**Per technology**	**Per unit**	**Per typical unit**	**Per technology**	**Per unit**	**Per typical unit**	**Per technology**
BIO	0.9701	0.6996	0.6416	0.9991	0.9983	0.9750	0.9749	0.9195	0.9130
GAS	0.9594	0.9453	0.7442	0.9793	0.9804	0.7271	0.9585	0.9226	0.9097
HRD	0.9255	0.8736	0.6670	0.9673	0.8702	0.6073	0.9968	0.9446	0.9498
LIG	0.9502	0.8765	0.8542	0.9392	0.9111	0.8550	0.9526	0.9390	0.9280

NUC	0.9845	0.9800	0.9861	0.9877	0.9840	0.9821	-	-	-
OIL	-	-	-	-	-	-	-	-	-
OTH	-	-	-	0.4963	0.4498	0.3114	0.7414	0.6866	0.6926
WAT	0.9306	0.8483	0.8495	0.9369	0.9228	0.8926	0.9827	0.9685	0.9576
WST	-	-	-	0.2154	-0.0012		0.8465	0.7546	0.2206

4.2.5. Merit order

The main purpose of this analysis is the investigation of potential mismatches between the actual power dispatch computed in the baseline and alternative formulations. Those differences have been analyzed as the errors of individual fuel types and have been computed for every time period in the optimization horizon. Merit order error curves from all three scenarios and all three alternative formulations are presented in Figure 10. It is clear that the highest errors are produced by the Per technology formulation. This is especially noticeable for lignite and hydro production where merit orders can, in some instances, differ up to 40% from the baseline one. Other two formulations also tend to compute different merit order curves, but they are less extreme. The lack of start-up and shut-down constraints in the Per technology formulation can be clearly seen for BIO units in the 2030 scenario. Those units are continuously turned on and off and are operating mostly below minimum stable power output. There is a clear correlation between lignite and hydro units as in this particular system configuration and they tend to complement each other as two main energy sources. This means that in most cases different clustering formulations tend to either overestimate or underestimate flexibility of the lignite-fired powerplants. This phenomenon needs compensation which in this particular case study can only be covered by flexible HDAM or HPHS units. Another reason for such high mismatch is the aggregation of HPHS and HDAM units. This aggregation either sums or aggregates original units and minimum reservoir levels into a single unit which consequently leads to the loss of accuracy in the generation. Generation is also limited by the production in every single optimization horizon and although total energy outputs of all four formulations in all simulation horizons fall within 1%, each formulation produces a different dispatch solution.

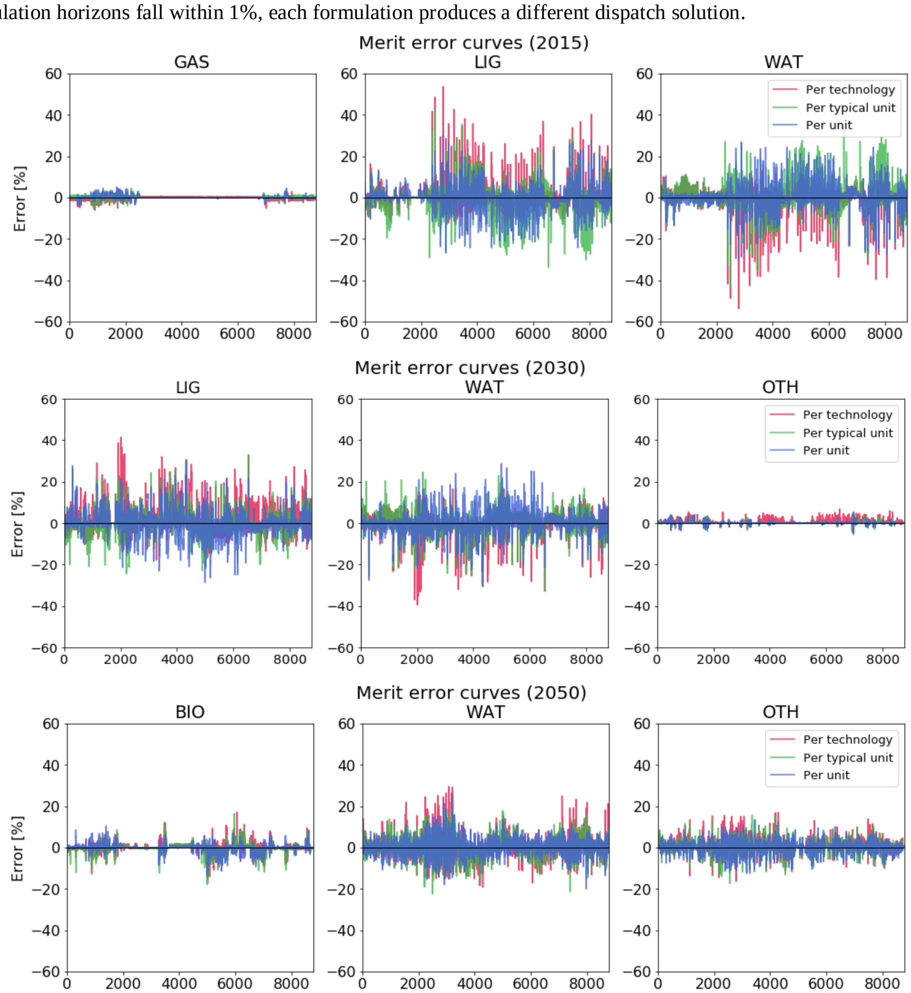

Figure 10 Merit order error curves of most influential non-RES-E fuel types from all scenarios. On average Per technology clustering (red) deviates the most from the No clustering in all three scenarios. Deviations from other two formulations, Per unit (blue) and Per typical unit (green), are (although still significant) less pronounced.

Figure 11 and Figure 12 represent the power dispatch plots from all four formulations in scenarios 2015 and 2030 for Bosna and Herzegovina. These mismatches in production from different energy sources are clearly visible between No clustering and Per technology clustering formulations in 2015 scenario. High RES-E 2050 scenario deviates less but has the tendency to overestimate the need and use of storage.

