

Intelligentes Entscheidungsunterstützungssystem für Ladevorgänge an Stromtankstellen

Abstract. Die deutsche Bundesregierung zielt darauf ab, bis 2022 eine Million Elektroautos auf die Straßen zu bringen. Bisher erscheint dieses Ziel jedoch unerreichbar, da zahlreiche Autofahrer auf den Kauf von Elektroautos verzichten. Als Grund nennen sie dabei primär die unzureichend ausgebaute Ladeinfrastruktur, die unter anderem aus der Unrentabilität des Betriebens von Stromtankstellen resultiert. Eine Möglichkeit, die Profitabilität solcher Investments zu steigern, ist, die Verweildauer der Ladekunden zu nutzen, um die Ladevorgänge auf monetär günstige Zeitpunkte zu legen. Intelligente Decision Support Systems können die aggregierte Berücksichtigung aller relevanten Einflussfaktoren unterstützen. Bisherige Lösungsansätze aus dem Bereich Green IS setzen beispielsweise auf die Reinforcement Learning Methode Q-Learning. Aufgrund der geringen Skalierbarkeit ist sie allerdings nicht auf größere Stromtankstellen anwendbar. Um auf diese Herausforderung einzugehen, wird in dieser Arbeit ein Deep Reinforcement Learning Ansatz verfolgt. Die Evaluation in einem Realwelt-Setting zeigt, dass die Profitabilität von Stromtankstellen durch den Einsatz dieses Modells deutlich steigt.

Keywords: Decision Support Systems, Elektromobilität, Stromtankstelle, Green IS, Deep Reinforcement Learning

1 Einleitung

Die deutsche Bundesregierung hat sich als Ziel gesetzt, bis 2020 eine Million Elektroautos auf die deutschen Straßen zu bringen [1]. Die potenziellen Käufer reagieren bislang jedoch mit Zurückhaltung [2] und nennen als Hauptgrund unter anderem die schlecht ausgebaute Ladeinfrastruktur und, eng damit verbunden, die Angst auf der Strecke stehen zu bleiben [3].

Mittlerweile wurde das ursprüngliche eine Million-Ziel auf das Jahr 2022 verschoben. Der Fokus der Bundesregierung liegt zunächst verstärkt darauf, die Ladeinfrastruktur auszubauen und somit verbesserte Bedingungen für den Einsatz von Elektroautos in Deutschland zu schaffen [4]. In diesem Rahmen fördern Bund und Länder den Bau von Stromtankstellen [5] sowie die Durchführung innovativer Großprojekte, die eine Ladeversorgung für Elektrofahrzeuge durch eine nachhaltige regionale Stromerzeugung garantieren [6].

Die staatliche Förderung für den Ausbau der Ladeinfrastruktur ist notwendig, um Privatinvestitionen in Stromtankstellen attraktiver zu gestalten [7]. Denn bisher ist das Betreiben einer Stromtankstelle aufgrund der Vielzahl an zu berücksichtigenden stochastischen Einflussfaktoren nicht profitabel [8]. So ist es beispielsweise durch das schwankende Kundenaufkommen und die volatile erneuerbare Stromerzeugung schwierig, alle Kunden zu jedem Zeitpunkt ausschließlich mit kostengünstigem, selbst erzeugtem Strom zu bedienen. Durch den ebenfalls variierenden Börsenstrompreis können daher hohe Strombezugskosten entstehen. Um diese Kosten möglichst gering zu halten, ist es für die Betreiber von Stromtankstellen erstrebenswert, viel Strom abzugeben, wenn er selbst gewonnen wird oder der externe Bezugspreis gering ist. Um diese Situation schaffen zu können, benötigen die Betreiber Flexibilität in der Entscheidung, zu welchem Zeitpunkt die Fahrzeuge der Kunden geladen werden. Bei der Bestimmung des optimalen Ladezeitpunktes spielen weitere Faktoren, wie unter anderem die Anforderungen der Kunden sowie die Vermeidung von Lastspitzen, eine Rolle [9].

Bei der Vielzahl und Komplexität der zu berücksichtigenden Einflussfaktoren können intelligente IT-Systeme, sogenannte *Decision Support Systems* (DSS), als Entscheidungsunterstützung für Stromtankstellenbetreiber dienen. In dem Bereich Green IS existieren analoge Problemstellungen für intelligente Stromnetze, sogenannte *smart grids*. Mit Hilfe von Reinforcement Learning Methoden werden hier tragfähige Lösungen entwickelt [10]. Auch einige existierende Forschungsansätze mit Fokus auf Privathaushalte und Gewerbe setzen Reinforcement Learning Methoden, wie *Q-Learning* ein. Für Privathaushalte ermöglichen DSS es beispielsweise, Elektroautos zu günstigen Zeitpunkten zu laden, um die Stromkosten zu reduzieren und gleichzeitig eine gleichmäßige Stromnutzung ohne Lastspitzen zu erzielen [11, 12]. Für eine optimale Ladesteuerung bei Stromtankstellen ist analog dazu ebenfalls die integrierte Betrachtung beider Ziele notwendig [13].

