

# BETRIEBSFÜHRUNG VON BATTERIESYSTEMEN IN INDUSTRIEBETRIEBEN MIT REINFORCEMENT LEARNING

Anna ROTHENHÄUSLER<sup>1(\*)</sup>, Arne GROSS<sup>2</sup>, Matthias KÜHNBACH<sup>1</sup>

## Einleitung

Batteriespeicher können einen substanziellen Beitrag zur Integration volatiler erneuerbarer Energien leisten. Eine Voraussetzung für die Diffusion von Batteriespeichern im Gewerbebereich ist, dass diese in ökonomisch attraktive Nutzungskonzepte eingebettet sind. Die Rentabilität von Batteriespeichern kann erhöht werden, indem verschiedene Nutzungskonzepte kombiniert werden (sog. Multi-Use/Revenue Stacking). Hierbei ist ein strategisches Energiemanagement maßgeblich.

Energiemanagementsysteme für Multi-Use-Konzepte werden in der Literatur häufig mit modellprädiktiver Regelung (MPC) umgesetzt [1, 3]. Diese Methode benötigt zum einen ein simples, jedoch präzises Modell, das die Auswirkungen einer Regelmaßnahme vorhersagt. Zum anderen sind Vorhersagen des Lastgangs, der PV-Produktion sowie der Energiepreise nötig.

Aufgrund der hohen Komplexität der modellprädiktiven Batteriesteuerung – besonders bei Betrachtung mehrerer paralleler Anwendungsfälle – stellt Reinforcement Learning (RL) eine vielversprechende Alternative zu MPC dar. RL kann hierbei unabhängig von einem Modell durch wiederholte Interaktionen mit der Umgebung komplexe Zusammenhänge erkennen und daraus eine Strategie ableiten.

In diesem Paper wird das ökonomische Potential eines RL-Algorithmus als Batteriesteuerungsstrategie für Unternehmen untersucht. Dazu wird ein RL-Modell entwickelt und mit einem Datensatz eines Industriebetriebs trainiert. Um die Performance eines RL-gesteuerten Batteriespeichers zu evaluieren, werden die Stromkosten eines solchen Betriebs denen eines ungesteuerten Betriebs ohne Speicher gegenübergestellt.

## Methodik & Case Study

In diesem Paper verwenden wir ein Soft-Actor Critic Modell (SAC) aus zwei Gründen. Bei SAC handelt es sich um einen model-free RL-Algorithmus, der in der Lage ist, auch auf kleinen Datensätzen eine gute Leistung zu erzielen [2]. Da der vorliegende Datensatz des Industriebetriebs nur die PV- und Lastdaten für das Jahr 2021 beinhaltet, ist ein effizienter RL-Algorithmus umso wichtiger. Zusätzlich dazu unterstützt SAC eine Batteriesteuerung auf einer kontinuierlichen Skala anstatt diskreter Werte, wie es bei anderen RL-Algorithmen der Fall ist.

Der RL-Algorithmus wurde auf einem Datensatz eines Industriebetriebs mit einem Jahresenergieverbrauch von etwa 5 MWh trainiert und getestet. Das Stromkosteneinsparpotential wurde anhand von zwei Use-Cases untersucht:

- 1) Energiearbitrage
- 2) Atypische Netznutzung

Das Ziel von Energiearbitrage ist die Senkung der Energiekosten. Hierfür werden die Strompreisschwankungen ausgenutzt und die bezogene Energie aus dem Netz an den entsprechenden Preis (Intraday-Preis für DE, 2021) angepasst.

Wenn Bedingungen für eine atypischen Netznutzung nach StromNEV §19 erfüllt sind, können die jährlich anfallenden Leistungskosten gesenkt werden. Hierfür muss die Lastspitze in den sogenannten Hochlastzeitfenstern im Vergleich zu den Lastspitzen außerhalb der Hochlastzeitfenster um einen gewissen Prozentsatz reduziert sein.

