

Bachelorarbeit

**Entwicklung einer statistischen Methodik
zur vergleichenden Ergebnisanalyse von
Open Source Energy System Simulations
Software**

Arne Bo Wilckens

Oktober 2021

Bachelorarbeit**Entwicklung einer statistischen Methodik
zur vergleichenden Ergebnisanalyse von
Open Source Energy System Simulations
Software****Arne Bo Wilckens**

Matr.-Nr.: 21599655

Erstprüfer: Dr.-Ing. Kristin Abel-Günther

Zweitprüfer: M.Sc.Mathias Ammon

Betreuer: M.Sc.Mathias Ammon

Hamburg, 06. Oktober 2021

Ich erkläre hiermit, dass die vorliegende Bachelorarbeit ohne fremde Hilfe selbstständig verfasst wurde und nur die angegebenen Quellen und Hilfsmittel benutzt worden sind. Wörtlich oder sinngemäß aus anderen Werken entnommene Stellen sind unter Angabe der Quelle kenntlich gemacht.

Alle Quellen, die dem World Wide Web entnommen oder in einer sonstigen digitalen Form verwendet wurden, sind der Arbeit beigefügt.

Diese Arbeit ist nach bestem Wissen erstellt worden. Für den Inhalt kann jedoch keine Gewähr übernommen werden.

Hamburg, 06. Oktober 2021

(Arne Bo Wilckens)

Inhaltsverzeichnis

Abbildungsverzeichnis	II
Tabellenverzeichnis	III
Symbolverzeichnis	IV
1 Einleitung	1
1.1 Ziel der Arbeit	1
2 Verwendete Simulationssoftware	3
2.1 Open Energy Modeling Framework (oemof)	4
2.2 Python for Power System Analysis (PyPSA)	4
3 Mathematische Grundlagen	5
3.1 Korrelationen	5
3.1.1 Arithmetischer Mittelwert	5
3.1.2 Varianz und Standardabweichung	5
3.1.3 Kovarianz	6
3.1.4 Pearson Korrelation	6
3.1.5 Kreuzkorrelation	7
3.2 Gleitender Mittelwert	8
3.3 Modellfehler	11
3.3.1 Mittlerer einfacher Fehler	11
3.3.2 Mittlerer absoluter Fehler	12
3.3.3 Mittlerer quadratischer Fehler	12
3.3.4 Normalisierung der Fehler	13
4 Methodisches Vorgehen	14
4.1 Schritt 1: Pearson Korrelation	15
4.2 Schritt 2: Graphische Gesamtübersicht	16
4.3 Schritt 3: Zeitverzögerung	16
4.4 Schritt 4: Modellfehler	17
4.5 Schritt 5: Graphische Detailanalyse	18
5 Fallbeispiele für die Methodik	19
5.1 Szenario 1: Das einfache Energiesystem	19
5.1.1 Keine Unterschiede (24 Zeitschritte)	20
5.1.2 Kaum Unterschiede (168 Zeitschritte)	21
5.1.3 Große Unterschiede (720 Zeitschritte)	24
5.2 Szenario 2: Das komplexe Energiesystem	29
6 Fazit und Ausblick	36
Literatur	41

Abbildungsverzeichnis

2.1	Beispiel eines Energiesystems nach der Graphentheorie	3
3.1	Graphische Veranschaulichung der Höhe der Pearson Korrelation [11] . . .	7
3.2	Beispiel einer Kreuzkorrelation mit unterschiedlichen Zeitfenstern	8
3.3	Beispiel eines one- und tow sided moving average	9
3.4	Beispiel eines one- und two sided moving average mit dynamischem Fenster	10
4.1	Methodisches Vorgehen	14
5.1	Graphische Darstellung des einfachen Energiesystems	20
5.2	Szenario 1: Graphische Darstellung aller Komponenten bei 24 Zeitschritten	21
5.3	Szenario 1: Graphische Darstellung aller Komponenten bei 168 Zeitschritten	22
5.4	Szenario 1: Die Modellfehler des Flusses „gas power plant“ zu „powerline“ bei 168 Zeitschritten	23
5.5	Szenario 1: Graphische Darstellung des Flusses von „gas power plant“ zu „powerline“ bei 168 Zeitschritten	24
5.6	Szenario 1: Graphische Darstellung aller Komponenten bei 720 Zeitschritten	25
5.7	Szenario 1: Die Modellfehler des Flusses von „gas power plant“ zu „po- werline“ bei 720 Zeitschritten	26
5.8	Szenario 1: Die Modellfehler des Flusses von „imported electricity“ zu „powerline“ bei 720 Zeitschritten	26
5.9	Szenario 1: Graphische Darstellung des Flusses von „gas power plant“ zu „powerline“ bei 720 Zeitschritten	28
5.10	Szenario 1: Graphische Darstellung des Flusses von „imported electricity“ zu „powerline“ bei 720 Zeitschritten	28
5.11	Graphische Darstellung des komplexen Energiesystems	29
5.12	Szenario 2: Graphische Darstellung aller Komponenten bei 720 Zeitschritten	31
5.13	Szenario 2: Die Modellfehler von „gas pipeline“ zu „gas power plant“ original	33
5.14	Szenario 2: Die Modellfehler von „gas pipeline“ zu „gas power plant“ unverzögert	33
5.15	Szenario 2: Ausschnitt der Abbildung 5.17 originale Zeitreihe	34
5.16	Szenario 2: Ausschnitt der Abbildung 5.18 unverzögerte Zeitreihe	34
5.17	Szenario 2: Graphische Darstellung des Flusses von „gas pipeline“ zu „gas power plant“ orginal	35
5.18	Senario 2: Graphische Darstellung des Flusses von „gas pipeline“ zu „gas power plant“ unverzögert	35
6.1	Die Modellfehler des Beispiels	39
6.2	Graphische Darstellung des Beispiels	40

Tabellenverzeichnis

3.1	Pearson Korrelation Werte: [17]	7
3.2	Beispiel eines dynamischen Fensters der Größe 3	10
5.1	Szenario 1: Ergebnisse der Pearson Korrelation für 24 Zeitschritte	20
5.2	Szenario 1: Ergebnisse der Pearson Korrelation für 168 Zeitschritte	21
5.3	Szenario 1: Ergebnis der Kreuzkorrelation des Flusses „gas power plant“ zu „powerline“ bei 168 Zeitschritten	22
5.4	Szenario 1: Ergebnisse der Pearson Korrelation für 720 Zeitschritte	24
5.5	Szenario 1: Ergebnis der Kreuzkorrelation der Flüsse bei 720 Zeitschritten	25
5.6	Szenario 2: Ergebnisse der Pearson Korrelation für 720 Zeitschritte	30
5.7	Szenario 2: Ergebnis der Kreuzkorrelationen des untersuchten Flusses	32
5.8	Szenario 2: Ergebnisse der Kreuzkorrelationen der untersuchten Flüsse	32
6.1	Beispiel eines zu betrachtenden Flusses	38

Symbolverzeichnis

Abkürzungen

Symbol	Definition
ESS	Energy Systems Simulation
ocfz	one component is filled with zeros
oemof	Open Energy Modeling Framework
PyPSA	Python for Power System Analysis
Tessif	Transform Energy Simulation Systems Framework

Griechische Symbole

Symbol	Definition	Einheit
τ	Zeitverzögerung	h

Indizes

Symbol	Definition	Einheit
i	Laufvariable	
j	Laufvariable	

Lateinische Symbole

Symbol	Definition	Einheit
\bar{x}	arithmetischer Mittelwert der Zeitreihe x	
\bar{y}	arithmetischer Mittelwert der Zeitreihe y	
\bar{e}	arithmetischer Mittelwert des mittleren Verlaufes	
a	Zeitreihe eines beliebigen Modells	
b	Betrachtungsfenster / Stützbereich	
cov	Kovarianz	
Cv	Coefficient of variation	
e	durchschnittlicher Verlauf aller Modelle	
k	Stützbereichskonstante	
K	Kreuzkorrelation	
m	Anzahl der Modelle	
ma	neue Zeitreihe des gleitenden Mittelwertes	
MAE	Mean Absolute Error	
MBE	Mean Bias Error	
MSE	Mean Square Error	
n	Anzahl	
NMAE	Normalized Mean Absolute Error	
NMBE	Normalized Mean Bias Error	
NRMSE	Normalized Root Mean Square Error	
p	ursprüngliche Zeitreihe	
PE	Anzahl der zu berechneten Pearson Korrelationen	
r	Pearson Korrelation	
RMSE	Root Mean Square Error	
s	Standardabweichung	
t	Zeit	h
var	Varianz	
x	beliebige Zeitreihe	
y	beliebige Zeitreihe	

1 Einleitung

Um den weltweiten Wandel zu nachhaltigeren Energiesystemen zu bewältigen, werden zuverlässige und präzise Prognosen benötigt. Mit diesen können Entscheidungsprozesse auf verschiedenen Ebenen unterstützt werden, von der internationalen Politikgestaltung bis zur lokalen Energieplanung [1].

Eine der maßgeblichen Untersuchungsmethoden zukünftiger Energiesysteme ist die Szenario-Analyse. Sie wird anhand von Energiesystemmodellen (eng: Energy Systems Simulation, ESS) realisiert [2]. Um so ein Energiesystem zu simulieren, gibt es zahlreiche Software-Modelle. Doch nur wenige davon sind Open Source.

In den letzten Jahren wurde nicht nur die fehlende Transparenz, sondern auch die begrenzte Reproduzierbarkeit von Energiesystemanalysen bemängelt [3]. Um das zu ändern und damit einen Beitrag zur Open Science zu leisten, ist die Entwicklung und Etablierung von Open Source Simulationsmodellen erstrebenswert.

Ein wesentliches Problem der ESS stellt die Vergleichbarkeit der Modelle da. So wird jedes Programm von verschiedenen Organisationen oder Personengruppen entwickelt, zwischen denen es bei der Entwicklung wenig bis gar keine Koordination gibt. Um diesem Problem weitestgehend entgegenzuwirken, wurde das Projekt Tessif ins Leben gerufen. Es wird am Institut für Energietechnik der Technischen Universität Hamburg entwickelt.

Dabei verfolgt das Projekt zwei grundsätzliche Ziele:

1. Die Senkung der Schwelle für Ingenieure, Open Source Simulationssoftware für Energiesysteme zu verwenden, indem mehrere dieser Modelle über eine Schnittstelle verwendet werden.
2. Einen Vergleich zwischen verschiedenen Modellen zu ermöglichen. Um anhand dessen das Fähigste auszuwählen und einen Überblick über die Bandbreite der möglichen Ergebnisse und Lösungsansätze der verschiedenen Modelle zu bekommen.

Im Rahmen dieses Projektes verlangt es nach einer Methodik, die aus verschiedenen Modellen Unterschiede herausarbeitet und sichtbar macht.

1.1 Ziel der Arbeit

Ziel dieser Arbeit ist die Entwicklung einer Analysemethode, die Unterschiede zwischen Energiesimulationsmodellen erkennt und die Genauigkeit des einzelnen Modells bewertet.

Dazu wird zunächst auf das Projekt Tessif noch einmal detailliert eingegangen. Hierbei werden kurz die beiden Simulationsprogramme oemof und PyPSA vorgestellt, die dem Analyseverfahren als Testmodelle dienen.

Weiter werden in Kapitel 3 die statistischen Grundlagen, die für die Analyse benötigt werden, erläutert.

Im 4. Kapitel wird die Methodik, welche der Analyse zugrunde liegt und in die die statistischen Verfahren eingebettet sind, beschrieben.

Die in Kapitel 4 erläuterte Methodik wurde mit Python implementiert und in Tessif integriert und anschließend anhand zweier Beispielsimulationen getestet.

Kapitel 5 beschreibt diese Tests und die dabei auftretenden Ergebnisse.

Im 6. Kapitel wird schließlich ein Fazit gezogen und ein Ausblick für eine zukünftige Verbesserung der Analyse von Energiesystemsimulationen in Tessif gegeben.

2 Verwendete Simulationssoftware

Tessif steht für „Transform Energy Simulation Systems Framework“ und kann als Plattform für Energiesystemsimulationen verstanden werden. Diese ermöglicht die Simulation eines Energiesystems mit mehreren ESSs. Tessif wird mit der Open Source Programmiersprache Python als projektorientierte Software geschrieben. Alle implementierten ESS müssen auf Python basieren [4].

Die in dieser Arbeit entwickelte Methodik zur statistischen Analyse wird ebenfalls in Python implementiert und in Tessif integriert, um die Analysefähigkeiten zu erweitern.

Die in Tessif verwendeten Energiesystemsimulationen können als Graphen betrachtet werden. Der Aufbau eines ESS nach der Graphentheorie wird in Abbildung 2.1 gezeigt: Es gibt Knoten und Kanten. Die Knoten bilden abstrakte und generische Komponenten eines Energiesystems, z.B. einen Erzeuger (Kraftwerk), einen Verbraucher (von der Industrie benötigter Strom) oder einen Verteiler. Die Kanten repräsentieren die Flüsse zwischen den Knoten bzw. Komponenten des Systems. Diese sind in der Regel Energieflüsse.

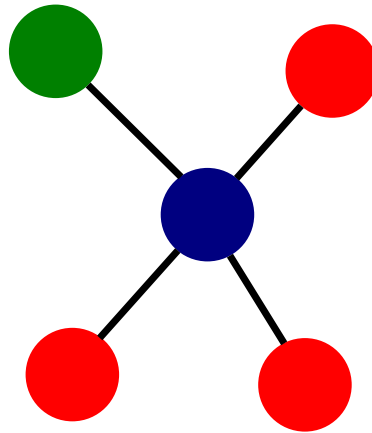


Abbildung 2.1: Beispiel eines Energiesystems nach der Graphentheorie

In dem Beispiel werden die Energieerzeuger als rote Knoten dargestellt. Diese laufen in den Verteiler (blau) zusammen und werden dort gebündelt, um anschließend in den Verbraucher (grün) überzugehen. Die Knoten sind Bilanzierungspunkte. Das bedeutet, dass alle Flüsse, die in die Knoten einfließen, auch wieder herausfließen müssen. Wenn es sich bei den Knoten um einen Erzeuger oder Verbraucher handelt, müssen die zu- oder abgeführten Flüsse in der Bilanz ausgeglichen werden.

Ein Energiesystem hat zwei verschiedene Optimierungsparameter. Der erste Parameter ist die Optimierung der Kosten eines Systems. Das heißt, das Energiesystem wird über den betrachteten Zeitraum auf minimale Kosten optimiert. So lassen sich für die einzelnen Komponenten und Flüsse Kosten individuell einstellen.

Die zweite Möglichkeit hat ein System durch die Optimierung der Emissionen. Hier werden über den Zeitraum der Simulation die Komponenten ausgebaut bzw. diejenigen vorrangig verwendet, die am Ende dann dem vorgegebenen Emissionsziel entsprechen. So können die Emissionswerte der einzelnen Komponenten und Flüsse einzeln angepasst werden.

Es stellen sich folgende Probleme bei der Simulation mit verschiedenen Modellen, die Tessif zu lösen versucht [4]:

1. Keine einheitliche Dateneingabe:

Jedes ESS hat auf sich zugeschnittene Datensätze und Datenstrukturen, da jedes Modell meist für einen speziellen Zweck geschrieben wurde. Dieses Problem wird dadurch verstärkt, dass jeder Entwickler eigene Vorstellungen, bzw. Lösungsansätze umsetzt.

2. Keine einheitliche Datenausgabe:

Jedes ESS kann zwar auf dem Modellkonzept der Graphentheorie, also ein Konstrukt aus Knoten und Kanten, dargestellt werden, es gibt aber keine einheitliche Darstellung, die universell auf jedes System angewendet werden kann.

Die bisherigen Energiesystemsimulationsmodelle, die in Tessif implementiert wurden, sind oemof und PyPSA. Mit diesen beiden Simulationsprogrammen werden im späteren Kapitel 5 die Simulationen durchgeführt und verglichen. Hier folgt eine kurze Vorstellung der beiden Programme:

2.1 Open Energy Modeling Framework (oemof)

Das Akronym oemof steht für „Open Energy Modeling Framework“. Die Programmiersprache Python bildet dafür die Basis. Oemof ermöglicht dem Benutzer die Modellierung und Optimierung von Energiesystemen. Die erste Version des Programms wurde 2015 von der Europa-Universität Flensburg (EUF), der Flensburg University of Applied Sciences (HFL), des Reiner Lemoine Instituts (RLI) und der Otto-von-Guericke-Universität Magdeburg (OVGU) entwickelt [5]. Seitdem wird das Programm von der oemof community via GitHub aktualisiert und weiterentwickelt [6]. Aufgrund der Vielzahl der Entwickler ist oemof breit aufgestellt.

2.2 Python for Power System Analysis (PyPSA)

Python for Power System Analysis kurz PyPSA wurde 2016 am Frankfurter Institut für fortgeschrittene Studien entwickelt und ist ein Programm zur Simulation und Optimierung moderner Stromversorgungssysteme [7]. Mittlerweile wird das Projekt von der Gruppe „Energiesystemmodellierung“ am Institut für Automatisierung und Angewandte Informatik des Karlsruher Instituts für Technologie weitergeführt [8].