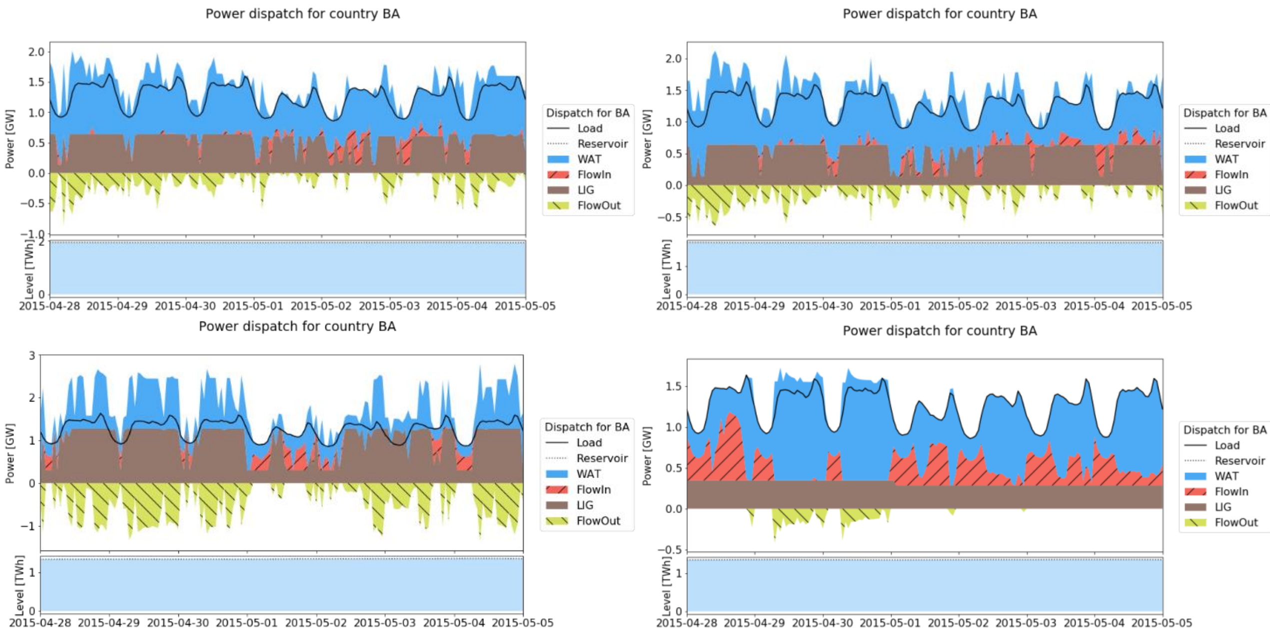

Figure 11 Power dispatch for Bosna and Herzegovina in scenario 2015. No clustering (top), Per unit (upper middle), Per typical unit (lower middle) and Per technology (bottom)

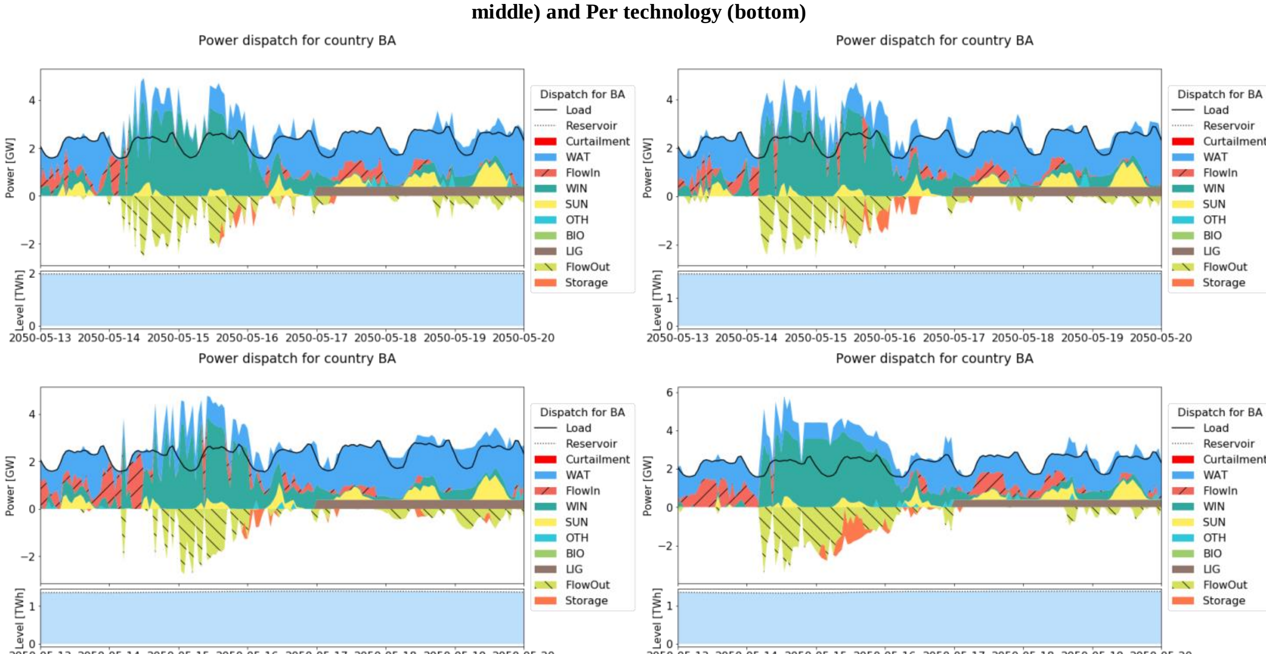

Figure 12 Power dispatch for Bosna and Herzegovina in scenario 2050. No clustering (top), Per unit (upper middle), Per typical unit (lower middle) and Per technology (bottom)

4.2.6. Startups

The number of startups is another important metrics affecting the operation and ageing of power generation units. This metrics is only applicable for No clustering, Per unit and Per typical unit clustering formulations. Per technology formulation, on the other hand, cannot be analyzed with these metrics since it doesn't have binary and integer variables build into the model. Results, presented in Table 8, have shown that in scenario 2015, which has low shares of RES-E Per unit formulation tends to produce more accurate results. The number of startups for all fuel types is less or equally different than the ones computed by the No clustering one. Per unit formulation underestimates the number of startups from lignite-fired powerplants by 17.9%, while Integer formulation tends to compute 71.4% less. In 2030 scenario RES-E penetration is increased to 46%. In this case, Per unit formulation still performs better and computes a lower number of startups than the No clustering one. In 2050 scenario, where RES-E share amounts to 92% Per typical unit formulation tend to perform slightly better than Per unit one. Those results, while they cannot be safely generalized to all cases, indicate a trend where the number of startups in systems with high shares of RES-E is affected more than in configurations with relatively low shares of RES-E.

