Pitfalls of Power Systems Modelling Metrics

Rogier Hans Wuijts
Department of Information and Computing Sciences & Copernicus Institute for Sustainable Development
Utrecht University
Utrecht, The Netherlands
R.H.Wuijts@uu.nl

William Zappa
TenneT TSO BV
Arnhem, The Netherlands
William.Zappa@tennet.eu

Marjan van den Akker
Department of Information and Computing Sciences
Utrecht University
Utrecht, The Netherlands
J.M.vandenAkker@uu.nl

Machteld van den Broek
Energy Research Institute Groningen (ESRIG) & Integrated Research on Energy, Environment and Society (IREES)
University of Groningen
Groningen, The Netherlands
M.A.van.den.Broek@rug.nl

2022 18th International Conference on the European Energy Market (EEM) | 978-1-6654-0896-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/EEM54602.2022.9920976

***Abstract*— In power system modelling the unit commitment problem is used to simulate the wholesale electricity market. A solution to the unit commitment problem is a least-cost schedule that contains information regarding the capacity factors of each generator, the total CO_2 emissions, and unserved energy per hour. However, since there might be a large variety of (sub)-optimal solutions, these characteristics might be arbitrary and conclusions about them may be presumptuous.**

In this article, we illustrate this by running multiple experiments on a future European power system. Each scenario was run multiple times by adding additional terms to the objective function such as the minimization and maximization of generator capacity factors, carbon emissions, and loss of load hours. The results showed that schedules can be equivalent in terms of cost, but that relative capacity factors, emissions, and loss of load hours could differ by large factors.

***Index Terms*--Power System Modelling, Unit Commitment Problem, Modelling Metrics**

I. Introduction

The unit commitment and economic dispatch (UC) problem is commonly used by academic researchers, Transmission and Distribution System Operators, energy companies, and policy makers to evaluate the resource adequacy of current and future electricity generation portfolios, and to support investment and operational decision-making ([6]-[11]). As a result of the energy transition and increasing reliance on variable renewable energy sources, the need for robust power system modelling is becoming increasingly important.

A UC model finds the least-cost schedule for a generation portfolio under a set of hard and soft constraints. Outputs provide specific information regarding the schedule such as capacity factors of each generator, CO_2 emissions, and unserved energy per hour. However, conclusions about these characteristics are often presumptuous. This is caused by the fact that they are not explicitly included in the objective function, which implies that they can take an arbitrary value so long as they do not violate any hard constraints, or negatively influence the objective function.

For example, many countries use the Loss of Load Expectation (LOLE) metric, i.e., the average number of hours unserved energy occurred in a year based on a Monte Carlo simulation over many years to assess the system adequacy. The LOLE is one of the two key reliability indicators which must be calculated in the European Union (EU) as part of the European Resource Adequacy Assessment (ERAA) [12]. A second indicator is expected Energy Not Served (ENS), which represents the energy which is not supplied over a given period due to insufficient capacity resources to meet the demand. LOLE is an important adequacy indicator as any EU member state wishing to implement a capacity mechanism to safeguard security of supply can only do so if the ERAA shows that the national LOLE is above a specified reliability standard [14,15], calculated according to a strict methodology and typically in the range of 3–8 h/year [12].

When a UC problem is solved to simulate the operation of the current European power system, the resulting power generation schedule has minimal generation cost and minimal amount of unserved energy. If a schedule has unserved energy, then there are specific hours in which loss of load occurs. These specific hours, however, might be arbitrary. For example, one schedule with all unserved energy in a single time step may cost as much as another schedule with the unserved energy distributed over several time steps [13]. However, these two schedules will have a different impact on the LOLE metric which in turn could imply divergent policy decisions.

Which (sub)-optimal solution you get out of an optimization can depend on random choices a solver or algorithm makes, the fact that an algorithm produces a heuristic solution, the stopping criteria of sub optimality, how solvers are configured, how the input data is structured, and much more. Two real-world examples we experienced of this are as follows:

- When adding valid inequalities to our Mixed Integer Linear Programming UC model to guide the solver by improving the relaxation bound, the optimal solution produced by the solver gave back very different capacity factors [4]. However, these valid inequalities were not expected to change the solution as they do not reduce the search space of the solution.
- The standard method of solving a LP in Gurobi is by concurrently running multiple algorithms such as Simplex, Dual Simplex, and the Barrier Method. When one of these methods finds a solution, the process is stopped, and the solution found by that method is returned. However, the other algorithms would have also resulted in an optimal solution. Thus, multiple runs of the same scenario on the same computer with the same program may randomly result in different solutions ([5] p. 748) with, for example, different LOLEs depending on the fastest algorithm in each run.

These are just a few examples of when the modelling metrics of a solution can be arbitrarily changed, and potentially lead analysts to draw non-robust conclusions from the results.

In this article, we illustrate this arbitrary effect by running multiple experiments on three scenarios of the European power system in the years 2030 and 2040 [3]. For these scenarios we ran the UC problem to find least-cost schedules that meet the electricity demand with unserved energy rated at a value of loss of load of 10,000 Euro/MWh. Each scenario was run multiple times by adding additional objectives such as the minimization and maximization of capacity factors of generators, emissions and loss of load hours, while keeping the total dispatch costs equivalent. In this way, we demonstrate the maximum and minimum values of these indicators which could technically arise from the UC model, and hence the possible range of arbitrariness in the indicator results.

II. METHODS

We want to test how arbitrary important metrics of power system modelling are. The metrics we investigate are Loss of Load Hours (LOLH), capacity factors (CF) of generators and total CO_2 emissions. First, we will give a definition for every metric. Then, we will describe the power system instances on which we run the UC. Finally, we will describe how we maximize the model metric differences between solutions of equivalent cost to analyse the different possible solutions.