3 Mathematische Grundlagen

In diesem Kapitel werden die in der entwickelten Methodik verwendeten statistischen Bewertungskriterien vorgestellt. Zunächst erfolgt eine Erläuterung der Korrelationen, ihrer Berechnung und der dafür nötigen Grundlagen. Im Anschluss wird der gleitende Mittelwert beschrieben. Zuletzt werden verschiedene Modellanalyse-Fehler und deren Eigenschaften erklärt.

In dieser Arbeit werden ausschließlich diskrete Zeitreihen betrachtet. Eine diskrete Zeitreihe ist dabei eine Folge von oftmals reellen Werten, die jeweils einem bestimmten Zeitpunkt zugeordnet sind [9].

3.1 Korrelationen

Dieser Abschnitt beschreibt die notwendigen statistischen Grundlagen zur Ermittlung der Pearson- und der Kreuzkorrelation. Für ihre Berechnung sind zunächst Definitionen der folgenden statistischen Größen notwendig: Arithmetischer Mittelwert, Standardabweichung, Varianz und Kovarianz. Danach werden die Pearson- und die Kreuzkorrelation eingeführt.

3.1.1 Arithmetischer Mittelwert

Der arithmetische Mittelwert (statistischer Durchschnitt) wird gemäß folgender Gleichung definiert:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.1)$$

\bar{x}_i bezeichnet die beobachteten Werte. Alle Daten x_1 bis x_n werden addiert und durch die Anzahl der Werte n dividiert [10].

3.1.2 Varianz und Standardabweichung

Die Varianz ist definiert als die mittlere quadratische Abweichung der Messergebnisse vom arithmetischen Mittelwert. Die Varianz informiert darüber, wie weit die Werte einer Verteilung vom arithmetischen Mittel entfernt sind.

Für die Berechnung der Varianz werden die Abstände aller Werte (x) zum Mittelwert (\bar{x}) quadriert, aufsummiert und anschließend durch die Anzahl der Werte (n) minus eins geteilt [11].

$$var = s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3.2)$$

Aufgrund der Quadratur unterscheidet sich die Maßeinheit der Varianz von der des arithmetischen Mittelwertes. Um die gleiche Maßeinheit herzustellen, wird die Wurzel gezogen. Dieser Wert wird als Standardabweichung (s) bezeichnet. Sowohl die Varianz, als auch die Standardabweichung sind, aufgrund der Quadratur, anfällig für Ausreißer [12].

Als Ausreißer werden große, aber seltene Abweichungen, die nicht den Erwartungen entsprechen, bezeichnet [13].

$$s = \sqrt{var} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.3)$$

3.1.3 Kovarianz

Unter der Kovarianz wird die wechselseitige Varianz zweier Variablen verstanden. Es wird der Abstand der jeweiligen Variablen (x, y) zu den Mittelwerten (\bar{x} bzw. \bar{y}) festgestellt und anschließend miteinander multipliziert.

Dies wird als Kreuzprodukt bezeichnet. Alle Kreuzprodukte werden aufsummiert und anschließend durch die Anzahl der Werte (n) einer Zeitreihe minus eins dividiert [14][11].

$$cov_{x,y} = \frac{1}{n-1} \sum_{i=1}^n ((x_i - \bar{x}) \cdot (y_i - \bar{y})) \quad (3.4)$$

Es können Werte von $-\infty$ und ∞ angenommen werden.

Eine positive Kovarianz bedeutet, dass größere Werte von x größere Werte von y implizieren. Ein negatives Vorzeichen der Kovarianz bedeutet, dass aus größeren Werten von x kleinere Werte von y folgen [15].

3.1.4 Pearson Korrelation

Die Pearson Korrelation ist ein Maß für den Grad des linearen Zusammenhangs zwischen zwei Variablen (x, y). Diese wird wie folgt berechnet [16]:

$$r = \frac{\frac{1}{n-1} \sum_{i=1}^n ((x_i - \bar{x}) \cdot (y_i - \bar{y}))}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.5)$$

Die Pearson Korrelation ist die Normierung der Kovarianz. Diese wird berechnet, indem die Kovarianz durch das Produkt der Standardabweichungen der beiden Zeitreihen x und y geteilt wird [11].

$$r = \frac{cov(x, y)}{s_x \cdot s_y} \quad (3.6)$$

Durch die Normierung liegt der Wert der Pearson Korrelation immer im Bereich von $r \in [-1, 1]$. Ein Korrelationswert von $r = +1$ steht dabei für einen perfekten linearen Zusammenhang. Je höher der Korrelationswert ist, desto ausgeprägter ist der lineare Zusammenhang [11].

Ein Wert von $r = 0$ signalisiert, dass keine lineare Korrelation zwischen den Variablen vorliegt und ein Wert von $r = -1$ steht für eine perfekte negative Korrelation. Das Vorzeichen der Pearson Korrelation r zeigt die Richtung des Zusammenhangs an, die Höhe gibt die Ausprägung der Korrelation wieder. Siehe Tabelle 3.1 und Abbildung 3.1. Durch die Normierung entfällt der Einfluss von Einheiten, so dass auch Variablen mit unterschiedlichen Einheiten verglichen werden können [11][16].

Betrag von r	Stärke des Zusammenhang
$0,00 \leq r < 0,10$	kein Zusammenhang
$0,10 \leq r < 0,30$	geringer Zusammenhang
$0,30 \leq r < 0,50$	mittlerer Zusammenhang
$0,50 \leq r < 0,70$	hoher Zusammenhang
$0,70 \leq r < 1,00$	sehr hoher Zusammenhang

Tabelle 3.1: Pearson Korrelation Werte: [17]

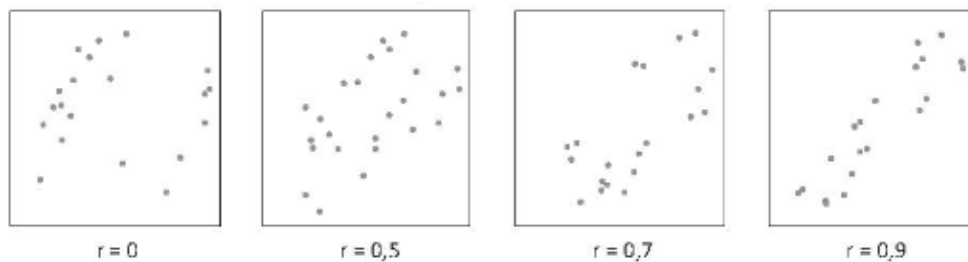


Abbildung 3.1: Graphische Veranschaulichung der Höhe der Pearson Korrelation [11]

3.1.5 Kreuzkorrelation

Um Zeitverzögerungen in den Daten zwischen zwei Zeitreihen zu erkennen, wird die Kreuzkorrelation verwendet. So ist die Kreuzkorrelation zweier Zeitfunktionen $x(t)$ und $y(t)$ definiert als [18]:

$$K(\tau) = \int_{-\infty}^{\infty} x(t)y(t + \tau)dt \quad (3.7)$$

Eine Funktion wird zeitlich um den Wert τ verschoben, dann mit der anderen Zeitfunktion multipliziert um anschließend integriert zu werden.

Wenn beide Funktionen sich in ihrem Verlauf ähneln und nur um den Wert τl zeitlich verschoben sind, wird die Korrelationsfunktion bei $\tau = \tau l$ ein Maximum aufweisen [18][19].

Da in dieser Arbeit die Zeitreihen nicht als zeitabhängige Funktionen, sondern als diskrete Zeitpunkte vorliegen, gilt für die Kreuzkorrelation [19][20] :

$$K_{x,y}(t_1, t_2) = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i(t_1) - \bar{x}(t_1)) \cdot (y_i(t_2) - \bar{y}(t_2))}{s_x \cdot s_y} \quad (3.8)$$

So wird die Pearson Korrelation der beiden Zeitreihen mit der Zeitdifferenz $\tau = t_2 - t_1$ berechnet.

Um nun eine Zeitdifferenz beziehungsweise Zeitverschiebung festzustellen, wird ein Zeitfenster beliebiger Größe bestimmt, indem die beiden Zeitreihen zeitlich zueinander verschoben werden und für jede Verschiebung dann die Pearson Korrelation berechnet wird. Für jede dieser getesteten Verschiebungen ergibt sich somit ein Korrelationskoeffizient. Analog zum oben beschriebenen Vorgehen, liegt die tatsächliche Zeitverzögerung zwischen den beiden Zeitreihen dort, wo die höchste Pearsonkorrelation auftritt. Das Zeitfenster sollte allerdings nicht zu groß gewählt werden. Je größer die Verschiebung ist, desto kleiner wird die Anzahl an Datenpunkten aus den Zeitreihen, die für die Berechnung der Pearson Korrelation verwendet werden. So steigt bei geringeren Datenmengen die Wahrscheinlichkeit der Übereinstimmung der beiden Zeitreihen.

Abbildung 3.2 zeigt den Einfluss, den unterschiedliche Zeitfenstergrößen auf die Anzahl, der für die Pearson Korrelation verwendeten Daten hat.

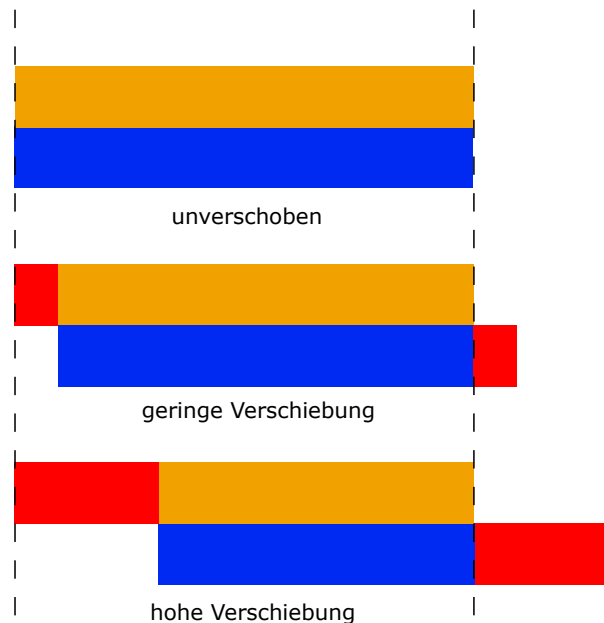


Abbildung 3.2: Beispiel einer Kreuzkorrelation mit unterschiedlichen Zeitfenstern

3.2 Gleitender Mittelwert

Der gleitende Mittelwert oder gleitender Durchschnitt (engl.: moving average) ist eine Methode zur „Glättung“ von Zeitreihen, die Schwankungen unterliegen. Nach der „Glättung“ wird der Trend einer Zeitreihe sichtbar, d.h. ob die Zeitreihe fällt oder steigt. Der gleitende Mittelwert lässt sich aber auch zur Vorhersage von Zeitreihen verwenden.

Der gleitende Mittelwert ist eine Folge von arithmetischen Mittelwerten, die über ein festgelegtes Betrachtungsfenster (Stützbereich) ermittelt werden. Das Fenster wird solange überlappend verschoben, bis die Größe des Fensters erreicht ist. Somit gleitet der arithmetische Mittelwert durch eine Zeitreihe. Je größer, bzw. kleiner das Fenster gewählt wird, umso mehr, bzw. weniger ist die Zeitreihe geglättet [21].

Es gibt viele verschiedene Arten des gleitenden Mittelwertes. Im Rahmen dieser Arbeit wird der einfache, gleitende Durchschnitt betrachtet.

Dieser lässt sich noch in zwei Unterkategorien aufteilen. Einmal den *one-sided moving average*, der in der Regel zu Vorhersagen benutzt wird, und dem *two-sided moving average*, der sich für eine Trendanalyse eignet. Sie unterscheiden sich in den Ausrichtungen des Betrachtungsfensters. So liegt es beim *one-sided moving average* nur in einer Richtung vom Startwert. Im Gegensatz zum *two-sided moving average*, wo jeweils Werte vor und nach dem Startwert genommen werden. Folglich weist der *one-sided moving average* immer eine Zeitverzögerung zur originalen Zeitreihe auf [22].

In Abbildung 3.3 ist jeweils der *one-sided moving average* und der *two-sided moving average* für eine Zeitreihe dargestellt. Deutlich zu sehen ist, dass diese beiden bis auf einen Zeitverzug identisch sind.

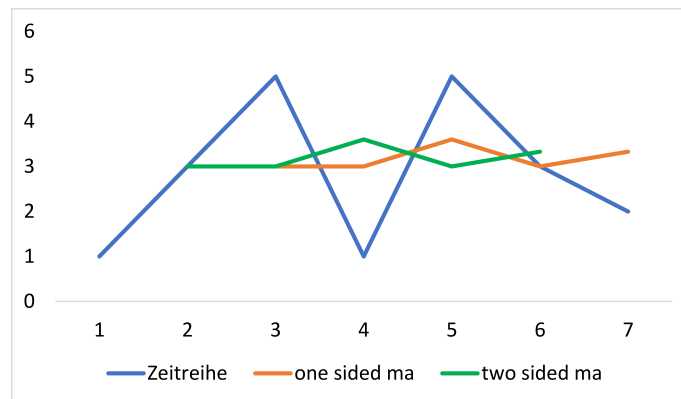


Abbildung 3.3: Beispiel eines one- und tow sided moving average

Aufgrund der Zeitverzögerung wird in dieser Arbeit nur der *two-sided moving average* betrachtet.

Dieser berechnet sich wie folgt: [21] [22]:

$$ma^{(b)}(t) = \frac{1}{2k+1} \sum_{j=-k}^k p_{j+t}, \quad t = k+1, k+2, \dots, k-n \quad (3.9)$$

$b = 2k + 1$ beschreibt die Größe des Betrachtungsfensters bzw. des Stützbereiches und wird häufig auch Ordnung genannt. p ist die originale Zeitreihe und k ist die Stützbereichskonstante.

Für den Fall, dass die Größe des Betrachtungsfensters auf eine gerade Zahl fällt, gibt es zwei mögliche Berechnungen. Dies wird hier anhand eines Beispiels mit der Ordnung $b = 4$ gezeigt [22]:

$$ma^4(t - 0.5) = \frac{p_{t-2} + p_{t-1} + p_t + p_{t+1}}{4} \quad (3.10)$$

oder

$$ma^4(t + 0.5) = \frac{p_{t-1} + p_t + p_{t+1} + p_{t+2}}{4} \tag{3.11}$$

Um eine Priorisierung in den positiven bzw. negativen Bereich zu verhindern, wird von den beiden berechneten Varianten der Mittelwert genommen. Dieser wird auch *centered moving average* genannt [22]. Somit gilt für einen *two-sided moving average* für Betrachtungsfenster mit einer gerade Zahl [22]:

$$ma^4(t) = \frac{1}{2} \cdot \left(\frac{p_{t-2} + p_{t-1} + p_t + p_{t+1}}{4} + \frac{p_{t-1} + p_t + p_{t+1} + p_{t+2}}{4} \right) \tag{3.12}$$

Ein Problem des gleitenden Mittelwertes beruht darauf, dass dieser nur auf eine Größe des Betrachtungsfensters berechnet werden kann. So werden bei einem Fenster der Größe 3 die ersten und letzten Werte der Zeitreihe nicht berücksichtigt, bzw. die neue geglättete Zeitreihe ist um vier Einträge kürzer.

Eine mögliche Lösung stellen dynamische Fenster dar. Hier wird der Stützbereich bis auf die Größe Eins für den *one-sided moving average*, bzw. Zwei für den *two-sided moving average* verringert. Tabelle 3.2 zeigt die Berechnung eines *two-sided moving average* mit und ohne dynamischem Fenster der Größe 3. Die Rahmen zeigen jeweils die Fenstergröße.

Zeitreihe	normal	dynamisch
1		2
3	3	3
5	3	3
1	3.66	3.6
5		3

Tabelle 3.2: Beispiel eines dynamischen Fensters der Größe 3

Graphisch würden die gleitenden Mittelwerte mit dynamischen Fenstern aus der vorherigen Abbildung 3.3 wie folgt aussehen:

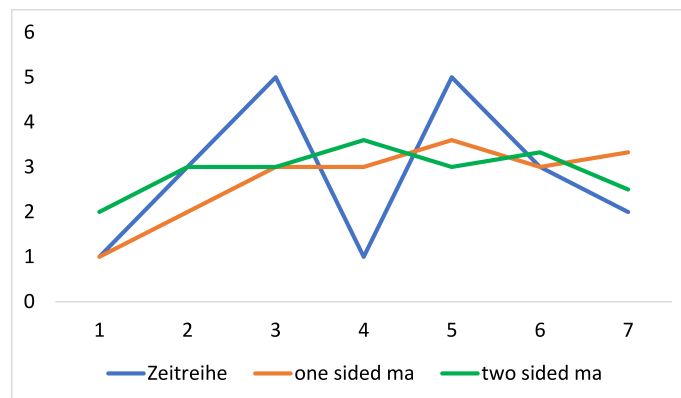


Abbildung 3.4: Beispiel eines one- und two sided moving average mit dynamischem Fenster

Der Nachteil dieser Darstellung ist, dass die Ränder weniger geglättet sind als der Rest der Zeitreihe.

3.3 Modellfehler

Im folgenden Teil wird auf die verschiedenen Modellfehlerarten und deren Bedeutung eingegangen. Anschließend werden die Gleichungen zur Berechnung der Fehlergrößen und die verwendete Normalisierung vorgestellt.

