

# Reinforcement Learning zur Ableitung autonomer Handelsstrategien in Elektrizitätsmärkten

M.Sc. Philipp Reuber, M.Sc. Jacob Tran, Dr. Simon Krahl, Prof. Albert Moser

Aachen, 16.02.2023



# Einleitung

## Motivation & Forschungsfrage

### Aktuelle Entwicklungen des Strommarktes

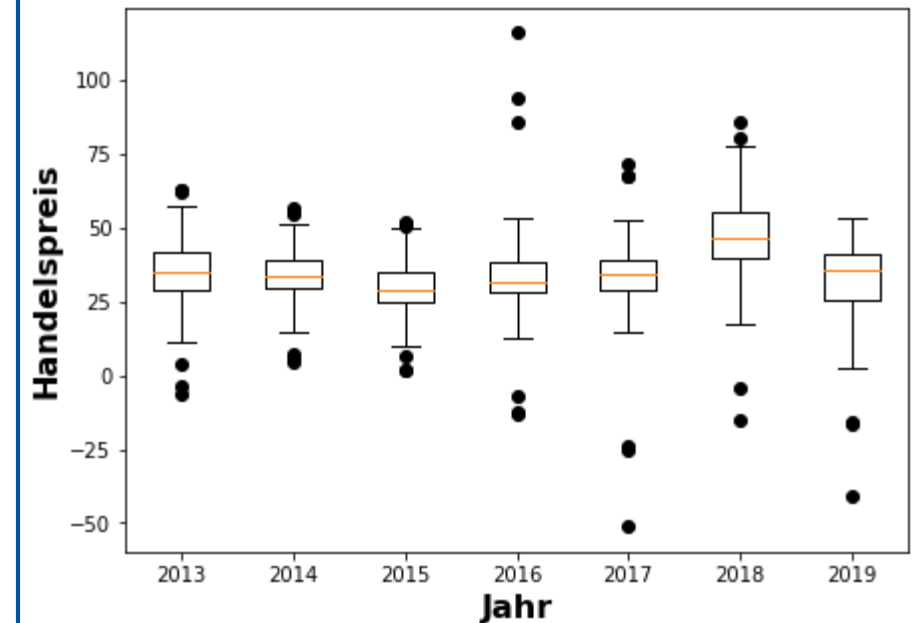
- Steigender Anteil von Erzeugungsanlagen auf Basis erneuerbarer Energien (EE) und neuartiger Verbraucher
- Zeitliche Ungleichgewichte zwischen Erzeugung und Last
- Steigende Volatilität des Strompreises, untertäglich und zwischen zwei Handelstagen
- Kurzfristiger Stromhandel gewinnt an Bedeutung
- Automatisierte Handelssysteme höheres Potential zur Gewinnerzielung

### Modelle des Reinforcement Learning (RL)

- Vereinigung von Prognosen und Ableitung von Entscheidungen
- RL als Möglichkeit zum automatisierten Stromhandel

Ziel: Einsatz des RL zur Ableitung von Handelsstrategien in Elektrizitätsmärkten

### Entwicklung der Preisschwankungen



## Analyse (1/3)

# Kurzfristiger Stromhandel – Day-Ahead Spotmarkt

- Handel elektrischer Energie in Deutschland im Wesentlichen über Over-the-Counter (OTC)-Markt und European Energy Exchange (EEX)

### OTC-Markt

- Keine festen Handelszeiten
- Käufer und Verkäufer sind Vertragspartner
- Bilateraler Handel
- Individuelle Preisgestaltung
- Intransparent