A. Metrics

Loss of load occurs when generation capacity is insufficient in a given hour to meet the demand. This can be a result of insufficient generation on the supply side (e.g. due to power plant outage, or a shortage in renewable generation), or insufficient flexibility on the demand side. The metric of Loss of Load Hours (LOLH) is relevant to determine the system adequacy of a certain power system. The relationship with LOLE is that LOLE is the expected LOLH in a year which is the result of a Monte Carlo simulation over many years. Let's define a loss of load hours, $LOLH_{nt}$, as a binary variable that is 1 when the Energy Not Served, ENS_{nt}, is above zero:

$$LOLH_{nt} = \begin{cases} 1 & ENS_{nt} > 0 \\ 0 & otherwise \end{cases}$$

where ENS_{nt} is the demand of node n and time t that cannot be provided.

Every generator has a capacity factor that indicates how much a generator is used. The capacity factor is 100%, if the generator is always producing at maximum capacity and it is lower otherwise. This metric is relevant for studies looking at the economic viability of power plants, which depends on the number of hours they run. The capacity factor of generator g is defined as:

$$CF_g = \frac{\sum_{t \in T} p_{gt}}{|T|\overline{P}_g}$$

where p_{gt} is the power production of generator g at time t. |T| is the number of timesteps and $\overline{P}_g$ is the maximum generation of generator g.

Many conventional thermal generators produce CO_2 when burning fossil fuels to generate power. Due to the negative greenhouse effect of CO_2, it is often of interest to investigate how much CO_2 is likely to result from electricity generation. The total CO_2 produced by a UC schedule is defined as:

$$Total\ CO_2 = \sum_{g \in G, t \in T} FCP_g u_{gt} + VCP_g p_{gt}$$

where p_{gt} is the power production of generator g at time t and FCP_g is the fixed and VCP_g is the variable CO_2 production of generator g.

B. Unit Commitment Instances

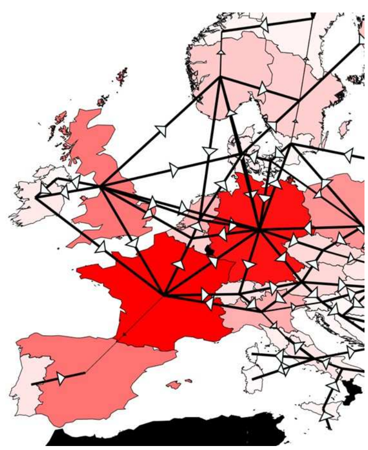

Figure 1, The bidding zones regions modelled in the TYNDP2020 scenarios

	Capacity (GW)				
Scenario	Thermal	Storage	RES	DSR (GW)	Demand (TWh)
GA2030	576	125	794	38	4038
GA2040	544	148	1089	44	4296
DE2030	567	130	944	38	4214
DE2040	535	221	1480	44	5075
NT2030	587	121	818	26	3968
NT2040	554	170	1093	31	4402

Table 1, Summary of the capacity resources scenarios from TYNDP [3]

The UC instances we performed the experiments on are based on future European power system scenarios of the "TYNDP2020" created by the European Network of Transmission System Operators for Electricity (ENTSO-E) [3]. These scenarios consist of 55 'nodes' corresponding roughly to the current bidding zone configuration of Europe. In practice, most bidding zones correspond to a single country (Figure 1) but some countries such as Sweden and Italy are divided into multiple bidding zones. Table 1 gives an overview of the total generation capacity for each scenario. For our study we used six scenarios of the TYNDP2020 study, namely the three different pathways of National Trends (NT), Global Ambition (GA) and Distributed Energy (DE) for target years 2030 and 2040.

These scenarios created by ENTSO provide insights into the possible energy system of the future and the effects of changes in supply and demand on the energy system [3].

In our experiments we used the climate years from 1950 till 2019 to simulate the variation in renewable energy generation. We used historic renewable energy availability factors for PV, onshore and offshore wind based on the ERA5 data set [16].

For the hydro inflow data, we used the scaled inflow data from another European power system study [1] which were obtained from the RESTORE 2050 project [2].

For our study we used a time horizon of 72 hours sliced from the yearly demand and renewable patterns of the scenarios. The start of these 72 hours periods where at the start of hour 0, 2400, 4800 and 7200 of the climate years 1950 to 2019 and from one of the 6 TYNDP scenarios. In total we used $4 * 60 * 6 = 1440$ UC instances of 72 timesteps.

For the other experiments with CF and CO_2 we used a smaller subset. Here we only looked at the years 1950, 1960, …, 2010 and in total we used $4 * 7 * 6 = 168$ UC instances of 72 timesteps.

The difference in the number of instances is because we expect not every UC instance to have ENS, leading to trivial results, but we do except that every instance has some CFs and CO_2 emissions.

C. Procedure

We want to maximise the difference of a specific model metric between two solutions with equivalent costs. In order to calculate the maximal difference, we first find the optimal solution of the UC problem. This solution has a certain objective value (e.g., a total cost of $\$1 \cdot 10^8$) and a metric value (e.g., 2 hours LOLH). From this original solution, we want to find other solutions with the same objective function value, but with different values for a particular metric, e.g., a solution with a total cost of $\$1 \cdot 10^8$ but with 10 LOLH. In order to see the range of potential solutions which give rise to an equivalent objective function value but different metric values, we find a minimum and maximum metric using optimization.