Für eine grundlegende Bewertung von Simulationsmodellen im Bereich der Natur- und Ingenieurwissenschaften stützt man sich auf die Berechnung von Modellfehlern, sogenannten Prognosegütemaßen. Das Prinzip jedes Modellfehlers ist ein paarweiser Vergleich des wahren Wertes mit dem prognostizierten Wert. Das Ergebnis der Modellfehler ist die Abweichung des Modells zum prognostizierten Wert, je kleiner der Fehler, desto näher liegt das Modell an dem Referenzwert [23][24].

Zu beachten ist, dass in dieser Arbeit keine wahren Werte vorliegen, sondern unterschiedliche Modelle miteinander verglichen werden sollen. So wird der wahre Wert durch einen zu erwartenden Wert ersetzt.

In der Stochastik ist der Erwartungswert als Mittelwert eines durchgeführten Zufallsexperiments definiert. Der Wert ist vergleichbar mit dem empirischen arithmetischen Mittel einer Häufigkeitsverteilung in der deskriptiven Statistik [25][26].

In Anlehnung an die Stochastik wird der zu erwartende wahre Wert als arithmetischer Mittelwert eines Zeitschrittes von allen (m) Modellen (a) angenommen,

$$e_i = \frac{1}{m} \sum_{i=1}^m a_i \quad (3.13)$$

Somit bezieht sich die Berechnung der Modellfehler immer auf den durchschnittlichen Verlauf (Mittelwert) der Zeitreihen aller Modelle. Wenn in den unten aufgeführten Abschnitten von einem tatsächlichen bzw. wahren Wert gesprochen wird, ist der, in Gleichung 3.13 ermittelte erwartete Wert gemeint.

Es werden die in der Literatur am häufigsten verwendeten Fehler für die entwickelte Methodik angewendet [27][28]:

3.3.1 Mittlerer einfacher Fehler

Der mittlere einfache Fehler (engl.: Mean Bias Error, MBE) eignet sich um die durchschnittliche Verzerrung (Bias) eines Modells zum wahren Wert anzuzeigen. Die Berechnung erfolgt über die Aufsummierung der Beträge der Differenz des wahren Werts \bar{e}_i zum Modellwert x_i dividiert durch die Anzahl der Zeitschritte n [29].

$$MBE = \frac{1}{n} \sum_{t=1}^n x_i - e_i \quad (3.14)$$

Ein MBE von fast Null kann nicht als eine perfekte Prognose interpretiert werden. Da sich Über- und Unterschätzungen ausgleichen. So gibt der MBE nur Auskunft über die Tendenz eines Modells, die beobachtete Größe zu unter- (MBE < 0) bzw. zu überschätzen (MBE > 0) [24][30].

3.3.2 Mittlerer absoluter Fehler

Der mittlere absolute Fehler (engl.: Mean Absolute Error, MAE), ist die durchschnittliche absolute Abweichung des prognostizierten Werts vom wahren Wert [23]. Die Berechnung erfolgt analog zum MBE mit dem Unterschied, dass hier der Betrag der Differenz aufsummiert wird [24].

$$MAE = \frac{1}{n} \sum_{t=1}^n |x_i - e_i| \quad (3.15)$$

Im Gegensatz zum MBE, liegt ein perfektes Modell vor, wenn die Berechnung des MAE gleich null ergibt.[24]

3.3.3 Mittlerer quadratischer Fehler

Der mittlere quadratische Fehler (engl.: Mean Square Error, MSE) ist der erwartete quadratische Abstand des Schätzwertes. Er ist die quadratische Form des MBE [23] und berechnet sich wie folgt [31]:

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{e}_i)^2 \quad (3.16)$$

Es liegt eine perfekte Prognose vor, wenn der MSE gleich null ist. Aufgrund der Quadratur ist der MSE schwierig zu interpretieren. Um die ursprüngliche Einheitsdimension wieder herzustellen, wird die Wurzel des MSE gezogen. Dieser wird als Root Mean Square Error (RMSE) bezeichnet [23]. Der RMSE wird oft verwendet, um eine Bewertungsgröße für die Qualität des Modells zu erhalten. Auch hier gilt, je kleiner der Wert des RMSE, desto genauer das Modell[24].

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - e_i)^2} \quad (3.17)$$

Durch die Quadratur bekommen Ausreißer beim MSE und MAE eine überproportionale Gewichtung im Vergleich zum MBE. Deshalb steigt die Aussagekraft des MSE und RMSE bei zunehmenden Datenmengen an. Wenn jedoch nur kleine Datenmengen betrachtet werden, muss der dort vorhandene stärkere Einfluss der Ausreißer bei der Interpretation berücksichtigt werden [32].

Der MAE kann mit dem RMSE verglichen werden, um festzustellen, ob ein Modell Ausreißer besitzt. Dazu wird die Differenz der beiden Fehlerwerte betrachtet [32][24].

3.3.4 Normalisierung der Fehler

Um einen Vergleich der mittleren Fehler (RMSE, MAE und MBE) mit verschiedenen Skalierungen zu ermöglichen, können diese Fehler normalisiert werden. Die Normalisierung erfolgt über den arithmetischen Mittelwert des prognostizierten \bar{x}_i bzw. des wahren Wertes \bar{e}_i [30][33].

Die Normalisierung über den Mittelwert hat den Vorteil, dass der RMSE im Verhältnis zum Mittelwert den sogenannten Variationskoeffizient (engl.: Coefficient of variation, CV) bildet. Dieser kann als relative Ungenauigkeit der Modellwerte in Prozent dargestellt werden [23][33]. Auch die Werte des NMAE können als prozentuale Werte interpretiert werden [34]. Der NMBE kann als prozentuale Abweichung des Modells vom mittleren Verlauf betrachtet werden [30].

Für den normalisierten, mittleren, quadratischen Fehler NRMSE, den normalisierten, mittleren, absoluten Fehler NMAE und den normalisierten, mittleren, systematischen Fehler NMBE gelten somit:

$$NRMSE = \frac{RMSE}{\bar{x}_i} \quad \text{oder} \quad NRMSE = \frac{RMSE}{\bar{e}_i} \quad (3.18)$$

$$NMAE = \frac{MAE}{\bar{x}_i} \quad \text{oder} \quad NMAE = \frac{MAE}{\bar{e}_i} \quad (3.19)$$

$$NMBE = \frac{MBE}{\bar{x}_i} \quad \text{oder} \quad NMBE = \frac{MBE}{\bar{e}_i} \quad (3.20)$$

4 Methodisches Vorgehen

In diesem Kapitel wird schrittweise die Vorgehensweise zur statistischen Analyse von unterschiedlichen Simulationsmodellen erläutert. Für die Analyse stehen die Zeitreihen, die von den verschiedenen Simulationsmodellen generiert wurden, zur Verfügung.

Da die verwendeten Simulationsmodelle Unterschiede in ihrem Vorgehen aufweisen, kann es in den erzeugten Zeitreihen zu Unterschieden kommen. Ziel der Methodik ist es, die Flüsse des untersuchten Energiesystems zu ermitteln, bei denen die Modelle unterschiedliche Ergebnisse aufweisen.

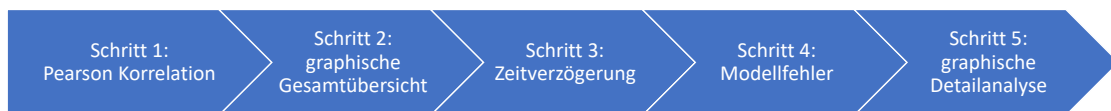


Abbildung 4.1: Methodisches Vorgehen

Die Methodik ist in fünf Schritte eingeteilt (siehe Abbildung 4.1). Der erste Schritt dient dazu einen Überblick über die Zusammenhänge zwischen den Verläufen der Zeitreihen aus den unterschiedlichen Modellen zu bekommen. Dazu wird die Pearson Korrelation verwendet.

Die Unterschiede, die durch diesen Schritt sichtbar werden, sind relative Unterschiede. Da allerdings auch die absoluten Unterschiede zwischen den Modellen in einer Komponente zu Unterschieden bei anderen Komponenten führen können, müssen auch die absoluten Unterschiede betrachtet werden.

Dies geschieht im zweiten Schritt. Hier werden alle Flüsse zwischen den Komponenten graphisch dargestellt. Unterschiede können aus der Graphik abgelesen werden. Zusätzlich können die in diesem Schritt erzeugten Graphiken später auch bei der Interpretation der Unterschiede zwischen den Simulationen verwendet werden, da hier sämtliche Flüsse dargestellt sind.

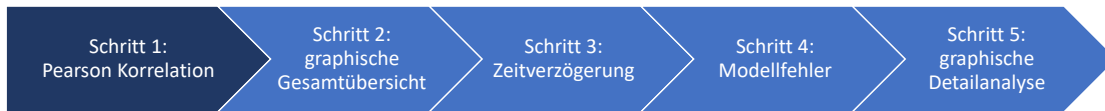
Aus den Schritten 1 und 2 werden die weiter zu betrachtenden Flüsse ausgewählt.

Schritt 3 dient dazu, die erkannten Unterschiede in den ausgewählten Flüssen daraufhin zu untersuchen, ob eine Zeitverzögerung die Ursache für die Unterschiede sein könnte.

Der vierte Schritt hat zwei Funktionen, zum einen soll eine Aussage über die Größe der Abweichung der Modelle zu einem Referenzwert getroffen werden. Zum anderen werden Ausreißer in den Flüssen der jeweiligen Modelle identifiziert. Hierfür werden die Modellfehler herangezogen.

Im letzten und fünften Schritt werden die in Schritt 3 und 4 getroffenen Aussagen graphisch im Detail dargestellt. Die Darstellung dient dazu, die Ausreißer zeitlich einzuordnen und die Ergebnisse der vorherigen Schritte zu analysieren. Die einzelnen Schritte werden im Folgenden näher beschrieben.

4.1 Schritt 1: Pearson Korrelation



Hier wird die in Kapitel 3.1.4 hergeleitete Pearson Korrelation zwischen zwei Simulationsmodellen angewendet. Es wird eine Liste aller Flüsse zwischen den Komponenten aus unterschiedlichen Modellen mit der jeweiligen Pearson Korrelation erstellt. Diese beschreibt das Maß des linearen Zusammenhangs zwischen den Modellen. Es dient dazu, diejenigen Flüsse zwischen den Komponenten ausfindig zu machen, die nicht das gleiche Ergebnis aufweisen. In der Regel kommen für eine genauere Betrachtung Werte infrage, die unter $r \leq 0,7$ liegen, da hier der Grenzfall zwischen einem hohen und sehr hohem Zusammenhang vorliegt [17]. Hierfür wird die Liste der Güte der Pearson Korrelation aus Kapitel 3 Tabelle 3.1 herangezogen.

Die Pearson Korrelation gibt in drei Fällen nicht interpretierbare Ergebnisse heraus. In allen drei Fällen entsteht das Problem dadurch, dass durch Null geteilt wird.

Im ersten Fall kann es vorkommen, dass ein Modell einen Fluss simuliert, welcher in dem anderen Modell nicht simuliert wird.

Der zweite Fall tritt ein, wenn beide Flüsse simuliert werden, aber in beiden Modellen die Zeitreihen mit Nullen gefüllt sind.

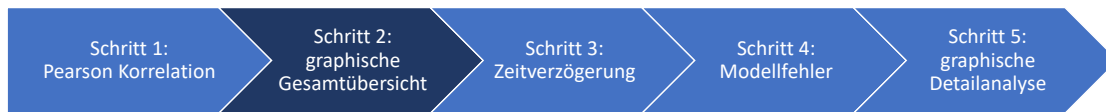
Im dritten Fall werden, wie in Fall Zwei, beide Flüsse simuliert, aber nur von einem der Modelle ein Fluss mit Null simuliert.

Von diesen drei Fällen kann der zweite Fall von einer weiteren Betrachtung ausgeschlossen werden. Da hier beide Modelle den Fluss mit Null simulieren und sie somit keine Unterschiede aufweisen.

Der erste Fall erlaubt keine Vergleichbarkeit der Modelle, da die Komponente nur von einem Modell simuliert wird. Die Information, dass hier ein Unterschied bei den Simulationen aufgetreten ist, wird mitgeteilt, aber in den späteren Schritten nicht weiter behandelt.

Der dritte Fall kann für einen Vergleich herangezogen werden, da beide Modelle den Fluss simulieren, die Ergebnisse sich aber unterscheiden. Sollte Fall Drei eintreten, wird dies mit „ocfz“ (engl.: one component is filled with zeros) abgekürzt.

4.2 Schritt 2: Graphische Gesamtübersicht



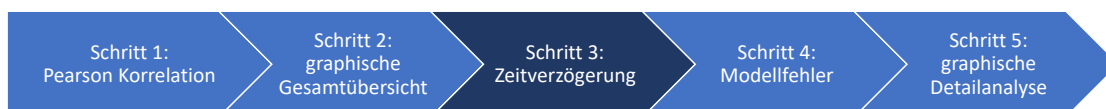
Schritt 2 ist ergänzend zu Schritt 1 zu sehen. Hier werden alle in Schritt 1 aufgeführten Flüsse in Form einer Graphik dargestellt. Jeder einzelne Fluss bekommt sein eigenes Fenster, in dem die Ergebnisse der Simulationsmodelle als geglättete Werte abgebildet sind. Die Glättung erfolgt über den, in Kapitel 3.2 beschriebenen gleitenden Mittelwert. Die Ordnung und somit auch die Stärke der Glättung kann variabel festgelegt werden.

Ein verstärktes Augenmerk sollte auf die Komponenten aus Schritt 1 gelegt werden, die durch ihre nicht vorhandene, bzw. niedrige Korrelation aufgefallen sind. Möglicherweise lässt sich hier schon ihre niedrige lineare Ähnlichkeit erklären.

In Folge findet eine erste graphische Analyse statt, die zwei Funktionen erfüllt. Erstens wird ein Aspekt beleuchtet, der bei der Pearson Korrelation nicht berücksichtigt ist. Mit der Pearson Korrelation wird nur eine relative Gleichheit dargestellt, wohingegen die Graphik absolute Unterschiede sichtbar macht. Als zweite Funktion lässt sich diese Darstellung der Werte als Trendanalyse für das Gesamtsystem benutzen. Die Trendanalyse dient zum allgemeinen Verständnis des Energiesystems. So können kausale Zusammenhänge zwischen den Komponenten schneller erkannt werden.

Für die nachfolgenden Schritte werden nach Schritt 2 die Flüsse ausgewählt, die weiter analysiert werden sollen (vorzugsweise die, mit den niedrigsten Korrelationen).

4.3 Schritt 3: Zeitverzögerung



Hier wird untersucht, ob bei den ausgewählten Flüssen eine Zeitverzögerung zwischen den Modellen vorliegt, da dies eine Ursache für eine niedrige Pearson Korrelation sein kann.

Die Zeitverzögerung wird mit Hilfe der Kreuzkorrelation über ein vorher festgelegtes Zeitfenster ermittelt. Das Zeitfenster begrenzt den zu untersuchenden Abschnitt. Die Größe des Fensters muss zu der Anzahl der untersuchten Zeitschritte passend festgelegt werden. Bei der Auswahl der Fenstergröße sollte immer der in Kapitel 3 in Abbildung 3.2 gezeigte Effekt berücksichtigt werden.

Nach der Festlegung des Zeitfensters kann mit dem Wert der höchsten Pearson Korrelation der Zeitpunkt der Verschiebung bestimmt werden. Wird eine Zeitverzögerung erkannt, sollte zunächst festgestellt werden, ob das Zeitfenster eventuell zu groß gewählt wurde. Es ist zu empfehlen, ein zweites Mal die Funktion anzuwenden, dieses Mal mit einem wesentlich größeren Zeitfenster. Wenn das selbe Ergebnis vorliegt, kann wahrscheinlich

davon ausgegangen werden, dass eine geeignete Fenstergröße gewählt wurde. Falls nicht das selbe Ergebnis vorliegt, sollte das verwendete Fenster verkleinert werden.

Damit eine Zeitverzögerung vorliegt, sollten folgende Kriterien erfüllt sein: Der Wert der Pearson Korrelation der verschobenen, nun unverzögerten Zeitreihe, sollte in einem Bereich von 0,7 bis 1,0 liegen [18]. Darüber hinaus sollte die Differenz der Pearson Korrelation von den ursprünglichen Zeitreihen und der Pearson Korrelation von den verschobenen Zeitreihen ungefähr bei 0,2 liegen.

Diese Rahmenbedingungen sind allerdings nur als vom Autor festgelegte Richtwerte zu verstehen und jeder Fall sollte einzeln betrachtet werden. Ein weiteres Indiz für eine Zeitverzögerung ist gegeben, wenn diese sich durch das System zieht.

4.4 Schritt 4: Modellfehler



In Schritt 4 werden die in Kapitel 3.3 beschriebenen statistischen Fehler zur Analyse der Modelle berechnet. Die Fehler MBE, MAE und der RMSE werden für die betrachteten Flüsse berechnet. Die normalisierte Form der Fehler (NMBE, NMAE und NRMSE) ermöglicht den Vergleich verschiedener Komponenten.

Die Analyse der Modellfehler ist in zwei Bereiche zu unterteilen. Zum einen geht es um die Beziehung zwischen den Modellfehlern untereinander und zum anderen um den Vergleich der Fehler verschiedener Modelle.

