

Übung 1 – Regressionsanalyse

VU Energiemodelle und Analysen
Theresia Perger

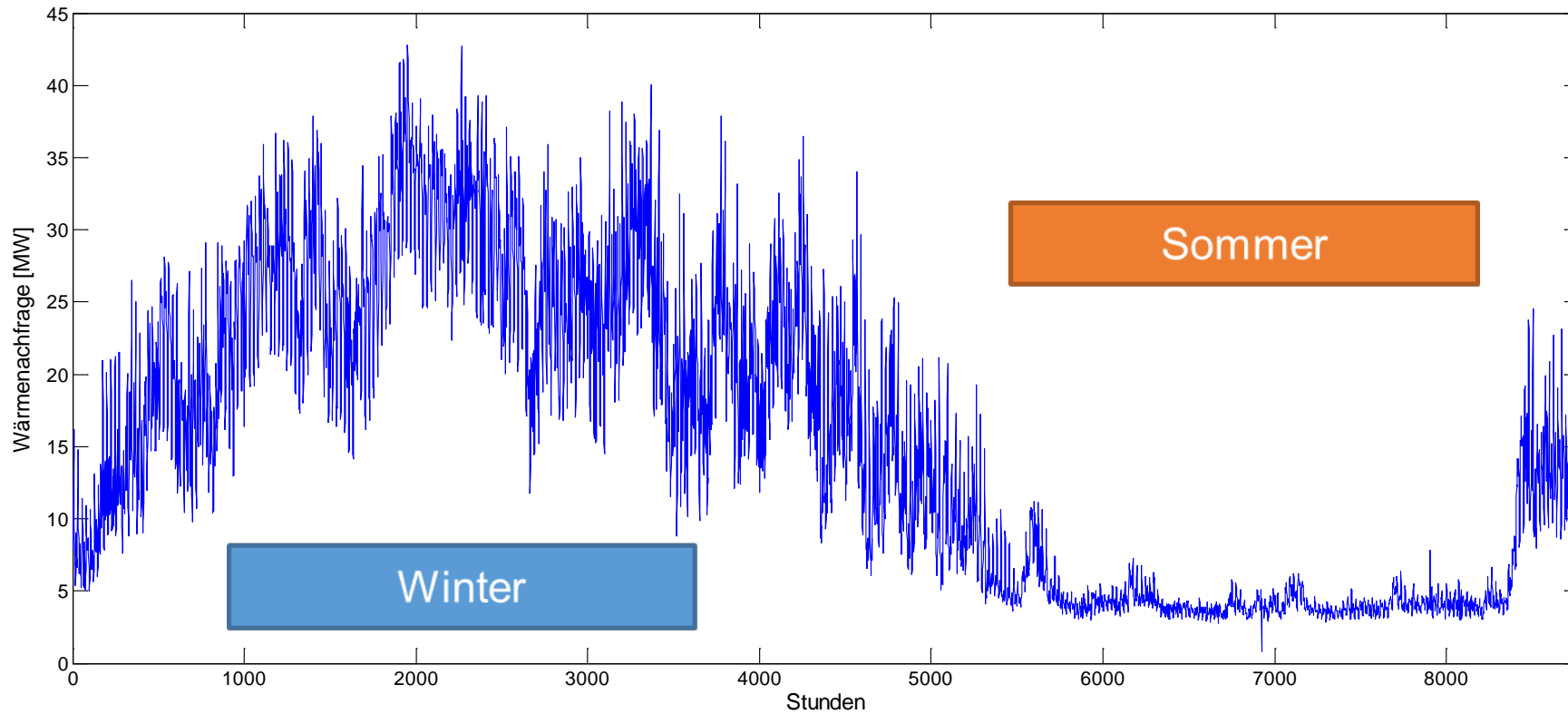
- Aufgabe 1.1 – Modellierung der Wärmenachfrage eines Fernwärmenetzes mittels linearer Regression
- Aufgabe 1.2 – Modellierung des Strompreises mittels linearer Regression

Aufgabe 1.1 – Modellierung der Wärmenachfrage eines Fernwärmenetzes mittels linearer Regression

Aufgabe 1.1 – Modellierung der Wärmenachfrage eines Fernwärmenetzes mittels linearer Regression

- Gegeben ist die stündliche Nachfrage nach Wärme (Raumwärme und Warmwasser) in einem Fernwärmenetz (gemessene Leistungsmittelwerte der Einspeisung - stündliche) und die dazugehörige Umgebungstemperatur. Die Nachfrager sind hauptsächlich Haushalte, zum Teil aber auch Gewerbebetriebe.
- Vergleichen und interpretieren Sie unterschiedliche Modellansätze zur Abschätzung der Nachfrage in Abhängigkeit von der Temperatur und der Tageszeit.