Table 8 Start-up difference between No clustering, Per unit and Per typical unit clustering formulations for all three scenarios

	2015		2030		2050	
	Per unit	Per typical unit	Per unit	Per typical unit	Per unit	Per typical unit
OTH	-	-	12.7%	12.9%	1.2%	-5.1%
BIO	0.0%	50.0%	0.0%	266.7%	-37.1%	-7.1%
LIG	-17.9%	-71.4%	-6.8%	-22.7%	-13.3%	-13.3%
GAS	-37.1%	71.4%	45.7%	-60.9%	-2.3%	11.6%
OIL	-	-	-100%	-100%	-	-
WST	-	-	-	-	50.0%	50.0%
HRD	-66.7%	-66.7%	-10.0%	50.0%	0.0%	-11.1%
WAT	-25.2%	-33.2%	-35.6%	-48.6%	-31.0%	-62.8%
NUC	0.0%	0.0%	0.0%	0.0%	-	-

4.2.7. Congestion

Congestion on the interconnection lines is another useful accuracy metrics that can be used to identify potential bottlenecks in the system. All scenarios have independent interconnection capacities. They represent future configurations of the system that could integrate 92% of RES-E, and thus cannot be compared. The congestion metrics is therefore computed individually for all three scenarios. Figure 13 provides the percentage of hours in which the interconnection line is congested. It can be seen that in scenario 2015, the interconnection between Serbia and Montenegro and Serbia and North Macedonia is quite critical as lines are maximally loaded almost 4000 hours per year. In this scenario, Per unit and Per typical unit formulations tend to mimic the results from the No clustering one. Per unit formulation tends to compute two bottlenecks more, and Per typical unit formulation three more than the No clustering one. Per technology formulation tends to underestimate the usage of NTC's. In the 2030 scenario, all NTC capacities have been increased. Interconnection capacities from Serbia are still problematic and the congestion is even higher and ranges from 54 to 56%. It is important to note that in this case Per unit and Per technology formulations are closest to the No clustering one. Per typical unit formulation still points out to the potential bottlenecks but still tends to underestimate others. In scenario 2050, with high shares of RES-E, results are even more stable than before. All four formulations are reliable. The main reason behind this is a more linear correlation between production units, as almost 92% of the demand is covered by VRES technologies.

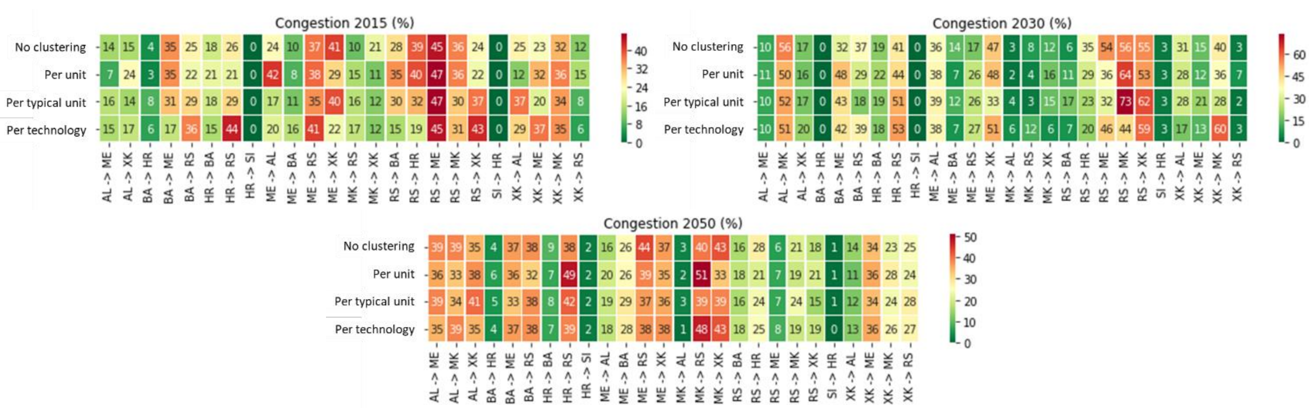

Figure 13 Proportion of congested hours in all interconnection lines from all three scenarios. The red-yellow-green color scheme is used for easier identification of potential bottlenecks (dark red) and well interconnected zones (dark green) in the system.

5. Conclusions

In this work, a comprehensive, free and open source model of the Western Balkans power system is developed for the years 2015, 2030 and 2050. The proposed years correspond to three scenarios (one historical and two future ones) that are used for testing and comparing various modelling formulations and simplifying hypotheses. In total, four different formulations are analyzed: "No clustering", "Per unit", "Per typical unit" and "Per technology" clustering.

Analysis has shown that the current Western Balkans power system mostly relies on lignite and hydro capacities as generation from these two energy carriers amounts to more than 90% (around 45% each). The energy mix in individual countries varies significantly. On one side Kosovo is dominated by lignite powerplants and on the other side, Albania is entirely powered by hydro units. Some neighboring countries like Slovenia and Croatia, and Bosna and Herzegovina and Serbia are well interconnected, while other ones like Albania and North Macedonia had, in 2015, no interconnections at all. RES-E capacities such as wind and solar are present only in the two EU member states, Croatia and Slovenia, and are almost nonexistent in the Western Balkans power system. This lack of RES-E capacities, together with relatively high amounts of flexible HDAM and HPHS units and to some extent already well-developed transmission system could potentially integrate up to 30% renewables without compromising the stability and integrity of the system. If all future transmission extension projects in the region are realized, additional 17% of renewables could be integrated into the system by the year 2030. If the transition of the whole region from mostly fossil based to more than 90% VRES based system is to be achieved by the year 2050, transmission capacities will have to be expanded significantly. Lack of flexibility in such future scenarios would require at least 7,087 MW of storage (1,502 MW HPHS and 5,585 MW BEVS). Results have shown that this transition could be easily achieved if 70% of all vehicles in the Western Balkans were electrified and used for balancing and flexibility purposes.