Ein Ansatz, der die Minimierung externer Stromkosten verfolgt [14], überträgt ein für Privathaushalte konzipiertes Modell [12] auf ein vereinfacht dargestelltes Szenario einer Stromtankstelle. Es ist darauf ausgelegt, für eine gegebene Nachfrage stündlich den Strom aus eigenen und externen Quellen so zu verteilen, dass der Gewinn über den gesamten Tag maximiert wird. Dabei werden die Ladekunden durch eine präferierte

Ladegeschwindigkeit, den Ladebedarf sowie das zum Laden zur Verfügung stehende Zeitfenster charakterisiert. Zur Deckung der Ladebedarfe stehen lokale erneuerbare Stromquellen und der externe Strommarkt zur Verfügung. Mit Hilfe von Q-Learning wird für dieses Szenario bereits eine erhebliche Gewinnsteigerung erzielt.

Jedoch basiert dieser Ansatz auf einer vereinfachten Darstellung der Realität, was zu einer reduzierten Anzahl an möglichen zu berücksichtigenden Zuständen führt [14]. Werden die Parameter allerdings auf ein realitätsnäheres Szenario erweitert, steigt der Zustandsraum exponentiell [15]. Diese Steigerung des Zustandsraumes resultiert in einem enormen Rechenaufwand und einer geringen Anpassbarkeit des Modells, was eine Problemlösung durch klassisches Reinforcement Learning mit Q-Learning ineffizient macht. Doch mit der fortschreitenden Entwicklung der Elektromobilität [16] ist von einer steigenden Kundenanzahl auszugehen, was mit einer erhöhten Komplexität an Stromtankstellen einhergeht. Daher sind skalierbare und anpassbare Ansätze für DSS bei Stromtankstellen zwingend erforderlich. Die Forschungsfrage, die dieser Arbeit zugrunde liegt, lautet daher:

Wie kann ein realitätsnäheres skalierbares Modell zur Steuerung der Ladevorgänge, die Rentabilität von Stromtankstellen steigern?

Eng damit verbunden ist die Frage nach dem richtigen Umgang mit großen Zustandsräumen bei Entscheidungsproblemen. Dafür wird in dem Bereich smart grid erfolgreich *Deep Reinforcement Learning (DRL)* angewendet [10]. Entsprechend stellt diese Arbeit einen Ansatz zur Realitätsannäherung bei Stromtankstellen durch DRL vor. Die Arbeit orientiert sich bei der Modellbewertung an der momentan größten Stromtankstelle der Welt (GSW), die sich gerade im Bau befindet und 144 Ladestationen umfasst [6].

Im Weiteren gibt der nachfolgende Abschnitt einen Überblick über die Modellstruktur sowie die relevanten Komponenten. Aufbauend wird in der Evaluation die Simulation und deren Ergebnisse vorgestellt. In Kapitel 4 werden diese Ergebnisse bewertet und das Modell evaluiert. Kapitel 5 fasst abschließend die Limitationen der Arbeit zusammen und gibt einen Ausblick über weitere mögliche Forschungsmöglichkeiten.

2 Modell

Durch das Zusammenspiel vieler stochastischer Einflussfaktoren, wie beispielsweise dem Kundenaufkommen, der Verweildauer der Kunden an der Tankstelle oder dem Börsenstrompreis, ist ein rentables Management von Stromflüssen an einer Stromtankstelle ein äußerst komplexes Gefüge. Analog zu [14] ist auch in dieser Arbeit die Anwendung eines Machine Learning Ansatzes anderen statistischen Ansätzen aufgrund der Komplexität bevorzugt worden. Zudem sind in dieser Arbeit, im Gegensatz zu [14], realitätsnähere Parameter gewählt worden, was die Komplexität verstärkt.

Das Ziel dieser Arbeit ist, die genannte Entscheidungskomplexität zu reduzieren, indem ein Modell den Tankstellenbetreibern zu jedem betrachteten Zeitpunkt Ladeempfehlungen bereitstellt. Die ankommenden Ladekunden sind dabei durch ihre Lade-

bedarfe charakterisiert. Diese werden im Modell durch den *State of Charge (SOC)* abgebildet, der dem prozentualen Ladestand des Akkus entspricht. Außerdem ist für jeden Kunden eine individuelle *Time to Leave (TTL)* bekannt, die definiert, zu welchem Zeitpunkt der Kunde die Tankstelle spätestens wieder verlässt. Um für jeden der zahlreichen möglichen Zustände die voraussichtlich gewinnbringendste Aktion zu bestimmen, wird in der Literatur unter anderem die Reinforcement Learning Methode *Q-Learning* eingesetzt [14]. Diese Methode eignet sich jedoch nur stark bedingt, da sie jedes Zustand-Aktions-Paar testet, wodurch ein Modell mit starren Entscheidungsvorlagen entsteht [15]. Bei kleinen Veränderungen in der Umgebung ist ein hoher Aufwand notwendig, um das Modell entsprechend darauf anzupassen. Eine Alternative bietet DRL, das die Q-Matrix durch ein oder mehrere neuronale Netze ersetzt [15]. Neuronale Netze können den Zusammenhang zwischen Ein- und Ausgabe erlernen, wodurch sie auch in der Lage sind, Q-Werte von bisher unbekanntem Zustands-Aktions-Paaren fundiert zu approximieren.