Aufgabe 1.1

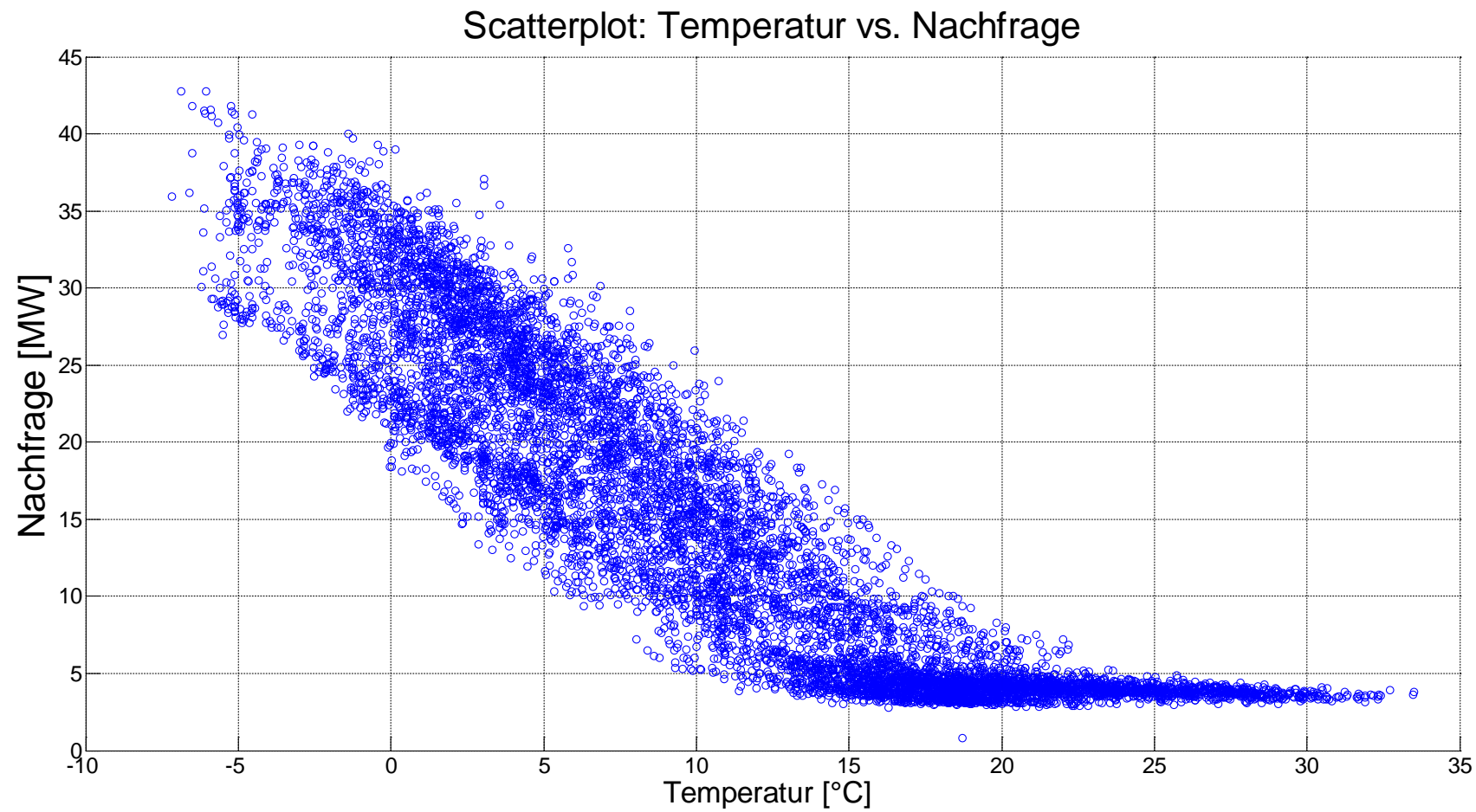


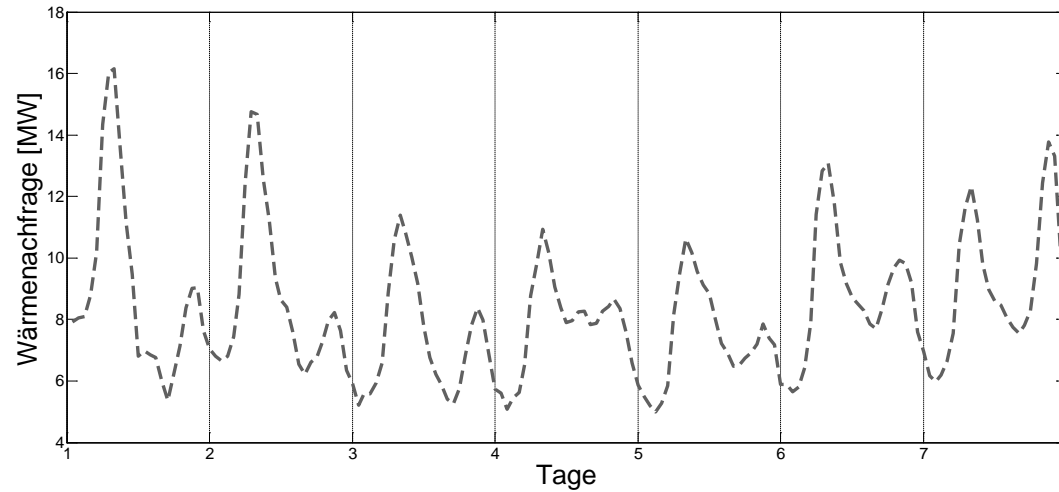
Beobachtungszeitraum:

1.10.2007 bis 1.10.2008

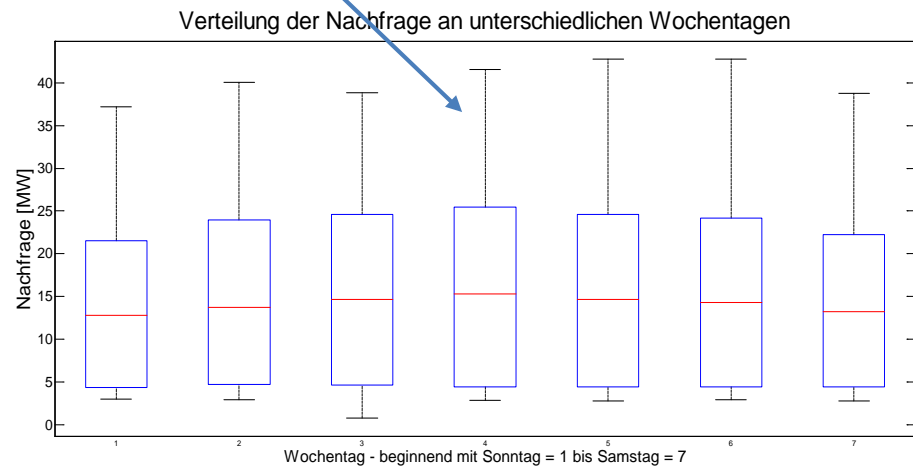
Stündliche Leistungsmittelwerte $t \in \{1 \dots 8774\}$

Aufgabe 1.1



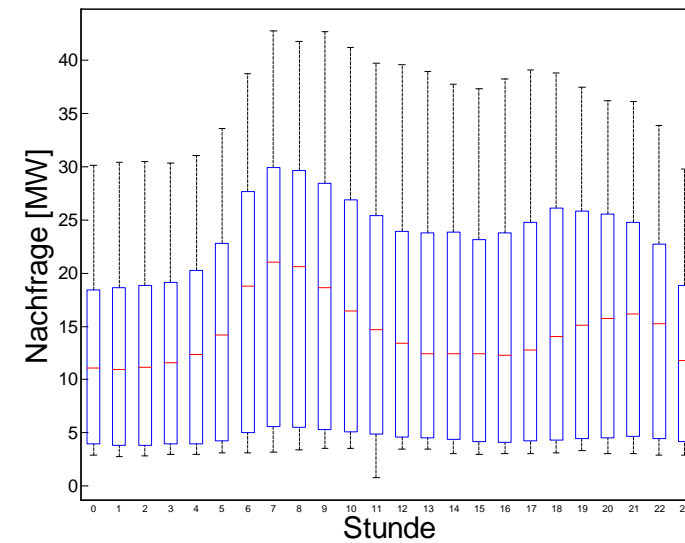


Wochentagsabhängigkeit



Tageszeitabhängigkeit

Verteilung der Nachfrage an Stunden des Tages



Aufgabe 1.1 – Fragestellung und Relevanz

Fragestellungen

- Wie können die Einflussfaktoren in einem Modell abgebildet werden?
- Wie wird das Modell mathematisch formuliert?
- Was bedeuten die einzelnen Koeffizienten?