### EEX

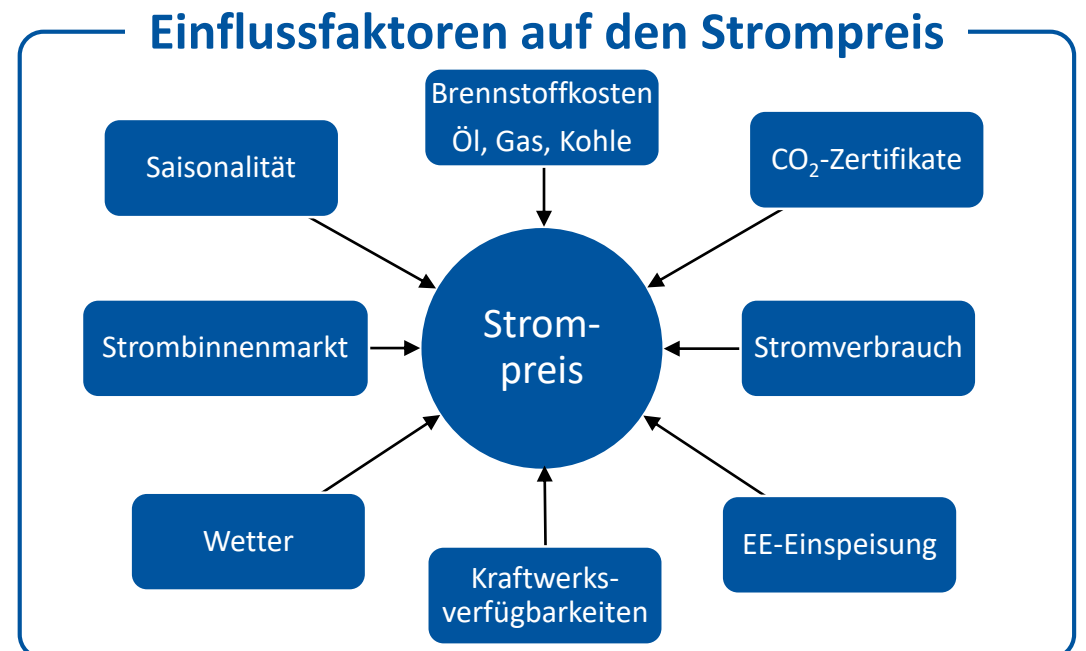
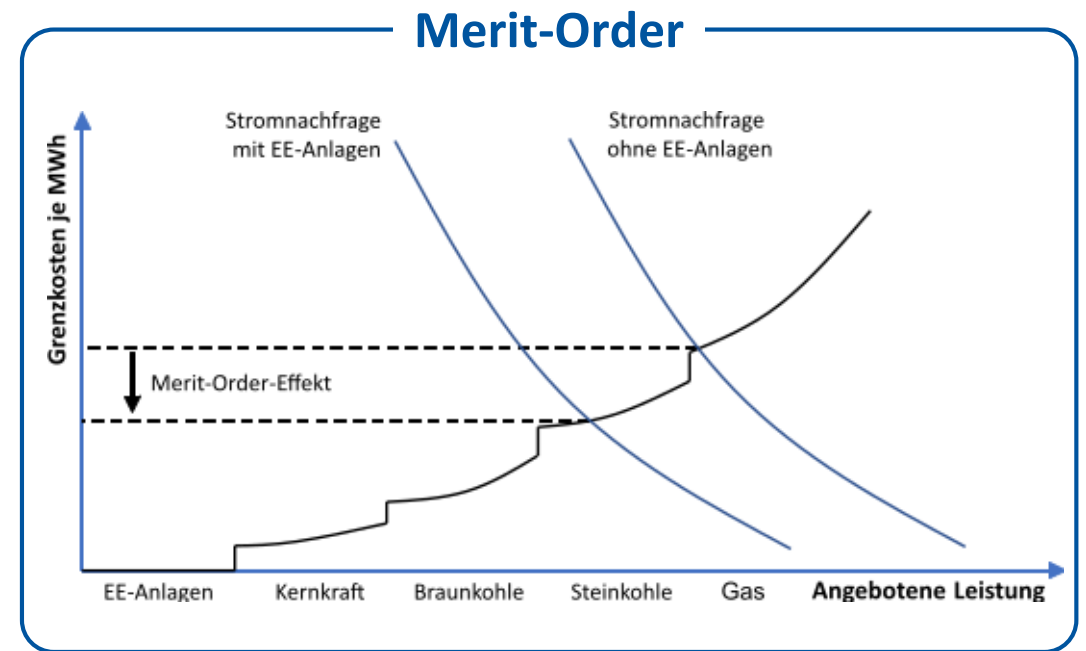
- Feste Handelszeiten (9 Uhr bis 12 Uhr)
- Börse ist Vertragspartner
- Hohe Anzahl an Marktteilnehmern
- Einheitlicher Markträumungspreis
- Hohes Maß an Transparenz