Orientiert am Framework QCON [15] wird in diesem Modell ein DRL Ansatz verfolgt und für jede mögliche Aktion ein künstliches neuronales Netz eingesetzt. Eine Aktion, die zu einem Zeitpunkt ausgeführt wird, entspricht dabei einer Ladeentscheidung für die einzelnen Ladesäulen. Pro Ladesäule stehen dabei drei mögliche Handlungsalternativen zur Verfügung: Das Fahrzeug an der Ladesäule wird nicht (I), zu zehn Prozent des Akkuvolumens (II) oder vollständig geladen (III).

Die Berechnungen zur Approximation der Q-Werte für verschiedene Handlungsalternativen können sehr unterschiedlich sein [15], da beispielsweise ein Einflussfaktor für Alternative (I) eine große Rolle spielt, während er für Alternative (III) irrelevant ist. Demnach müsste ein einziges neuronales Netz die Fähigkeit erlernen, exakt zwischen den verschiedenen Handlungsalternativen zu differenzieren, um eine gute Vorhersagequalität zu erzielen. Dafür sind entsprechend viele Trainingsdurchläufe notwendig [15]. Dagegen wird bei der Verwendung von nur einem neuronalen Netz pro Alternative jedes Netz auf die spezifischen Zusammenhänge der Einflussfaktoren trainiert und kann somit bei weniger Trainingsdurchläufen eine höhere Prognosegüte erzielen. Aus diesem Grund wird in diesem Modell der zuletzt genannte Ansatz verfolgt.

Vorausgehende Modellannahmen.

1. Zu Beginn jeder Iteration t können neue Ladekunden an der Tankstelle ankommen. Ein ankommender Kunde k begibt sich auf einen Ladeplatz und teilt seinen aktuellen SOC sowie seine Verweildauer, die anfängliche TTL $aTTL$, mit, die er strikt einhält: $TTL_k(t) = aTTL_k$
2. Der Ablauf ist iterativ und nicht deterministisch, wobei in jeder Iteration t über die Lademenge der Autos entschieden wird. Nach einer abgeschlossenen Iteration werden, abhängig von der gewählten Aktion, die SOCs der geladenen Elektroautos neu bestimmt und die TTL der bestehenden Kunden $k \in K(t)$ um eine Stunde reduziert: $TTL_k(t+1) = TTL_k(t) - 1$. Gilt für einen Kunden k $TTL_k(t+1) = 0$, dann verlässt er die Tankstelle, sodass $k \notin K(t+1)$.
3. Die Stromtankstelle stellt unbegrenzt viele Ladeplätze inklusive Ladesäulen für die Kunden bereit, allerdings können maximal n Autos zeitgleich Strom beziehen.
4. Alle Elektroautos gleichen sich und besitzen einen Akku mit derselben Kapazität.

Umwelt. Die Umwelt stellt in diesem Modell die Stromtankstelle und deren Umgebung dar. Dazu gehören Kunden, deren Ankunft als stochastisch angenommen und dementsprechend modelliert wird. Weiterhin berücksichtigt das Modell, wie viel Strom $e(t)$ die Tankstelle zum Zeitpunkt t lokal erzeugt. Die bei der Stromnutzung entstehenden Kosten, beispielsweise für den Transport von der lokalen Stromquelle zur Tankstelle oder für die Wartung, werden vernachlässigt. Zwar ist es nicht möglich, diesen Strom für einen späteren Zeitpunkt zu speichern oder ihn auf dem Strommarkt zu handeln, jedoch besteht im Modell die Möglichkeit, zum Zeitpunkt t zusätzliche Energie aus dem öffentlichen Stromnetz zu einem dynamischen Preis $p(t)$ zu beziehen.

Zustand. Ein Zustand $s \in S$ ist definiert durch einen Zeitpunkt t , die anwesenden Ladekunden $K(t)$, die verfügbare erneuerbare Energie $e(t)$ und den Strompreis pro zugekaufte Einheit $p(t)$. Dies wird wie folgt dargestellt.

$$s(t) = [t, K(t), e(t), p(t)] \quad (1)$$

Hervorzuheben ist, dass der Zeitpunkt t sowohl die Tagesstunde als auch den Wochentag berücksichtigt. Weiterhin ist jeder aktive Kunde $k \in K(t)$ durch die Attribute SOC_k , TTL_k und $aTTL_k$ definiert. Dabei gibt SOC_k den Ladestand von k zum Zeitpunkt t in zehn Prozentwert-Schritten an. Aus diesem Attribut wird abgeleitet, wie viele Kilowattstunden bis zu einem vollständig geladenen Auto benötigt werden. TTL_k beschreibt die verbleibende Verweildauer eines Kunden an der Tankstelle. Während sich diese beiden Attribute mit jeder Iteration ändern, gibt die statische $aTTL_k$ die zu Beginn festgelegte Verweildauer des Kunden an. Diese bildet die Grundlage zur Berechnung des Ladepreises des Kunden $b: \mathbb{N} \rightarrow \mathbb{R}$ (Gl. 6).