Relevanz

- Prognosen (falls Einflussfaktoren vorhersagbar)
- Analysen von Einflussfaktoren
- Passen die Ergebnisse mit der Idee des Modells zusammen?

Aufgabe 1.1.1 – Modell 1

Einfache lineare Regression

Abschätzung über die Temperatur

$$y_t = \beta_0 + \beta_1 \cdot T_t$$

In diesem Modellansatz hängt die Nachfrage nur von der Umgebungstemperatur ab (deshalb einfache Regression).

Die Koeffizienten ergeben sich jeweils aus der linearen Regression, wobei die Funktion **fitlm(...)** in Matlab zu verwenden ist. Der Koeffizient entspricht jeweils in der Ausgabe „Estimate“.

Aufgabe 1.1.1 – Modell 1

Einfache lineare Regression

Abschätzung über Temperatur

- a) Geben Sie die Koeffizienten (Matlab Ausgabe: Spalte „estimate“) und die zugehörige t-Statistik an, und interpretieren Sie die Ergebnisse.
- b) Vergleichen Sie die modellierten Werte für die Beobachtung von $t=1680:1775$ bzw. $t=6480:6575$. Erstellen Sie dazu eine Grafik. Wie interpretieren Sie die Abweichungen? Wodurch unterscheiden sich die Abweichungen in den beiden Beobachtungszeiträumen?

Aufgabe 1.1.2 – Modell 2

Versuch der zusätzlichen Modellierung des Tagesverlaufs

Abschätzung über Temperatur und Stunde (als Polynom 3.Grades)

$$y_t = \beta_0 + \beta_1 \cdot T_t + \beta_2 \cdot h_t + \beta_3 \cdot h_t^2 + \beta_4 \cdot h_t^3$$

- Hier wird versucht, den typischen Tagesverlauf der Nachfrage (der nicht von der Temperatur abhängt) in das Modell zu integrieren. Die Variable h entspricht dabei der Datenspalte „Stunde“ (bzw. `data_heat.Stunde` im Mat-File), weist also jeder Beobachtung die dazugehörige Stunde zu.
- Die Stunden gehen hier zusätzlich zur Temperatur als Polynom 3. Grades in das Modell ein. Die Daten müssen also dementsprechend aufbereitet werden bevor die Regression durchgeführt wird.
- Die jeweiligen Koeffizienten werden wiederum über die lineare Regression mit der Funktion **fitlm(...)** geschätzt.

Aufgabe 1.1.2 – Modell 2

Versuch der zusätzlichen Modellierung des Tagesverlaufs

Abschätzung über Temperatur und Stunde (als Polynom 3.Grades)

$$y_t = \beta_0 + \beta_1 \cdot T_t + \beta_2 \cdot h_t + \beta_3 \cdot h_t^2 + \beta_4 \cdot h_t^3$$

- a) Geben Sie die Koeffizienten und die zugehörige t-Statistik an. Wie interpretieren Sie die Ergebnisse?
- b) Vergleichen Sie die modellierten Werte für die Beobachtung von t=1680:1775 bzw. t=6480:6575. Erstellen Sie dazu eine Grafik. Wie interpretieren Sie die Abweichungen? Wodurch unterscheiden sich die Abweichungen in den beiden Beobachtungszeiträumen?

Aufgabe 1.1.3 – Modell 3

Modellierung der Nachfrage getrennt für einzelne Stunden

Stündliche Abschätzung über die Temperatur

$$y_t^j = \beta_0 + \beta_1 \cdot T_t^j$$

- a) Vergleichen Sie die beiden Konstanten β_0 sowie die Koeffizienten β_1 des Temperatureinflusses aus den Ergebnissen der Regression für Stunde 7 und Stunde 23. Wie interpretieren Sie diese und wie interpretieren Sie die Unterschiede zwischen den beiden Stunden?
- b) Vergleichen Sie das jeweilige Bestimmtheitsmaß R^2 aus beiden Modellen für Stunde 7 und Stunde 23 nach Modellansatz 3 mit dem Bestimmtheitsmaß aus Modell 1. Wie würden Sie die Qualität der beiden Modellansätze beurteilen? Woraus ergeben sich mögliche Unterschiede?

Modellierung der Nachfrage getrennt für einzelne Stunden

Stündliche Abschätzung über die Temperatur

$$y_t^j = \beta_0 + \beta_1 \cdot T_t^j$$

- c) Vergleichen Sie die modellierten Werte für die Beobachtungen für Stunde 7 und für Stunde 23 mit den gemessenen Werten der jeweiligen Stunde über alle 366 Tage. Erstellen Sie dazu eine Grafik. Was beobachten Sie? Wieso schwanken zu einer bestimmten Zeit im Jahr die modellierten Werte um die relativ konstante gemessene Nachfrage?

Vergleichen Sie dazu den modellierten Zusammenhang zwischen Temperatur und Nachfrage (Skizze) mit dem **Scatterplot** Temperatur vs. Nachfrage aus der Angabe. (Scatterplot siehe Matlab Skript oder Folien zur Übungsangabe)

Aufgabe 1.1.4 – Verbesserung des Modells

Verbesserung des Modells

- a) Wie würden Sie vorgehen, um das Problem, das im letzten Punkt von Beispiel 1.1.3c beobachtet wurde, zu beheben und damit bessere Vorhersagen für diesen Zeitraum zu erhalten?

Welche Verbesserungsvorschläge für weitere Modellansätze fallen Ihnen ein? Gehen Sie dabei allerdings weiter davon aus, dass Ihnen nur die gegebenen Daten bzw. allgemein zugängliche Daten zur Verfügung stehen.