Earlier publications have already proven that the binary formulation is the state-of-the-art modeling framework that can accurately represent real-world systems [50]. This was the starting point of this analysis, in which the performance of alternative formulations has been compared to the baseline. Results have shown that computation time in a “Per unit” formulation was x1.4, “Per typical unit” x2.1 and “Per technology” x19 lower than the baseline. Computed prices and energy mixes are quite similar, and they do not point out the potential errors that alternative formulations tend to produce. This is especially true in well interconnected systems with high shares of RES-E. The load duration curves, power dispatch and merit order curves in the proposed system show that the LP formulation is the least reliable formulation as dispatch, in some cases, tends to be off by up to 54%. Other two alternative formulations have produced a more reliable results but were still unable to replicate the number of start-up events and congestion bottlenecks in the transmission lines. Results also tend to show that these alternative formulations are reliable for estimating the energy and power mixes in future energy systems. This is especially true for systems with high and extremely high shares of RES-E as the deviation from the baseline formulations decreases significantly with the number of conventional and inflexible units.

This work should be seen as an attempt to compare various model formulations and extract general guidelines. However, the comparison is performed for a specific case only and cannot be considered as a comprehensive and exhaustive model comparison work, which would be out of the scope of this paper. The results from this analysis are not a guarantee that the same conclusion could be drawn for different system configurations. Future work should focus on applying the same type of comparison methodology on different energy systems (e.g. higher RES-E shares, inclusion of battery storage, inclusion of power-to-X technologies, etc.), different spatial resolutions (e.g. with more than one node per country) and temporal resolutions (e.g. sub-hourly).

Finally, it is worthwhile to note that, in order to ensure good transparency and reproducibility of the work [68], the model source code and the input data are provided using open licenses. They can be download from the model web page [69] and from the Dispa-SET Balkans - Dataset repository [52].

6. Annex A – Model formulation

6.1. Model formulation

Traditional formulations of unit commitment problem are dedicated to the minimization of operating costs of two or more generator units committed to meet the power demand from the network.

6.1.1. Objective function

In most simple case the objective function can be formulated as sum of fixed and variable costs. This paper expands the classical formulation with additional costs and is formulated as follows:

$$C^{tot} = min \sum_{n\in N} \sum_{u\in U} \sum_{t\in T} \left(C^{fix}_{u,t} + C^{start}_{u,t} + C^{var}_{u,t} + C^{ramp}_{u,t} + C^{trans}_{t,l} + C^{shed}_{n,t} + \sum_{chp} C^{CHP}_{chp,t} + C^{VOLL}_{n,t,u} \right) \tag{12}$$

where C^{tot} are the total operation costs (€); $C^{fix}_{u,t}$ are fixed costs (€) of running the unit, u, in all time periods, t; $C^{start}_{u,t}$ are the start-up costs (€) and $C^{var}_{u,t}$ are variable costs of all units, u, and all time periods, t; $C^{trans}_{t,l}$ are transmission costs (€) directly depending by the flow transmitted through the lines, l; $C^{CHP}_{chp,t}$ are costs (€) associated to the CHP plants, chp; $C^{shed}_{n,t}$ are load costs (€) associated to the necessary load shedding and $C^{VOLL}_{n,t,u}$ are costs of lost load (€) associated to each zone, n.

6.1.2. Fixed costs

Fixed costs represent operation and maintenance (O&M) costs and other costs associated for running the unit. They are formulated as follows:

$$C^{fix}_{u,t} = U_{u,t} \cdot c^{fix}_{u} \tag{13}$$

where $U_{u,t}$ represents the commitment (on/off) of each unit and is usually set to 1 if running and 0 if shut down and c^{fix}_{u} are fixed costs (€) of operating unit u.

6.1.3. Start-up costs

Start-up and shut-down costs represent the costs of (de)committing the unit (on/off state) and can be expanded by the additional fixed costs such as personnel and maintenance. They approximate the fuel consumption of a start-up event and are assumed to be a fixed value. They are formulated as follows:

$$C_{u,t}^{start} = S_{u,t} \cdot f_u^{start} \cdot c_u^{fuel} \tag{14}$$

where $S_{u,t}$ ($D_{u,t}$) represents the start-up and shot-down events (-), f_u^{start} is a fuel use per start-up (MWh_f/start), c_u^{fuel} are additional fixed costs per start (€/MWh_f) and c_u^{fix} are all additional fixed costs (€). This formulation is appropriate for long-term unit commitment problems and does not take into the account the warm and cold start-up costs [29]. Number of start-up events can be formulated as follows:

$$U_{u,t} - U_{u,t-1} = S_{u,t} - D_{u,t} \quad \forall\, U_{u,t}, S_{u,t}, D_{u,t} \in \{0,1\} \tag{15}$$

6.1.4. Variable costs

Variable costs include fuel costs and variable CO_2 emission costs and are formulated as follows:

$$C_{u,t}^{var} = \sum_{n\in \boldsymbol{N}} \sum_{f\in \boldsymbol{F}} \left(\frac{F_{u,f} \cdot c_{n,f,t}^{fuel} \cdot l_{u,n}}{\eta_u} \right) + \sum_{p\in \boldsymbol{P}} \left(EM_{u,p} \cdot c_p^{polut} \right) \tag{16}$$

where $F_{u,t}(P_{u,t})$ is the fuel usage (MWh_f) given as a function of the power output $P_{u,t}$ (MW_e). $c_{u,t}^{fuel}$ are fuels costs in any given time period (€/MWh_f); $l_{u,n}$ is the location, n, of the units, u; η_u, electrical efficiency of any given unit (-), $EM_{u,p}$ are the emission rates (t) of individual units given as function of fuel consumption and technology-specific pollutants p; and c_p^{polut} is fixed price (€/t) of each individual pollutant, p.