Aktion. Eine Aktion entspricht in diesem Modell der zu einem Zeitpunkt t von einer Ladesäule i abgegebenen Strommenge $l_i(t) \in L$. Der SOC des Elektroautos, das an der entsprechenden Ladesäule steht, kann sich entweder um 0%, 10% oder auf 100% erhöhen. Demzufolge kann jede Ladung $l_i(t)$ nur einen der Werte $\{0; 10; 100 - SOC_i\}$ annehmen. Aufgrund der Restriktion, dass maximal n Säulen gleichzeitig Strom beziehen können, gilt: $|l_i(t) \in L: l_i(t) > 0| \leq n$. Entsprechend wird eine Aktion als Kombination der Lademöglichkeiten für alle n zulässigen Ladesäulen betrachtet, woraus sich 3^n unterschiedliche Aktionen ergeben. Nach der Ausführung einer Aktion berechnet sich der Ladezustand eines Elektroautos wie folgt: $SOC_k(t + 1) = SOC_k(t) + l_k(t)$.

Belohnung. Die Belohnung entspricht dem Gewinn ϕ_t (Gl. 2), der aus der Durchführung von Aktionen resultiert. Die Einnahmen $w(t)$ berechnen sich dabei aus der Strommenge, die den Elektroautos zugeführt wird. Die Bepreisung des Stroms erfolgt auf Basis der Zahlungsbereitschaft der Kunden (Gl. 3). Die Ausgaben $c(t)$ entstehen im Modell (Gl. 4), indem Strom vom öffentlichen Netz hinzugekauft wird, falls die abgegebene Strommenge die selbst erzeugte Strommenge $e(t)$ übersteigt und somit die Differenz $\Delta(t) = \sum_{i=1}^{|K|} l_i(t) - e(t)$ positiv ist. Für dieses Delta fallen pro kWh Kosten in Höhe des aktuellen Strompreises $p(t)$ an. Zusätzlich wird überverhältnismäßiger ex-

terner Strombezug mehrstufig bestraft, was eine Reduzierung der Lastspitzen ermöglicht. Eine Strafgrenze g_j gibt an, ab wann ein Malus m_j pro zusätzlich bezogener kWh fällig ist. Weiterhin bestimmt die Kontrollvariable u_j , ob eine Strafgrenze überschritten wurde $\Delta(t) > g_j$ und somit eine Strafe zu zahlen ist ($u_j = 1$) oder nicht ($u_j = 0$).

$$\phi_t = w(t) - c(t) \quad (2)$$

$$w(t) = \sum_{i=1}^{|K|} l_i(t) b(aTTL_i) \quad (3)$$

$$c(t) = \begin{cases} \Delta(t)p(t) + \sum_j (\Delta(t) - g_j)m_j u_j & \text{falls } \Delta(t) \leq 0 \\ 0 & \text{sonst} \end{cases} \quad (4)$$

Aktionsauswahl. Um den Gewinn der Tankstelle zu steigern, berücksichtigt das Modell bei der Auswahl einer Aktion sowohl den direkt resultierenden als auch den zukünftig erzielbaren Gewinn. Aufgrund der hohen Komplexität ist die zu maximierende Zielfunktion $z(t)$ dabei auf fünf Zeitschritte beschränkt (Gl. 5). Neben der Gewinnorientierung wird bei der Auswahl der durchführenden Aktion auch auf die Garantie eines Mindestmaßes an Service geachtet. Demnach ist berücksichtigt, dass der SOC eines Kunden, der die Tankstelle verlässt, mindestens 30% beträgt und um mindestens 10% während seiner Verweildauer gestiegen ist.

$$z(t) = \sum_{i=0}^5 0,9^i \phi(t + i) \quad (5)$$

Zur Bestimmung der voraussichtlich besten Aktion prognostizieren die neuronalen Netze für jede Aktion zunächst den zu erwartenden Gewinn über die nächsten fünf Iterationen. Ein anschließender Vergleich dieser Ergebnisse gibt Auskunft darüber, welche Aktion den größten Gewinn verspricht und durchgeführt werden sollte. Nach fünf Iterationen ist es dann möglich, das neuronale Netz der durchgeführten Aktion mittels dem nun eindeutig bestimmten $z(t)$ rückwirkend zu trainieren und somit für zukünftige Prognosen zu verbessern.

3 Evaluation

Ein DSS, welches das vorgestellte Modell abbildet und auf die GSW angepasst ist, dient der Evaluation des Modells. Das DSS wird in einer 52-Wochen Simulation getestet und die Ergebnisse, mit denen einer simplen und einer zufälligen Ladestrategie verglichen. Die Simulation und das DSS sind in Python (Version 3.6.2) umgesetzt. Dazu werden insbesondere bekannte Unterstützungspakete für Datenstrukturen, *pandas* (0.24.2), und mathematische Operationen, *numpy* (1.13.1), verwendet. Weiterhin sind die neuronalen Netze mithilfe des Machine Learning Paketes, *sklearn* (0.19.0), programmiert. Dazu werden *fully connected Multilayer Perceptrons* mit der Schichtenarchitektur 14-100-100-1 sowie dem stochastischen, gradientenbasierten Lernalgorithmus *adam* [17] verwendet. Dieser erlaubt kontinuierliches Online Learning, welches durch Mini-Batches der Größe 3 umgesetzt wird.

3.1 Fallbeispiel: Größte Stromtankstelle der Welt

Mit der sich aktuell in Bau befindlichen Stromtankstelle entsteht an der A8 bei Augsburg für Elektrofahrzeuge der größte und innovativste Schnellladestandort der Welt [6]. Aus dem Projektförderungsantrag [18] geht die nachfolgende Beschreibung hervor.