UC can be formulated compactly as follows (The full and precise formulation of the UC is in the appendix (9) - (28)):

$$min\, f(x)$$

$$Ax = b$$

where x is the UC solution that contain all the relevant information that satisfies all the constraints, $Ax = b$. The objective $f(x)$ includes the generation cost, CO_2 emission certificate cost, start-up cost and system wide costs such as value of lost load multiplied by the ENS.

Suppose that the optimal solution of the optimization problem is x^*. Let g be our metric function that we want to minimize and maximize in order to find the largest difference. The following program gives an equivalent solution in terms of cost but minimizes the modeling metric g:

$$min\, g(x)$$

$$Ax = b$$

$$f(x) \leq \alpha f(x^*)$$

if α is set to 1 then the solution to this problem has optimal cost. If $\alpha = 1.001$ then the solution is within 0.1% of optimality.

In this article we want to focus on the three metrics LOLH, CF and CO_2. For each of the three metrics we performed various experiments. For the LOLH experiments we kept α at 1, i.e. we only looked at optimal solutions, but at three levels of detail of the UC problem. The first level is a simplified *ENS-Model*, where the original objective is minimizing the total ENS and not the production cost of generators. Here we do not include flexibility constraints such as minimum up and down time and ramping limits. The second level is the *Cost-Model* in which production cost are minimized, but still without flexibility constraints. In the third level, the *Full-Model,* we take all the cost and flexibility constraints into account. Here, we also included a wheeling charge on the transmission lines, which implies that exporting and importing power between bidding zones has a small cost.

For the CF experiments we set $\alpha \in \{1.001, 1.005, 1.01\}$ and minimized and maximized the capacity factor of Gas, Coal and Nuclear-powered electrical generators.

For the CO_2 experiments we set $\alpha \in \{1.001, 1.005, 1.01\}$ and minimized and maximized the total CO_2 produced by the power system. Table 5 in the appendix presents the precise setup of the experiments, the definition of the original objective function *f* and the model metric function *g* we want to minimize or maximize.

III. RESULTS

	Minimizing		Original		Maximizing		Factor difference		
	Avg	σ	Avg	σ	Avg	σ	Avg	σ	max
ENS-Model	16.0	11.4	27.1	14.8	45.3	13.4	5.1	5.4	38.0
Cost-Model	17.7	13.5	27.1	14.4	40.2	16.4	3.4	2.6	19.5
Full-Model	28.7	20.2	34.8	17.0	43.3	17.2	2.3	1.7	11.1

Table 2, Summary of the LOLH metric for the ENS-, Cost- and Full-Model

A. Loss of Load Hours

The results of our experiments with minimising and maximising the LOLH are presented in Table 2, Figure 2 and Figure 4, and additional figures in the appendix.

Table 2 presents the average LOLH of the three different models. The second column is the LOLH from the initial solution when we minimize the objective *f*. The first and third column are the minimized and maximised LOLH. Note that the original LOLH falls somewhere between the minimized and maximized variant. For some UC instances there was no ENS in any region or at any timestep. We removed those instances from our figures and tables.

For the *ENS-Model* the total difference of the average LOLH is a factor of $45.3/16 \approx 2.8$, for the *Cost-Model* it is ≈2.3 and for the Full-Model this is a factor of ≈1.5. The difference with single UC instances is much higher. If we look at individual instances, we can see that the average factor between the minimized and maximized LOLH is 5.1 for the *ENS-Model* and 2.3 for the *Full-Model*. Furthermore, some specific runs have even more extreme differences such as one case with the *ENS-Model* where the maximum LOLH was 38 times higher than the minimum one.

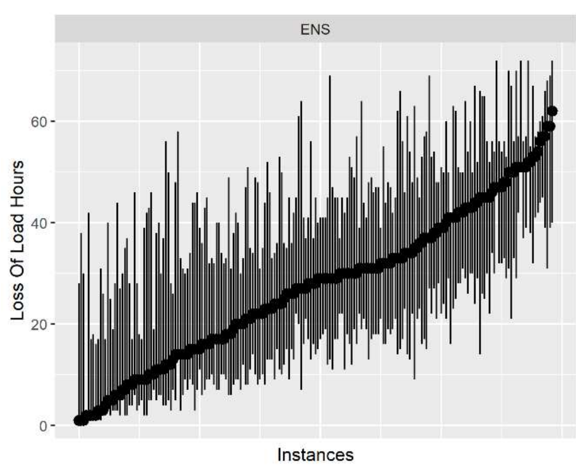

Figure 2, The LOLH for all instances, ordered by the LOLH, for the ENS-Model. (See Figure 3 for clarification)

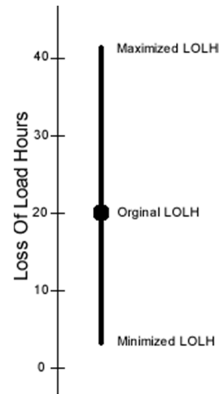

Figure 3, Explanation of the range graphs, the middle circle is the LOLH of the original optimization. The upper and lower value are the minimized and maximized LOLH. The whole range are the possible values of LOLH an equivalent solution can take on.