Der Vergleich der Fehler untereinander ergibt folgendes: Aufgrund der Ähnlichkeit des NMAE und des NRMSE hinsichtlich ihrer Aussage (beide geben die mittlere Abweichung des Modells zum durchschnittlichen Verlauf aller Modelle an) und der Anfälligkeit des NRMSE für Ausreißer in den Zeitreihen, kann die Differenz vom NRMSE und NMAE betrachtet werden um Ausreißer zu prognostizieren. Wenn die Differenz von NRMSE und NMAE sehr groß ist, deutet das auf Ausreißer in der Zeitreihe hin.

Für NMAE und NRMSE gilt, je näher die Werte an Null Prozent sind, desto genauer bildet das Modell den zu erwartenden Wert ab.

Der NMBE gibt im Vergleich nur die Tendenz des Modells an, den durchschnittlichen Verlauf zu unter- bzw. zu überschätzen.

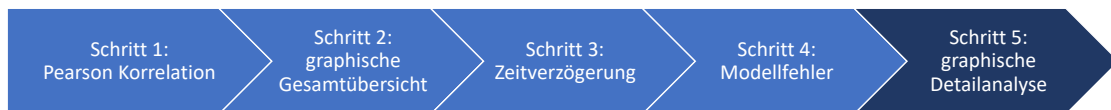
Da jedes Modell den durchschnittlichen Verlauf beeinflusst, ist dies bei der Interpretation der resultierenden Fehler zu berücksichtigen. Bei einer geringen Anzahl von zu vergleichenden Modellen ist der Einfluss eines einzelnen Modells auf den durchschnittlichen Verlauf größer, als bei einer großen Anzahl. So sollten die einzelnen Fehler immer in Relation zueinander gesetzt werden.

Das folgende Beispiel soll dies verdeutlichen: Angenommen wird eine Anzahl von drei Modellen, von denen zwei sich sehr ähneln, das dritte aber extrem abweicht. Das hätte einen spürbaren Einfluss auf den mittleren Verlauf. Die beiden ersten Modelle werden dadurch eine hohe Abweichung von z.B. $NRMSE = 25\%$ zum durchschnittlichen Verlauf

zeigen. Wenn das dritte Modell aber durch ein weniger abweichendes Modell ersetzt würde, wäre der NRMSE geringer. Die Modelle Eins und Zwei haben sich aber nicht verändert.

Dadurch, dass die Fehlerwerte immer von der relativen Lage der Modelle zueinander abhängen, kann kein fester prozentualer Wert für die Fehler angegeben werden, der eine Bewertung der Modelle zulässt. Bei großen Prozentzahlen ist lediglich davon auszugehen, dass sich die Modelle untereinander stark voneinander unterscheiden.

4.5 Schritt 5: Graphische Detailanalyse



In Schritt 5 werden die Flüsse, die über die Schritte 1 und 2 ausgewählt wurden, graphisch detailliert dargestellt.

So wird der durchschnittliche Verlauf, der in Schritt 4 den Fehlern als Referenz dient, gezeichnet. Der Benutzer kann über einen vorweg festgelegten prozentualen Grenzwert (den threshold) zusätzlich die Abweichungen der Modelle, die oberhalb des Grenzwertes liegen, in Ergänzung zum durchschnittlichen Verlauf, zeichnen lassen.

Schritt 5 erfüllt zwei Funktionen.

1. Der komplette ungeglättete Verlauf der Flüsse kann analysiert werden.
2. Wenn nur die Ausreißer untersucht werden sollen, werden die aus Schritt 4 ermittelten NMAE benutzt, um den prozentualen Grenzwert festzulegen, ab dem die Abweichungen der Modelle graphisch dargestellt werden. Da die Differenz zwischen NRMSE und NMAE auf Ausreißer hindeutet, sollten alle Werte, die über den prozentualen Grenzwert hinausgehen, weiter untersucht werden.

5 Fallbeispiele für die Methodik

In diesem Kapitel wird die in Kapitel 4 erläuterte Methodik, anhand eines einfachen und eines komplexen Energiesystems vorgestellt. Die Energiesysteme wurden in Tessif erzeugt und dann jeweils mit oemof und PyPSA simuliert.

Alle Daten, die in das komplexe bzw. einfache Energiesystem eingelesen werden, sind Wetter- bzw. Lastdaten, die den folgenden Quellen entstammen:

1. Wetterdaten für Wind und Solar: [35] [36] [37]
2. Stromlastdaten: [38]
3. Wärmelastdaten: [39]

Das einfache Energiesystem soll auf Unterschiede in drei verschiedenen Zeitabschnitten untersucht werden. Die Untersuchung der drei verschiedenen Zeiträume soll die unterschiedlichen Fälle, die bei der späteren Verwendung dieser Methodik auftreten können, verdeutlichen. Für das einfache Energiesystem treten im ersten Fall (ein Tag) keine Unterschiede auf. Bei einem größeren Zeitraum, wie in Fall 2 (eine Woche), werden geringe Unterschiede zwischen den Modellen sichtbar. Wird der Zeitraum noch erweitert, wie in Fall 3 (ein Monat), treten größere Unterschiede auf.

In dem komplexen Energiesystem soll gezeigt werden, dass sich die Methodik auch bei komplexeren Systemen anwenden lässt. Zudem soll mit diesem System verdeutlicht werden, wie die Methodik anzuwenden ist, wenn eine mögliche Zeitverzögerung auftritt. Das komplexe Energiesystem wird, aufgrund des zu großen Datenvolumens, hier nur in einem Zeitbereich betrachtet.

5.1 Szenario 1: Das einfache Energiesystem

Bei dem einfachen Energiesystem handelt es sich um ein sehr minimalistisches System. Es gibt drei primäre Energiequellen: „wind“, „solar“ und „gas power plant“, dazu kommt die Komponente „imported electricity“. Hier kann Strom von den Modellen importiert (gekauft) werden. Weiter existiert noch eine Komponente „excess“, diese dient zum Export von überschüssiger Energie (Verkauf). Als letztes gibt es noch den „demand“, dieser fungiert als Verbraucher und legt fest, wie viel Energie das System benötigt. Alle Komponenten laufen in der, als Verteiler fungierenden Komponente „powerline“, zusammen.

Bei der Methodik werden nicht die Komponenten, sondern die Flüsse zwischen den Komponenten betrachtet. Abbildung 5.1 stellt das einfache Energiesystem schematisch dar.

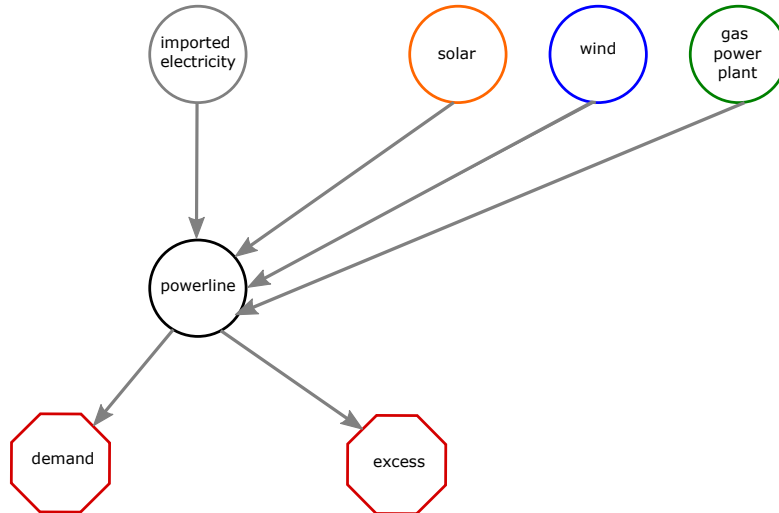


Abbildung 5.1: Graphische Darstellung des einfachen Energiesystems

In vorherigen Tests wurde festgestellt, dass es sehr schwierig ist, in einem einfachen Energiesystem Unterschiede zwischen oemof und PyPSA zu erzeugen. Daher werden durch künstlich festgelegte Rahmenbedingungen, die in die Modelle eingebaut sind, Unterschiede erzwungen.

Infolge dessen verkauft PyPSA Solaranlagen, hingegen baut oemof diese aus. Dieser Unterschied in der Stromerzeugung führt dazu, dass PyPSA andere Quellen für die Stromerzeugung erschließen muss.

5.1.1 Keine Unterschiede (24 Zeitschritte)

Aus dem ersten Schritt der Methodik resultiert Tabelle 5.1. Hier kann an der Pearson Korrelation eine Korrelation von 1 abgelesen werden. Daraus lässt sich ableiten: Es liegen keine Unterschiede vor.

Komponente/ Fluss	Pearson Korrelation
powerline to demand	1
imported electricity to powerline	1
gas power plant to powerline	1
solar to powerline	1
wind to powerline	1

Tabelle 5.1: Szenario 1: Ergebnisse der Pearson Korrelation für 24 Zeitschritte

Für die Betrachtung des zweiten Schritts in Abbildung 5.2 wird das Betrachtungsfenster des gleitenden Mittelwertes auf eine Größe von 10 Schritten festgelegt. Ein kurzer Abgleich bestätigt die aus Schritt 1 abgeleitete These. Somit müssen hier auch keine weiteren Untersuchungen mehr vorgenommen werden.

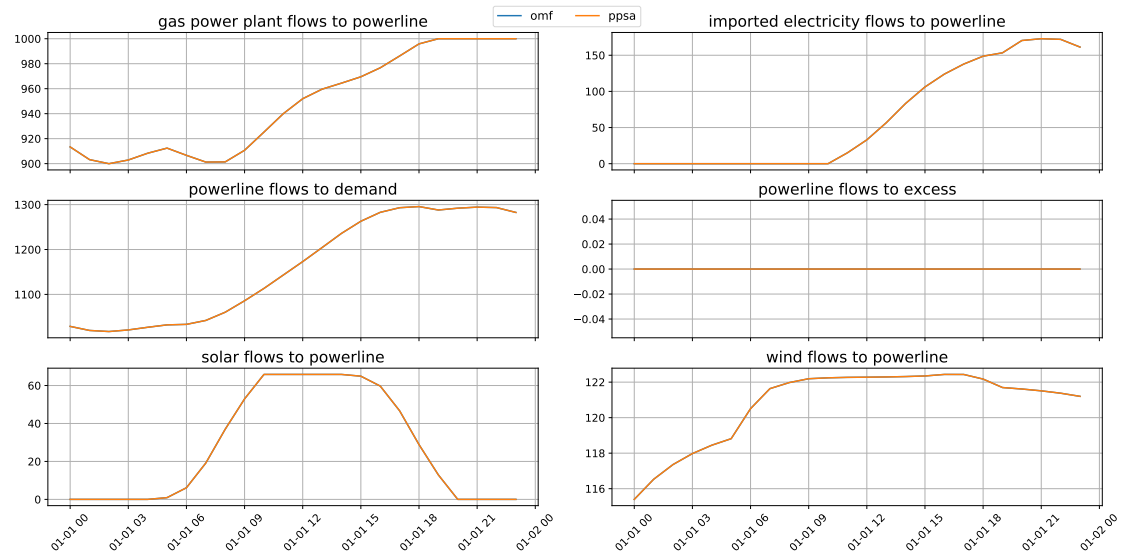


Abbildung 5.2: Szenario 1: Graphische Darstellung aller Komponenten bei 24 Zeitschritten

5.1.2 Kaum Unterschiede (168 Zeitschritte)

Bei der Analyse der Pearson Korrelation in Tabelle 5.2 fallen folgende Flüsse auf: „powerline“ zu „excess“, „imported electricity“ zu „powerline“ und „gas power plant“ zu „powerline“. Besonders auffällig ist der Fluss von „powerline“ zu „excess“ durch das Kürzel „ocfz“.

Hier wird nur von einem Simulationsprogramm Strom exportiert. Die anderen beiden Flüsse weisen mit einer Korrelation von $r \geq 0.7$ eine sehr hohe Ähnlichkeit auf. Laut der in Kapitel 3.1 eingeführten Tabelle, liegt bei einem Wert von $r = 0,7039$ bei „gas power plant“ zu „powerline“ ein Grenzfall zwischen hohem und sehr hohem Zusammenhang vor.

Komponente/ Fluss	Pearson Korrelation
powerline to demand	1
imported electricity to powerline	0.7962
gas power plant to powerline	0.7039
solar to powerline	1
wind to powerline	1
powerline to excess	ocfz

Tabelle 5.2: Szenario 1: Ergebnisse der Pearson Korrelation für 168 Zeitschritte

Im zweiten Schritt wird aufgrund der größeren Anzahl von Zeitschritten das Betrachtungsfenster des gleitenden Mittelwertes verdoppelt. Wie schon in der Einleitung des Szenarios erwähnt, fällt beim Betrachten von Schritt 2 in Abbildung 5.3 sofort auf, dass

oemof viel mehr Solarenergie erzeugt als PyPSA. Die Korrelationen von $r = 1$ bei „wind“ und „solar“ resultieren aus den gleichen eingelesenen Wetterdaten, so dass sich ein identisch linearer Verlauf ergibt. Ähnliches gilt auch für den „demand“ (Verbrauch), der bei beiden Systemen identisch sein muss, um überhaupt eine Vergleichbarkeit herzustellen. Diese Überproduktion sorgt dafür, dass oemof den überschüssigen Strom exportiert. Die Ausschläge der Solarproduktion decken sich mit dem des Energieexports.

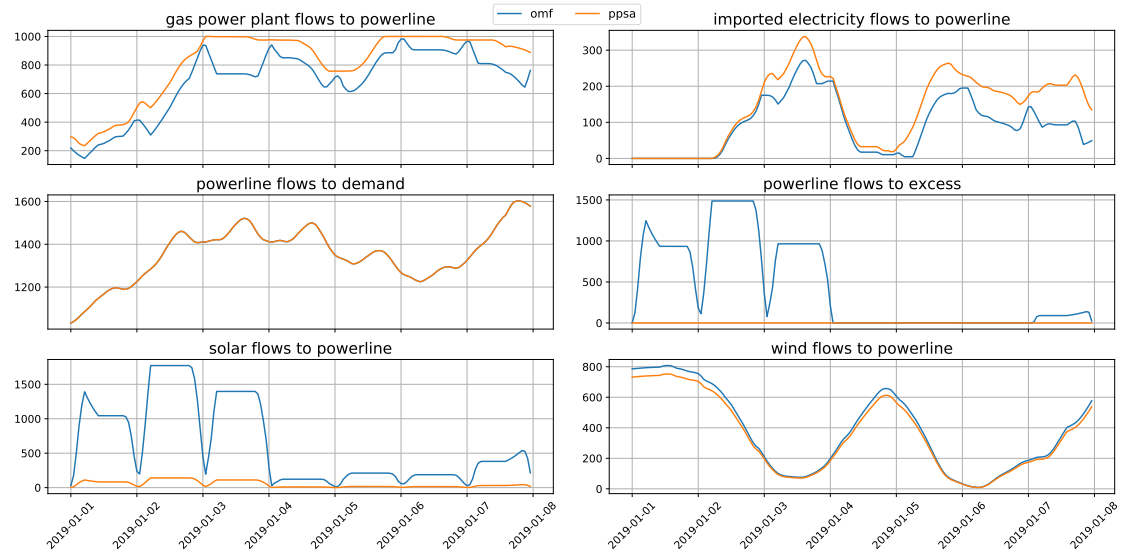


Abbildung 5.3: Szenario 1: Graphische Darstellung aller Komponenten bei 168 Zeitschritten

Bei der Betrachtung des Flusses „imported electricity“ zu „powerline“ ist der Verlauf der Modelle sehr ähnlich und erfordert daher keine weitere Beachtung. Dies deckt sich mit der Pearson Korrelation von fast $r = 0,8$.

Aus der Anschauung des Grenzfalles, „gas power plant“ zu „powerline“ ist ebenfalls eine hohe Übereinstimmung von oemof und PyPSA zu erkennen. Über die Spitzen des Graphen von oemof ist, aufgrund der geglätteten Zeitreihe, ein sehr unruhiger Verlauf dieses Modells zu vermuten. In den folgenden Schritten 3 bis 5 wird deshalb „gas power plant“ zu „powerline“ untersucht.

Die Kreuzkorrelation soll eine Verschiebung von -15 auf +15 Zeitschritte untersuchen. Dies deckt die Verschiebung von ungefähr einem Tag ab.

Das Ergebnis: Es liegt keine Verschiebung vor.

Komponente/ Fluss	Verschiebung	Pearson Korrelation
gas pipeline zu powerline	0	0.7039

Tabelle 5.3: Szenario 1: Ergebnis der Kreuzkorrelation des Flusses „gas power plant“ zu „powerline“ bei 168 Zeitschritten

In Schritt 4 der Modellfehleranalyse fallen mehrere Dinge auf. Zum einen weisen die Fehler der verschiedenen Modelle alle den gleichen Betrag auf, zum anderen sind die Beträge des NMAE und des NMBE gleich. Die Fehler sind in Abbildung 5.4 dargestellt.

Die erste Beobachtung beruht darauf, dass hier nur zwei Modelle vorliegen und somit beide Modelle die gleiche Entfernung zum Bezugswert (dem durchschnittlichen Verlauf beider Modelle) haben.