Die Stromtankstelle soll den Kunden 144 Ladesäulen sowie Rast-, Einkaufs- und Parkmöglichkeiten bieten. Integrierte Photovoltaikanlagen erzeugen lokal Strom, der bei maximaler Erzeugung circa 63 der 144 Ladestationen gleichzeitig versorgen kann. Weiterhin können umliegende Unternehmen und private Haushalte von der GSW Strom beziehen oder ihre überschüssige erneuerbare Energie ins lokale Stromnetz einspeisen. Ein weiterer Stromabnehmer ist die firmeneigene Produktionshalle für Fahrzeugausstattungen. Die vorhandene Energieinfrastruktur soll dabei so gesteuert werden, dass die Einspeisung aus den Photovoltaikanlagen voll genutzt werden kann und nicht abgeriegelt werden muss.

3.2 Simulation der Tankstelle und der Ladevorgänge

In der Simulation entspricht ein Zeitschritt einer Stunde. Durch die Differenzierung der Wochentage im Modell ergeben sich $24 \cdot 7 = 168$ unterschiedliche Zeitpunkte.

Stromtankstelle. Im Sinne der Generalisierbarkeit bildet die Simulation keine weiteren Stromabnehmer neben der Tankstelle und keine weiteren Stromquellen abseits der Photovoltaikanlage und dem Strommarkt ab. Weiterhin sind nur drei, anstelle von 144 Ladesäulen vorhanden, um die Ladekunden zu bedienen. Im entsprechenden Verhältnis ist die produzierte Strommenge adjustiert. In Anlehnung an die Angaben zur GSW können mittels Photovoltaikanlagen in einer Stunde bei maximaler Erzeugung 1,31 Elektroautos geladen werden. Bei einer angenommenen Akkukapazität von 50 kWh pro Fahrzeug entspricht das einer Leistung von 65,5 kW. Zur stochastischen Approximation des erzeugten Solarstromes pro Stunde werden historische Wetterdaten¹ für die Sonneneinstrahlung in der Region der GSW (NUTS-2 Region 27) verwendet.

Strompreis. Der Strombezugspreis wird auf Basis von historischen Daten² des deutschen Strommarktes durch eine Normalverteilung pro Wochentag und Tagesstunde für jeden Zeitpunkt stochastisch approximiert. Aufgrund der großen Preisschwankungen ist der Strombezugspreis mittels der 2- σ -Methode auf das Intervall $[\mu - 2\sigma; \mu + 2\sigma]$ beschränkt. Zusätzlich zum stochastischen Strombezugspreis sind Netzentgelte in Höhe von 10 Cent pro kWh berücksichtigt. Weiterhin sind zur Minimierung der Lastspitzen Strafkosten auf zwei Ebenen instanziiert; 3 Cent pro zusätzlicher kWh ab einem Bezug von mindestens 50 kWh und weitere 2 Cent pro kWh ab einem Bezug von 100 kWh.

¹ Datenquelle: Stündliche Wetterdaten von 1980 bis 2016 der NUTS-2 Region 27 [19]

² Datenquelle: Deutscher Strommarkt von 2005 bis 2018 [20]

Kunden. In der Simulation sind Ladekunden durch unterschiedliche Kundengruppen abgebildet. In Abstimmung mit den Projektmitarbeitern der GSW werden sechs Gruppen berücksichtigt. Diese unterscheiden sich in der erwarteten Ankunftszeit, dem SOC und der TTL. Für den Tankstellenbetreiber rücken zwei relevante Kundengruppen in den Fokus. Erstere sind die *Mitarbeiter*, die während der Arbeitszeit ihre privaten Elektroautos laden können und für gewöhnlich eine lange Verweildauer aufweisen. Außerdem profitiert die *Firmenflotte*, welche größtenteils von den Vertriebsmitarbeitern genutzt wird, vorwiegend abends an den Ladepunkten ankommen und über Nacht geladen werden können. Die Gruppe *Vorbeifahrer Business* fasst die Ladekunden zusammen, die tagsüber berufsbedingt mit dem Auto unterwegs sind und in der Ladezeit beispielsweise die vorhandenen Besprechungsräume nutzen. Wochentags, insbesondere morgens und abends, am Freitag auch mittags, sind vermehrt *Pendler* unterwegs. Analog zu den Geschäftsleuten ist auch hier von einer geringeren TTL auszugehen. Da die GSW außerdem über einen großen Einkaufsbereich verfügt, sind über den Tag verteilt und am Wochenende ebenfalls *Einkäufer* anzutreffen, deren TTL je nach Art und Dauer des Einkaufs variiert. Am Wochenende sind *Familien und Reisende* unterwegs, die auf dem Weg zu ihrem Reiseziel eine Tankmöglichkeit an der A8 suchen.

Die Anzahl ankommender Ladekunden zu jedem Zeitpunkt ist normalverteilt. Durch eine Konfigurationsdatei ist der Parameter für jede einzelne Stunde in der Woche an veränderte Umweltbedingungen anpassbar. Die Simulation berücksichtigt zudem, wie wahrscheinlich es ist, dass ein Kunde aus einer bestimmten Kundengruppe stammt. Die Werte für SOC und TTL werden über gruppenspezifisch parametrisierte Gleichverteilungen stochastisch bestimmt.