- b) Formulieren Sie einen verbesserten Modellansatz mathematisch. Ermitteln Sie die Koeffizienten Ihres Modellansatzes und bewerten Sie die Güte Ihres Modells im Vergleich zu den vorhergehenden Varianten.

Dokumentieren Sie Ihren Modellansatz (Formeln, Skizzen, Grafiken). Ihre Lösung kann auch aus mehreren Modellen bestehen!

Checkliste Regressionsanalyse

- Mathematische Formulierung des Modells
 - Definition der Variablen
 - Koeffizienten (deren Wert und Vorzeichen) schätzen und interpretieren
 - Zugehörige t-Statistiken angeben und interpretieren
 - Güte des Modells, Sinnhaftigkeit!
 - Fehlerterme graphisch darstellen und beschreiben/interpretieren
-
- Nicht nur auf Zahlenwerte achten
 - Sinnvoll nachdenken, was passiert und warum!
 - Unterschiede versuchen zu erklären (eventuell näher analysieren)

Aufgabe 1.2 – Modellierung des Strompreises mittels linearer Regression

Aufgabe 1.2.1 – Erstellung Strompreismodell

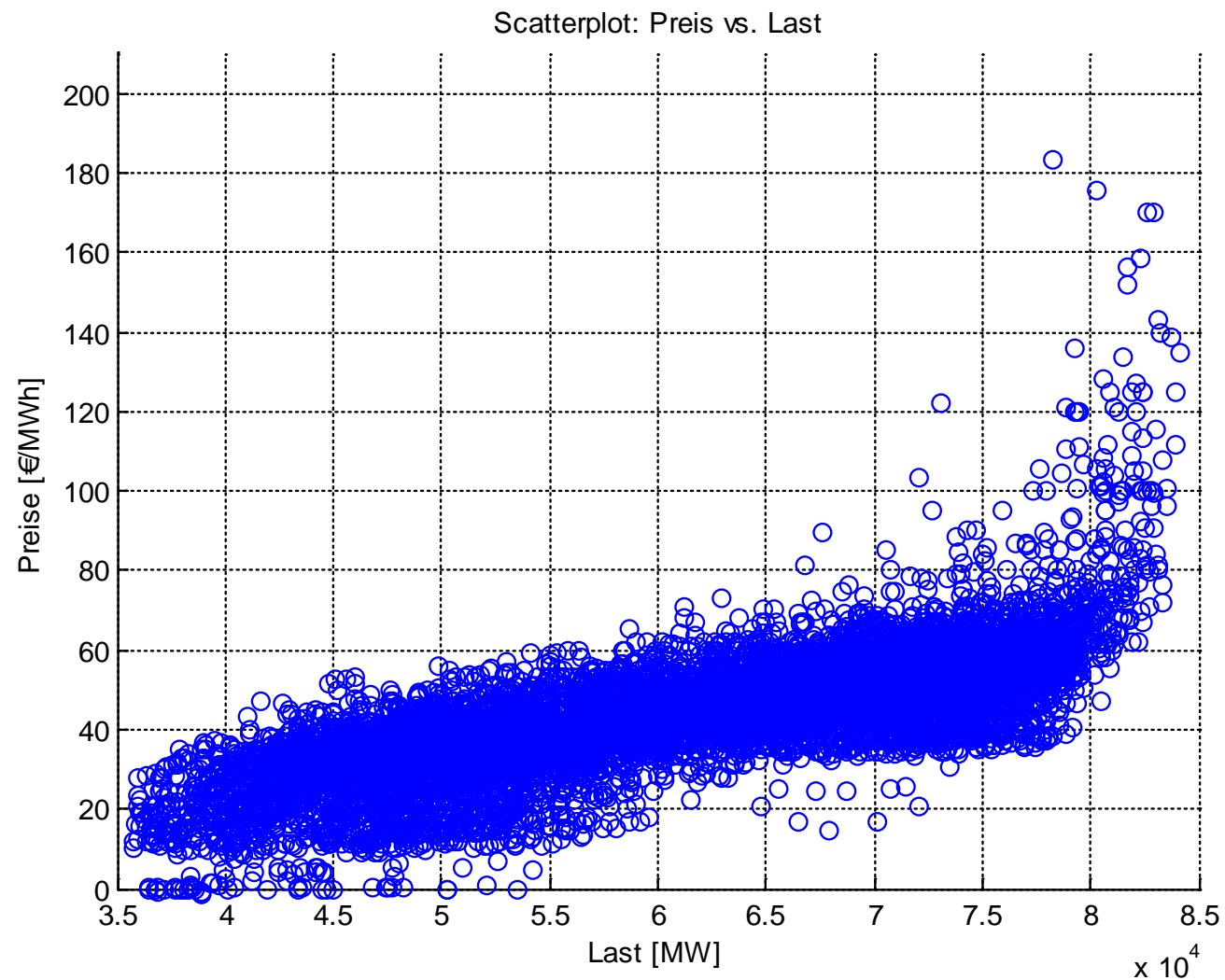
Erstellung Strompreismodell

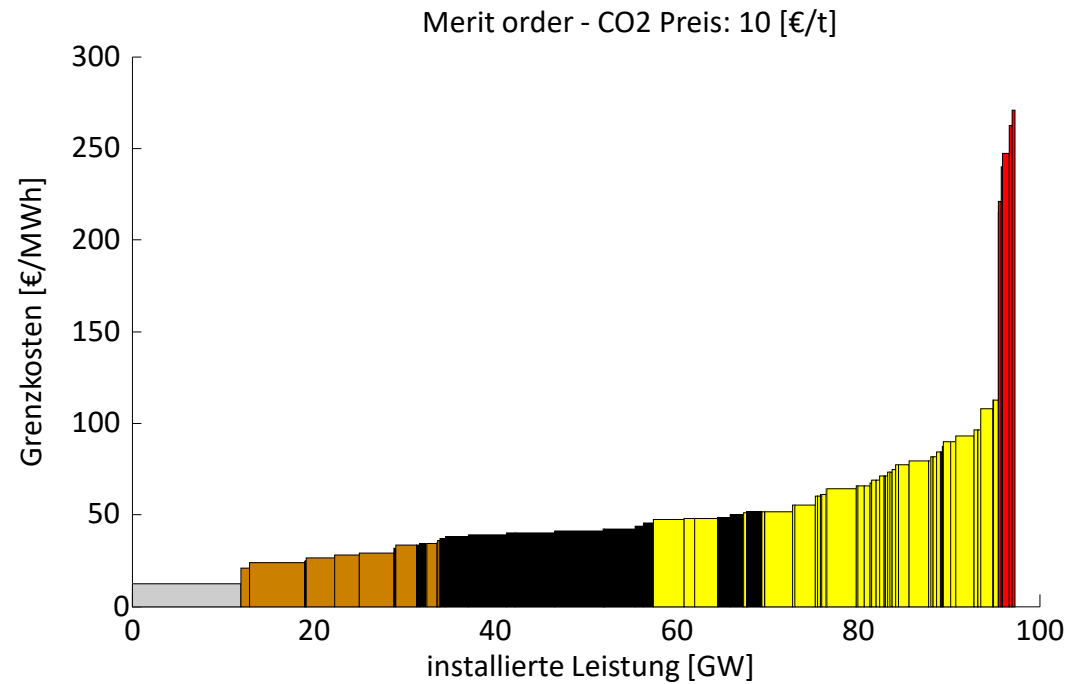
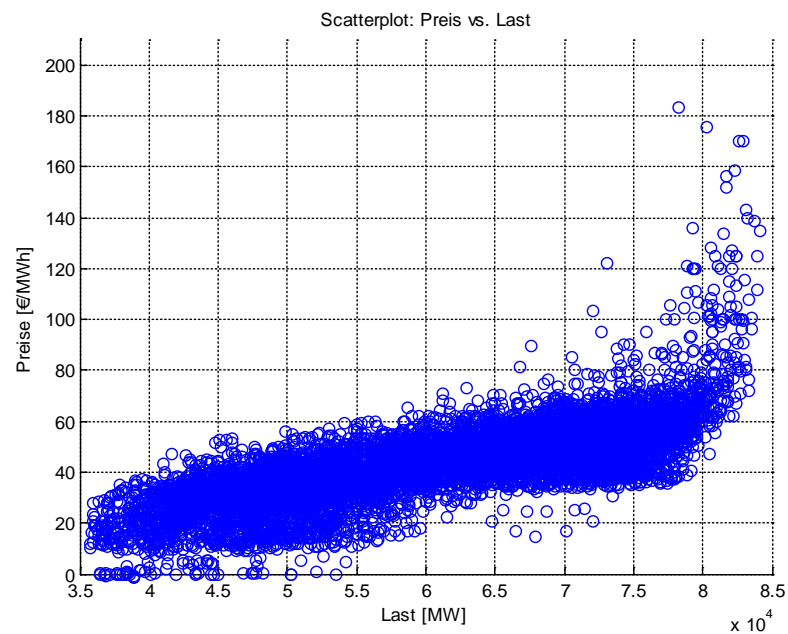
Erstellen Sie 3 möglichst gute Modellansätze zur Abschätzung des stündlichen Strompreises in Abhängigkeit der Netzlast und der Einspeisung erneuerbarer Energieträger für die Jahre 2012 und 2015. Siehe „*Regression_Strompreise.pdf*“.

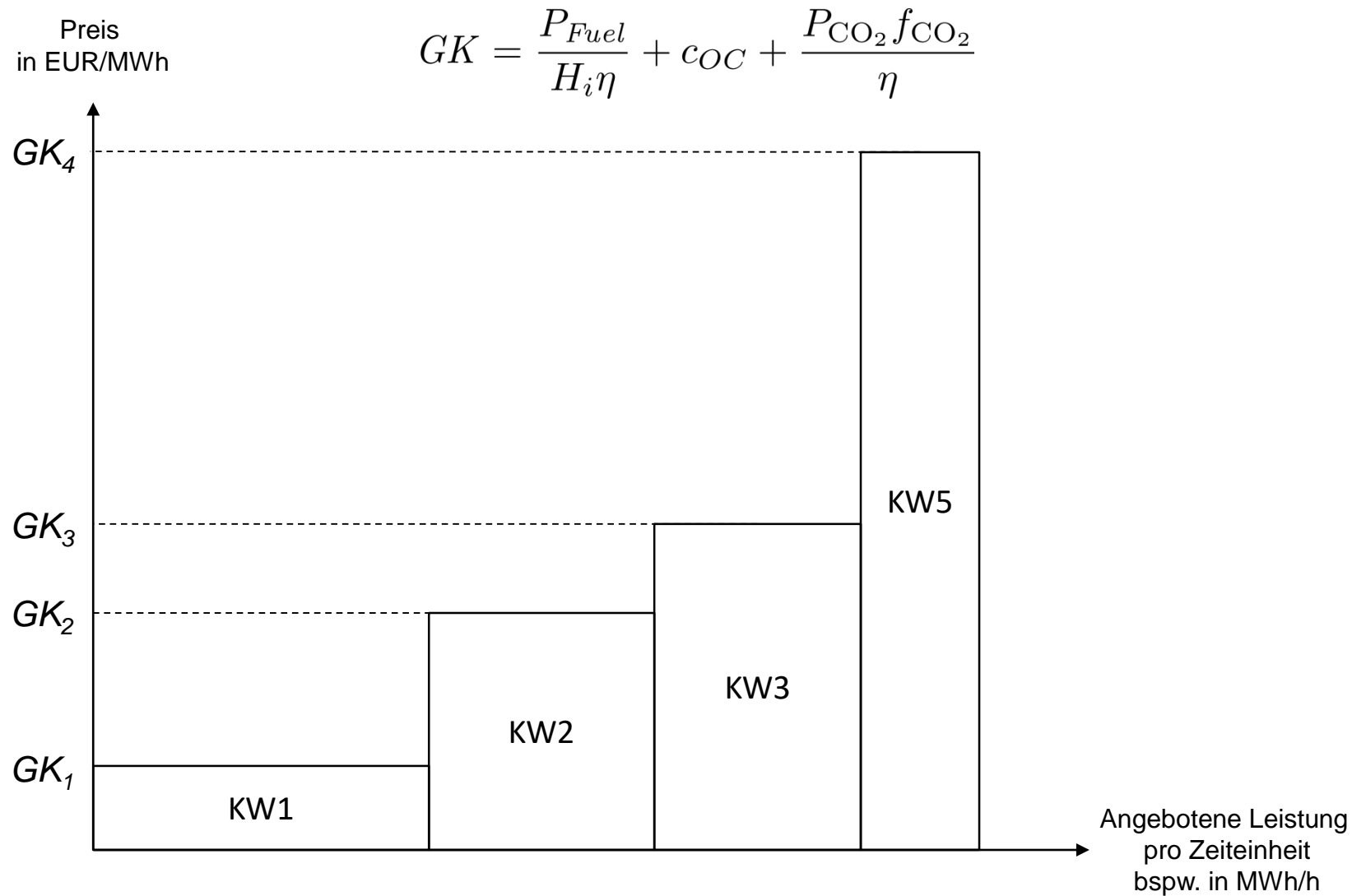
Daten finden Sie in „*Daten_Preise_Last_2012.xlsx*“, „*Daten_Preise_Last_2015.xlsx*“