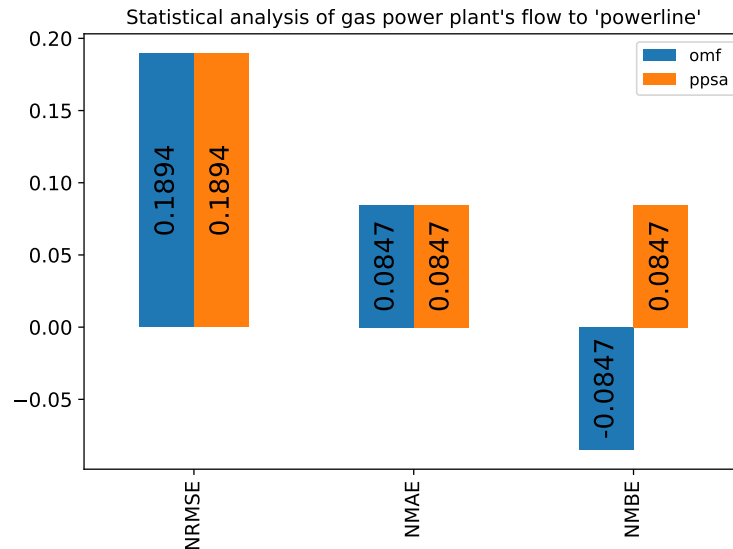


Abbildung 5.4: Szenario 1: Die Modellfehler des Flusses „gas power plant“ zu „powerline“ bei 168 Zeitschritten

Die zweite Beobachtung basiert auf der Beschaffenheit der Daten, die von Tessif ausgegeben werden. So sind die Flüsse zwischen den Komponenten in Zu- und Abströme aufgeteilt. Das bedeutet: die Vorzeichen der Daten in einer Zeitreihe sind immer gleich, immer positiv, bzw. immer negativ. Daraus ergibt sich, dass der Betrag des NMAE gleich dem Betrag des NMBE sein muss. Ist dies nicht der Fall, deutet das auf einen schwerwiegenden Fehler in einem Simulationsmodell hin.

Somit verbindet der NMBE gleichzeitig die Eigenschaft des NMAE. Das bedeutet, dass ein kleiner negativer, bzw. positiver MBE nicht nur die Tendenz des Modells, den Mittelwert zu unter- bzw. zu überschätzen, angibt, es bedeutet auch, dass es genauer abgebildet wird.

Diese Beobachtung gilt für alle Szenarien, die nur mit zwei Modellen durchgeführt werden und bei denen der durchschnittliche Verlauf als Bezugswert der Fehler gewählt wird.

Aus dem NMBE wird geschlossen, dass PyPSA den durchschnittlichen Verlauf überschätzt, oemof ihn unterschätzt.

Der Vergleich des RMSE mit dem NMAE kommt zu dem Ergebnis, dass der NRMSE mehr als doppelt so groß wie der NMAE (19% zu 8%) ist. Das deutet auf Ausreißer in einem der Modelle hin. Die Fehler von 19% bzw. 8% lassen darauf schließen, dass die Modelle, bis auf die oben beschriebenen Ausreißer, wenig voneinander abweichen. Um die prognostizierten Ausreißer zu bestätigen, wird nun Schritt 5 vollzogen.

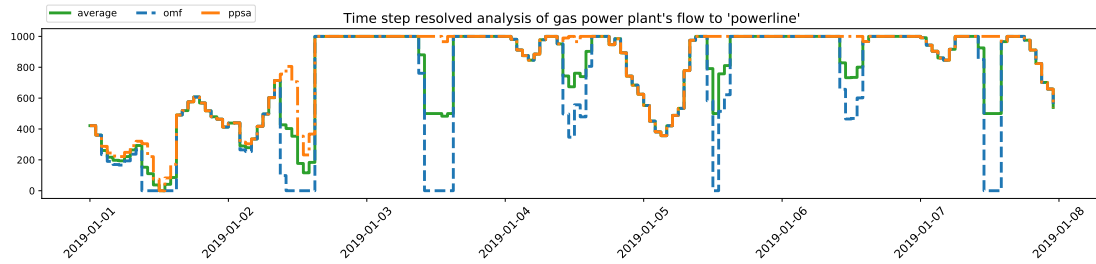


Abbildung 5.5: Szenario 1: Graphische Darstellung des Flusses von „gas power plant“ zu „powerline“ bei 168 Zeitschritten

Für die Auswertung der graphischen Detailanalyse in Abbildung 5.5 wird zunächst die prozentuale Abweichung festgelegt, ab welchem die originalen Werte zum mittleren Verlauf gezeichnet werden sollen. Hier wird dieser Wert auf 8% festgelegt. Dieser Wert wurde dem NMAE entnommen.

Die Betrachtung der detaillierten Zeitreihe zeigt eindeutig die Ausreißer von oemof. Der ähnliche Verlauf der beiden Modelle ist auch sonst gut zu erkennen.

5.1.3 Große Unterschiede (720 Zeitschritte)

Als letztes wird das System bei 720 Zeitschritten analysiert. Bei der Betrachtung der Ergebnisse der Pearson Korrelation in Tabelle 5.4 fallen wieder die gleichen Flüsse der vorherigen Betrachtung mit 168 Zeitschritten auf. Auch hier gibt es eine komplette Übereinstimmung für die Werte „solar“, „wind“ und „demand“. Bei den anderen Werten sinkt lediglich die Korrelation, bzw. zeigt immer noch „ocfz“.

Die neuen Werte der Pearson Korrelation sinken unter $r = 0,7$. Laut Tabelle 3.1 liegt nur noch ein hoher Zusammenhang, statt eines sehr hohen Zusammenhangs, der beiden Modelle vor. Weiter richtet sich in Schritt 2 die Betrachtung auf alle drei Flüsse „powerline“ zu „excess“, „imported electricity“ zu „powerline“ und „gas power plant“ zu „powerline“.

Komponente/ Fluss	Pearson Korrelation
powerline to demand	1
imported electricity to powerline	0.6646
gas power plant to powerline	0.5662
solar to powerline	1
wind to powerline	1
powerline to excess	ocfz

Tabelle 5.4: Szenario 1: Ergebnisse der Pearson Korrelation für 720 Zeitschritte

Aufgrund der nochmals erhöhten Anzahl von Zeitschritten im Vergleich zur Wochenanalyse, wird in Schritt 2 das Betrachtungsfenster des gleitenden Mittelwertes auch hier verdoppelt (von 20 auf 40). Die Betrachtung der Abbildung 5.6 ergibt: Bei „solar“ ist die Differenz zwischen PyPSA und oemof immer noch deutlich. Anders als bei der vorherigen Analyse, sind nun auch bei der Komponente „wind“ deutliche Unterschiede erkennbar.

Wie auch bei 168 Zeitschritten, ist aus dem Fluss „powerline“ zu „excess“ deutlich abzulesen, dass oemof wesentlich mehr Energie als PyPSA produziert und diese exportiert.

Bei der Analyse von „imported electricity“ zu „powerline“ verlaufen die Graphen relativ gleich. Das deckt sich mit der Beobachtung von der Wochenanalyse. Erkennbar sind aber nun die großen Differenzen zwischen den Modellen.

Beim Fluss von „gas power plant“ zu „powerline“ sind die Verläufe zwischen PyPSA und oemof in der ersten Hälfte des Monats ebenfalls sehr ähnlich, variieren aber danach. In der zweiten Hälfte unterliegt oemof stärkeren Schwankungen. Zwischen den Komponenten „gas power plant“ und „wind“ wird ein Zusammenhang deutlich. So decken sich die Ausschläge von „wind“ mit den Stillstandszeiten von „gas power plant“. Auch hier wird bestätigt, dass nur oemof Strom exportiert.

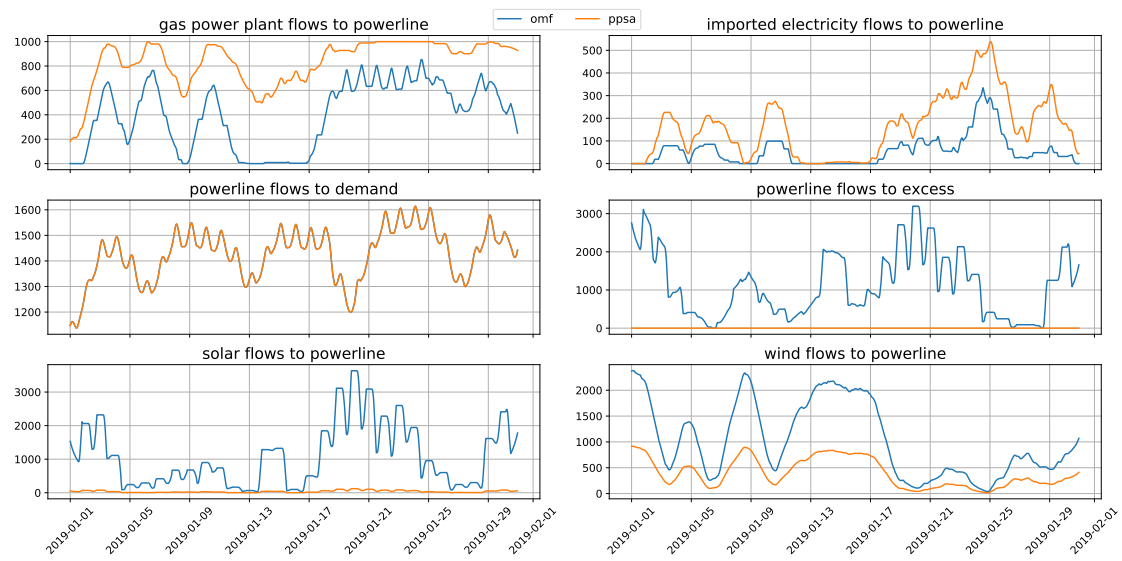


Abbildung 5.6: Szenario 1: Graphische Darstellung aller Komponenten bei 720 Zeitschritten

Die Kreuzkorrelation soll eine Verschiebung von -36 auf +36 Zeitschritte untersuchen. Dies entspricht einem betrachteten Verschiebungszeitraum von drei Tagen. Das Ergebnis der untersuchten Flüsse ist, dass keine Verschiebung vorliegt.

Komponente/ Fluss	Verschiebung	Pearson Korrelation
gas power plant zu powerline	0	0.5662
imported electricity zu powerline	0	0.664

Tabelle 5.5: Szenario 1: Ergebnis der Kreuzkorrelation der Flüsse bei 720 Zeitschritten

Die Fehleranalyse zeigt hier folgende Ergebnisse: Zunächst wird der Fluss von „gas power plant“ zu „powerline“ (siehe Abbildung 5.7) betrachtet. Der NMBE gibt für oemof einen negativen, für PyPSA einen positiven Wert an. Daraus resultiert für oemof eine Unterschätzung des durchschnittlichen Verlaufes und für PyPSA eine Überschätzung. Nun wird der NMAE mit dem NRMSE verglichen. Sie weisen einen Unterschied von 36% zu 46% auf. Aus der Differenz kann man auf eine geringe Anzahl von Ausreißern schließen.

Über die Größe des NRMSE und NMAE lässt sich feststellen, dass die Modelle in ihren absoluten Ergebnissen voneinander sichtbar abweichen.

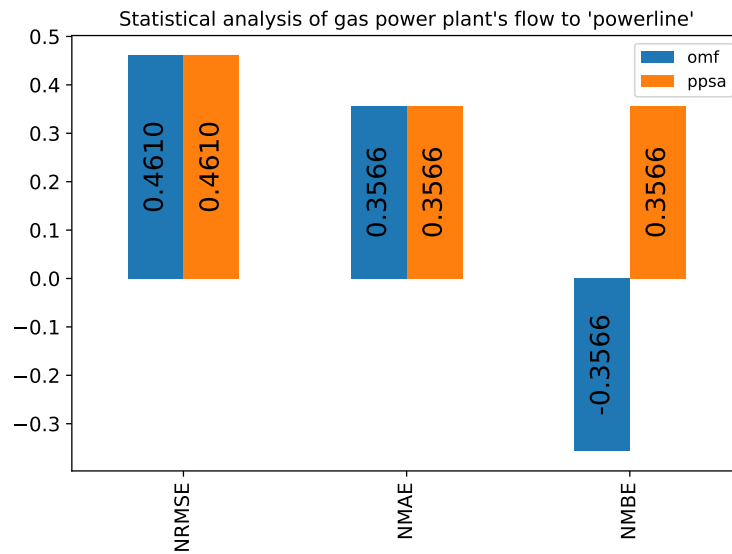


Abbildung 5.7: Szenario 1: Die Modellfehler des Flusses von „gas power plant“ zu „powerline“ bei 720 Zeitschritten

Bei der Analyse von „imported electricity“ zu „powerline“ wird ersichtlich (siehe Abbildung 5.8):

Wie zuvor bei „gas power plant“ zu „powerline“ ist das Ergebnis des NMBE, dass oemof den durchschnittlichen Verlauf unterschätzt und PyPSA diesen überschätzt. Beim Vergleich des NMAE zum NRMSE zeigt sich ein deutlicher Unterschied von 49% zu 85%. Weiter ist aus der Höhe dieser Fehlergrößen abzuleiten, dass die absoluten Werte der Modelle deutlicher voneinander abweichen, als im vorherigen Fluss.

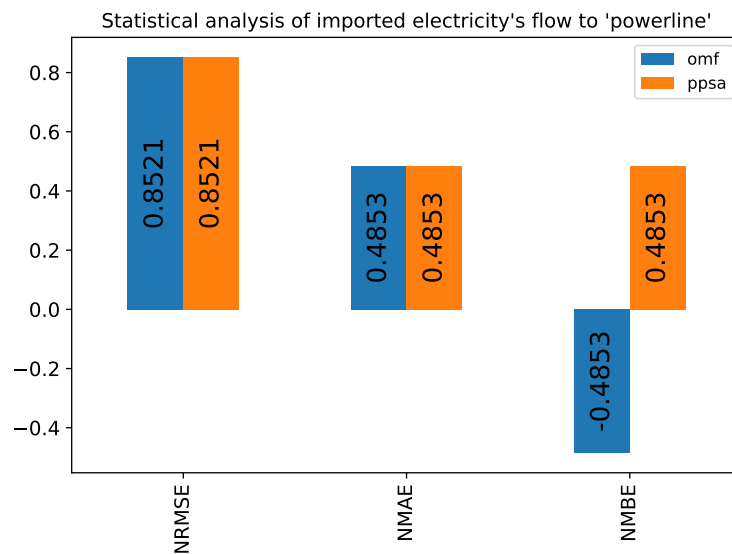


Abbildung 5.8: Szenario 1: Die Modellfehler des Flusses von „imported electricity“ zu „powerline“ bei 720 Zeitschritten

Für Schritt 5 wird zuerst die prozentuale Abweichung festgelegt, ab wann der Graph der Modelle gezeichnet werden soll. Bezugspunkt ist der in Schritt 4 angezeigte NMAE. Der Wert von „gas power plant“ zu „powerline“ ist somit 36% und bei „imported electricity“ zu „powerline“ beträgt er 49%.

Bei der detaillierten graphischen Darstellung in den Abbildungen 5.9 und 5.10 kann abgelesen werden, dass oemof den durchschnittlichen Verlauf unterschätzt, während PyPSA ihn überschätzt. Diese Beobachtung gilt für beide untersuchten Flüsse („imported electricity“ zu „powerline“ und „gas power plant“ zu „powerline“). Weiter ist festzustellen, dass bei „gas power plant“ zu „powerline“ oemof sehr viele Stillstandszeiten hat. Diese erklären den hohen NRMSE bzw. NMAE. Die Stillstandszeiten treten mit einer hohen Konstanz auf, so dass diese kaum noch als Ausreißer bezeichnet werden können. Daher ist die Differenz zwischen NRMSE und NMAE nicht besonders groß.

Im Gegensatz dazu wird bei „imported electricity“ zu „powerline“ diese Differenz sehr deutlich. Die hohen Werte des NRMSE und NMAE lassen sich durch die Art der Normalisierung erklären. In der graphischen Darstellung zeigen sich deutlich lange Stillstandszeiten bei beiden Modellen. Dadurch wird der Mittelwert des durchschnittlichen Verlaufs sehr klein. Wenn, wie hier, die Zeitreihe große Schwankungen aufweist, entstehen solche großen Fehler.