Zahlungsbereitschaft. Die Bepreisung des Ladestromes richtet sich nach der durchschnittlichen Leistung. Entsprechend zahlen Kunden mit einer geringeren $aTTL$ einen höheren Preis pro kWh als Kunden, die länger an der GSW verweilen und der Tankstelle somit ein größeres Zeitfenster zur gewinnsteigernden Koordination der Stromflüsse bieten. Da sich beispielsweise die Flexibilität bei der Erhöhung der TTL von einer auf zwei Stunden verdoppelt, allerdings von zehn auf elf Stunden lediglich um 10% erhöht, verläuft die Nutzenfunktion nicht linear. Daran angepasst ist die Bepreisung des Ladestroms (Gl. 6) simuliert.

$$b(aTTL) = 60 * e^{-0,06 aTTL} \quad (6)$$

3.3 Ergebnisse

Im Lernprozess werden 500.000 Stunden bei einer konstanten Lernrate (hier: 0,001) und einer steigenden Exploitation Rate simuliert. Nach etwa 200.000 Stunden liegt die durchschnittliche Abweichung der Prognose der Q-Werte gegenüber dem tatsächlichen Wert bei etwa 50%. Nach etwa 400.000 Stunden ist zu erkennen, dass diese Abweichung gegen 45% konvergiert. Dennoch zeigen die Ergebnisse, dass der vorgestellte Ansatz einer zufälligen und einer simplen Ladestrategie überlegen ist. Die nachfolgende Tabelle veranschaulicht die Gegenüberstellung dieser drei Ladestrategien. Dazu werden dreimal 52 Wochen simuliert, in welchen der Agent, die Stromtankstelle, die

ankommenden Autos nach drei unterschiedlichen Ladestrategien (*Zukunft 5*, *Zufall*, *Simpel*) bedient. Mit *Zukunft 5* empfiehlt das DSS jeweils die Aktion, die gemäß den neuronalen Netzen über die nächsten fünf Stunden den höchsten Gesamtgewinn verspricht. Die Strategie *Zufall* bezeichnet die Auswahl einer zufälligen, aber zulässigen Aktion. *Simpel* dagegen beschreibt, dass die Elektroautos primär mit der lokal erzeugten Solarenergie geladen werden. Sollte der externe Strompreis jedoch geringer sein als im Erwartungswert, oder es aufgrund von Restriktionen nötig sein, wird externer Strom hinzugekauft.

Anhand dieser 52-Wochen-Simulation wird deutlich, dass die Strategie *Simpel* zwar etwas besser als *Zufall* ist, *Zukunft 5* jedoch den besten Ansatz im Hinblick auf die Gewinnsteigerung darstellt. Sie erzielt im Vergleich zu *Simpel* im Schnitt pro Tag ~25% und zu *Zufall* ~35% mehr Gewinn. Weiterhin zeigt eine Sensitivitätsanalyse, dass neuronale Netze, die 3, 5, 6, 7 oder 8 zukünftige Stunden in der Prognose berücksichtigen, im Schnitt bessere Ergebnisse erzielen als der Ansatz *Simpel*. Lediglich *Zukunft 4* unterliegt dieser sehr simplen Strategie marginal.

Tabelle 1. Durchschnittlicher Tagesgewinn in 52 Wochen in Euro

Strategie		Durchschnitt	Wochentag						
			Mo	Di	Mi	Do	Fr	Sa	So
Evaluation	Zukunft 5	2.419	2.373	2.420	2.408	2.407	2.625	2.367	2.331
	Simpel	1.930	1.791	1.721	1.734	1.743	2.061	2.175	2.284
	Zufall	1.787	1.706	1.674	1.697	1.773	2.003	1.789	1.869
Sensitivitätstest	Zukunft 3	2.316	2.355	2.338	2.296	2.322	2.576	2.126	2.199
	Zukunft 4	1.905	1.828	1.840	1.730	1.866	2.041	2.016	2.017
	Zukunft 6	2.133	2.017	2.059	2.082	2.100	2.274	2.199	2.197
	Zukunft 7	1.964	1.873	1.952	1.958	1.916	2.152	1.978	1.920
	Zukunft 8	2.249	2.147	2.320	2.200	2.187	2.450	2.197	2.244

4 Diskussion

In dieser Arbeit wird ein intelligentes DSS für das Management von Stromflüssen vorgestellt, welches auf einem DRL Ansatz basiert. Das zugrundeliegende Modell baut auf existierenden Ansätzen aus der Literatur [14] auf und überführt diese mittels des Frameworks QCON in eine skalierbare Lösung. Als zusätzliche Erweiterung sind Zeitschritte modelliert, welche auch im DSS weitergeführt werden können. Somit kann sich das DSS kontinuierlich verbessern und sich an verändernde Umwelteinflüsse anpassen. Dadurch ist es weniger von der Ausgangsinanziierung abhängig und kann als online Learning System angesehen werden. Zudem ist die Implementierung des Modells stark konfigurierbar, wodurch beispielsweise Schranken zu Lastspitzenvermeidung oder die zu erwartende Anzahl und Art ankommender Kunden pro Zeitschritt berücksichtigt

werden können. Weiterhin ist es möglich, in diesen Zeitschritten zwischen der Tageszeit und dem Wochentag zu unterscheiden.