- Markträumungspreis erst nach Auktionsende bekannt, OTC-Handelsgeschäfte bereits vorher möglich
- Unsicherheiten in der Preisbildung im OTC-Markt
- Ausnutzung der Preisdifferenzen zwischen beiden Märkten zur Gewinnerzielung möglich

## Analyse (2/3)

# Strompreisbildung

- Preisbildung im Day-Ahead Spotmarkt an der EEX per Merit-Order für jede Stunde getrennt
- Sortierung der Kraftwerksgebote nach ihren jeweiligen Grenzkosten
- Brennstoffkosten und Kosten für Zertifikate entscheidend für Reihenfolge der Kraftwerke
- Markträumungspreis: Schnittpunkt zwischen Angebots- und Nachfragekurve
- EE-Anlagen verringern Menge an konventionellen Kraftwerken zur Deckung der Nachfrage (Merit-Order-Effekt)
- Volatilität von EE-Anlagen führt zu starker Unsicherheit in der Preisbildung
- Erschwerte Quantifizierung des Preises

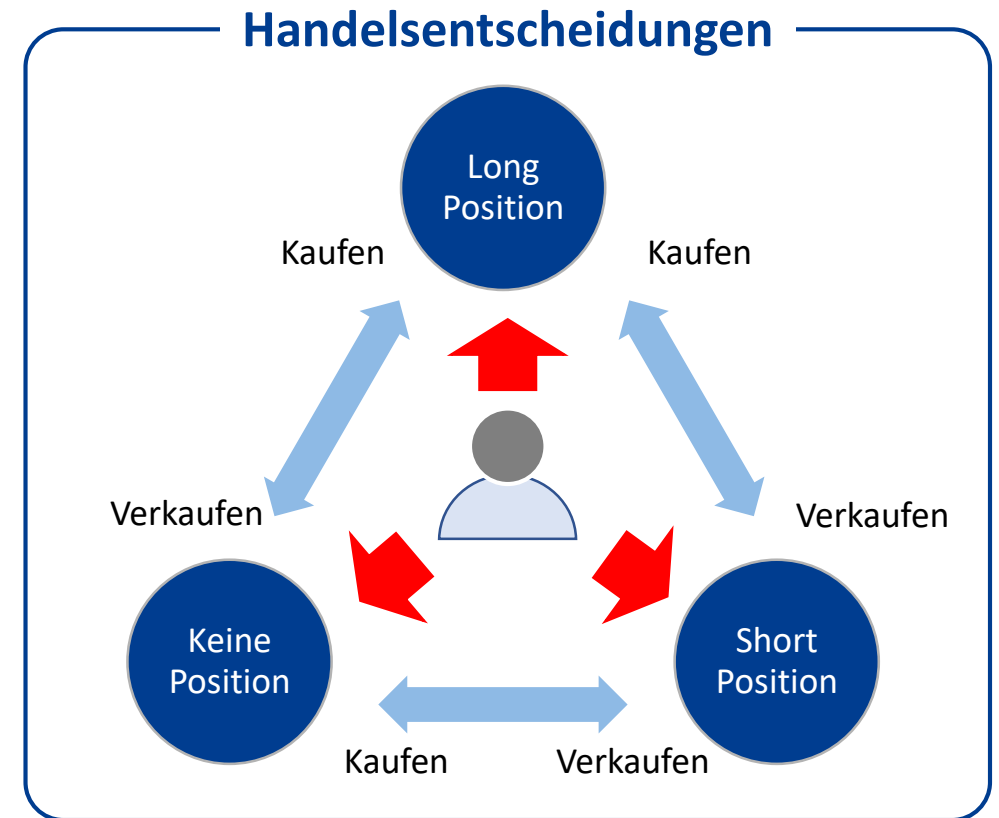


## Analyse (3/3)

# Ableitung von Handelsstrategien

### Entscheidung im Stromhandel

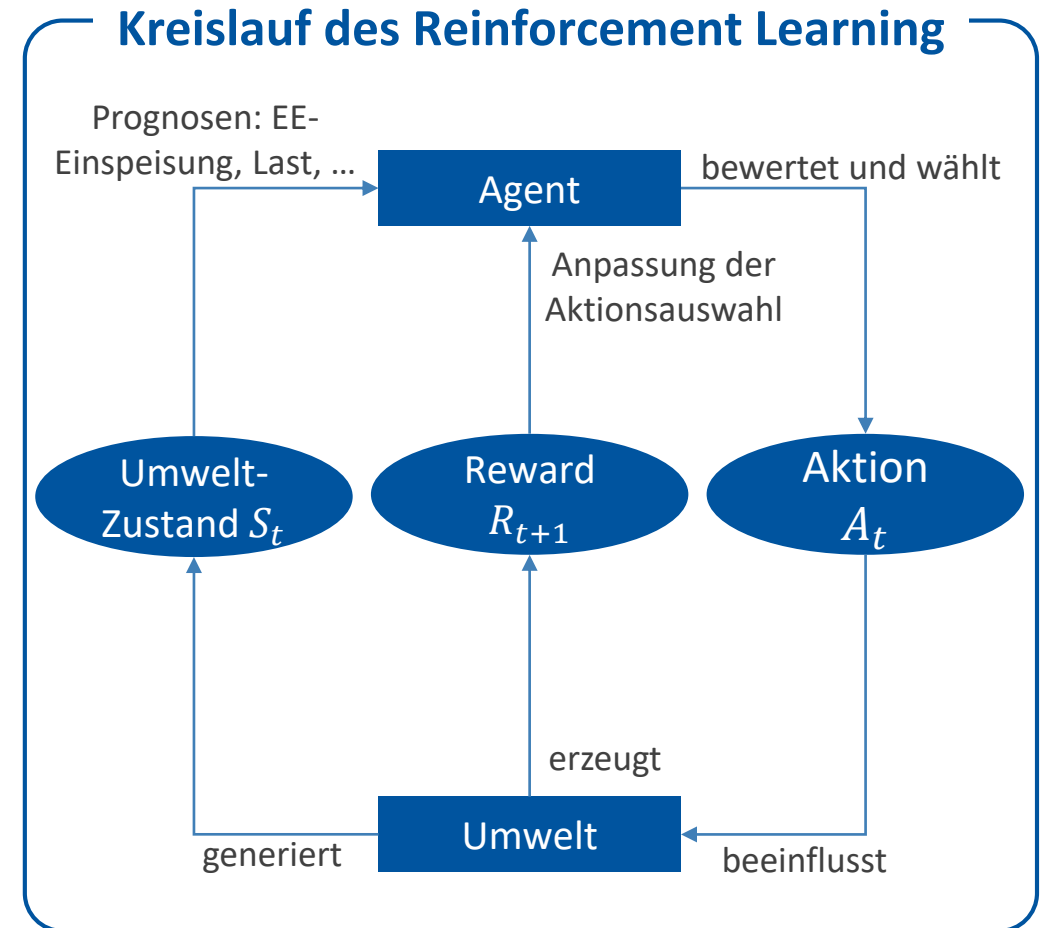
- Kauf/Verkauf von Strom in beiden Märkten möglich
- Gewinnerzielung durch Nutzung von Preisbewegungen innerhalb des OTC-Marktes oder Preisdifferenz zwischen den Märkten
- Eingehen einer Position im OTC-Handel, schließen an der EEX möglich
- Gewinnmaximierung durch finden von optimalen Handelszeitpunkten
- Eingehen optimaler Positionen erfordert Prognose des Strompreises
- Erlernen der Zusammenhänge innerhalb historischer Daten durch RL



