International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: 450

INFORMATION THEORETICAL EVALUATION OF AGGREGATION METHODS IN THE MATHEMATICAL OPTIMIZATION OF THE UNIT COMMITMENT PROBLEM

Thomas Weber*[1], Thomas Kohne[1], Nina Strobel[1], Jakob Wolber[1], Marlene Sachs[1], Eberhard Abele[1]

[1] Institute of Production Management, Technology and Machine Tools, Technische Universität Darmstadt, Germany

* Corresponding Author: t.weber@ptw.tu-darmstadt.de

ABSTRACT

While energy systems become more and more complex mathematical optimization can be used to generate good operating strategies for power generators and consumers. In order to reduce the complexity and thus the computing time for solving the Unit Commitment Problem several approaches for abstracting the original problem are introduced in literature. This paper focuses on a theory for the information theoretical evaluation of aggregation approaches in operating strategy optimization of energy systems. The theory is validated via two aggregation approaches, which aim at reducing computing time for solving the optimization problem, while maintaining the solution quality of the original problem.

Keywords: industrial microgrid, unit commitment problem, optimization, information theory

NONMENCLATURE

Abbreviations	
CHP	*Combined heat and power plant*
MILP	*Mixed integer linear programming*
SOC	*State of Charge*
UCP	*Unit Commitment Problem*

1. INTRODUCTION

Climate change has been a topic of world politics for several decades [1]. In order to reduce CO2 emissions, one of the most important measures is the gradual substitution of electricity supply from conventional power plant technologies by renewable energies such as wind power and photovoltaics [1]. Energy-efficient technologies, i.e. consumer-oriented cogeneration are increasingly integrated into the electrical energy supply. In contrast to conventional power plants, the renewable energy sources wind and photovoltaics are weather dependent and show a volatile feed-in behavior [2]. Since generation and consumption of electrical energy must always be in balance, more and more decentralized stakeholder must take over system responsibility. The scheduling of many decentralized plants on the energy market can be ensured by mathematical optimization procedures, such as Mixed Integer Linear Programming (MILP). The Unit Commitment Problem (UCP) in the electrical energy market mainly focuses on operating strategies of plants to fit a certain energy demand by minimizing cost or maximizing the system's revenue in energy supply [3]. As the number of plants increases so does the complexity of the resulting optimization problem. This can lead to high computational effort and out-of-memory problems [4].

In literature, there are several approaches with the aim of handling the complexity of the UCP [3]. In addition to decomposition methods, like Lagrange Decomposition or problem-specific heuristics, the reduction of the problem size is intended in order to solve complex optimization problems while maintaining optimality [5]. Aggregation and subsequent disaggregation techniques provide an efficient way to reduce the problem size while still addressing all components of the original optimization problem. For this purpose, parts of the original problem are merged into a smaller, less complex surrogate problem, which preferably provides a good approximation solution [6].

In all the approaches considered in this paper [4,6–10], a good approximate solution is calculated by simplifying the original problem. This leads to the questions, which parts of an optimization model actually add value to the calculated solution and which simplifications can be applied to optimization problems with negligible loss of optimality. To answer these questions, a strategy is presented, how optimization problems can be analyzed regarding the added value of model components to the actually executed operating strategy. The following section firstly introduces the core idea of information theory within mathematical optimization for solving the UCP. Afterwards two aggregation approaches, time-step aggregation and plant aggregation, are used to evaluate the presented information theoretical assumptions for mathematical programming. Lastly, a conclusion about the results is made.

2. INFORMATION THEORY AND OPTIMIZATION

Russell and Norvig describe the value of information in terms of operating strategies by the following [11]: *“In sum, information has value to the extent that it is likely to cause a change of plan and to the extent that the new plan will be significantly better than the old plan.”* This is an important basic hypothesis for the qualitative evaluation of information in optimization problems. A quantitative measure for the information content of a random variable can be found in information theory with Shannon's entropy [12]. If decision variables of an optimization problem are regarded as continuous stochastic variables, it is possible to evaluate their information content according to the information-theoretical measure of information entropy. Before the optimization problem is solved, there is only an uncertain range of solutions for the variables optimal outcome. If the decision variable itself can be predicted depending on other decision variables (without solving the optimization problem), it may be possible to remove the variable from the optimization problem without changing the overall problem structure. For reducing the complexity, simplified model formulations must be defined, which summarize several predictable decision variables in such a way, that the original and the simplified model will have an identical or similar resulting schedule as an optimal solution. In summary, decision variables can be aggregated with little loss of information if the following requirements are met:

- *The decision variable shows low conditional entropy*
- *Neighboring or dependent decision variables show low conditional entropy*
- *A simplified model formulation exists, which represents the decision variables outcome considering its conditional entropy of the original problem*

This perspective makes it possible to identify and eliminate redundancies in the optimization problem in order to reduce the problem size, while preserving information of the original problem.

In order to formulate the UCP as a mixed integer linear optimization problem, technical and economical characteristics are mapped via constraints and an objective function. In the following an optimization model based on [13] providing a cost-optimized operating strategy for a CHP in combination with a thermal storage is used for illustration purposes. Plant characteristics are modelled time-step and plant-specifically. As all time-step dependent constraints, time-step overlapping constraints are also repeated for each individual time-step. This results in large optimization problems and repeated symmetrical model formulations in the optimization model, as can be seen in Fig. 1. In this paper, symmetries in the context of mathematical optimization describe identical, repeating model formulations. If decision variables in related, symmetrical model formulations are very likely to be occupied similarly one can speak of redundancies in the optimization problem. Redundancies are therefore a subset of symmetries. As shown in Fig. 1, redundancies can be reduced by aggregating symmetrical model formulations. With an energy content E_t, the Power P_t^a and an energy demand P_t^d the energy content for the following time-step E_{t+1} is calculated by $\boldsymbol{E_{t+1} = E_t + \Delta t \cdot (P_t^a - P_t^d)}$. If P_1^a and P_2^a are probably similarly assigned (a), this redundancy can be eliminated by combining time-step 1 and time-step 2 (b).

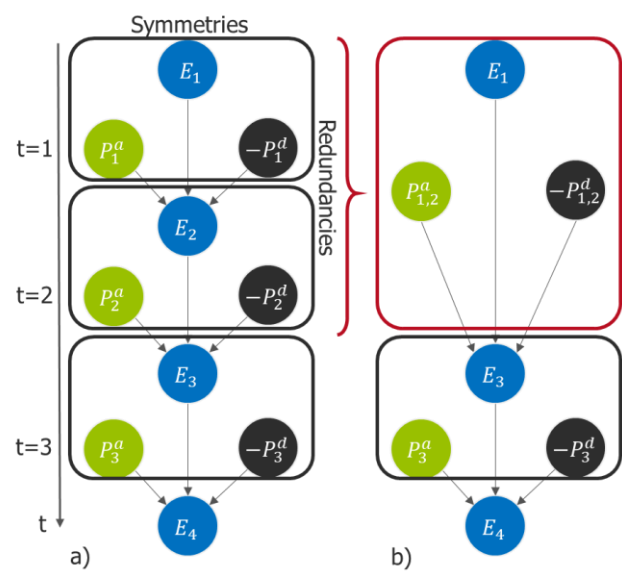

Fig. 1: Visualization of symmetries and redundancies arising from repeating time-steps in a typical UCP

Besides of time-steps, symmetries can also arise from the modelling of similar plants. If the operating strategy of several identical plants is optimized, this leads to a duplicate implementation of all decision variables, objectives and constraints within a time-step. This results not only in unnecessarily large problems, but also in problems that are difficult to handle for commercial solvers based on *Branch and Bound* algorithms like IBMs CPLEX, as could be shown in [4].

One way to reduce the complexity of the problem is to reduce the time horizon under consideration. In order to still be able to optimize over long periods of time, the so-called "rolling horizon" approach has been established in the optimization via model predictive control. A short time period of the UCP is repeatedly optimized, whereby only the first calculated actions are executed [14]. When using rolling horizon algorithm, the optimization of all time-steps that are not realized as actions only aims at obtaining a foresighted plant schedule. This in turn affects the degree of detail with which these time-steps have to be modelled in order to obtain an optimal schedule, since the influence of these variables on the first time-step(s) is more relevant. In most approaches optimizing the UCP, for determining the system state in the current time-step x_t, knowledge of the state in the preceding time-step x_{t-1} is sufficient, so the problem satisfies the first Markov property [15]. In the case of the operating strategy optimization, this means that all information of past time-steps is represented by state variables, such as storage filling levels, and thus dependencies only have to be mapped between two adjacent time-steps. This automatically leads to the fact that specifications of future time-steps are less and less relevant for the actual actions to be implemented. Aggregation methods can take advantage of this, since in future time-steps higher inaccuracies are allowed for simplifying the problem.