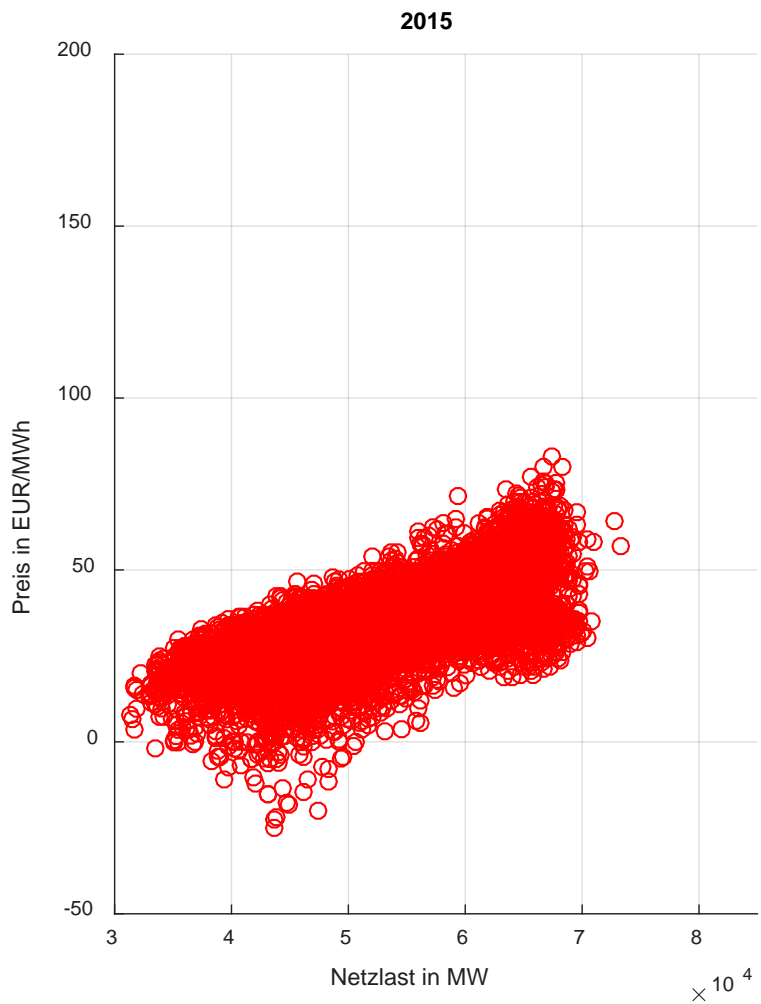
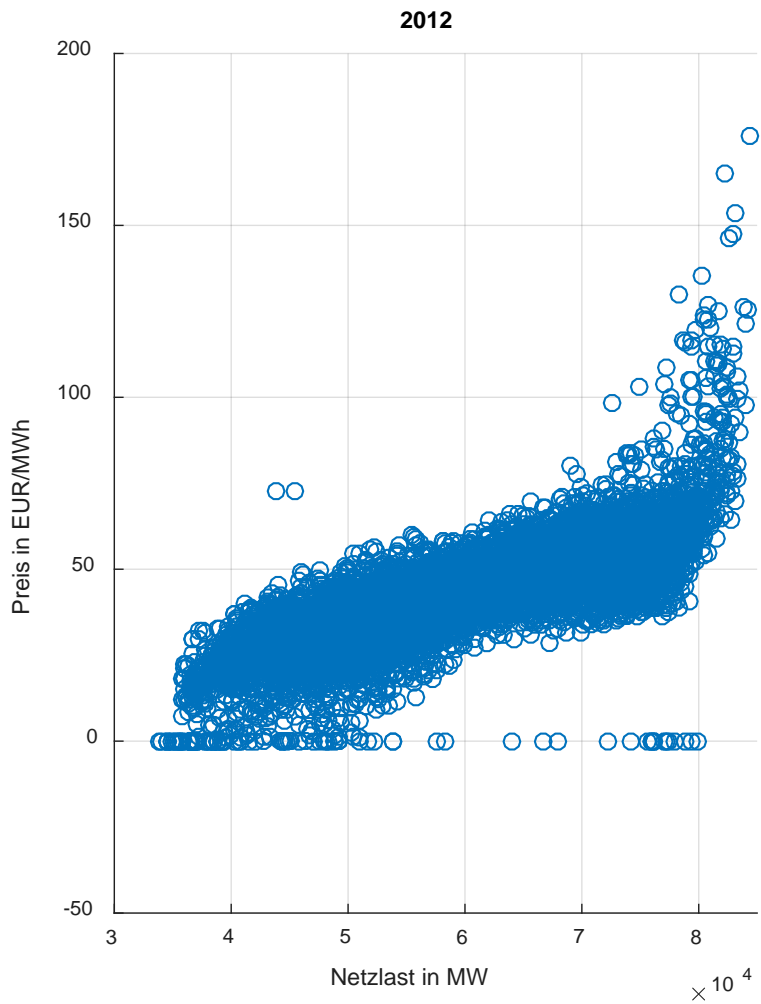
- a) Dokumentieren Sie die Modellansätze mathematisch.
- b) Führen Sie die Regression in Matlab durch und geben Sie die Werte für die geschätzten Koeffizienten, t-Statistiken und das adjustierte Bestimmtheitsmaß an.