Zur Evaluation des Modells dient ein DSS, das die Logik des entwickelten Modells abbildet und an die Bedingungen der GSW angepasst ist. Die Evaluationsergebnisse einer Simulation zeigen, dass der Gewinn an der simulierten Stromtankstelle durch den Einsatz des intelligenten DSS, verglichen mit der Anwendung einer simplen und zufälligen Ladestrategie, im Schnitt um ~25% und um ~35% steigt. Die anschließende Sensitivitätsanalyse bezüglich der betrachteten Zukunftsperioden zeigt weitere interessante Beobachtungen. So erzielt die Berücksichtigung des Gewinns der nächsten 4 oder 7 Stunden durchschnittlich deutlich schlechtere Resultate als die Betrachtung von 3 oder 5 aufeinanderfolgenden Stunden. Einem intuitiven Gedanken folgend wäre jedoch davon auszugehen, dass die Berücksichtigung mehrerer Stunden zu besseren Resultaten führt. Außerdem fällt auf, dass die Unterschiede unter der Woche höher ausfallen als am Wochenende. Eine mögliche Erklärung hierfür liefert die Simulation der ankommenden Kunden. Demnach erreichen am Wochenende vorwiegend zwei und unter der Woche vier unterschiedliche Kundengruppen die Tankstelle.

Verglichen mit bisherigen Ergebnissen aus der Wissenschaft fällt die erzielte Gewinnsteigerung relativ gering aus. So kann durch einen reinen Reinforcement Learning Ansatz bereits eine Steigerung zwischen 40% und 80%, im Vergleich zu einer zufälligen Strategie, erzielt werden [14]. Ein Grund hierfür kann die deutlich komplexere und realitätsnähere Simulation sein, die in dieser Arbeit durchgeführt wird. Dazu zählt unter anderem, dass jeder Kunde eine Ladegarantie erhält und somit kein Kunde die Tankstelle verlässt, ohne dass das Fahrzeug geladen wurde. Durch die adaptive Sortierung wird dabei für jede Iteration neu bestimmt, welche Fahrzeuge geladen werden können. Als weitere Realitätsannäherung berücksichtigt die Simulation die stündlich erzeugte Strommenge im Solarpark sowie den stündlichen Preis für extern bezogenen Strom auf Basis historischer Daten. Das Modell bezieht somit zusätzliche Einflussfaktoren ein, die bei Verwendung globaler Durchschnitte vernachlässigt werden. Zudem orientiert sich das Modell bei der Bepreisung des von den Kunden geladenen Stroms an der Verweildauer, der *aTTL*, anstelle der Lademenge in Form des *SOC*. Denn je höher die *aTTL* eines Kunden ist, desto mehr Zeit steht dem Betreiber der Stromtankstelle für den Ladevorgang zur Verfügung. Diese zusätzliche Flexibilität kann der Betreiber durch das Ausnutzen günstiger Wetterbedingungen und niedriger Strompreise in einen monetären Zugewinn umwandeln.

5 Limitationen und Ausblick

Der in dieser Arbeit entwickelte Lösungsansatz dient dem nachhaltigen und intelligenten Management von Stromtankstellen sowie, durch die Bereitstellung entsprechender Ladeempfehlungen, der Komplexitätsreduktion für die Betreiber. Das Modell stellt dabei eine realitätsnähere Erweiterung zu bestehenden Ansätzen dar.

Das Modell und die der Evaluation zugrundeliegende Implementierung stellen die Realwelt dennoch durch einige Annahmen vereinfacht dar, weshalb Limitationen existieren. Diese Limitationen sollten in anknüpfenden Forschungsarbeiten adressiert werden.

Beispielsweise setzt das Modell voraus, dass die Kunden die aTTL strikt einhalten, wodurch ein vorzeitiges Verlassen der Stromtankstelle nicht vorgesehen ist. Zudem beschränkt sich das Modell darauf, dass alle Kunden ein baugleiches Elektroauto mit derselben Akkukapazität besitzen. Somit berücksichtigt es nicht die Diversität der Elektroautos auf dem Markt mit unterschiedlichen Akkukapazitäten.

Der Fokus der Evaluation liegt auf der GSW, sodass unterschiedliche Konfigurationen, wie beispielsweise andere Kundengruppen, nicht validiert wurden. Zudem sind zunächst lediglich drei Ladepunkte implementiert, die jeweils eine Maximalleistung von 50 kW erzielen. Wird die Anzahl der Ladepunkte erhöht, so steigt damit zwar die Komplexität nur marginal, jedoch erhöht sich Anzahl benötigter neuronaler Netze deutlich. Eine mögliche Lösung bietet die Zusammenfassung mehrerer Netze zu einem umfangreicheren Netz, wobei dabei zu beachten ist, inwiefern die Aktionen gegenläufig zueinander sind. Bei den herangezogenen Stromdaten bleiben Saisonalitäten, die Auswirkungen auf die Intensität der Sonneneinstrahlung und die Höhe des Strompreises haben können, unberücksichtigt. Des Weiteren sind die Kundengruppen und deren Ankunftszeiten, die für die Implementierung angenommen werden, als subjektiv zu bewerten. Realdaten, wie beispielsweise das Verkehrsaufkommen auf der A8 oder speziell von Elektroautos, könnten die Ergebnisse weiter verbessern.

