Comments on: A comparative study of time aggregation techniques in relation to power capacity-expansion modeling

Kris Poncelet & Erik Delarue

TOP

An Official Journal of the Spanish Society of Statistics and Operations Research

ISSN 1134-5764

TOP

DOI 10.1007/s11750-019-00526-0

Your article is protected by copyright and all rights are held exclusively by Sociedad de Estadística e Investigación Operativa. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Comments on: A comparative study of time aggregation techniques in relation to power capacity-expansion modeling

Kris Poncelet^{1,2} · Erik Delarue^{1,2}

© Sociedad de Estadística e Investigación Operativa 2019

Mathematics Subject Classification 90-02

1 Introduction

Planning models are widely used to provide insight in the transition of energy systems. With the increasing levels of intermittent renewables, the temporal representation in such models requires careful attention. However, representing all 8760 h in each modeled year while at the same time integrating detailed technical constraints of power plants might lead to an overall computationally intractable problem. In this context, several time aggregation methods to select a number of typical periods that together form an adequate representation of the full year have recently been developed.

In this context, we have very much enjoyed reading the paper by Buchholz et al. We would first of all like to give the authors credit for doing the effort of implementing and comparing different time aggregation methods (including several self-developed methods). The effort to (i) get an overview of different aggregation methods, (ii) understand and implement these and (iii) do a careful comparison of their performance cannot be underestimated.

In this comment, we would like to focus on the validation of time aggregation methods. In this sense, we very much applaud the emphasis the authors have put on the important, but not trivial, task of validating/evaluating time aggregation methods. In the remainder of this comment, we will present some reflections, challenges and proposals on how to validate time aggregation techniques. Following the convention in the paper, in this comment we distinguish between data and model validation.

This comment refers to the invited paper available at <https://doi.org/10.1007/s11750-019-00519-z>.

✉ Kris Poncelet
kris.poncelet@kuleuven.be

¹ KU Leuven, Leuven, Belgium

² EnergyVille, Genk, Belgium

2 Data validation

Data validation starts from the inherent assumption that one *knows* which elements of the input data exactly drive the results, and to what extent. Given the increasingly complex power system, acquiring this knowledge is not straightforward. In the context of long-term planning models, the primary focus is frequently on approximating the residual load duration curve (RLDC). The RLDC definitely holds some relevant information for planning models, such as the required capacity (peak residual load) and the amount of potential excess renewable generation (near-zero or negative residual load). It must be noted that in models in which investments in renewable energy sources are co-optimized, the RLDC is not known a priori. In addition, as discussed in our earlier work (Poncelet et al. 2017), the RLDC also lacks information, e.g., regarding chronology and the correlation between different time series. Chronological information is relevant for short-term technical constraints as well as for the arbitrage potential for different storage technologies or start-up behavior. Correlations between time series are relevant to consider when aggregating time as to capture the fact that the value of generating power in a specific moment in time depends on the demand for power, and the ability of other sources of power generation to generate electricity in that moment in time. In addition, also other relevant features, such as the distributions of ramps, minimum or maximum ramping events and maximum load periods can influence the results. Due to this complexity, it is highly difficult to formulate a single metric for data validation that encompasses all these different aspects and their relative importance. In summary, the main challenge for data validation is to find an appropriate metric(s) to evaluate the quality of the aggregated input data in a way that it reflects the elements that drive the results in the final model one deploys with the aggregated instance of the data.

3 Model validation

In contrast to data validation, model validation could feel a more natural and fair way of validating temporal aggregation techniques, as achieving a good approximation of the non-aggregated model outcome is in the end the goal of the aggregation technique. However, model validation faces also a number of challenges, namely:

1. Tractability of computing the solution of the non-aggregated problem;
2. Defining a suitable metric to evaluate model performance;
3. Dependence on the specific model and case used.

The issue related to the tractability of computing the solution of the non-aggregated problem is discussed well by the authors and is not further discussed in this comment.

3.1 Metrics for evaluating model performance

As the authors point out, different metrics have been used to evaluate the accuracy of the solution of the aggregated problem. Most popular metrics are the comparisons of the installed capacity and the system cost. However, metrics based on comparing

generation shares, curtailment of VRE, carbon emissions and number of start-ups have also been used in the literature. In this paper, the authors have combined metrics for the accuracy of approximating the maximum capacity, the generation mix, the investment decisions, the number of shut downs, and the overall system costs, into a single performance index. In general, it is common practice and feels intuitive to use such metrics that directly reflect those parts of the results in which users of investment models are typically interested in.

In this comment, we argue that one should be highly careful in employing such metrics. The reason is that an optimization problem can have multiple solutions that might be very different in terms of investment decisions, generation mix, etc., but almost identical in terms of the total system cost (which in the end forms the objective function of the optimization problem). As a result, a very slight bias introduced in an investment model by using a good time aggregation method can be sufficient to arrive in a solution that is very close in terms of total system cost, but very different in terms of capacity mix, generation mix, etc.). Moreover, a large bias introduced in an investment model by using a less good time aggregation method does not necessarily result in a solution that is (significantly) different from that found by the optimization model without aggregated temporal data. This because it could be possible that the bias happens to affect the optimal solution less or simply in a more favorable direction than it would affect other potential solutions.