6.1.5. Ramping costs

Conventional units are also characterized by hidden costs due to ramping. Ramping refers to how fast a thermal unit can adjust its power output. The importance of cycling costs is discussed in more detail by Keatley et.al. [70]. They represent the costs due to the ageing of the power plants, mainly caused by thermal stress due to the varying operating conditions of units which were sometimes designed to run at nominal capacity most of the time. Ramping costs are modeled as follows:

$$C_{u,t}^{ramp} = C_{u,t}^{ramp,up} + C_{u,t}^{ramp,down} \tag{17}$$

where $C_{u,t}^{ramp,up}$ and $C_{u,t}^{ramp,down}$ are ramping up and ramping down costs (€). Ramping costs are defined as positive variables (i.e. negative costs are not allowed) and are computed with the following equations:

$$\begin{aligned} C_{u,t}^{ramp,up} &\geq c_u^{up} \cdot \left(P_{u,t} - P_{u,t-1}\right) \\ C_{u,t}^{ramp,down} &\geq c_u^{down} \cdot \left(P_{u,t-1} - P_{u,t}\right) \end{aligned} \tag{18}$$

where $c_{\mathbf{u}}^{up}$ and $c_{\mathbf{u}}^{down}$ are ramping up and ramping down costs (€/MW).

6.1.6. Transmission costs

Transmission costs are costs associated to the energy flows through the cross-border interconnection lines and are given by the following expression:

$$C_{t,l}^{trans} = c_l^{trans} \cdot E_{l,t}^{trans} \tag{19}$$

$$E_l^{trans,min} \leq E_{l,t}^{trans} \leq E_l^{trans,max} \tag{20}$$

where $c_{\mathbf{l}}^{trans}$ are fixed costs (€/MW) of using the line l; $E_{l,t}^{trans}$ are the energy flows (MW) through the lines; $E_l^{trans,min}$ and $E_l^{trans,max}$ are minimal and maximal capacities (MW) of the transmission lines l. In this particular study transmission costs have no value (costs of energy exchange between different zones equals 0). This simplification eliminates the unequal pricing in different interconnections which forces the model to utilize them based on the energy needs rather than the costs.

6.1.7. Load Shedding costs

Load shedding costs are costs that occur due to the necessary load shedding in time periods when demand is higher than the sum of available generation capacities and cross-border interconnection flows and is expressed similarly to transmission costs:

$$C_{n,t}^{shed} = c_n^{shed} \cdot E_{n,t}^{shed} \tag{21}$$

$$E_{n,t}^{shed} \leq E_n^{shed,max} \tag{22}$$

where $c_{\mathbf{n}}^{shed}$ are fixed costs (€/MW) of load shedding in any particular zone n, and $E_{n,t}^{shed}$ is the amount of energy (MW) being shed and $E_n^{shed,max}$ is the maximum load shedding (MW) allowed in a particular zone n. In practice load shedding is an additional safety mechanism that can be enforced to prevent system blackouts. Their costs are of contractual nature between consumers that significantly impact the demand curves and the system operator. Usually load shedding contracts are signed by large industrial facilities which can decrease their production for a certain amount of time. These costs are significantly higher than the shadow price of additional MW generated by the system.

Load Shedding should be distinguished from “Lost Load” which is the load that cannot be supplied by the system, and which is defined in Equation (12) in order to avoid a solver failure.

6.1.8. CHP and storage costs

Costs of running the CHP plants are described in more detail by Jiménez Navarro et all. [32] who also use the Dispa-SET modeling tool to conduct the analysis on the joint effects of centralized cogeneration plants coupled with thermal storage on the efficiency and cost of the power system. These costs take into the account the costs of heat from other/backup heat sources that are not available in the model and the variable costs of producing heat from either backpressure or extraction turbines, or power to heat technologies. The same authors also describe the formulation of thermal storage with a high level of detail. Fernandez-Blanco et al. [39] quantifies the water-power linkage on hydrothermal power systems by using Dispa-SET model. In this publication a formulation of hydro storage units is described in more detail. Since those two publications already cover the formulation of storage and CHP units they wont be detailed here.

6.1.9. Costs of lost load

Lost load occurs when power exceeds the demand or is not able to match it due to ramping limitations and lack of reserve margins. It is formulated as follows:

$$C_{n,t,u}^{VOLL} = c_{n,t}^{pwr} \cdot \left(LL_{n,t}^{\overset{max}{pwr}} + LL_{n,t}^{\overset{min}{pwr}} \right) + c_{n,t}^{rsrv} \cdot \left(LL_{n,t}^{2U} + LL_{n,t}^{2D} + LL_{n,t}^{3U} \right) + c_{u,t}^{ramp} \cdot \left(LL_{u,t}^{\overset{ramp}{up}} + LL_{u,t}^{\overset{ramp}{down}} \right) \tag{23}$$

where $c_{n,t}^{pwr}$ are the costs (€/MW) of lost load due to maximal and minimal power; $LL_{n,t}^{\overset{max}{pwr}}$ is deficit in terms of maximum power (MW), $LL_{n,t}^{\overset{min}{pwr}}$ suficit when power is exceeding the demand (MW), $c_{n,t}^{rsrv}$ are the costs (€/MWh) of lost load due to lack of up and down reserves, $LL_{n,t}^{2U}$ is deficit in reserve up (MWh), $LL_{n,t}^{2D}$ is deficit in reserve down (MWh), $LL_{n,t}^{3U}$ deficit in non-spinning reserve up (MWh), $c_{u,t}^{ramp}$ are costs of lost load due to lack of ramping capacities (€/MWh), $LL_{u,t}^{\overset{ramp}{up}}$ and $LL_{u,t}^{\overset{ramp}{down}}$ are deficits in terms of ramping up and down for each plant, u.

6.1.10. Energy balance

Solving the unit commitment problem results in optimal dispatch of given units they are dedicated to covering the demand. Thus, the following system balance constraint is mandatory in all unit commitment formulations and ensures that the sum of all power outputs $P_{u,t}$, (MW) is equal to the sum of all the demands $D_{n,h}$, (MW), at all time periods:

$$\sum_{u \in \boldsymbol{U}} \left(P_{u,t} \cdot l_{u,n} \right) + \sum_{l \in \boldsymbol{L}} \left(E_{l,t}^{trans} \cdot l_{l,n}^{node} \right) = D_{n,h} + \sum_{r \in R} \left(Stor_{s,h} \cdot l_{s,n} \right) - E_{n,t}^{shed} - LL_{n,t}^{\overset{max}{pwr}} + LL_{n,t}^{\overset{min}{pwr}} \tag{24}$$

where $Stor_{s,h}$ is storage input (MW) of storage unit s located in $l_{s,n}$ zone.