Strompreisbildung









Aufgabe 1.2.2 – Interpretation der Ergebnisse

- a) Erstellen Sie eine Grafik/Skizze zu dem von Ihnen modellierten Zusammenhang zwischen Netzlast und Strompreis für alle 3 Modellansätze. Beurteilen Sie wie gut dieser Zusammenhang die Realität in Bezug auf Vorgänge im Stromsystem beschreibt.
- b) Vergleichen Sie die Ergebnisse der Modellansätze des Jahres 2012 (Daten von Punkt a) mit denen des Jahres 2015. Was können Sie daraus schließen?
- c) Beschreiben Sie den von Ihnen modellierten Zusammenhang zwischen der Einspeisung erneuerbarer Energien und dem Strompreis (2012 und 2015). Interpretieren Sie die Bedeutung der Werte der von Ihnen ermittelten Koeffizienten.

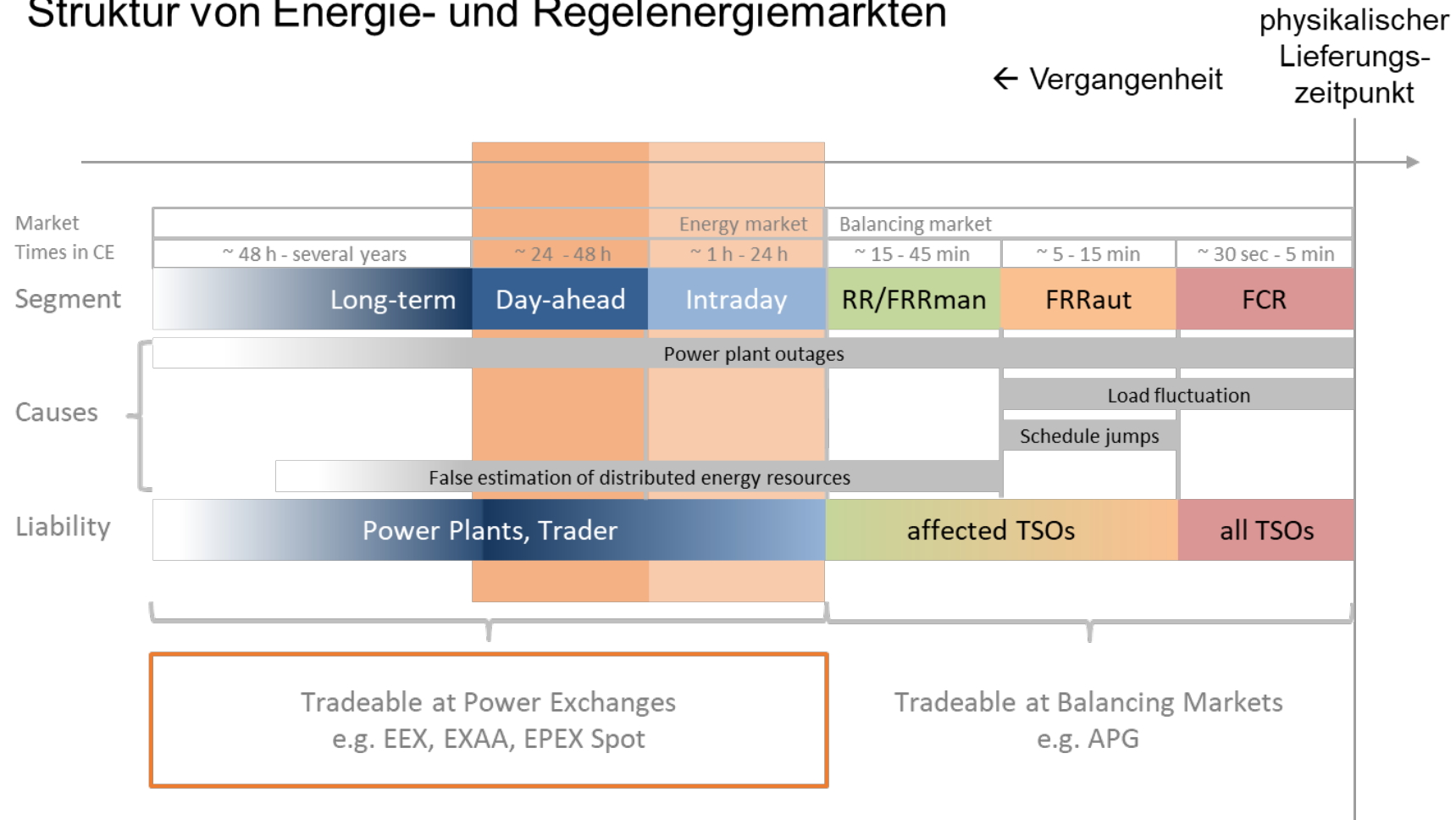
Aufgabe 1.2.3 – Einfluss von Prognosedaten

- Da der Anteil an erneuerbarer Erzeugung immer größer wird, besteht durch den vorhandenen Prognosefehler von Erzeugung und Verbrauch die Anforderung, die über- oder überschüssige Energie noch vermarkten zu können.
- Für den kurzfristigen Handel (Spotmarkt) stehen mehrere Möglichkeiten zur Verfügung, die gebräuchlichsten sind:
 - *Day-Ahead-Märkte (DA)*: Handel von Strom für den nächsten Tag
 - *Intraday-Märkte (ID)*: Handel von Strom für den gleichen (oder nächsten) Tag
- In dem nachfolgenden Beispiel wird eruiert, welche Daten eine Auswirkung auf den jeweiligen Preis (DA, ID) haben.

Siehe u.a. <http://www.epexspot.com>

Aufgabe 1.2.3 – Einfluss von Prognosedaten

■ Struktur von Energie- und Regelenergiemärkten



Aufgabe 1.2.3 – Einfluss von Prognosedaten

Erstellen Sie mit den Daten des Jahres 2016 ein lineares Modell um einerseits den Day-Ahead-Preis und andererseits den Intraday-Preis zu beschreiben. Sie können auf folgende (Input) Daten zurückgreifen („*Daten_Preise_Last_2016.xlsx*“):

