DOI: 10.58895/ksp/1000162754-4 erschienen in:
Proceedings. 33. Workshop Computational Intelligence - Berlin, 23. – 24. November 2023
DOI: 10.5445/KSP/1000162754 | https://www.ksp.kit.edu/site/books/e/10.5445/KSP/1000162754/

Machine learning based model fitting concept for energy system components in energy management

Martin Winkelkotte[1], Steffi Naumann[2], Sebastian Flemming[2], Peter Bretschneider[1]

[1]Technische Universität Ilmenau
Gustav-Kirchhoff-Str. 5, 98693 Ilmenau
E-Mail: martin.winkelkotte@tu-ilmenau.de
E-Mail: peter.bretschneider@tu-ilmenau.de

[2]Fraunhofer IOSB-AST
Am Vogelherd 90, 98693 Ilmenau
E-Mail: steffi.naumann@iosb-ast.fraunhofer.de
E-Mail: sebastian.flemming@iosb-ast.fraunhofer.de

1 Introduction

The heterogenous nature of energy systems is a big challenge for the large-scale implementation of energy management systems (EMS). The different energy systems can be small energy communities, factories, or large-scale supply areas. A consistent challenge throughout the various energy system forms is the accurate modelling of the generation and storage units as well as the prediction of consumer behaviour, all what is needed to optimize which energy source to use. While significant work has been put into the automated prediction of consumption timeseries, the modelling of generation and storage units is still mostly done “by hand” and is reliant on external information sources like datasheets. Models created using this data basis are extremely vulnerable to discrepancies between the external information they are based on and the actual unit properties. Such inaccuracies often go unnoticed in the initial model generation, until they cause large errors in the optimization which then require time consuming analysis and finetuning to correct. In this paper an initial approach to automate the modelling of generation and storage units is presented. After that the approach is tested on its sensitivity to different training data sets since a dependency on high quality training data would severely limit its transferability.

2 Concept

The idea of the approach is to create a library of simplified generic models for the most common types of generation and storage technologies. These models can then be fitted to match the properties of the target units using key parameters whose values are identified by an artificial neural network (ANN). For this purpose, the ANN is trained on timeseries data that has been created using numerus simulated units with different properties of one unit type. The defining properties of the simulated units have been distributed to reflect the entire feasible parameter range. The trained ANN is then used on measured data of the target unit to identify the relevant parameter values to modify the generic model. The resulting fitted model should more accurately reflect the properties and performance limits of the target than a model based on generic datasheet information. Additionally, it would be easy to implement a periodic refitting of the model on the most recently measured data. This would allow the model to reflect changes due to aging and could even be used as an indicator for maintenance scheduling.

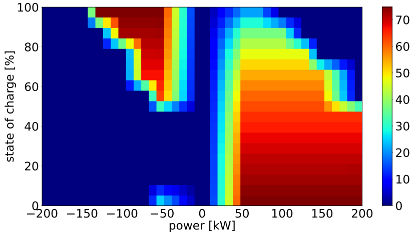
3 Generation of training data

The performance of ANN is dependent on the relation between the data used for its training and the target. Since the goal is the recognition of the physical parameters describing the behaviour of the target unit, the input timeseries must be related to / influenced by, said parameters. Unfortunately, high resolution timeseries measurements of the target unit are rarely available which forces engineers to install temporary measurement equipment which then can only provide data for a short time range.

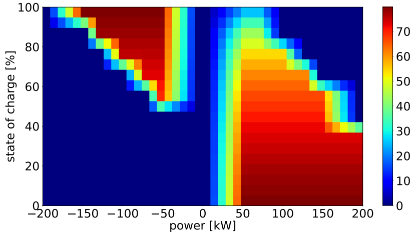
For this paper a vanadium redox flow battery storage, build as part of the research project: "Smart Region Pellworm" [1], was used as target unit. First a model for this unit type was chosen from the available literature [2]. This model was further simplified until its behaviour was defined by 6 variable parameters. In the next step plausible parameter ranges were set and discretized into steps, resulting in 125 000 possible parameter combinations. These combinations provide the labels for the training data. As input variables, timeseries for the voltage (U), current (I) and state of charge (SoC) were chosen.
In the next step an initial SoC was selected for each model configuration after which they were run through specific load profiles to generate the required timeseries for I, U and SoC. Here 3 different sets of load profiles were used to test the ANN's dependency on data that is highly correlated to the target unit. In set 1, each model configuration was given one of 4 historic measured load profiles and related initial SoC's of the target unit to create the training data. In set 2, 10 historic measured load profiles and SoC's were used and in the last, set 3, each model configuration was given a randomly generated load profile and starting SoC. This resulted in 3 separate training data sets.