6.1.11. Reserve constraints

Besides the production/demand balance, the reserve requirements (upwards and downwards) in each node must be met as well. In Dispa-SET, three types of reserve requirements are taken into account: 2U and 2D reserve that can be covered by spinning units and 3U reserve that can be covered either by spinning units or by quick start offline units. Those reserve requirements and formulations are described in more detail in the Dispa-SET online documentation [55].

6.1.12. Power output

The minimum power output is determined by the must-run constrained where unit, when committed, must operate at the stable generation level. Thus, it is necessary to introduce minimum and maximum output constraints that can be modelled as follows:

$$U_{u,t} \cdot P_u^{min} \leq P_{u,t} \leq U_{u,t} \cdot P_u^{max} \tag{25}$$

where P_u^{min} is the minimum power output (MW) and P_u^{max} maximum power output (MW) of unit g.

6.1.13. Ramping constraints

All thermal units are characterized by a maximum ramp-up and ramp-down capability. This are inequality constraints:

$$P_{u,t-1} - P_{u,t} \leq \left(U_{u,t} - D_{u,t} \right) \cdot \Delta P_u^{down,max} + D_{u,t} \cdot \Delta S_u^{down,max} - S_{u,t} \cdot P_u^{min} + LL_{u,t}^{\overset{ramp}{down}} \tag{26}$$

$$P_{u,t} - P_{u,t-1} \leq \left(U_{u,t} - S_{u,t} \right) \cdot \Delta P_u^{up,max} + S_{u,t} \cdot \Delta S_u^{up,max} - D_{u,t} \cdot P_u^{min} + LL_{u,t}^{\overset{ramp}{up}} \tag{27}$$

where $\Delta P_u^{down,max}$ and $\Delta P_u^{up,max}$ are ramp-up and ramp-down limits (MW), $\Delta S_u^{up,max}$ and $\Delta S_u^{down,max}$ are ramp-up and ramp-down limits at the startups.

6.1.14. Minimum up and down times

This are the minimum up and down times that limit the operation of the units by the amount of time the unit has been running or stopped. They are formulated as follows:

$$U_{u,t} \geq \sum_{\tau=t-a_u^{min\,up}}^{t} S_{u,\tau}$$
$$1 - U_{u,t} \geq \sum_{\tau=t-a_u^{min\,down}}^{t} D_{u,\tau} \quad (28)$$

where $a_u^{\min up}$ and $a_u^{\min down}$ are minimum up and down times (h).

REFERENCES

[1] Papakostas K, Mavromatis T, Kyriakis N. Impact of the ambient temperature rise on the energy consumption for heating and cooling in residential buildings of Greece. Renewable Energy 2010;35:1376–9. doi:10.1016/j.renene.2009.11.012.

[2] de Bruin K, Dellink RB, Ruijs A, Bolwidt L, van Buuren A, Graveland J, et al. Adapting to climate change in The Netherlands: an inventory of climate adaptation options and ranking of alternatives. Climatic Change 2009;95:23–45. doi:10.1007/s10584-009-9576-4.

[3] Dewulf A. Contrasting frames in policy debates on climate change adaptation: Contrasting frames on climate change adaptation. Wiley Interdisciplinary Reviews: Climate Change 2013;4:321–30. doi:10.1002/wcc.227.

[4] Vijaya Venkata Raman S, Iniyan S, Goic R. A review of climate change, mitigation and adaptation. Renewable and Sustainable Energy Reviews 2012;16:878–97. doi:10.1016/j.rser.2011.09.009.

[5] Intergovernmental Panel on Climate Change, editor. Summary for Policymakers. Climate Change 2013 - The Physical Science Basis, Cambridge: Cambridge University Press; 2014, p. 1–30. doi:10.1017/CBO9781107415324.004.

[6] European Commission. EU - 2030 climate & energy framework. Climate Action - European Commission 2016. https://ec.europa.eu/clima/policies/strategies/2030_en (accessed January 29, 2019).

[7] European Commission. EU - 2050 long-term strategy. Climate Action - European Commission 2016. https://ec.europa.eu/clima/policies/strategies/2050_en (accessed January 29, 2019).

[8] Pietrapertosa F, Khokhlov V, Salvia M, Cosmi C. Climate change adaptation policies and plans: A survey in 11 South East European countries. Renewable and Sustainable Energy Reviews 2018;81:3041–50. doi:10.1016/j.rser.2017.06.116.

[9] Energy Community Homepage n.d. https://www.energy-community.org/ (accessed January 20, 2019).

[10] Poncelet K, Delarue E, Six D, Duerinck J, D'haeseleer W. Impact of the level of temporal and operational detail in energy-system planning models. Applied Energy 2016;162:631–43. doi:10.1016/j.apenergy.2015.10.100.

[11] Koltsaklis NE, Dagoumas AS. State-of-the-art generation expansion planning: A review. Applied Energy 2018;230:563–89. doi:10.1016/j.apenergy.2018.08.087.

[12] Han X, Chen X, McElroy MB, Liao S, Nielsen CP, Wen J. Modeling formulation and validation for accelerated simulation and flexibility assessment on large scale power systems under higher renewable penetrations. Applied Energy 2019;237:145–54. doi:10.1016/j.apenergy.2018.12.047.

[13] Hua B, Baldick R, Wang J. Representing Operational Flexibility in Generation Expansion Planning Through Convex Relaxation of Unit Commitment. IEEE Transactions on Power Systems 2018;33:2272–81. doi:10.1109/TPWRS.2017.2735026.

[14] Pfenninger S, Hawkes A, Keirstead J. Energy systems modeling for twenty-first century energy challenges. Renewable and Sustainable Energy Reviews 2014;33:74–86. doi:10.1016/j.rser.2014.02.003.

[15] Heuberger CF, Rubin ES, Staffell I, Shah N, Mac Dowell N. Power capacity expansion planning considering endogenous technology cost learning. Applied Energy 2017;204:831–45. doi:10.1016/j.apenergy.2017.07.075.