3. AGGREGATION APPROACHES

In this section we will validate the presented theory by a qualitative approximation of information entropy.

3.1: Time-step aggregation

As discussed in section 2, the complexity and thus symmetries and redundancies in UCP models increase with an increasing number of time-steps. Although only the first time-step is operated in rolling horizon approaches, the quality of optimization results increases by integrating longer forecast series. This raises the question, which time interval Δt and time horizon T is necessary for optimizing operating strategies in the UCP. Changing parameters over time-steps in UCP are fluctuating energy prices as well as energy demands while plant characteristics usually remain unchanged. Since energy is traded as products on German energy markets at fixed 15-minute intervals, the time interval Δt for the optimization of smart grids is usually set to 15 minutes (0.25 hours). For each 0.25-hours interval, the model formulation is repeated, whereby a simplification of the problem can easily be achieved by combining individual time-steps into time-steps with a larger time interval. If the conditional entropy of decision variables is relatively low due to similar parameterization of individual, adjacent time-steps, the optimal allocation of the decision variables is more likely predictable and aggregation can be performed with little loss of optimality. Hypothesis 1 is introduced to estimate the conditional entropy of decision variables:

Hypothesis 1: *The more similar adjacent energy prices and demands in forecast series are, the more similar is the operating strategy of a plant over different time-steps and the less information and optimality is lost when these time-steps are aggregated.*

In the following, an approach with three adjustable parameters is introduced to qualitatively analyze the hypothesis and find time-steps of the UCP that can be aggregated with negligible information and thus optimality loss. Since not only the environmental conditions, but also the respective state at the beginning of a time-step, such as the energy content of the thermal storage, have an effect on the operating strategy during the time-step, hypothesis 1 is only an approximation to the real conditional entropy of the time-steps. For this reason, aggregation on the basis of hypothesis 1 is only implemented after a defined time-step t_{agg}, in order to reduce the negative effects of the resulting error on the first time-step. The second parameter is a list of potential time intervals $\mathrm{L}^{\Delta \mathrm{t}}$ to which the former time-steps are aggregated. If, for example, the original time interval Δt is 0.25 h and the list $\mathrm{L}^{\Delta \mathrm{t}}$ contains {0.5,1}, then 0.25 h time-steps may be aggregated to time-steps with intervals of 0.5 h or 1 h. For adjacent time-steps a low standard deviation σ describes little fluctuation of energy prices and energy demand and thus following hypothesis 1 the similarity of the resulting optimal operating strategy for different time-steps. Therefore a standard deviation filter (SDF) is introduced. The task of the SDF is to find time-steps of the UCP that can be aggregated with negligible information and thus optimality loss to reduce redundancies in existing symmetries of the problem.

The standard deviation of each input parameter is calculated for the adjacent time-steps depending on the intervals in list $\mathrm{L}^{\Delta \mathrm{t}}$. To standardize σ, the individual σ of each parameter must be normalized and then multiplied. This normalized standard deviation σ^{norm} can be compared to a manually defined threshold value σ^{thresh}:

$$\sigma_t^{\text{norm}} = \prod_{j \in J} \frac{\sigma_{t,j}}{\sigma_j^{\text{ave}}} \leq \sigma^{\text{thresh}} \qquad \forall t \qquad (1)$$

σ_j^{ave} is the mean standard deviation of parameter j. To find time-steps which can be aggregated the following procedure is carried out. Firstly, the three parameters are set and the normalized σ^{norm} is calculated for the highest time interval in list $L^{\Delta t}$. Then, the time-steps with the lowest σ^{norm} which fits the threshold are aggregated to one time-step. This is repeated until all time-steps which meet the threshold are considered. After that the procedure is operated for the next time interval in $L^{\Delta t}$ until all time intervals are used. For each aggregated time-step, the new energy price and demand is calculated by the mean value of the former time-steps. Finally, the optimization for the aggregated UCP is operated. This procedure is repeated for every plant with different energy demands.