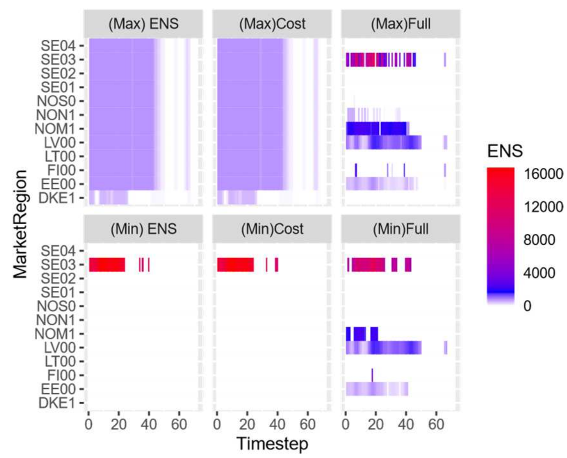

Figure 4, The difference in solutions where we maximize (upper figures) and minimize (lower figures) LOLH with the ENS-, Cost- and Full-Model

The difference of the LOLH in the different models of individual runs becomes more apparent in Figure 2. It shows the individual model runs (black dots) that had ENS sorted on the original LOLH. The black bar around the dot shows the minimized and maximised LOLH. This black bar can be interpreted as the possible LOLH that could come out of the model given the same ENS (Figure 3).

The experiments with the *Cost-Model* and the *Full-Model* show a similar pattern as Figure 2 (Figure 7 in the appendix). However, the black bar representing the possible range of LOLH given the same ENS and the same cost is a little smaller for the *Cost-Model* and even smaller for the *Full-Model*. This makes sense, as with more detail the model has less leeway for a solution to change while the objective function stays the same. It is to be expected that if we added even more detail to the *Full-Model* that this range would become even smaller.

Figure 4 shows the cause of such a discrepancy of LOLH between UC solutions. All the solutions have exactly the same ENS between the minimization and maximization run and the same primary objective function *f*. However, the total LOLH is different. When the LOLH is minimized the ENS is concentrated in a single node but when the LOLH is maximized it is spread out as evenly as possible between nodes and timesteps.

For the UC instance used in Figure 4, we can see that the amount of leeway in the solution space is almost the same for the *ENS-Model* and the *Cost-Model* but that the *Full-Model* has significantly less ability to spread out the ENS.

		Minimizing		Original		Maximizing		Factor difference		
	α	Avg	σ	Avg	σ	Avg	σ	Avg	σ	max
Gas	0.001	8.3	6.0	9.6	6.8	10.8	7.5	1.7	1.8	21.0
	0.005	6.2	4.8	9.6	6.8	12.3	8.1	4.6	9.4	83.5
	0.01	4.9	4.0	9.6	6.8	13.6	8.7	-	-	-
Coal	0.001	8.5	5.6	11.1	6.9	14.2	8.7	-	-	-
	0.005	5.9	4.9	11.1	6.9	19.5	11.6	-	-	-
	0.01	4.3	4.1	11.1	6.9	23.0	13.7	-	-	-
Nuc	0.001	64.1	19.2	64.9	19.3	65.8	19.3	1	0	1.2
	0.005	62.6	19.1	64.9	19.3	68.0	19.4	1.1	0.1	1.9
	0.01	61.2	19.1	64.9	19.3	70.5	19.7	1.2	0.2	2.7

Table 3, Summary of the experiments with capacity factor

B. Capacity Factor

The results of our experiments with maximizing the difference in capacity factors are summarized in Table 3, Figure 5, and figures in the appendix.

Table 3 shows that equivalent solutions for within a margin of 1% of the optimum on average can have a Coal CF that is either 4.3% or 23.0%, a factor 5.3 different. In some cases, the CF of coal became 0% when minimizing this metric (Figure 5).

For Gas the difference is still high but slightly lower than the coal generators. It seems that nuclear in these instances is the least flexible and on average could be a factor of 1.2 higher. If we reduce α to be closer to optimality, we can see that the difference reduces for all generator types. However, for gas and coal this difference at 0.1% removed from optimum is still relatively high. Figure 8 in the appendix shows the differences for all instances and all generator types. The figure shows that on average the difference is high but that also a lot of outliers exist where the difference is even higher. Even at 0.1% removed from optimum this difference sometimes is between zero production and some production for the coal-fired generators.

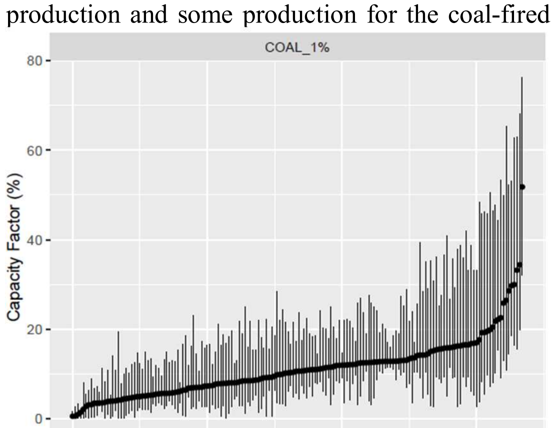

Figure 5, The CFs for all instances, ordered by the original total CF, for the solutions within 1% of the optimum.

C. CO_2 Emissions

		Minimizing		Original		Maximizing		Factor difference		
	α	Avg	σ	Avg	σ	Avg	σ	Avg	σ	max
CO_2	0.001	1.99	1.06	2.08	1.11	2.18	1.16	1.1	0.1	1.5
	0.005	1.88	1.02	2.08	1.11	2.35	1.26	1.3	0.2	2.1
	0.01	1.79	0.98	2.08	1.11	2.47	1.34	1.4	0.2	2.5

Table 4 Summary of the experiments with CO_2

The results of our experiments with maximizing the difference in CO_2 emissions are summarized in Table 4 and Figure 6.

Table 4 shows that equivalent solutions within a margin of 1% of the optimum on average are a factor of 1.4 different and this factor can be as high as 2.5.

Figure 9 in the appendix shows the different percentages removed from optimum. The difference of total CO_2 is almost negligible when the solution is 0.1% removed from optimum. When LOLH, CF and CO_2 are compared, we can see that the difference in the CO_2 metric are much lower than those of the LOLH and CF metric. This difference might be attributed to the fact that the CO_2 is a part of the objective function as there is a tax on CO_2 emissions.