Der Umfang, der für die Evaluation des Modells durchgeführten Sensitivitätsanalyse sollte in einem nächsten Schritt erweitert werden. Zudem sollten die erzielten absoluten Ergebnisse in Form des Gewinns mit denen weiterer Steuerungsstrategien verglichen werden, um die Aussagekraft zu erhöhen. Gleichzeitig können so weitere Stärken und Schwächen des Modells aufgezeigt werden, welche wichtige Erkenntnisse für mögliche Erweiterung liefern könnten. Weitere Aufmerksamkeit verdient außerdem das abweichende Ergebnis der Sensitivitätsanalyse für die Gewinnprognose bei vier betrachteten Perioden in der Zukunft.

Literaturverzeichnis

1. Bundesregierung: Aufgabe des Elektrofahrzeuge-Ziels der Bundesregierung (Drucksache 18/12670). Drucksache, <http://dip21.bundestag.de/dip21/btd/18/130/1813034.pdf> (2017)
2. Ahlswede, A.: Anzahl der Elektroautos in Deutschland von 2006 bis 2019, <https://de.statista.com/statistik/daten/studie/265995/umfrage/anzahl-der-elektroautos-in-deutschland/>
3. Ahlswede, A.: Was spricht aus Ihrer Sicht gegen die Anschaffung eines Elektroautos?, <https://de.statista.com/statistik/daten/studie/164474/umfrage/top-10-gruende-fuer-die-ablehnung-von-elektroautos/>
4. Presse- und Informationsamt der Bundesregierung: Rede von Bundeskanzlerin Merkel beim Symposium „10 Jahre Elektromobilität – Zukunft wird Gegenwart“ der Mennekes GmbH & Co. KG am 27. September 2018 (2018)

5. Presse- und Informationsamt der Bundesregierung: Neue Kraftstoffe und Antriebe, <https://www.bundesregierung.de/breg-de/themen/neue-kraftstoffe-und-antriebe-994216>
6. Sortimo Innovationspark Zusmarshausen. Zukunftsweisendes Projekt wird umgesetzt. Zusmarshausen (2017)
7. Bundesministerium für Verkehr und digitale Infrastruktur: Förderrichtlinie Ladeinfrastruktur für Elektrofahrzeuge, <https://www.bmvi.de/SharedDocs/DE/Artikel/G/foerderrichtlinie-ladeinfrastruktur-elektrofahrzeuge.html>
8. Vetter, P.: Was der Strom für ein Elektroauto pro Jahr kosten wird, <https://www.welt.de/wirtschaft/article172139579/Elektroautos-Ladesaeulen-kosten-mindestens-360-Milliarden-Dollar-bis-2025.html>
9. Uddin, M., Romlie, M.F., Abdullah, M.F., Halim, S.A., Bakar, A.H.A., Kwang, T.C.: A review on peak load shaving strategies. *Renewable and Sustainable Energy Reviews*, 3323–3332 (2018)
10. Glavic, M., Fonteneau, R., Ernst, D.: Reinforcement Learning for Electric Power System Decision and Control: Past Considerations and Perspectives. *IFAC-PapersOnLine* 50, 6918–6927 (2017)
11. Valogianni, K., Ketter, W., Collins, J.: Smart Charging of Electric Vehicles Using Reinforcement Learning. In: *Proceedings of the 15th AAAI Conference on Trading Agent Design and Analysis*, pp. 41–48. AAAI Press (2013)
12. O'Neill, D., Levorato, M., Goldsmith, A., Mitra, U.: Residential Demand Response Using Reinforcement Learning. *IEEE SmartGridComm* (2010)
13. Veldman, E., Verzijlbergh, R.A.: Distribution Grid Impacts of Smart Electric Vehicle Charging From Different Perspectives. *IEEE Transactions on Smart Grid* 6, 333–342 (2015)
14. Dimitrov, S., Lguensat, R.: Reinforcement Learning Based Algorithm for the Maximization of EV Charging Station Revenue. In: *International Conference on Mathematics and Computers in Sciences and in Industry (MCSI), 2014*. Varna, Bulgaria, 13 - 15 September 2014, pp. 235–239. IEEE, Piscataway, NJ (2014)
15. Lin, L.-J.: Self-Improving Reactive Agents Based On Reinforcement Learning, Planning and Teaching. *Machine Learning* 1992, 293–321 (1992)
16. Mönnig, A., Schneemann, C., Weber, E., Zika, G., Helmrich, R.: *Elektromobilität 2035: Effekte auf Wirtschaft und Erwerbstätigkeit durch die Elektrifizierung des Antriebsstrangs von Personenkraftwagen* (2018)
17. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization. In: Bengio, Y., LeCun, Y. (eds.) *3rd International Conference on Learning Representations. Conference Track Proceedings* (2015)
18. Projektskizze zum Sortimo Innovationspark Zusmarshausen. Offenes IKT-Ökosystem zur cross-sektoralen Steuerung und Betriebsführung der Energieinfrastruktur an einem multifunktionalen Hub für Mobilität, Produktion und Quartiersversorgung (2019)
19. Pfenninger, S., Staffell, I.: Data Platform. Weather Data. Open Power System Data (2019)
20. Muehlenpfordt, J.: Data Platform. Time Series. Open Power System Data (2019)