In order to aggregate certain time-steps in a UCP for complexity reduction with regard to their information theoretical value and following hypothesis 1, time-step aggregation is performed for optimizing a plant pool of three CHP systems with individual storage systems each. For this purpose, from a defined t_{agg}, time-steps with similar parameterization are combined (see Fig. 2.).

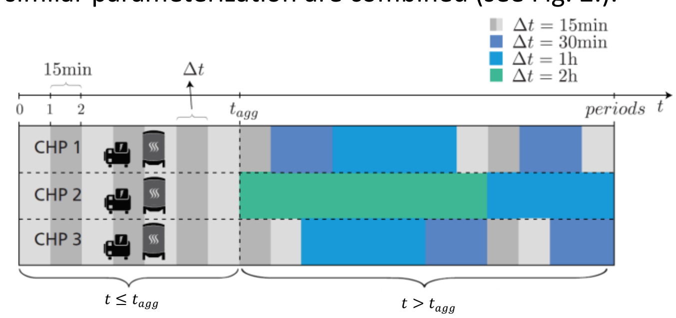

Fig. 2: Visualization of the time-step aggregation for three different CHP plants.

The procedure of time-step aggregation is validated by a comparison between the presented and a simple approach (Fig. 3). Both approaches are again related to a reference model without aggregation. In the simple aggregation approach time-steps are aggregated independently of the fluctuation of prices and heating demands starting from time-step $t_{\text{agg}} = 8$. Since the simple aggregation cannot combine different time-step sizes, $L^{\Delta t} = \{0.5\}$ contains only one time-step size. A comparison shows that the revenue of the presented approach is significantly higher than the one of the simple approach. By targeted combining time-steps, a significantly higher accuracy can be achieved than by randomly aggregation in the simple method. At the same time, the calculation time in the presented approach is reduced by almost 40% as shown in Fig. 3.

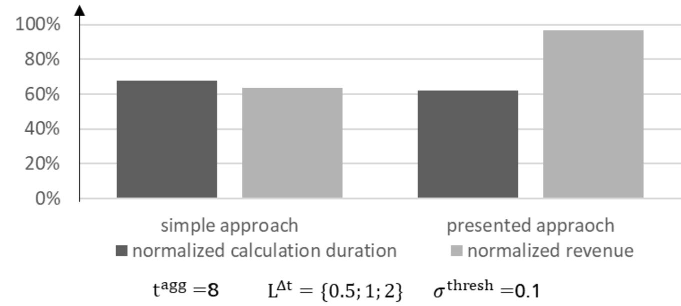

Fig. 3: Comparison of simple aggregation app vs. presented time-step-aggregation approach

Fig. 4 shows that the operating strategy with time-step aggregation is quite similar to the one without aggregation (small time horizon visualized). Therefore, the assumption from hypothesis 1 is confirmed, with which a good approximation of the conditional entropy can be achieved.

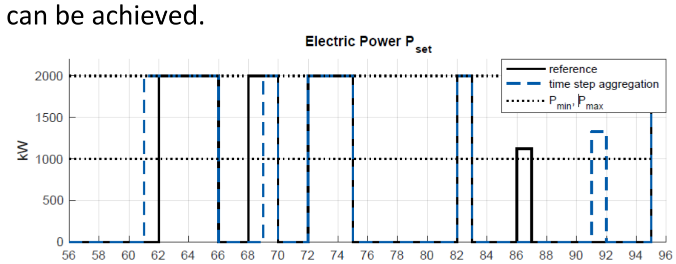

Fig. 4: Operation Strategy with and without time-step aggregation for time-steps 56 to 96

Although only a rough estimation of the conditional information entropy can be achieved via hypothesis 1, targeted simplifications can still be implemented in the optimization problem. By a more exact estimation of the conditional entropy between decision variables as well as by further simplified modelling approaches, a much better trade-off between computing time and solution quality can be achieved.