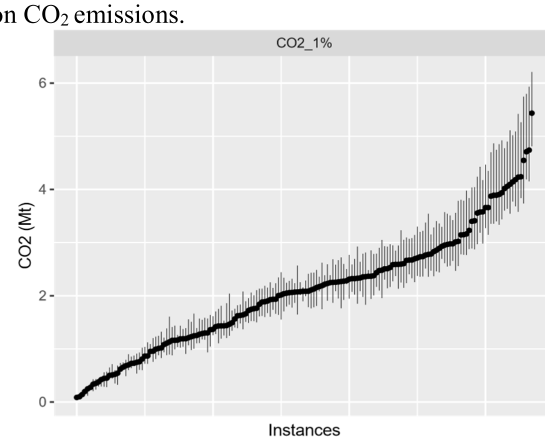

Figure 6, The total CO_2 emissions for all instances, ordered by the original total CO_2 emissions, for the solutions within 1% of the optimum.

IV. DICUSSSION

In the previous section we showed the difference in the modeling metrics, LOLH, CF and CO_2 for equivalent solutions that are close or equal to the optimal solution. The way we maximised those differences with optimization is not something you would normally do in power system modelling analysis, but it does show the maximal error these solutions might have. The next question that might come up is what the chance is, when running the model, one might get one solution or the other. The answer is that this is arbitrary. Or to be more specific: since the solutions have equivalent objective function values, it is equally valid for any solver or algorithm to produce either solution. Moreover, it would be an unintended consequence if they would prefer one solution over the other. If it was intended, then it should be explicitly modeled.

In this paper we showed that we could minimize the spread of LOLH in a more detailed model. The takeaway, however, from this article is that more detail should not be added to avoid arbitrariness in these metrics but to be conscious and careful when using any metric that is not explicitly optimized. These metrics are secondary characteristics of the solution, and they are not the goal and therefore should not be used as the main subject of analysis of these solutions. The analysis might be better on the objective itself. For example, the total ENS as a metric for adequacy is more robust than LOLE. ENS is explicitly modeled in the objective with a high coefficient, the value of lost load, and therefore (sub)optimal solutions have the same or similar ENS.

V. Conclusion

In this article we investigated the arbitrariness of important modelling metrics such as LOLH, CF and CO_2 in power system modelling. These metrics are used to analyse current and future power systems. However, often when simulating these power systems with optimization, these metrics are not explicitly optimized for. Therefore, conclusions from these modeling metrics should not be drawn and if they are drawn people should at least be conscious of the arbitrariness of these metrics.

Experimentally we showed the arbitrariness of the previously mentioned metrics. We used the 6 future European power systems with multiple climate years and simulated short term electricity market simulations by solving multiple 72-hour long UC instances. With these instances we re-optimized these different modeling metrics given equivalent objective functions.

We found that the difference in the LOLH metric could be as high as a factor 38 for individual instances and the total difference, depending on the model, varied from a factor of 1.5 to 2.8. For the capacity factor of coal, gas- and nuclear-powered generators these differences depend on how close they are to the optimal solution. For individual instances the CF could be 0% when minimizing but more than 20% when maximizing, while both are within 1% of optimality. And in total the CF over all experiments for coal was a factor of 5.3 different. For gas and nuclear this was 2.8 and 1.2 respectively. For the CO_2 the difference was less pronounced, on average single instances had a factor difference of 1.1 to 1.4.

Acknowledgment

This work is part of the research programme “Energie: Systeem Integratie en Big Data” with project number 647.003.005, which is financed by the Dutch Research Council (NWO).

References

[1] Pavicevi ̌ c, Matija, et al. “The potential of sector coupling in future ́ European energy systems: Soft linking between the Dispa-SET and JRCEU-TIMES models.” Applied Energy 267 (2020).

[2] Kies, A., L. von Bremen, and D. Heinemann. “Hydro energy inflow for power system studies.” (2017).

[3] ENTSOG, ENTSO-E. “TYNDP 2020 Scenario Report.” (2020).

[4] Rogier Hans Wuijts, Marjan van den Akker, and Machteld van den Broek. “Effect of Modelling Choices in the Unit Commitment Problem”. [Manuscript submitted for publication] (2022).

[5] Gurobi Optimization, LLC. “Gurobi Optimizer Reference Manual.” (2022).

[6] William Zappa, Martin Junginger, and Machteld van den Broek. “Is a 100% renewable European power system feasible by 2050?” Applied Energy 233 (2019), pp. 1027–1050

[7] Bas van Zuijlen et al. “Cost-optimal reliable power generation in a deep decarbonisation future”. In: Applied Energy 253 (2019).

[8] Anne Sjoerd Brouwer et al. “Least-cost options for integrating intermittent renewables in low-carbon power systems”. Applied Energy 161 (2016).

[9] Matija Paviˇcevi ́c et al. “The potential of sector coupling in future European energy systems: Soft linking between the Dispa-SET and JRC-EU-TIMES models”. Applied Energy 267 (2020).

[10] ENTSO-E. “Mid Term Adequacy Forecast (MAF)”. (2018).

[11] ENTSO-E. “Pentalateral Energy Forum Support Group 2 Generation Adequacy Assessment”. (2018).

[12] ENTSO-E, “European Resource Adequacy Assessment - 2021 Edition,” Brussels, (2021).

[13] Energy Systems Integration Group, “Redefining Resource Adequacy for Modern Power Systems”. (2021).