## Modellbildung (1/2)

# Mathematische Modellierung des Reinforcement Learning

- Aktion  $A_t$  beeinflusst Umwelt und Generierung des nächsten Umweltzustands
- Nach Durchführung der Aktion Erzeugung Reward  $R_{t+1}$
- Ziel: Maximierung der Summe der Rewards (Gewinn)
- Agent bewertet Aktionen und wählt auf Basis einer Policy die vielversprechendste Aktion aus
- Verbesserung der Aktionsauswahl durch Anpassung der Policy-Funktion (Policy-basiert), der Bewertungsfunktion (Wertebasiert) oder Anpassung beider Funktionen (Actor-Critic)
- RL erlernt Zustände und Auswirkungen der Aktionen durch Interaktionen mit der Umwelt (Trial-and-Error)
- Policy-Funktion und/oder Bewertungsfunktion wird durch Künstliche Neuronale Netze (KNN) modelliert



## Modellbildung (2/2)

# Modellierung der Umwelt und des Agenten

### Agent

- Vereinfachungen:
  - Eingehen einer Position nur mit einer festen Menge
  - Betrachtungen von Lieferungen für den gesamten Handelstag (Base)
- Aktionen: Long Position, Short Position und keine Position
- Verschiedene Verfahren zur Anpassung der Aktionsauswahl vorhanden

### Umwelt

- Ausführung der Handelsaktionen des Agenten
- Berechnung des Rewards:
  - Ausnutzung der Preisdifferenz zwischen OTC-Markt und Strombörse: Ermittlung durch Differenz zwischen beiden Preisen und eingegangener Position
  - Mehreren Handelsschritte im OTC-Markt: gegebenenfalls Reward-Engineering erforderlich (Credit-Assignment-Problem)
- Berücksichtigung weiterer Nebenbedingungen möglich

# Verfahren

## Verfahrensaufbau



### Datenvorverarbeitung

#### Bereinigung

- ⑩ Fehlende Daten
- ⑩ Fehlerhafte Daten

#### Normalisierung

- ⑩ Anpassung der Größenordnung
- ⑩ Standardisierung

#### Transformation

- ⑩ Nominale/Ordinale Attribute
- ⑩ Erstellung von Merkmale

### Parametrierung

- Voruntersuchung mit ausgewählten RL-Verfahren
  - Verwendung eines Tools zur automatisierten Parametrierung
  - Vorgabe von Suchbereichen für alle Parameter (Netzstrukturparameter des KNN und Parameter des Algorithmus)
  - Output: Parametersets mit Bewertung
- Bestimmung optimaler Trainings- und Netzstrukturparameter der KNN

### Training/Test

- Aufteilung in Trainings- und Testdatensatz
- Handelsstrategie zur Anwendung auf Testdatensatz