3.2: Plant aggregation

Similar plants, which are independent from each other, will have a similar operating strategy under similar boundary conditions, so information-theoretical redundancies can also be assumed here. To overcome these redundancies, with the approach described in this work several plants are merged into a single plant, which is modelled by only one set of decision variables and constraints (Fig. 5). In order to obtain an approximation of the conditional entropy between the decision variables of different plants, the following hypothesis is tested in this paper:

Hypothesis 2: *The more similar plants and their boundary conditions (e.g. utilization and prices) are the less information and thus optimality is lost when the facilities are aggregated in terms of operating strategy optimization.*

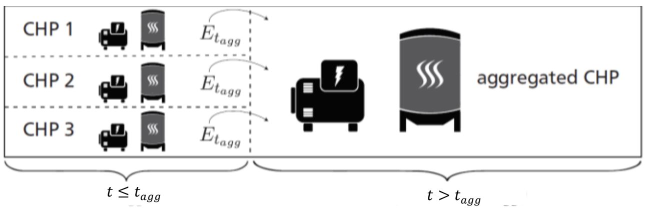

Fig. 5: Aggregation of plants in the UCP at time-step $\boldsymbol{t_{agg}}$

The aggregation itself is achieved by merging plants with similar characteristics into one aggregated plant. The parameters of the aggregated plant are calculated by summing up the power, the capacity, the inertia and the heat demand and taking the average efficiency of the individual systems. With aggregation of several individual plants to one virtual plant, however, some information on how the individual system actually operate is lost. Since only the first time-steps are relevant in the case of the repetitively executed operational strategy optimization, the detailed information of the optimization itself is also more important for current time-steps than for future ones. Accordingly, the described loss of information is more justifiable in future time-steps.

In the considered case of the CHP optimization problem with thermal energy storages, there are various characteristic values in which the systems can differ from each other. For example, the maximum electrical power output of the CHPs $P^{\mathrm{el,max}}$, the capacity of the heat storages $E^{\ \mathrm{max}}$ or the thermal demand profile can vary for different systems. In this paper the fact was outpointed that due to the rolling horizon optimization only the first time-step is actually executed. Therefore, in the plant-aggregation approach, plants are aggregated from a future time-step $\mathrm{t_{agg}}$ onwards, as it was implemented in time-step-aggregation. Here, $\mathrm{t_{agg}}$ is selected in such a way that the individual plants can reach any possible subsequent state in the range from t = 0 to t $=\mathrm{t_{agg}}$. In the case of cogeneration units, this subsequent state is determined by the energy content of the thermal storages, which in turn depends on the maximum thermal power production of the cogeneration $\mathrm{P^{th,max}}$, the heat demand $\mathrm{P^{demand}}$ and the capacity of the storages $\mathrm{E^{max}}$.

$$t_{\mathrm{agg}} = \max\left(\frac{E^{\mathrm{max}}}{(P^{\mathrm{th,max}} - P^{\mathrm{demand}})}; \frac{E^{\mathrm{max}}}{P^{\mathrm{demand}}}\right) \qquad (2)$$

The plants can therefore still react flexibly and, if necessary, compensate inaccuracies caused by the aggregation in previous calculations. Fig. 6 shows the idea of the presented approach considering a starting $\mathrm{SOC} = \mathrm{SOC^{start}}$.

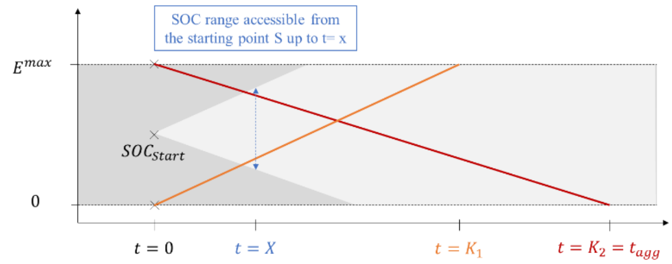

Fig. 6: SOC range accessible from t = 0 to t = $\mathrm{t_{agg}}$

Chosen results calculated via 2880 consecutive optimization runs are presented in Tab. *1*. The parameterization of the different systems was varied with respect to $\mathrm{P^{el,max}}$, $\mathrm{E^{max}}$and $\mathrm{P^{demand}}$ as well as the used $\mathrm{t_{agg}}$. In the following either a fixed $\mathrm{t_{agg}} = 8$ or a dynamic $\mathrm{t_{agg}}$ calculated via the formula mentioned in section 3.1. The results show that the method presented significantly reduces the computing time for solving the optimization problem while maintaining the quality of the solution. However, the dynamically calculated aggregation time-step $\mathrm{t_{agg}}$ only provides an orientation here, since results for a constant $\mathrm{t_{agg}}$ were significantly better in some cases.