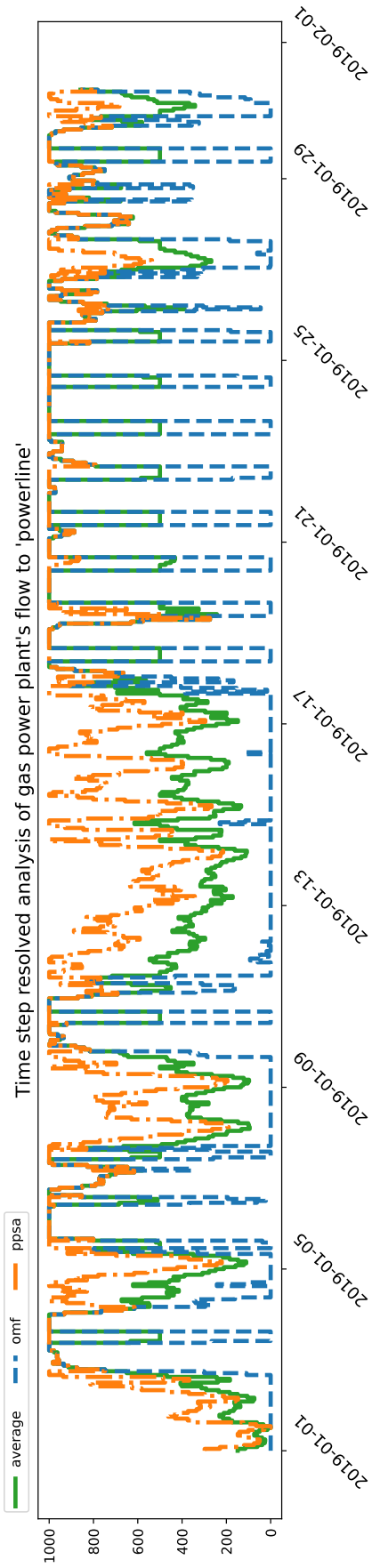


Abbildung 5.9: Szenario 1: Graphische Darstellung des Flusses von „gas power plant“ zu „powerline“ bei 720 Zeitschritten

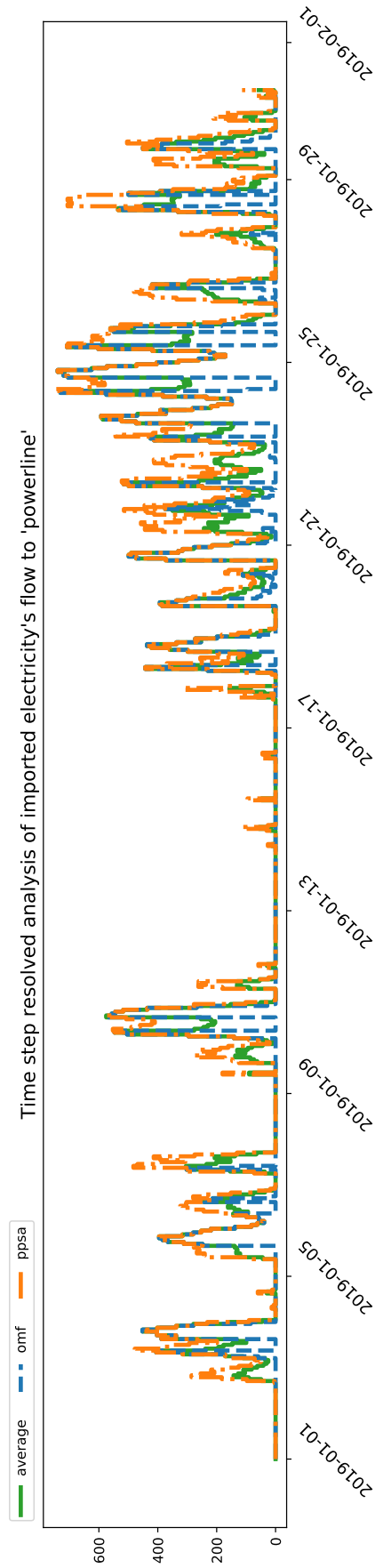


Abbildung 5.10: Szenario 1: Graphische Darstellung des Flusses von „imported electricity“ zu „powerline“ bei 720 Zeitschritten

5.2 Szenario 2: Das komplexe Energiesystem

Dies ist eine erweiterte Version des in Szenario 1 verwendeten einfachen Energiesystems. Die Auswahl der betrachteten Größe des gleitenden Mittelwertes und des Zeitfensters der Verschiebung in Schritt 3 gleichen dem in Szenario 1 untersuchten Fall mit 720 Zeitschritten.

Es wurden folgende Funktionen dem einfachen Energiesystem hinzugefügt: Die Möglichkeit zur Speicherung der elektrischen Energie („electrical storage“), d.h. überschüssige Energie muss nicht mehr exportiert werden. Zusätzlich wird die Gasversorgung des Heizkraftwerks dargestellt („gas supply“ und „gas pipeline“).

Neben der Versorgung mit Elektrizität, kommt bei Szenario 2 thermische Energie (Warmwasser) hinzu. Im Zuge der Erweiterung der thermischen Energie erhält diese auch ein eigenes Energienetz. So laufen die Energieströme in der Komponente „district heating pipeline“ zusammen. Es besteht die Möglichkeit, thermische Energie zu exportieren („excess th“), bzw. zu importieren („imported th“). Darüber hinaus ist es möglich, elektrische Energie in thermische Energie („power to heat“) umzuwandeln. Desweiteren wurde für diese Simulation ein Heizkraftwerk zur reinen Wärmeerzeugung hinzugefügt („heat power plant“).

Das komplexer Energiesystem ist in der folgenden Abbildung 5.11 dargestellt.

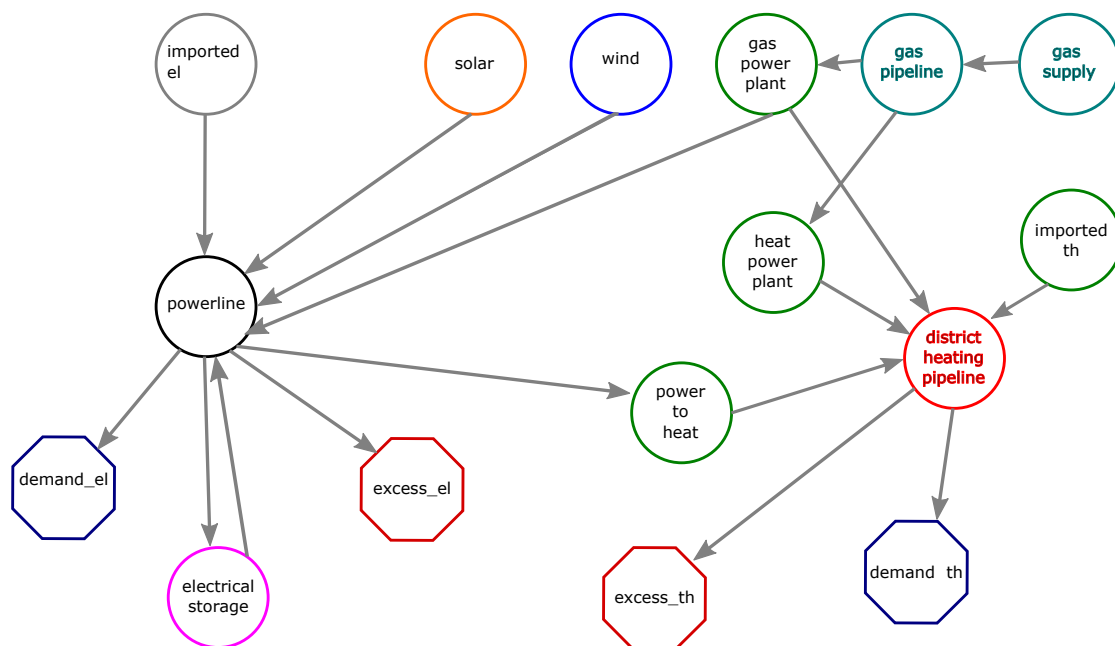


Abbildung 5.11: Graphische Darstellung des komplexen Energiesystems

Die künstlichen Rahmenbedingungen aus Szenario 1 werden in Szenario 2 nicht benötigt, um Unterschiede zu erzwingen. Das Gaskraftwerk „gas power plant“, ist die einzige Komponente, die zur Erzeugung von thermischer und elektrischer Energie genutzt wird. Da die Analyse des kompletten Systems sehr viel umfangreicher ist, wird hier nur eine einzelne Komponente analysiert.

Bei der Untersuchung der Pearson Korrelation in Tabelle 5.6 fällt auf, dass folgende Flüsse eine fast identische Korrelation aufweisen: „gas power plant“ zu „powerline“, „gas power plant“ zu „district heating pipeline“, „gas pipeline“ zu „gas power plant“ und „gas supply“ zu „gas pipeline“.

Zudem treten folgende Werte mit einem Ergebnis von ocfz hervor: „powerline“ zu „electrical storage“, „electrical storage“ zu „powerline“ und „district heating pipeline“ zu „excess th“. Die Methodik signalisiert bei der Analyse, dass PyPSA den Fluss „gas pipeline“ zu „heat power plant“ nicht simuliert. Anscheinend wird von keinem Modell thermische Energie importiert, Strom exportiert und Strom in thermische Energie umgewandelt. Besonders auffällig ist „heat power plant“ zu „district heating pipeline“ mit einer Korrelation von fast Null.

Komponente/ Fluss	Pearson Korrelation
gas power plant to district heating pipeline	0.4531
gas power plant to powerline	0.4531
district heating pipeline to demand th	1
gas pipeline to gas power plant	0.4531
gas supply to gas pipeline	0.4536
heat powerplant to district heating pipeline	-0.0036
powerline to demand el	1
solar to powerline	1
wind to powerline	1
district heating pipeline to excess th	ocfz
electrical storage to powerline	ocfz
powerline to electrical storage	ocfz

Tabelle 5.6: Szenario 2: Ergebnisse der Pearson Korrelation für 720 Zeitschritte

Bei der Analyse aller Komponenten im zweiten Schritt (Abbildung 5.12) wird ersichtlich, welche Ursache die niedrige Korrelation von „heat power plant“ zu „district heating pipeline“ hat. So steigt nur PyPSA gegen Ende des betrachteten Zeitintervalls, ansonsten gibt es noch einen kleinen Anstieg von oemof im mittleren Abschnitt der Zeitreihe. Eine weitere statistische Untersuchung dieses Flusses ist, wegen der nicht vorhandenen Daten, nicht möglich.

Auffällig bei diesem System ist, dass PyPSA viel mehr Strom aus Windenergie („wind“) produziert als oemof. Oemof hingegen priorisiert die Stromproduktion über das Kraftwerk „gas power plant“. Die Wahl der weiter zu betrachtenden Komponente, fällt, nach Ausschluss des Flusses „heat power plant“ zu „district heating pipeline“, auf den Fluss „gas pipeline“ zu „gas power plant“.

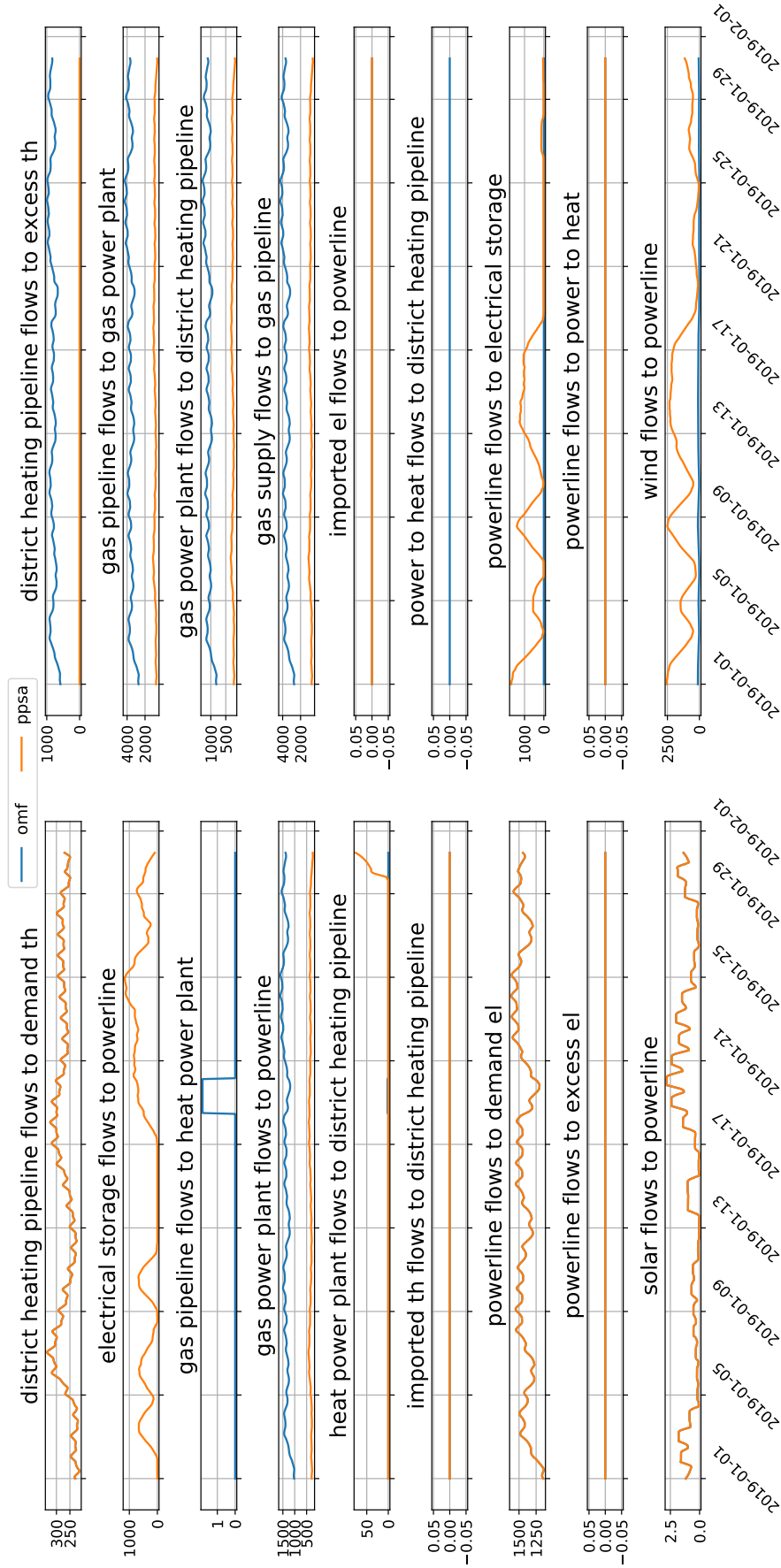


Abbildung 5.12: Szenario 2: Graphische Darstellung aller Komponenten bei 720 Zeitschritten

Die Analyse in Schritt 3 Tabelle 5.7 ergibt eine höhere Korrelation von oemof und PyPSA, wenn die Zeitreihe von PyPSA in positiver Richtung um 26 Zeitschritte verschoben wird. Die neue Pearson Korrelation würde einen Wert von $r = 0,6642$ ergeben. Nun stellt sich die Frage, ob tatsächlich eine Zeitverzögerung vorliegt. In Kapitel 4.3 wurde ein Rahmen festgelegt, welche Ergebnisse vorliegen sollten, damit es sich um eine Zeitverzögerung handelt. Die erste Bedingung, dass der neue Wert um 0.2 Punkte höher sein muss, ist erfüllt. Die zweite Voraussetzung, dass der Wert über $r = 0.7$ liegen sollte, hingegen nicht.

Komponente/ Fluss	Verschiebung	Pearson Korrelation
gas pipeline zu gas power plant	26	0.6642

Tabelle 5.7: Szenario 2: Ergebnis der Kreuzkorrelationen des untersuchten Flusses

Um ein zu groß gewähltes Zeitfenster auszuschließen, wird das zu untersuchende Fenster der Kreuzkorrelation auf 181 Zeitschritte vergrößert. Bei der erneuten Überprüfung kommt die Kreuzkorrelation auf das gleiche Ergebnis.

Bei der Untersuchung der Korrelationswerte ist aufgefallen, dass einige Komponenten eine identische Korrelation aufweisen. So könnten diese auch zeitlich zueinander verschoben sein. Zudem tritt eine zusätzliche Parallele in Erscheinung: Alle Komponenten haben einen Bezug zum Gaskraftwerk „gas power plant“. Dies kann auch anhand der graphischen Darstellung des System in Abbildung 5.11 nachvollzogen werden.

Nach einer kurzen Überprüfung dieser Flüsse (Tabelle 5.8) wird festgestellt, dass auch hier die mögliche Zeitverzögerung bei einer Verschiebung von 26 Zeitschritten vorliegt.

Komponente/ Fluss	Verschiebung	Pearson Korrelation
gas power plant zu district heating pipeline	26	0.6642
gas power plant zu powerline	26	0.6642
gas supply to gas pipeline	26	0.665

Tabelle 5.8: Szenario 2: Ergebnisse der Kreuzkorrelationen der untersuchten Flüsse

In den folgenden Schritten 4 und 5 wird jeweils die originale und die vermutlich unverzögerte Zeitreihe weiter analysiert. Der schon vorher betrachtete Fluss „gas pipeline“ zu „gas power plant“ wird weiter betrachtet.