4 Parameter recognition

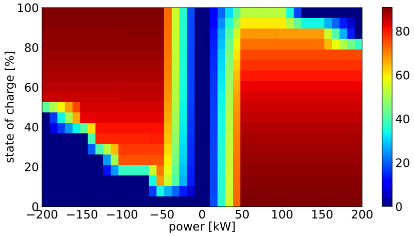
The ANN used in this work is a basic MLP [3] with 3 hidden layers. All layers used the tanh activation function except the output layer for which a linear activation function was used. The width of the layers was arranged in descending size with the intention of compressing the information contained in the input timeseries, with 3 times 288 timesteps, down to the 6 target parameters. The timeseries and labels were normalized between -1 to 1 and 0 to 1 respectively, before being used in the training. During training overfitting was limited by an early stopping function with a patience of 10 steps.
The resulting trained ANN were first tested solely on the historic data belonging to the load profiles used in the creation of their respective training data. It became apparent that the parameter values identified by the two ANN trained on historic load profiles of the target united, differed significantly form the values of the ANN trained on randomized load profiles. Especially the parameters total vanadium concentration and maximum flow speed of the electrolyte are



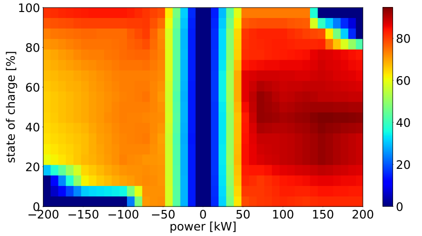
(a) Fitted by ANN trained on data based on 4 historic load profiles



(b) Fitted by ANN trained on data based on 10 historic load profiles



(c) Fitted by ANN trained on data based on random load profiles



(d) Generated based on detailed measurements of the target Battery

Figure 1: Heatmaps depicting the efficiency of the battery system models in realation to the SoC and charging (positiv) or discharging (negative) power

lower than expected in the historic data trained ANN. These parameters, in combination with the SoC, are responsible for determining how much usable vanadium ions can be pumped into the cell, which in turn defines the maximum charge and discharge power of the battery.

To allow for a better comparison between the parameter sets a second test was performed during which the ANN where each given historic data of 50 days and the resulting parameters where averaged and put into individual sets. To evaluate the accuracy of the identified parameters sets they were then used to fit 3 generic models to the target battery. The fitted models were then used to simulate a test scenario and from the resulting data the efficiency across the full range of possible operation situations was calculated.

The results are presented in the heatmaps of Figure 1 which show the efficiency across the full range of possible SoC's and all charge and discharge speeds.

The heatmap of Figure 1d was generated during experiments with the target unit and represent the real battery properties.

The heatmap in Figure 1a was generated by the model fitted with dataset 1 and shows a very early discharge power restriction, these restrictions can be seen in the heatmaps as 0 % efficiency zones. According to this model it would be impossible to discharge the battery further than roughly 50 %. Additionally, it shows an artifact near a SoC of 10 to 0 % which would indicate that a further discharge would suddenly become possible again. The charging side however is much better represented, while the charging power restriction near a full SoC is significantly overestimated, the remaining field is much closer to the reality than the discharge field.

The heatmap in Figure 1b, generated using data set 2, looks very similar to Figure 1a even though its training data was based on a larger set of load profiles. The most notable improvements are the removal of the artifact near the bottom of the discharge field and a slight widening of the possible discharge zone on the top. The charging side however shows no improvement and instead exhibits a slight expansion of the charging power restriction in the top right corner.

Figure 1c shows the results of the model fitted with data set 3, generated with the ANN using randomized training data. It produces by far the most accurate representation of the target properties. The largest improvement can be seen on the discharge side, that now, while still overestimating the discharge power restriction, shows a much more plausible range of usage. A similar improvement is visible on the charging side where the restricted zone has been significantly reduced and now closely resembles the restriction shown by the target unit. Besides the estimated power restrictions another modelling error is visualised in the heatmaps.

All 3 models exhibit a clear overestimation of the battery efficiency. This overestimation is more severe during discharge processes. This error would cause the model SoC to drift away from the actual SoC over time and thus require frequent refreshing with a measured value.

5 Conclusion

The paper proposed a data driven approach to fit generic models to the target generation or storage units. The approach was tested on the example of a vanadium redox flow battery system and further evaluated on its sensitivity regarding the required training data. The results show the best performance can be achieved using randomized load profiles and initial SoC's for the training data generation instead of historic load profiles of the target unit. This has some positive implications for the transferability of the approach on different target units of the same type. While the generation of training data using historic load profiles would have required a new set of training data and retraining of the ANN for each target unit, the randomized approach could allow the usage of one pretrained ANN for all target units of the same type, no retraining necessary. This would also justify putting more resources into the training of the ANN since it would be possible to not only create a library with generic models for each unit type but also create a library of pretrained, ready to use, ANN to fit them to the desired target. Thus significantly reducing the expertise required from users. Further research is required to determine if a single ANN can really fit any unit of a single type regardless of its scale and to test if other ANN architectures could offer higher accuracy. There is also the possibility that completely random load profiles might not be the best approach to training data generation since some of the load profiles might end up being nonsensical and their resulting data devoid of useful information for training.

References

[1] Forschungsprojekt: "Smart Region Pellworm". Gefördert durch BMWi, Förderkennzeichen: 0325498A-F

[2] C. Blanc. "Modeling of a Vanadium Redoxflow Battery Electricity Storage system". École Polytechniqe Fédérale de Lausanne, Suisse 2009.

[3] I. Goodfellow, Y. Bengio, A. Cousville “Deep Learning - das umfassende Handbuch”. mitp-Verlag, Frechen 2018.