Tab. 1: Summery of the results of the plant aggregation approach

Experiment	t_{agg}	Reduction of Computation time	Optimality gap
3 x CHP systems $P^{\mathrm{MAX}} = \mathbf{2000};\ E^{\mathrm{MAX}} = \mathbf{4400};$ $P^{\mathrm{demand}} = Q^{\mathrm{Reference}}$	8	91,7 %	1,64%
3 x CHP systems $\boldsymbol{E^{\mathrm{max}}} = \mathbf{4400}$; $\boldsymbol{P^{\mathrm{demand}}} = \boldsymbol{Q^{\mathrm{Reference}}}$; $\boldsymbol{P^{\mathrm{max,1}}} = \mathbf{2000}$; $\boldsymbol{P^{\mathrm{max,2}}} = \mathbf{200};\ \boldsymbol{P^{\mathrm{max,3}}} = \mathbf{20}$	8	91,5 %	10,44%
3 x CHP systems $\boldsymbol{P^{demand}} = \boldsymbol{Q^{\mathrm{Reference}}}$; $\boldsymbol{P^{max,1}} = \mathbf{2000}$; $\boldsymbol{E^{\mathrm{max,1}}} = \mathbf{4400}$; $\boldsymbol{E^{\mathrm{max,2}}} = \mathbf{4400};\ \boldsymbol{E^{\mathrm{max,3}}} = \mathbf{4400};$	8	97,4 %	8,72 %
3 x CHP systems $\boldsymbol{E^{\mathrm{max}}} = \mathbf{4400}$; $\boldsymbol{P^{\mathrm{max}}} = \mathbf{2000}$; $\boldsymbol{P^{\mathrm{demand,1}}} = \boldsymbol{Q^{\mathrm{Reference}}};\ \boldsymbol{P^{\mathrm{demand,2}}} = \boldsymbol{Q^{\mathrm{Reference}}} * \mathbf{0,5}$; $\boldsymbol{P^{\mathrm{demand,3}}} = \boldsymbol{Q^{\mathrm{Reference}}} * \mathbf{0,1}$	8	89,0 %	11,49 %
3 x CHP systems $\boldsymbol{E^{\mathrm{max}}} = \mathbf{4400};\ \boldsymbol{P^{\mathrm{max}}} = \mathbf{2000};$ $\boldsymbol{P^{\mathrm{demand,1}}} = \boldsymbol{Q^{\mathrm{Reference}}};\ \boldsymbol{P^{\mathrm{demand,2}}} = \boldsymbol{sine};\ \boldsymbol{P^{\mathrm{demand,3}}} = \boldsymbol{const}$	8 14,6	92,9 % 90,1 %	4,31 % 2,52 %
10 x CHP systems $\boldsymbol{P^{\mathrm{max}}} = \mathbf{2000}$; $\boldsymbol{E^{\mathrm{max}}} = \mathbf{4400}$; $\boldsymbol{P^{\mathrm{demand}}} = \boldsymbol{Q^{\mathrm{Reference}}}$	8 5,16	99,6 % 99,7 %	1,35 % 2,76 %

Fig. 7 shows that the operating strategy with plant aggregation is also similar to the one without aggregation (small time horizon visualized). Therefore, the assumption from hypothesis 2 is confirmed, with which an approximation of the conditional entropy can be achieved. In some cases the aggregated CHP is operated at P^{min}, so the power of one CHP unit equals $0{,}33 * P^{min}$, which actually means, that not all of the systems are used.