[14] ACER. “Methodology for Calculating the Value of Lost Load, the Cost of New Entry and the Reliability Standard.” (2020).

[15] The European Parliament and the Council of the European Union, “Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity,” Official Journal of the European Union, (2019).

[16] Hersbach H, Bell B, Berrisford P, et al. “The ERA5 global reanalysis.” QJRMeteorol (2020).

VI. Appendix

Our UC description contains generators, RES, storage and transmission lines. The entire UC formulation is presented in (9) - (28).

$$\min \sum_{t\in T}\sum_{g\in G} a_g u_{gt} + b_g p_{gt} + c_g p_{gt}^2 + v_{gt} cost_{start} \quad (9)$$

$$+ \sum_{t\in T}\sum_{n\in N} ENS_{nt} VOLL \quad (10)$$

$$+ \sum_{t\in T}\sum_{l\in L} f_{lt} WC \quad (11)$$

$s.t.$

$$p_{gt} \geq u_{gt}\underline{P}_g, g \in G, t \in T \quad (12)$$

$$p_{gt} \leq \overline{P}_g u_{gt}, g \in G, t \in T \quad (13)$$

$$\sum_{i=t-UT_g+1}^{t} v_{gi} \leq u_{gt}, t \in T, g \in G \quad (14)$$

$$\sum_{i=t-DT_g+1}^{t} w_{gi} \leq 1 - u_{gt}, t \in T, g \in G \quad (15)$$

$$p_{gt} - p_{gt-1} \leq (SU_g - RU_g)v_{gt} + RU_g u_{gt}, t \geq 2, g \in G \quad (16)$$

$$p_{gt-1} - p_{gt} \leq (SD_g - RD_g)w_{gt} + RD_g u_{gt-1}, t \geq 2, g \in G \quad (17)$$

$$p_{rt} \leq AF_{rt}\overline{P}_{rt}, r \in R, t \in T \quad (18)$$

$$0 \leq pc_{st} \leq \overline{PC}_s, t \in T, s \in S \quad (19)$$

$$0 \leq pd_{st} \leq \overline{PD}_s, t \in T, s \in S \quad (20)$$

$$p_{st} = pd_{st} - pc_{st}, t \in T, s \in S \quad (21)$$

$$\underline{PE}_s \leq pe_{st} \leq \overline{PE}_s, t \in T, s \in S \quad (22)$$

$$pe_{st} = pe_{st-1} + pc_{st} * \eta_{st}^c - \frac{pd_{st}}{\eta_{st}^d}, t \in T, s \in S \quad (23)$$

$$inj_{nt} = \sum_{l=(n'\rightarrow n), n'\in N} f_{lt}, t \in T, n \in N \quad (24)$$

$$\underline{f_l} \leq f_{lt} \leq \overline{f_l}, l \in L, t \in T \quad (25)$$

$$\sum_{g\in G_n} p_{gt} + \sum_{r\in R_n} p_{rt} + \sum_{s\in S_n} pd_{st} + inj_{nt} = \quad (26)$$

$$D_{nt} + \sum_{s\in S_n} pc_{st}, t \in T, n \in N \quad (27)$$

$$u_{gt} - u_{gt-1} = v_{gt} - w_{gt}, t \in T, g \in G \quad (28)$$

$$u_{gt}, v_{gt}, w_{gt} \in \{0,1\}, p_{gt}, p_{rt}, p_{st}, pe_{st}, inj_{nt}, f_{lt} \in \mathbb{R} \quad (29)$$

(9) is the objective function of the UC it includes the generation cost (including a price, 80 $\frac{\$}{ton}$, for emissions produced) and start cost. (10) is the system wide cost for energy that is not served, ENS_{nt} at node *n* and time *t* is multiplied by the value of lost load (VOLL) which is set at 10.000 $\frac{\$}{MWh}$. (11) is an additional wheel charge (WC) of 1 $\frac{\$}{MWh}$ on the power flow between countries. Constraint (12) and (13) ensures the minimum and maximum production of a generator. Constraint (14) and (15) ensures the minimum up and downtime of the generators. Constraint (16) and (17) ensures the ramping limits of generators between timesteps. Constraint (18) ensures that the RES production is lower that the availability at that hour. (19), (20) and (22) ensure the charge, discharge and energy storage limits for storage units. Equation (21) is the sum of charge and discharge i.e., the net storage production. Equation (23) describes the relation between the charge, discharge and net power production of a storage unit. Equation (27) describes the logic between the binary commitment, start and stop variables of the generators. Equation (24) describes the relation between the flow on transmission lines and the power injection at nodes. Constraint (25) ensures flow limits on transmission lines. Equation (26) ensures that the total generation meets the total demand at every node and timestep. At last, the commitment variables are binary while the production are real numbers (28).

		LOLH			CF		CO_2
Variant	**ENS**	**Cost**	**Full**	**Gas**	**Coal**	**Nuclear**	
Cost of unserved energy	X	X	X		X		X
Variable generation cost		X	X		X		X
Wheeling charge			X				
Flex constraints included			X		X		X
Function f	(10)	(9), (10)	(9), (10), (11)		(9), (10)		(9), (10)
Constraints	(12), (13), (18)-(28)	(12), (13), (18)-(28)	(12)- (28)		(12)- (28)		(12)- (28)
α		1			{1.001,1.005,1.01}		{1.001,1.005,1.01}
Function g_{max}		$min \sum_{t\in T, g\in G} ENS_{gt}^2$			$max \sum_{t\in T, g\in G} p_{gt}$		$max \sum_{t\in T, g\in G} FCP_g u_{gt} + VCP_g p_{gt}$
Function g_{min}		$min \sum_{t\in T, n\in N} LOLH_{nt}$			$max \sum_{t\in T, g\in G} p_{gt}$		$min \sum_{t\in T, g\in G} FCP_g u_{gt} + VCP_g p_{gt}$
Extra constraints	$LOLH_t \in \{0,1\}$, $\mathrm{ENS_{nt}} \leq \mathrm{D_{nt}}\,\mathrm{LOLH_{nt}}$						

Table 5, The explicit details of all the experiments.