---

<sup>1</sup> Fraunhofer-Institut für Solare Energiesysteme, Heidenhofstr. 2 79110 Freiburg, +49 761 4588 2220, e-mail: [anna.christin.rothenhaeusler@ise.fraunhofer.de](mailto:anna.christin.rothenhaeusler@ise.fraunhofer.de), <https://www.ise.fraunhofer.de/>,

<sup>2</sup> Universität Freiburg, [arne.gross@imtek.uni-freiburg.de](mailto:arne.gross@imtek.uni-freiburg.de)

## Ergebnisse & Schlussfolgerungen

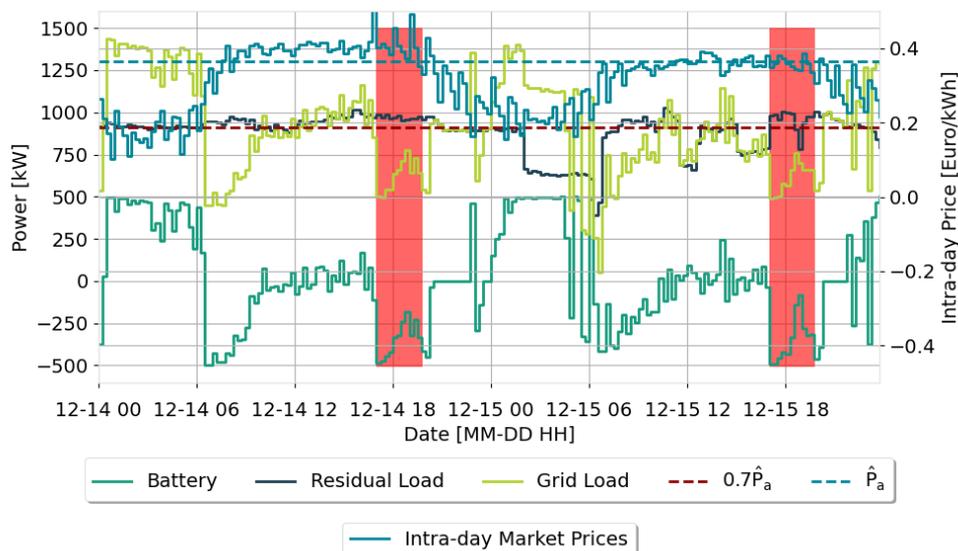


Abbildung 1: Leistung und Preis für zwei exemplarische Tage aus dem Validierungsdatensatz für den Multi-Use-Anwendungsfall Energiearbitrage/atypische Netznutzung. Die Hochlastzeitfenster sind hier durch einen roten Hintergrund veranschaulicht. Zusätzlich ist die maximale Leistungsspitze des Vorjahres  $\hat{P}_a$  dargestellt, sowie die maximale Leistungsgrenze für atypische Netznutzung  $0.7\hat{P}_a$  laut Netzbetreiber.

In Abbildung 1 ist anhand von zwei exemplarischen Tagen des Validierungsdatensatzes das aus der Batteriesteuerung resultierende Verhalten veranschaulicht. Wie die Abbildung zeigt, wird die Leistung im Fall von Netzbezug kontinuierlich unter der vom Netzbetreiber<sup>3</sup> geforderten 70% Grenze gehalten. Zusätzlich kann beobachtet werden, dass zu Niedrigpreisphasen die Batterie geladen wird und im Gegensatz dazu zu Hochpreisphasen die Batterie entladen wird, um die vom Netz bezogene Leistung zu reduzieren.

Die Kombination aus Energiearbitrage und atypischer Netznutzung über die RL-Batteriesteuerung führt zu einem jährlichen Energiekosteneinsparpotential von etwa 97.000 Euro. Dies entspricht ca. 12% der Jahresstromkosten.

Die Ergebnisse des Papiers deuten auf ein hohes Potential der Verwendung eines RL-Algorithmus als Batteriesteuerungsstrategie hin. Zusätzlich dazu bietet RL den Vorteil, dass der Algorithmus sich im Laufe der Zeit an eine sich ändernde Umgebung anpassen kann ohne explizite Intervention in den Algorithmus.

Um eine eindeutige Aussage zur Profitabilität der Anwendung von RL zur Batteriesteuerung treffen zu können, sind weitere Simulationen nötig, die Investitionskosten sowie zyklische und kalendarische Alterungsmechanismen in das Training des RL-Algorithmus berücksichtigen.

## Referenzen

- [1] Chen, C., Wang, J., Heo, Y., and Kishore, S. 2013. MPC-Based Appliance Scheduling for Residential Building Energy Management Controller. *IEEE Trans. Smart Grid* 4, 3, 1401–1410.
- [2] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. 2018. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. *International Conference on Machine Learning*, 1861–1870.
- [3] Xu, X., Li, G., and Zhang, H. 2020. Optimal Energy Management System Design Based on Dynamic Programming for Battery Electric Vehicles. *IFAC-PapersOnLine* 53, 5, 634–637.

<sup>3</sup> Im vorliegenden Papier werden Hochlastzeitfenster und Netzentgelte der BadenovaNETZE GmbH verwendet.