[16] Min D, Ryu J, Choi DG. A long-term capacity expansion planning model for an electric power system integrating large-size renewable energy technologies. Computers & Operations Research 2018;96:244–55. doi:10.1016/j.cor.2017.10.006.

[17] Teichgraeber H, Brandt AR. Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. Applied Energy 2019;239:1283–93. doi:10.1016/j.apenergy.2019.02.012.

[18] Koltsaklis NE, Georgiadis MC. A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints. Applied Energy 2015;158:310–31. doi:10.1016/j.apenergy.2015.08.054.

[19] Pineda S, Morales JM. Chronological Time-Period Clustering for Optimal Capacity Expansion Planning With Storage. IEEE Transactions on Power Systems 2018;33:7162–70. doi:10.1109/TPWRS.2018.2842093.

[20] Blanford GJ, Merrick JH, Bistline JET, Young DT. Simulating Annual Variation in Load, Wind, and Solar by

Representative Hour Selection. The Energy Journal 2018;39. doi:10.5547/01956574.39.3.gbla.
[21] Slednev V, Bertsch V, Ruppert M, Fichtner W. Highly resolved optimal renewable allocation planning in power systems under consideration of dynamic grid topology. Computers & Operations Research 2018;96:281–93. doi:10.1016/j.cor.2017.12.008.
[22] González IH, Ruiz P, Sgobbi A, Nijs W, Quoilin S, Zucker A, et al. Addressing flexibility in energy system models. Publications Office of the European Union 2015:80. doi:10.2790/925.
[23] Quoilin S, Nijs W, Zucker A. Evaluating flexibility and adequacy in future EU power systems: Model coupling and long-term forecasting. Proceedings of the 30th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems, San Diego: 2017.
[24] Pietzcker RC, Ueckerdt F, Carrara S, de Boer HS, Després J, Fujimori S, et al. System integration of wind and solar power in integrated assessment models: A cross-model evaluation of new approaches. Energy Economics 2017;64:583–99. doi:10.1016/j.eneco.2016.11.018.
[25] Pavičević M, Nijs W, Kavvadias KC, Quoilin S. Modelling flexible power demand and supply in the EU power system: soft-linking between JRC-EU-TIMES and the open-source Dispa-SET model. 32nd International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems, Wrocław, Poland: 2019.
[26] Bistline JE. Turn Down for What? The Economic Value of Operational Flexibility in Electricity Markets. IEEE Transactions on Power Systems 2019;34:527–34. doi:10.1109/TPWRS.2018.2856887.
[27] Palmintier BS, Webster MD. Impact of Operational Flexibility on Electricity Generation Planning with Renewable and Carbon Targets. IEEE Transactions on Sustainable Energy 2016;7:672–84. doi:10.1109/TSTE.2015.2498640.
[28] Collins S, Deane JP, Poncelet K, Panos E, Pietzcker RC, Delarue E, et al. Integrating short term variations of the power system into integrated energy system models: A methodological review. Renewable and Sustainable Energy Reviews 2017;76:839–56. doi:10.1016/j.rser.2017.03.090.
[29] Palmintier BS (Bryan S. Incorporating operational flexibility into electric generation planning : impacts and methods for system design and policy analysis. Thesis. Massachusetts Institute of Technology, 2013.
[30] El Kafazi I, Bannari R, Hernánde ACL. Optimization strategy considering Energy Storage Systems to minimize energy production cost of power systems. International Journal of Renewable Energy Research 2018;8:2199–209.
[31] Mehdinejad M, Mohammadi-Ivatloo B, Dadashzadeh-Bonab R. Energy production cost minimization in a combined heat and power generation systems using cuckoo optimization algorithm. Energy Efficiency 2017;10:81–96. doi:10.1007/s12053-016-9439-6.
[32] Jiménez Navarro JP, Kavvadias KC, Quoilin S, Zucker A. The joint effect of centralised cogeneration plants and thermal storage on the efficiency and cost of the power system. Energy 2018;149:535–49. doi:10.1016/j.energy.2018.02.025.
[33] Silvente J, Papageorgiou LG. An MILP formulation for the optimal management of microgrids with task interruptions. Applied Energy 2017;206:1131–46. doi:10.1016/j.apenergy.2017.08.147.
[34] Farhat IA, El-Hawary ME. Optimization methods applied for solving the short-term hydrothermal coordination problem. Electric Power Systems Research 2009;79:1308–20. doi:10.1016/j.epsr.2009.04.001.
[35] Quoilin S, Nijs W, Gonzalez IH, Zucker A, Thiel C. Evaluation of simplified flexibility evaluation tools using a unit commitment model. Proceedings of the 12th International Conference on the European Energy Market (EEM), vol. 2015- August, Lisbon, Portugal: Institute of Electrical and Electronics Engineers; 2015. doi:10.1109/EEM.2015.7216757.
[36] Zhou B, Geng G, Jiang Q. Hydro-Thermal-Wind Coordination in Day-Ahead Unit Commitment. IEEE Transactions on Power Systems 2016;31:4626–37. doi:10.1109/TPWRS.2016.2530689.
[37] He J, Hu Z, Liu Y. Establishment and solution of the large-scale multi-objective hydro-thermal-wind power coordination optimization dispatching model. Dianli Xitong Baohu Yu Kongzhi/Power System Protection and Control 2015;43:1–7.
[38] Nakamura M, Kumarawadu P, Yoshida A, Hatazaki H. Reliable Maintenance Scheduling of Pumps in Existing Thermal Power Stations. IFAC Proceedings Volumes 1997;30:165–8. doi:10.1016/S1474-6670(17)42249-X.
[39] Fernández-Blanco R, Kavvadias K, Hidalgo González I. Quantifying the water-power linkage on hydrothermal power systems: A Greek case study. Applied Energy 2017;203:240–53. doi:10.1016/j.apenergy.2017.06.013.
[40] Pandzic H, Qiu T, Kirschen DS. Comparison of state-of-the-art transmission constrained unit commitment formulations, 2013. doi:10.1109/PESMG.2013.6672719.
[41] Gentile C, Morales-España G, Ramos A. A tight MIP formulation of the unit commitment problem with start-up and shut-down constraints. EURO Journal on Computational Optimization 2017;5:177–201. doi:10.1007/s13675-016-0066-y.
[42] Morales-España G, Ramírez-Elizondo L, Hobbs BF. Hidden power system inflexibilities imposed by traditional unit