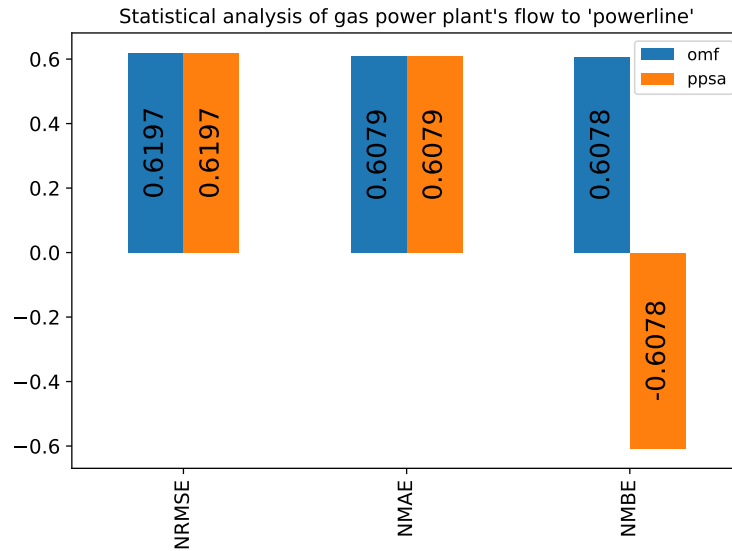


Abbildung 5.13: Szenario 2: Die Modellfehler von „gas pipeline“ zu „gas power plant“ original

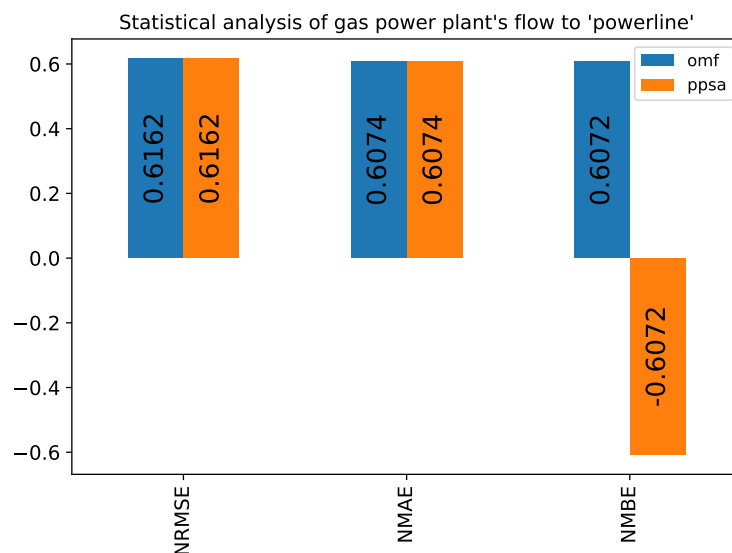


Abbildung 5.14: Szenario 2: Die Modellfehler von „gas pipeline“ zu „gas power plant“ unverzögert

Aus der Analyse der Modellfehler wird schnell ersichtlich, dass die originale (Abbildung 5.14) und die unverzögerte Zeitreihe keine Unterschiede aufweisen. Es hätte durchaus zu Unterschieden kommen können, da die verschobene Zeitreihe um 73 Zeitschritte kürzer ist. Der Fakt, dass die Fehler keinen signifikanten Unterschied aufweisen, lässt auf eine gleichmäßige Verteilung der Daten schließen.

Der NMBE ergibt für den durchschnittlichen Verlauf eine Überschätzung oemofs und eine Unterschätzung PyPSA. Aus dem Abgleich von NRMSE und NMAE lässt sich schließen, dass es keine Ausreißer gibt. Ein NRMSE und NMAE von 62% und 61% lässt auf deutliche Differenzen der Modelle und keine signifikanten Ausreißer schließen.

Bei der Betrachtung von Schritt 5 wird die prozentuale Abweichung auf 0.01% angesetzt, da bei dieser Untersuchung der Abgleich der Verschiebung im Vordergrund steht. Zudem ist die Differenz zwischen NRMSE und NMAE nur ein Prozentpunkt. Somit liegt sehr wahrscheinlich kein Ausreißer vor.

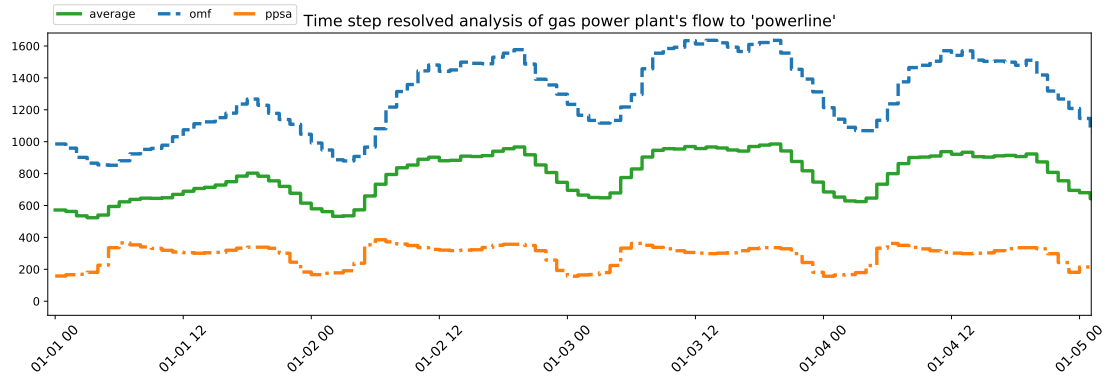


Abbildung 5.15: Szenario 2: Ausschnitt der Abbildung 5.17 originale Zeitreihe

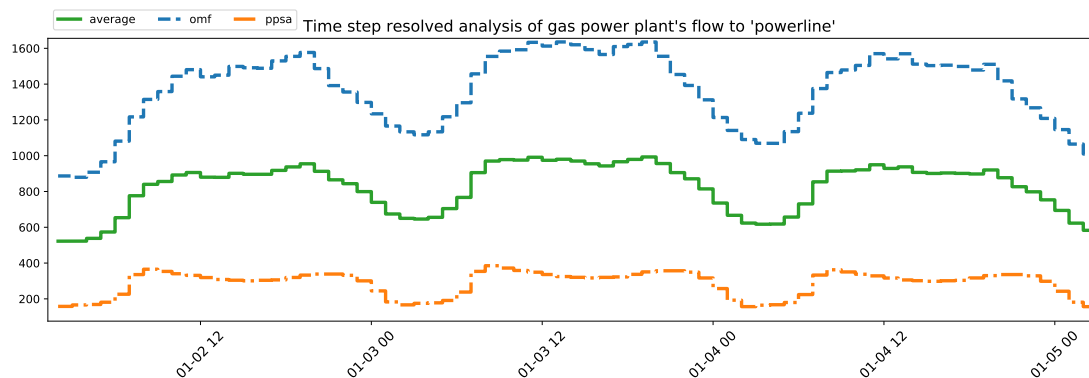


Abbildung 5.16: Szenario 2: Ausschnitt der Abbildung 5.18 unverzögerte Zeitreihe

Beim Vergleich der Abbildung 5.15 mit 5.16 ist zu erkennen, dass die verschobene Zeitreihe eine bessere Übereinstimmung der Modelle aufweist. Bei den Abbildungen handelt es sich um einen vergrößerten Ausschnitt der untersuchten Zeitreihen in den Abbildungen 5.17 und 5.18. Die Analyse der Abbildungen 5.17 und 5.18 zeigt einen sehr periodischen Verlauf der Modelle. Weiter ist ein Ausreißer im Verlauf von oemof im zweiten Drittel der Zeitreihe festzustellen. Dieser wurde nicht von den Fehlern erfasst, da der durchschnittlichen Verlauf nicht signifikant beeinflusst wurde. Ob tatsächlich eine Verschiebung vorliegt, ist nicht genau zu erkennen. Aufgrund des periodischen Verhaltens der beiden Zeitreihen, kann die größere Übereinstimmung auch zufällig entstanden sein. Ein Merkmal spricht zumindest für eine Zeitverzögerung. So zieht sich diese durch alle Komponenten, die mit Gas zu tun haben.

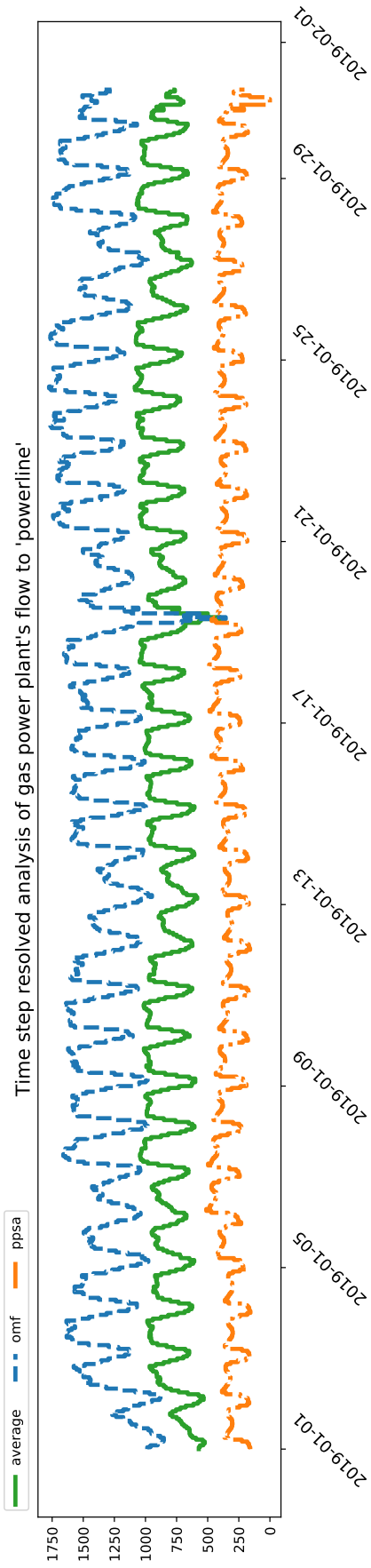


Abbildung 5.17: Szenario 2: Graphische Darstellung des Flusses von „gas pipeline“ zu „gas power plant“ original

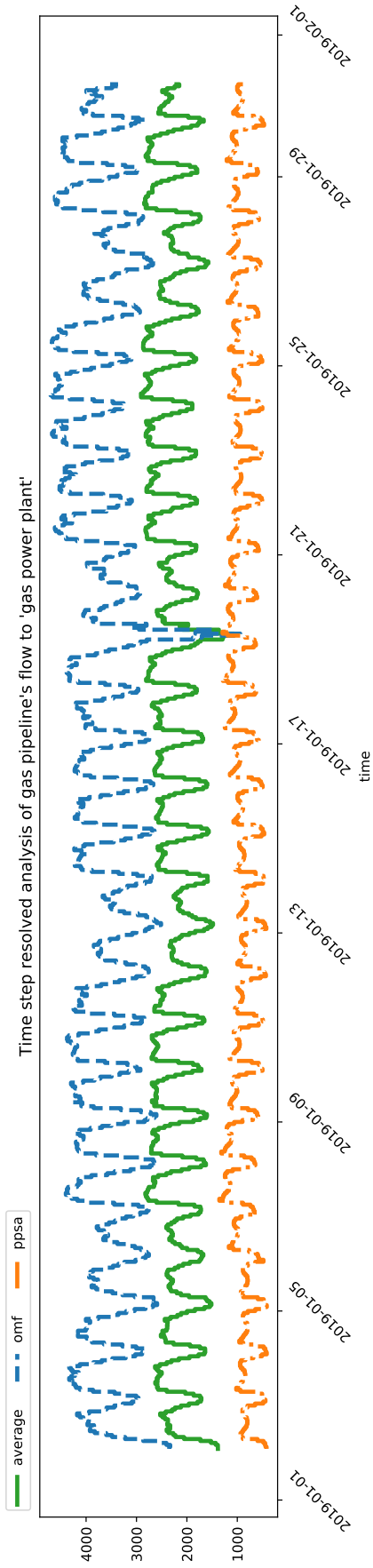


Abbildung 5.18: Szenario 2: Graphische Darstellung des Flusses von „gas pipeline“ zu „gas power plant“ unverzögert

6 Fazit und Ausblick

Fazit

Im Rahmen dieser Arbeit wurde eine allgemeine, statistische Methode entwickelt, mit der Simulationsmodelle in verschiedenen Szenarien verglichen werden können. Die Anwendung der Methodik in den betrachteten Szenarien hat gezeigt, dass diese zuverlässig die Unterschiede der Ergebnisse herausarbeiten kann.

Die Durchführung der Schritte 1 und 2 hat ergeben, dass durch diese Schritte diejenigen Flüsse heraus gefiltert werden konnten, die Unterschiede aufweisen.

Die Schritte haben sich optimal ergänzt. So liefert die Pearson Korrelation in Schritt 1 immer eine leicht zu interpretierende und übersichtliche Liste des linearen Zusammenhangs der betrachteten Modelle.

Schritt 2 ergänzt diese Liste, indem ein Einblick in alle Flüsse des Energiesystems ermöglicht wurde und damit die absoluten Unterschiede sichtbar wurden. Hingegen zeigt die Pearson Korrelation nur relative Änderungen an. So konnte in Szenario 1 beobachtet werden, dass oemof viel mehr Solarenergie produziert als PyPSA, was durch ein alleiniges Betrachten der Pearson Korrelation nicht aufgefallen wäre, da diese einen Wert von $r = 1$ ergab.

Aufgefallen ist, dass auch mit zunehmender Größe der Zeitschritte die Pearson Korrelation kleinere Werte aufweist. Dies liegt an der höheren Wahrscheinlichkeit von potenziellen Ausreißern. Eine mögliche Lösung dieses Problems wird im Ausblick erörtert.

Auch bei der Betrachtung des komplexen Systems zeigt sich, dass trotz der größeren Anzahl an Komponenten und der steigenden Komplexität, die Pearson Korrelation ihre Übersichtlichkeit behält. In Schritt 2 wird es mit zunehmender Komplexität der Systeme aufwendiger Zusammenhänge abzuleiten.

Schritt 3 wurde konzipiert, um niedrige Pearson Korrelationen durch mögliche Zeitverzögerungen zu erklären.

Die in Szenario 2 aufgetretene, vermutliche, Zeitverzögerung wurde erkannt. Ob tatsächlich eine reale Zeitverzögerung vorliegt, ist nicht die Aufgabe der Methodik, lediglich eine mögliche zu erkennen. Abschließend ist festzustellen: Die Modelle stimmen besser überein, wenn die erkannte Verschiebung berücksichtigt wird.

Die Schritte 4 und 5 konnten bei der Analyse der Szenarien nicht ihre volle Funktion erfüllen. So konnte in Schritt 4 keine Differenzierung der Abbildungsgenauigkeit der Modelle auf den durchschnittlichen Verlauf vollzogen werden.

Die Begründung liegt in der Tatsache, dass die Modellfehler ihren Bezug auf den durchschnittlichen Verlauf als Referenz nehmen. Bei der Betrachtung zweier Modelle sind diese immer gleich weit vom Mittelwert entfernt.

Es wurden lediglich modellbezogene Aussagen getroffen. Der Abgleich des MAE mit dem RMSE lieferte eine Aussage über Ausreißer in den Zeitreihen.

Über die Größe der Fehler konnte nur abgeleitet werden, wie weit die beiden Modelle sich voneinander unterscheiden.

Der Spezialfall des MBEs gibt hier eine zusätzliche Methode, um einen möglichen Fehler auszuschließen. Ansonsten erfüllte dieser Fehler als einziger seine Funktion vollständig. Er zeigte auf, welches der Modelle den mittleren Verlauf über- bzw. unterschätzt.

Bei Schritt 5 zeigte sich ein ähnlicher Sachverhalt. Hier konnten die Zeitreihen sehr gut im Detail analysiert werden. Bei der Betrachtung der Zeitverzögerung erwies sich Schritt 5 als besonders hilfreich, weil hier die Verschiebung zu erkennen war. Die Darstellung der Filterung der Ausreißer über einen prozentualen Grenzwert funktioniert, ist aber bei der Verwendung von zwei Simulationsmodellen unübersichtlich. Wenn ein Ausreißer eines Modells auftritt, wird auch immer das Modell ohne Ausreißer abgebildet.

Abschließend ist zu sagen, dass die in den Schritten 4 und 5 auftretenden Probleme gelöst werden, wenn ein weiteres Modell mit in die Betrachtung einbezogen wird. Die in dieser Arbeit entwickelte Methode hat zuverlässig alle Differenzen der Modelle erkannt.

Ausblick

Die vorgestellte Methodik wurde für den Vergleich mehrerer Modelle entwickelt. In diesem Abschnitt soll ein Ausblick gegeben werden, wie die Methodik anzuwenden ist, wenn drei oder mehr Modelle betrachtet werden.

Die Beobachtung, dass die Pearson Korrelation mit zunehmenden Zeitschritten kleinere Werte aufweist, könnte über eine Glättung der Daten oder das Entfernen der Ausreißer durch den vorherigen Wert gelöst werden. Dies hätte aber auch einen Qualitätsverlust der Aussagekraft zur Folge. Da nun nicht mehr die originalen Daten vorlägen und da in einem Energiesystem alle Komponenten zusammenwirken, könnten über eine solche Glättung der Zeitreihen wichtige Informationen aus dem Vergleich entfallen. Ein Beispiel für derartige Informationen wären Stillstandszeiten einer Komponente. Dennoch wäre es empfehlenswert, eine solche Funktion zur Glättung zu implementieren.

Mit der zusätzlichen Betrachtung eines dritten Modells würde die Interpretation der Ergebnisse der Pearson Korrelation erschwert werden, da diese immer nur zwei Modelle miteinander vergleichen kann. So müssten bei einer Anzahl von m Modellen, eine Anzahl PE von Pearson Korrelationen durchgeführt werden, um alle Modelle und Korrelationen abzudecken. Die Anzahl an Pearson Korrelationen lässt sich mit folgender Formel berechnen:

$$PE = \sum m - 1 \quad (6.1)$$

Bei drei Modellen müssten drei Korrelationstabellen betrachtet und interpretiert werden, bei vier Modellen wären es schon sechs.

Eine mögliche Lösung dieser Problematik wäre denkbar, indem ein Referenzmodell bestimmt würde. Dies würde die Menge der zu betrachtenden Tabellen auf die Anzahl der Modelle bzw. die Anzahl der Modelle minus eins reduzieren. Weil man die Modelle nicht mehr untereinander vergleicht, sondern immer mit dem Referenzmodell, lassen sich die resultierenden Tabellen zu einer einzigen übersichtlichen Tabelle zusammenfassen.