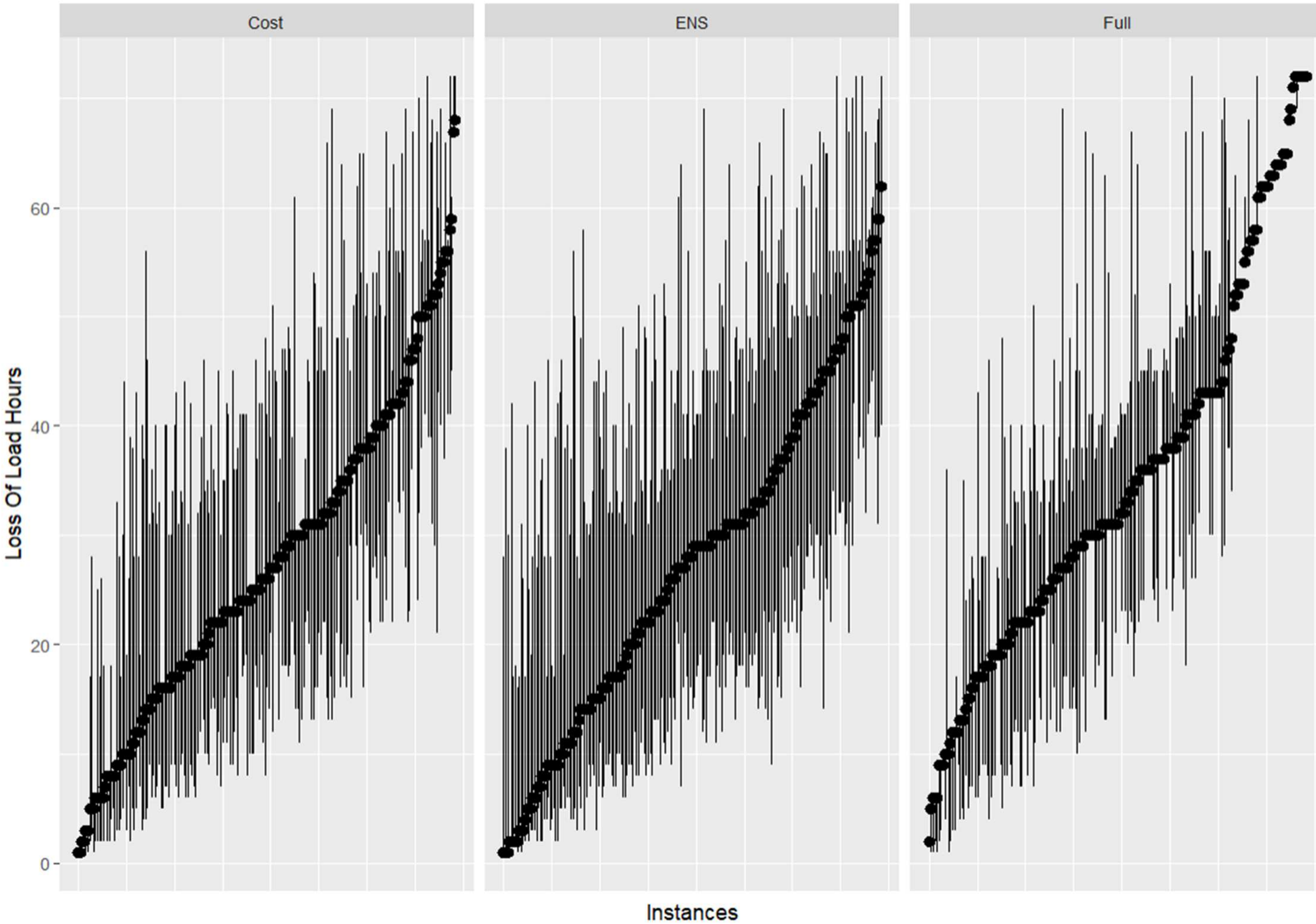

Figure 7, LOLH for all instances, ordered by the LOLH, for the Cost-Model, ENS-Model and Full-Model