commitment formulations. Applied Energy 2017;191:223–38. doi:10.1016/j.apenergy.2017.01.089.
[43] Yang L, Zhang C, Jian J, Meng K, Xu Y, Dong Z. A novel projected two-binary-variable formulation for unit commitment in power systems. Applied Energy 2017;187:732–45. doi:10.1016/j.apenergy.2016.11.096.
[44] Wang W, Li C, Liao X, Qin H. Study on unit commitment problem considering pumped storage and renewable energy via a novel binary artificial sheep algorithm. Applied Energy 2017;187:612–26. doi:10.1016/j.apenergy.2016.11.085.
[45] Yang Z, Li K, Niu Q, Xue Y. A novel parallel-series hybrid meta-heuristic method for solving a hybrid unit commitment problem. Knowledge-Based Systems 2017;134:13–30. doi:10.1016/j.knosys.2017.07.013.
[46] Alvarez GE, Marcovecchio MG, Aguirre PA. Security-constrained unit commitment problem including thermal and pumped storage units: An MILP formulation by the application of linear approximations techniques. Electric Power Systems Research 2018;154:67–74. doi:10.1016/j.epsr.2017.07.027.
[47] Alvarez GE, Marcovecchio MG, Aguirre PA. Unit commitment scheduling including transmission constraints: a MILP formulation. Computer Aided Chemical Engineering 2016;38:2157–62. doi:10.1016/B978-0-444-63428-3.50364-7.
[48] Hinojosa VH, Gutiérrez-Alcaraz G. A computational comparison of 2 mathematical formulations to handle transmission network constraints in the unit commitment problem. International Transactions on Electrical Energy Systems 2017;27. doi:10.1002/etep.2332.
[49] Meus J, Poncelet K, Delarue E. Applicability of a Clustered Unit Commitment Model in Power System Modeling. IEEE Transactions on Power Systems 2018;33:2195–204. doi:10.1109/TPWRS.2017.2736441.
[50] Palmintier BS, Webster MD. Heterogeneous unit clustering for efficient operational flexibility modeling. IEEE Transactions on Power Systems 2014;29:1089–98. doi:10.1109/TPWRS.2013.2293127.
[51] Palmintier B, Webster M. Impact of unit commitment constraints on generation expansion planning with renewables, 2011. doi:10.1109/PES.2011.6038963.
[52] Pavičević M, Quoilin S. Dispa-SET Balkans - Dataset 2019. doi:10.5281/zenodo.2551747.
[53] Open Energy Modelling Initiative (OPENMOD) - forum n.d. https://forum.openmod-initiative.org/.
[54] Alemany J, Kasprzyk L, Magnago F. Effects of binary variables in mixed integer linear programming based unit commitment in large-scale electricity markets. Electric Power Systems Research 2018;160:429–38. doi:10.1016/j.epsr.2018.03.019.
[55] Quoilin S. DispaSET Documentation n.d.:84.
[56] Quoilin S, Kavvadias KC, Pavičević M. energy-modelling-toolkit: Dispa-SET github repository n.d. https://github.com/energy-modelling-toolkit/Dispa-SET (accessed April 24, 2019).
[57] Welsch M. European energy markets and society: findings informing the european commission. 1st edition. Cambridge, MA: Elsevier; 2017.
[58] Myers L, Sirois MJ. Spearman Correlation Coefficients, Differences between. Encyclopedia of Statistical Sciences, American Cancer Society; 2006. doi:10.1002/0471667196.ess5050.pub2.
[59] Pavičević M, Quoilin S, Zucker A, Krajačić G, Pukšec T, Duić N. Applying the Dispa-SET Model on the Western Balkans Power System. Journal of Sustainable Development of Energy, Water and Environment Systems 2019;(In Press). doi:10.13044/j.sdewes.d7.0273.
[60] Szabó L, Kelemen Á, Mezősi A, Pató Z, Kácsor E, Resch G, et al. South East Europe electricity roadmap – modelling energy transition in the electricity sectors. Climate Policy 2018:1–16. doi:10.1080/14693062.2018.1532390.
[61] Beltramo A, Julea A, Refa N, Drossinos Y, Thiel C, Quoilin S. Using electric vehicles as flexible resource in power systems: A case study in the Netherlands. Proceedings of the 14th International Conference on the European Energy Market, Dresden: Technische Universität Dresden; 2017.
[62] Villavicencio M. A capacity expansion model dealing with balancing requirements, short.term operations and long-run dynamics. France: Chaire European Electricity Markets - CEEM; 2017.
[63] Med-TSO members. Mediterranean Project 1 (2015-2018) - Deliverable 2.1.3: Relevant Grid Map n.d.
[64] Yasuda Y, Årdal AR, Hernando DH, Carlini EM, Estanqueiro A, Flynn D, et al. Evaluation on diversity of flexibility in various areas. Proceedings of the 12th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants, London: 2013, p. 6.
[65] ENTSOE. TYNDP - Europe's Network Development Plan to 2025, 2030 and 2040 n.d. https://tyndp.entsoe.eu/ (accessed January 28, 2019).
[66] Staffell I, Pfenninger S. Using bias-corrected reanalysis to simulate current and future wind power output. Energy 2016;114:1224–39. doi:10.1016/j.energy.2016.08.068.
[67] Pfenninger S, Staffell I. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. Energy 2016;114:1251–65. doi:10.1016/j.energy.2016.08.060.
[68] Pfenninger S, DeCarolis J, Hirth L, Quoilin S, Staffell I. The importance of open data and software: Is energy research lagging behind? Energy Policy 2017;101:211–5. doi:10.1016/j.enpol.2016.11.046.

[69] Dispa-SET for the Balkans region — Documentation n.d. http://www.dispaset.eu/en/latest/casebalkans.html (accessed January 30, 2019).

[70] Keatley P, Shibli A, Hewitt NJ. Estimating power plant start costs in cyclic operation. Applied Energy 2013;111:550–7. doi:10.1016/j.apenergy.2013.05.033.