Für die Auswahl eines Referenzmodells gibt es mehrere Möglichkeiten. So könnte in Tessif ein globales Referenzmodell z.B. oemof festgelegt werden, welches eine unveränderbare, feste Bezugsgröße darstellt. Dies hätte den Vorteil, dass alle mit Tessif angefertigten Arbeiten und Studien untereinander einfacher verglichen werden können. Probleme könnten allerdings entstehen, wenn Arbeiten für spezifische Zwecke angefertigt werden.

So eignet sich PyPSA wesentlich besser als Referenzmodell für eine Simulation von Stromnetzen als oemof, da PyPSA eigens für solche Simulationen entwickelt wurde.

Die nächsten Möglichkeit eines Referenzmodells wäre ein vom Nutzer frei wählbares Referenzmodell zu verwenden. So kann der Nutzer selbst über das Referenzmodell entscheiden.

Die dritte Möglichkeit eines Referenzmodells wurde in dieser Arbeit bereits in Kapitel 3.3 vorgestellt: Die Bildung eines mittleren Energiesystemmodells, also aus allen Modellen ein mittleres Modell zu erzeugen. Ein Schwachpunkt bei diesem Referenzmodell ist die Implementierung neuer Modelle. Das würde das mittlere Modell immer wieder verändern und somit die Vergleichbarkeit von Arbeiten, die mit Tessif angefertigt werden, erschweren.

Die letzte und einfachste Option wäre es, echte Daten zu verwenden. Dies ist die beste und sollte immer in Betracht gezogen werden, wenn die Möglichkeit dazu besteht.

Die Referenzmodelle lassen sich in zwei Kategorien aufteilen: Ein internes Referenzmodell, das aus den vorhandenen Modellen ausgewählt und ein externes Referenzmodell, bei dem es sich um ein von außen hinzugefügtes Modell (echte Werte, mittleres Energiesystemmodell) handelt. Die Auswahl eines Referenzsystems hätte dementsprechend auch Auswirkungen auf die Schritte 4 und 5, wo momentan als Referenz der durchschnittliche Verlauf aller Modelle verwendet wird. Falls eine andere Referenz als diese festgelegt wird, empfiehlt es sich auch hier, diese neue Referenz in den beschriebenen Schritten anzuwenden.

So würden Schritt 4 und 5 auch an Relevanz gewinnen, wenn mehr als zwei Modelle analysiert werden. Wie schon im Fazit festgestellt, war es nicht möglich, die Modelle anhand ihrer unterschiedlichen Fehler zu bewerten, da diese immer den gleichen Wert des selben Fehlers ergaben.

Im folgendem Beispiel soll noch einmal verdeutlicht werden, wie die Betrachtung eines Flusses mit drei Modellen aussehen könnte.

So wird ein Fluss mit folgenden Werten gegeben:

Modell 1	Modell 2	Modell 3
500	510	490
550	545	505
525	530	550
490	480	490
480	459	470

Tabelle 6.1: Beispiel eines zu betrachtenden Flusses

Eine kurze Analyse der Modellfehler in Abbildung 6.1 ergibt: Modell 1 und Modell 2 überschätzen den Mittelwert und Modell 3 unterschätzt ihn. So wird nun der Vergleich

der Modelle in den einzelnen Fehlern möglich. Die Modelle bilden den mittleren Verlauf in absteigender Reihenfolge gut ab. Dieser Trend ist sowohl für den NMAE, als auch NRMSE zu beachten. Die Unterschiede zwischen Modell 1 und Modell 2 sind aber minimal.

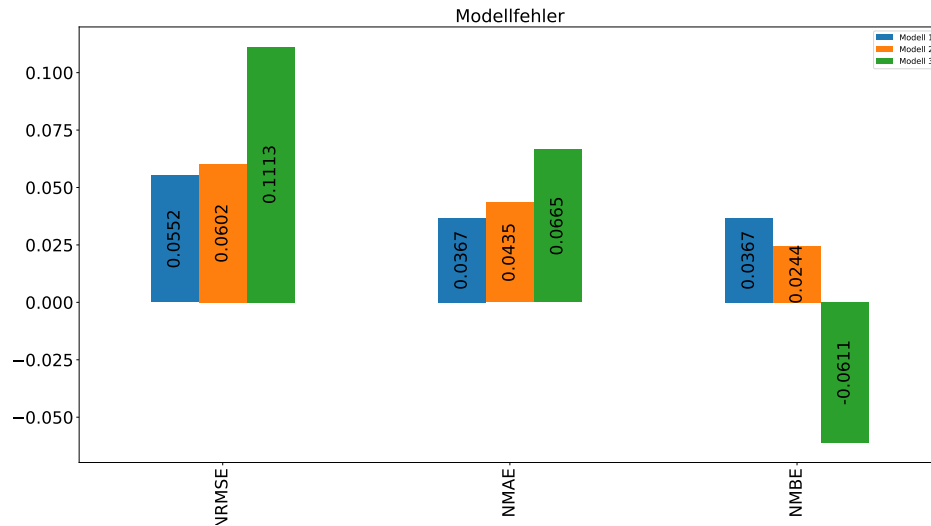


Abbildung 6.1: Die Modellfehler des Beispiels

Bei der Betrachtung des NRMSE und des NMAE fällt auf, dass Modell 1 und Modell 2 ziemlich ähnliche Werte von 5,52% bzw. 6,02% für den NRMSE aufweisen. Auch die Werte des NMAE unterscheiden sich nur um einen Prozentpunkt. Modell 3 hingegen fällt ziemlich aus dem Muster mit einem NRMSE von 11,13% und NMAE 6,65% .

Über das Verhältnis der Fehlergrößen zueinander ist klar zu erkennen, dass Modell 3 stärker vom durchschnittlichen Verlauf abweicht als die Modelle 1 und 2. Der Vergleich des NMAE und NRMSE ergibt, dass Modell 3 einen Ausreißer besitzt, Modell 1 und 2 aber nicht.

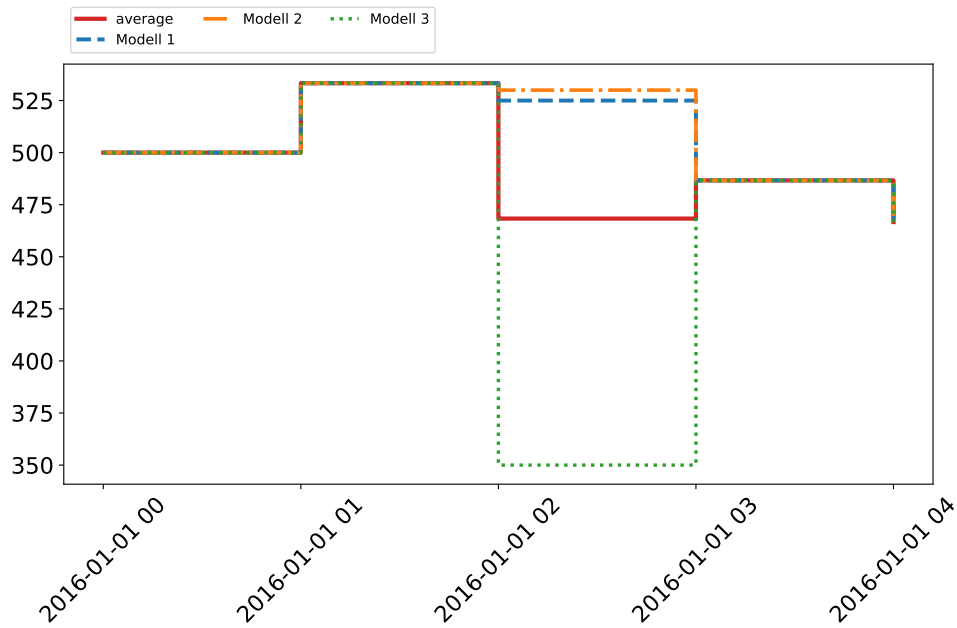


Abbildung 6.2: Graphische Darstellung des Beispiels

Für die in Schritt 5 generierte, graphische Darstellung in Abbildung 6.2 werden Abweichungen ab 7% gezeichnet. Dies entspricht dem Wert des NMAE der Modelle 1 und 2.

Die Ergebnisse des NMBE werden bestätigt und auch der Ausreißer von Modell 3 ist deutlich zu sehen.

In dem Beispiel wird verdeutlicht, dass die Aussagekraft der Schritte 4 und 5 steigt, wenn ein zusätzliches Modell in die Betrachtung einbezogen wird. So kann ein Modell mit großen Abweichungen über die Fehler in Schritt 4 identifiziert werden. In Schritt 5 wird die Übersichtlichkeit verbessert, da nun nicht mehr die Ausreißer den durchschnittlichen Verlauf so stark beeinflussen, wie es bei zwei verwendeten Modellen der Fall ist.

Die Umsetzung der im Ausblick vorgestellten Punkte, wie die Wahl eines Referenzmodells oder eine Möglichkeit die Datensätze für die Berechnung der Pearson Korrelation zu glätten, würden die von der Methodik gelieferten Analysen in ihrer Aussagekraft verbessern und somit den Nutzen weiter steigern.

Literatur

- [1] S. PFENNINGER, A. HAWKES und J. KEIRSTEAD: „Energy systems modeling for twenty-first century energy challenges“. In: *Renewable and Sustainable Energy Reviews* 33 (2014), S. 74–86.
- [2] S. OBERLE und R. ELSLAND: „Are open access models able to assess today’s energy scenarios?“ In: *Energy Strategy Reviews* 26 (2019), S. 49–63.
- [3] R. BEHN und S. BYFIELD: *Acatech: Consulting with energy scenarios*. 2016. URL: <https://energiesysteme-zukunft.de/en/position-paper/consulting-with-energy-scenarios> (aufgerufen am 04.10.2021).
- [4] MATHIAS AMMON: „*Transforming Energy System Simulations Framework*“. *Dokumentation (unveröffentlicht)*, Institut für Energietechnik, Technische Universität Hamburg, 2021.
- [5] S. HILPERT et al.: „The Open Energy Modelling Framework (oemof) - A new approach to facilitate open science in energy system modelling“. In: *Energy Strategy Reviews* 22 (2018), S. 16–25.
- [6] GITHUB: *oemof community*. 2021. URL: <https://github.com/oemof> (aufgerufen am 04.10.2021).
- [7] T. BROWN, J. HÖRSCH und D. SCHLACHTBERGER: „PyPSA: Python for Power System Analysis“. In: *Journal of Open Research Software* 6 (2018).
- [8] T. BROWN: *PyPSA*. 2021. URL: <https://pypsa.org/> (aufgerufen am 04.10.2021).
- [9] J.-P. KREISS und G. NEUHAUS, Hrsg.: *Einführung in die Zeitreihenanalyse*. Statistik und ihre Anwendungen. Berlin und Heidelberg: Springer, 2006.
- [10] W. KOHN und R. ÖZTÜRK: „Mittelwert“. In: *Statistik für Ökonomen*. Hrsg. von W. KOHN und R. ÖZTÜRK. Springer-Lehrbuch. Berlin und Heidelberg: Springer Gabler, 2013, S. 35–41.
- [11] U. KUCKARTZ et al., Hrsg.: *Statistik: Eine verständliche Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden, 2010.
- [12] R.-D. HILGERS, N. HEUSSEN und S. STANZEL: „Standardabweichung“. In: *Lexikon der Medizinischen Laboratoriumsdiagnostik*. Hrsg. von A. M. GRESSNER und T. ARNDT. Springer Reference Medizin. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 2206.
- [13] R.-D. HILGERS, N. HEUSSEN und S. STANZEL: „Ausreißer, statistischer“. In: *Lexikon der Medizinischen Laboratoriumsdiagnostik*. Hrsg. von A. M. GRESSNER und T. ARNDT. Springer Reference Medizin. Berlin, Heidelberg: Springer Berlin Heidelberg, 2019, S. 234.
- [14] N. HENZE: „Kovarianz und Korrelation“. In: *Stochastik für Einsteiger*. Hrsg. von N. HENZE. Wiesbaden: Springer Fachmedien Wiesbaden, 2018, S. 166–178.

- [15] W. KOHN und R. ÖZTÜRK: „Kovarianz und Korrelationskoeffizient“. In: *Statistik für Ökonomen*. Hrsg. von W. KOHN und R. ÖZTÜRK. Springer-Lehrbuch. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2011, S. 91–95.
- [16] K. SIEBERTZ, D. VAN BEBBER und T. HOCHKIRCHEN: „Korrelationsanalyse“. In: *Statistische Versuchsplanung*. Hrsg. von K. SIEBERTZ, D. VAN BEBBER und T. HOCHKIRCHEN. VDI-Buch. Berlin, Heidelberg: Springer Vieweg, 2017, S. 381–394.
- [17] U. KUCKARTZ et al.: „Korrelation: Zusammenhänge identifizieren“. In: *Statistik*. Hrsg. von U. KUCKARTZ et al. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden, 2010, S. 189–213.
- [18] TU-FREIBURG: *DIE KREUZKORRELATION*. 2021. URL: <https://tu-freiberg.de/sites/default/files/media/institut-fuer-mechanik-und-fluidynamik-15832/Lehre/lehrveranstaltungen/fluid/MT/kreuzkorrelation.pdf> (aufgerufen am 04.10.2021).
- [19] M. MEYER: *Signalverarbeitung Anhang: B: Analoge und digitale Signale, Systeme und Filter*. 7., verb. Aufl. Wiesbaden: Springer, 2014.
- [20] J. A. GUBNER: „Introduction to random processes“. In: *Probability and Random Processes for Electrical and Computer Engineers*. Hrsg. von J. A. GUBNER. Cambridge: Cambridge University Press, 2006, S. 383–442.
- [21] P. P. ECKSTEIN: „Zeitreihenanalyse“. In: *Statistik für Wirtschaftswissenschaftler*. Hrsg. von P. P. ECKSTEIN. Lehrbuch. Wiesbaden: Gabler, 2010, S. 341–380.
- [22] R. J. HYNDMAN: „Moving Averages“. In: *International Encyclopedia of Statistical Science*. Springer, Berlin, Heidelberg, 2011, S. 866–869.
- [23] C. BARROT: „Prognosegütemaße“. In: *Methodik der empirischen Forschung*. Hrsg. von S. ALBERS et al. Wiesbaden: Gabler, März 2007, S. 417–430.
- [24] C. J. WILLMOTT: „Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance“. In: *Climate Research* 30.1 (2005), S. 79.
- [25] H.-F. ECKEY, R. KOSFELD und M. TÜRCK: *Wahrscheinlichkeitsrechnung und Induktive Statistik: Grundlagen - Methoden - Beispiele*. Wiesbaden: Gabler, 2005.
- [26] J. A. GUBNER: „Introduction to discrete random variables“. In: *Probability and Random Processes for Electrical and Computer Engineers*. Hrsg. von J. A. GUBNER. Cambridge: Cambridge University Press, 2006, S. 63–107.
- [27] C. J. WILLMOTT: „Some Comments on the Evaluation of Model Performance“. In: *Bulletin of the American Meteorological Society* 63.11 (1982), S. 1309–1313.
- [28] G. RUIZ und C. BANDERA: „Validation of Calibrated Energy Models: Common Errors“. In: *Energies* 10.10 (2017), S. 1587.
- [29] R. PAL: „Chapter 4 - Validation methodologies“. In: *Predictive modeling of drug sensitivity*. Hrsg. von R. PAL. London, United Kingdom: Academic Press, 2016], 2017, S. 83–107.
- [30] J. GRANDERSON et al.: *Assessment of Automated Measurement and Verification (M&V) Methods*. Office of Scientific and Technical Information (OSTI), 2015.
- [31] C. CZADO und T. SCHMIDT: „Vergleich von Schätzern: Optimalitätstheorie“. In: *Mathematische Statistik*. Hrsg. von C. CZADO und T. SCHMIDT. Statistik und ihre Anwendungen. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2011, S. 103–137.

-
- [32] T. CHAI und R. R. DRAXLER: „Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature“. In: *Geoscientific Model Development* 7.3 (2014), S. 1247–1250.
- [33] AMERICAN SOCIETY OF HEATING, REFRIGERATING AND AIR CONDITIONING ENGINEERS: *Ashrae Guideline 14-2014: Measurement of Energy, Demand and Water Savings*. ASHRAE guideline. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, 2014.
- [34] IÑIGO PAGOLA et al.: „A Combination of Heliosat-1 and Heliosat-2 Methods for Deriving Solar Radiation from Satellite Images“. In: *Energy Procedia* 57 (2014).
- [35] *Atlite: Convert weather data to energy systems data — atlite 0.2.6.dev8+g228530b documentation*. 2021. URL: <https://atlite.readthedocs.io/en/latest/index.html> (aufgerufen am 04.10.2021).
- [36] *ERA5-Land hourly data from 1950 to present*. 2021. URL: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview> (aufgerufen am 04.10.2021).
- [37] *Aktuelle Einheitenübersicht | MaStR*. 2021. URL: <https://www.marktstammdatenregister.de/MaStR/Einheit/Einheiten/OeffentlicheEinheitenuebersicht> (aufgerufen am 04.10.2021).
- [38] *Energieportal Hamburg*. 2021. URL: <https://www.energieportal-hamburg.de/> (aufgerufen am 04.10.2021).
- [39] TOBIAS ZIMMERMANN: „*Beitrag des Wärmesektors zur Reduzierung der Emissionen in sektorgekoppelten Energiesystemen*“. *Dissertation (unveröffentlicht)*. Institut für Energietechnik, Technische Universität Hamburg, 2021.