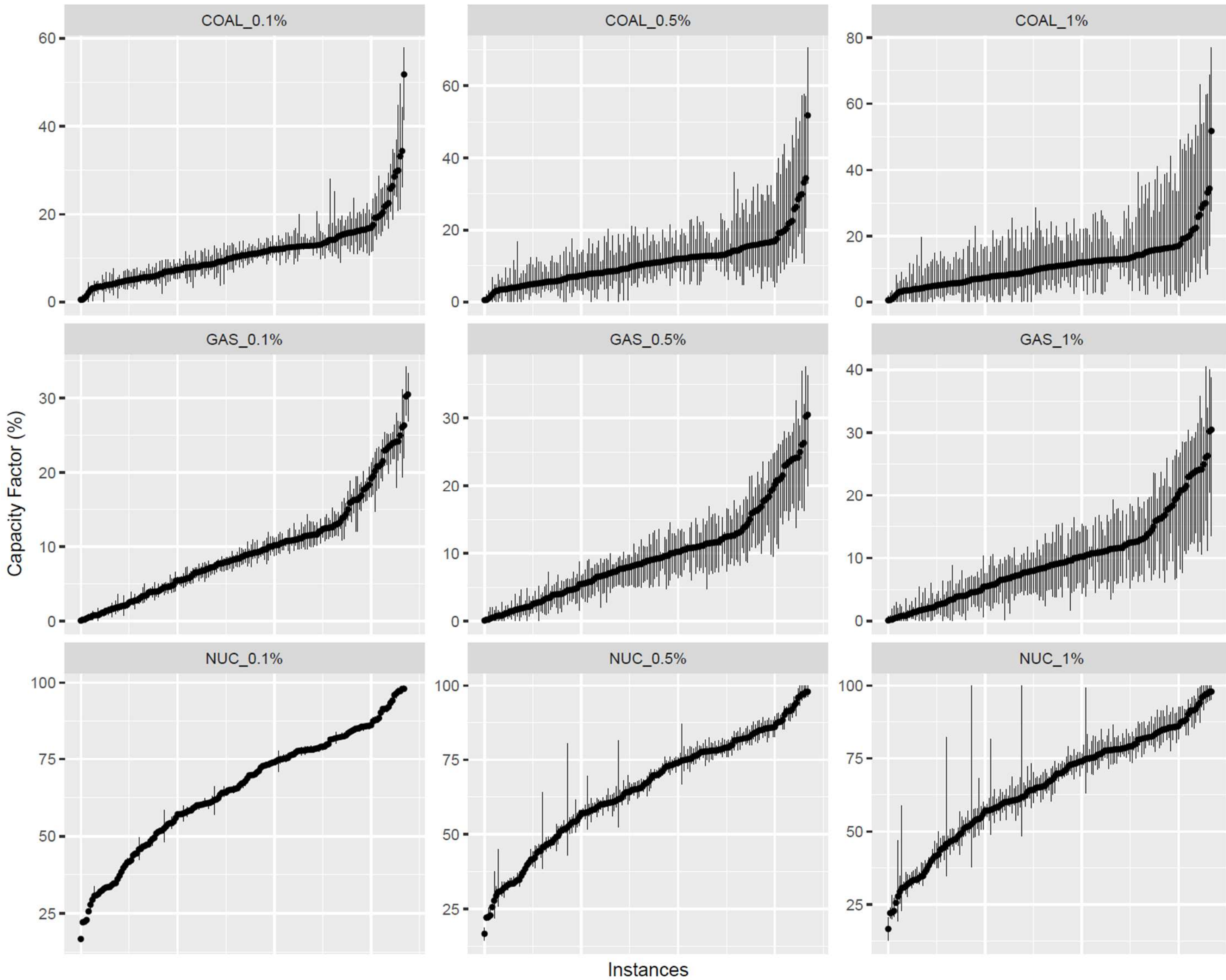

Figure 8, CF for all instances, ordered by the CF, for the Coal, Gas and Nuclear for different values of α.

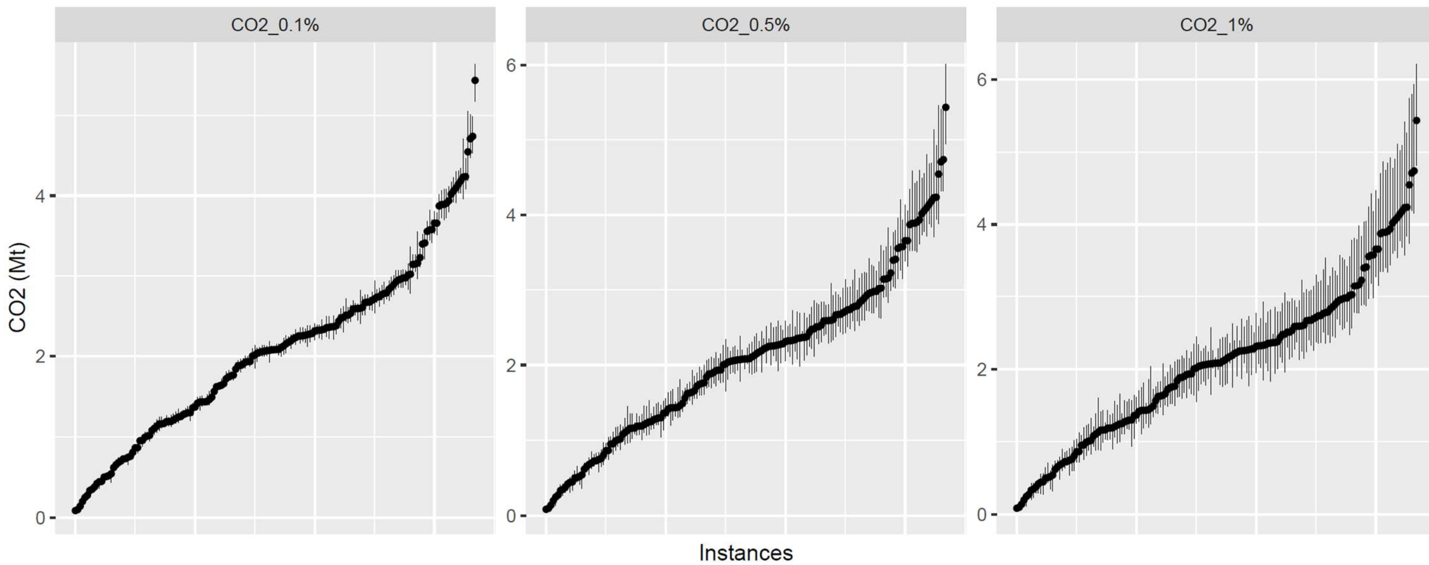

Figure 9, CO2 for all instances, ordered by the CO2, for different values of α