# Exemplarische Untersuchungen (1/2)

## Datengrundlage

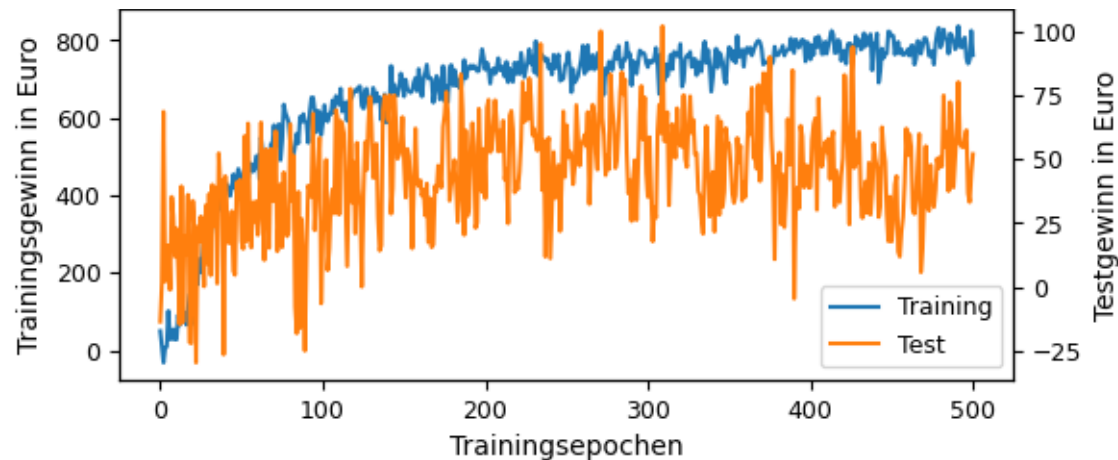
- Prognosen von EE-Anlagen, Temperatur, Einspeisung Kernkraft und Verbrauch
- Kraftwerksverfügbarkeiten, Brennstoffkosten, Kosten Zertifikate, Datum, Klassifikation des Tages
- Historische Daten im Zeitraum vom 01.05.2013 bis 31.05.2020
  - Untersuchungszeitraum: 01.01.2018 – 31.05.2020 (877 Tage) aufgrund der Gebotszonentrennung Deutschland/Österreich
- Testzeitraum: Zufällige Auswahl von 20% aller Handelstage (175 Tage)
- Transformation stündlicher Daten zu einem einzigen Datenpunkt für den gesamten Handelstag
- Unterschiedliche Anzahl an OTC-Preisen für Base-Lieferungen während den Handelszeiten der Börse pro Handelstag
- Vereinfachung: Beschränkung auf einen einzigen OTC-Preis pro Tag
- Ausnutzung der Preisdifferenzen zwischen beiden Märkten

# Exemplarische Untersuchungen (2/2)

## Validierung des Verfahrens

- Parametrierung des Modells für PPO<sup>1</sup>, Bewertung anhand des Trainingsrewards
- Schnellerer Anstieg des Trainingsgewinns im parametrisierten Zustand
- Hohe Volatilität des Testgewinns
- Spitzenwerte des Testgewinns überschreiten Gewinn des Random Forest Modells
- **Die Ergebnisse zeigen eine grundsätzliche Eignung des Verfahrens zur Ableitung von Handelsstrategien in Elektrizitätsmärkten**

### Vergleich der unterschiedlichen RL-Verfahren



### Vergleich der Gewinne für Referenzstrategien

	Trainingsdatensatz	Testdatensatz
OTC-Kaufen	-35,34€	22,86€
OTC-Verkaufen	11,34€	-28,57€
Maximal möglicher Gewinn	922,22€	248,44€
Random Forest Modell	689,31€	90,84€
Anzahl Tage	702	175

# Zusammenfassung & Ausblick

## Zusammenfassung & weiterer Forschungsbedarf

### Motivation und Forschungsfrage

- Zunehmende Volatilität der Strompreise durch zunehmende Dargebotsabhängigkeit und neuartige Verbraucher
- Automatisierte Handelssysteme höheres Potential zur Gewinnerzielung
- Vielfältige Einflussfaktoren erschweren Quantifizierung des Strompreises
- Untersuchung des Reinforcement Learning zur Ableitung von Handelsstrategien in Elektrizitätsmärkten

### Exemplarische Untersuchungen

- Grundsätzliche Eignung des Verfahrens, jedoch hohe Volatilität des Testgewinns

### Weitergehende Fragestellungen

- Wie lässt sich die hohe Volatilität des Testgewinns reduzieren?
- Lässt sich dieses Verfahren zur Gewinnerzielung durch Preisbewegungen im OTC-Markt einsetzen?

# Vielen Dank für Ihre Aufmerksamkeit

## Reinforcement Learning zur Ableitung autonomer Handelsstrategien in Elektrizitätsmärkten

### Kontakt

**M.Sc. Philipp Reuber**

Tel.: +49 241 – 997857 148

[philipp.reuber@fgh-ma.de](mailto:philipp.reuber@fgh-ma.de)

Forschungsgemeinschaft für Elektrische Anlagen und  
Stromwirtschaft e.V.

Roermonder Straße 199  
52072 Aachen

[www.fgh-ma.de](http://www.fgh-ma.de)