We argue in this comment that a suitable way to evaluate the model performance of an aggregation technique is to compare the difference between the projected system cost (as estimated by the investment model using the aggregated data), and the effective system cost (i.e., the cost an investment model using non-aggregated data would find if the same investment decisions would be made). Moreover, we propose to do this for a range of different potential solutions. The steps involved in the proposed methodology are presented below:

1. Set up a number of cases used for the evaluation: select different capacity mixes by hand and combine with different data sets;
2. Derive the aggregated data sets by using an aggregation technique;
3. For each case, run the investment model using the aggregated data with fixed installed capacities corresponding to the capacity mix of that case (i.e., basically run the dispatch part of the investment model using the aggregated data for each case);
4. For each case, run the investment model using the full data set with fixed installed capacities corresponding to the capacity mix of that case;
5. For each case, compare the total system costs resulting from steps 3 and 4.

Using the above methodology, it would be possible to have a good measure of how accurate a model using aggregated data approximates the actual costs for a wide range of potential solutions. If an aggregation technique systematically achieves a high accuracy in terms of projecting the total system costs, it can be ensured that the investment model using that aggregated data would not find highly suboptimal solutions. Moreover, this methodology allows gaining insights into when (i.e., for which type of systems) certain aggregation techniques perform poorly, which is essential information for improving the performance of aggregation techniques. Finally, an additional

benefit of this methodology is that one can calculate a good reference solution without facing computational issues, as one basically only needs to solve operational problems (or investment problems in which all investment decisions are fixed).

3.2 Model(s) and case(s) used for evaluating model performance

As the authors very validly suggest at the very end of the paper, specific features of the model and the cases used for the evaluation of an aggregation technique can potentially play an important role.

In terms of models, it is relevant to note that to capture the challenges related to integrating intermittent RES, not only an accurate temporal representation is required, but the model would also require an appropriate level of technical and spatial detail. Particularly regarding the representation of technical constraints in planning models, a lot can be said.

First, to maintain computational tractability, most generation expansion planning models do not integrate the full set of constraints typically adopted in unit commitment (UC) models [for examples of papers in which UC constraints are integrated with different levels of detail, we refer to Kirschen et al. (2011) and Palminier and Webster (2016)]. As a result, the validity of the reference simulation can be questioned. Again, the validation methodology we propose overcomes this problem.

Second, when using aggregated data in a planning model, there are typically some decisions to be made regarding how to couple the different representative periods (e.g., hours, days or weeks). Can storage technologies arbitrage only within or also between different representative periods? Are ramping and other intertemporal constraints between different representative periods considered or not? The evaluation of the accuracy of the aggregation technique might depend on the decisions made here. Regarding the choice of case(s), important choices relate to the scope and characteristics of the system under consideration.

Regarding the scope of the system, using a small-scale test system has the advantage that it would be feasible from a computational perspective to compute a reference solution using the non-aggregated data (although, with our proposed methodology, computation times should not be a limiting factor). A drawback of using a small system is that one does not capture the flexibility that the network could provide in an integrated electricity system through spatial smoothing. In addition, care should be taken when using integer investment decisions in a small system. This is because the distance between two different solutions is inherently big as a single additional/different power plant is already significant. As a result, it might be difficult to fully see the subtle differences between time aggregation techniques.

Regarding the characteristics of the system, there can be significant differences between the available flexibility and the need for flexibility. These differences can come from the system being considered (e.g., potential for basin hydro or pumped hydro-storage), the scenario being considered (e.g., share of variable renewables in the generation mix, costs assumptions of batteries), as well as specific data assumptions (e.g., the sizing of reserves, the cost of batteries, the flexibility of conventional power plants). The characteristics of the system can hence have a significant impact on the

evaluation of the time aggregation technique. In a system with a lot of basin hydro or in a system in which batteries are assumed relatively inexpensive, big ramping events might be easily overcome, and hence it would be of less relevance of an aggregation technique to capture such events.

4 Concluding remarks

The authors of the paper rightfully point out the importance of the validation/evaluation of time aggregation techniques. However, it is not a trivial task and different papers on time aggregation techniques use different approaches. In this comment, we have added certain suggestions and reflections regarding the methodology to use for the evaluation of aggregation techniques. Most importantly, we argue to evaluate aggregation techniques by their ability to accurately approximate the total system cost of the non-aggregated system. In addition, we propose to do this for a variety of different system compositions (imposed capacity mixes). One of the benefits of fixing the system composition is that the difficulty of calculating a reference solution using non-aggregated data is avoided. An additional advantage is that this analysis provides useful insights regarding which temporal features drive the results of the planning models. As such, it can form the basis for better metrics for data validation of time aggregation techniques. Finally, we warn for certain assumptions taken in the model or the data that might have an impact on the evaluation of the time aggregation technique.

As a last note, we see certain open issues for data aggregation techniques requiring further research. These include a good approximation of the arbitrage opportunities for storage technologies and the performance and application of temporal aggregation methods for multi-regional models (e.g., European-scale models). In addition, there might be opportunities to combine different time aggregation techniques. An example would be to select representative periods using one method followed by an efficient way of reducing the temporal resolution via consecutive time blocks using another method. Another option would be to complement a heuristic selection of extreme points (hours/days/weeks) with a method to select representative periods.

References

- Kirschen DS, Ma J, Silva V, Belhomme R (2011) Optimizing the flexibility of a portfolio of generating plants to deal with wind generation. In: 2011 IEEE power and energy society general meeting. Institute of Electrical & Electronics Engineers (IEEE). <https://doi.org/10.1109/pes.2011.6039157>
- Palminier BS, Webster MD (2016) Impact of operational flexibility on electricity generation planning with renewable and carbon targets. *IEEE Trans Sustain Energy* 7(2):672–684. <https://doi.org/10.1109/TSTE.2015.2498640>
- Poncelet K, Höschle H, Delarue E, Virag A, D'haeseleer W (2017) Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems. *IEEE Trans Power Syst* 32(3):1936–1948. <https://doi.org/10.1109/TPWRS.2016.2596803>