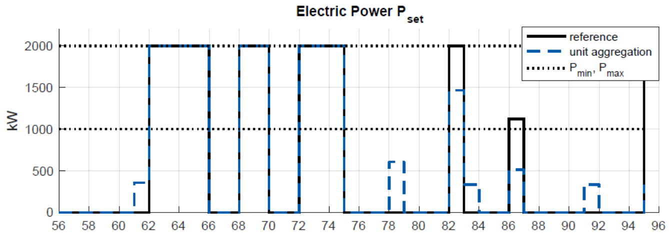

Fig. 7: Operation Strategy with and without plant aggregation for one CHP for time-steps 56 to 96

4. CONCLUSION AND OUTLOOK

In this paper a theory for the evaluation of aggregation approaches in the UCP according to information-theoretical criteria is presented. Decision variables are seen as stochastic variables whose optimal allocation can be predicted with a certain probability. The measure for evaluating the predictability of decision variables is information entropy. If the optimal allocation of decision variables can be predicted and the entropy is low, this can be used for a simplified model formulation. An information-theoretical evaluation of an optimization model is carried out by a qualitative approximation of information entropy in this paper. The established theory is validated and confirmed in first instance by two aggregation procedures, the time-step and the plant-aggregation in the UCP. However, an information-theoretical evaluation of the optimization model has been carried out by a qualitative evaluation of information entropy. Since optimization problems with different parameterizations are repeatedly executed using the rolling horizon procedure and thus all decision variables of the problem are optimally assigned, the conditional entropy of the optimization problem can be approximated quantitatively using this assignment. With this even more precise information, the aggregation procedure can also be carried out more precisely and even better results can be expected. However, the quantitative evaluation of the presented procedure will be the task of future research work.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support of the Kopernikus-project "SynErgie" by the Federal Ministry of Education and Research (BMBF) and the project supervision of the project management organization Projektträger Jülich (PtJ).

REFERENCES

[1] International Renewable Energy Agency. A New World - The Geopolitics of Energy Fransformation; 2019.

[2] Mobius T, Musgens F. The effect of variable renewable energy sources on the volatility of wholesale electricity prices — A stylized full cost approach. In: 2015 12th International Conference on the European Energy Market (EEM): IEEE; 2015 - 2015, p. 1–5.

[3] Bhardwaj A, Kamboj VK, Shukla VK, Singh B, Khurana P. Unit commitment in electrical power system-a literature review. In: 2012 IEEE International Power Engineering and Optimization Conference: IEEE; 2012 - 2012, p. 275–280.

[4] Knueven B, Ostrowski J, Watson J-P. Exploiting Identical Generators in Unit Commitment. IEEE Trans. Power Syst. 2018;33(4):4496–507.

[5] Grossmann IE. Advances in mathematical programming models for enterprise-wide optimization. Computers & Chemical Engineering 2012;47:2–18.

[6] Rogers DF, Plante RD, Wong RT, Evans JR. Aggregation and Disaggregation Techniques and Methodology in Optimization. Operations Research 1991;39(4):553–82.

[7] Deml S, Ulbig A, Borsche T, Andersson G. The role of aggregation in power system simulation. In: 2015 IEEE Eindhoven PowerTech, p. 1–6.

[8] Abgottspon H, Schumann R, Epiney L, Werlen K. Scaling: managing a large number of distributed battery energy storage systems. Energy Inform 2018;1(S1):13.

[9] Bahl B, Söhler T, Hennen M, Bardow A. Typical Periods for Two-Stage Synthesis by Time-Series Aggregation with Bounded Error in Objective Function. Front. Energy Res. 2018;5:333.

[10] Bahl B, Lützow J, Shu D, Hollermann DE, Lampe M, Hennen M et al. Rigorous synthesis of energy systems by decomposition via time-series aggregation. Computers & Chemical Engineering 2018;112:70–81.

[11] Russell SJ, Norvig P. Artificial intelligence: A modern approach. 3rd ed. Upper Saddle River: Pearson; 2016.

[12] Shannon CE. A Mathematical Theory of Communication. Bell System Technical Journal 1948;27(3):379–423.

[13] Weber T, Strobel N, Kohne T, Wolber J, Abele E. Realistic modeling of a combined heat and power plant in the context of mixed integer linear programming. Energy Informatics (Energy Informatics) 2018;1(S1):1098.

[14] Costley M, Feizollahi MJ, Ahmed S, Grijalva S. A rolling-horizon unit commitment framework with flexible periodicity. International Journal of Electrical Power & Energy Systems 2017;90:280–91.

[15] Rønn-Nielsen A, Hansen E. Conditioning and Markov properties. Cph.: Department of Mathematical Sciences; 2014.