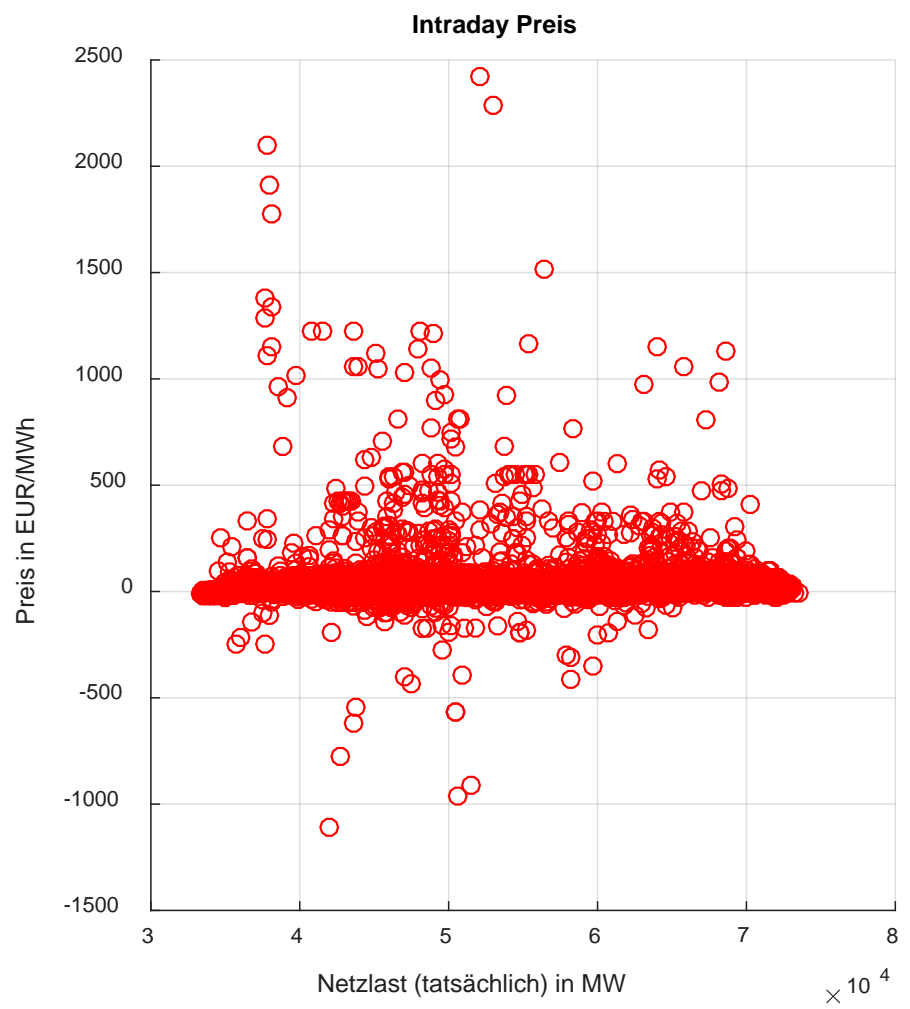
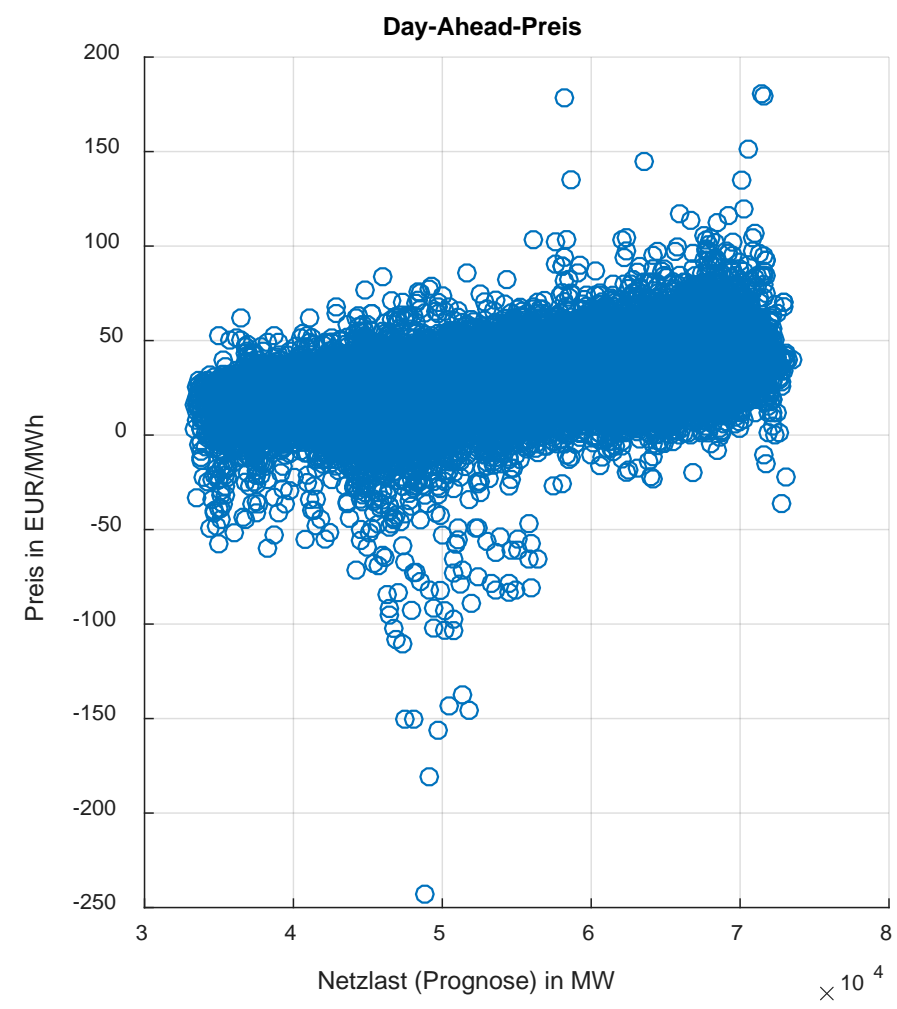
- Last Day Ahead Prognose [MW]
- Last Actual [MW]
- Wind Day Ahead Prognose [MW]
- Wind Actual [MW]
- PV Day Ahead Prognose [MW]
- PV Actual [MW]

um die Preise

- Day Ahead Spotpreis [Euro/MWh]
- Intraday Spotpreis [Euro/MWh]

zu beschreiben. Die “*Day Ahead Prognosen*” werden jeweils einen Tag vor der Lieferung erstellt, die “*Actual*” Daten repräsentieren die tatsächlich gemessenen Lasten bzw. Produktion.

Aufgabe 1.2.3 – Einfluss von Prognosedaten



Aufgabe 1.2.3 – Einfluss von Prognosedaten

a) Erstellen Sie ein Modell in der Form:

$$y_t = \beta_0 + \beta_1 * Last_{DAPrognose} + \beta_2 * Last_{Actual} + \beta_3 * Wind_{DAPrognose} \\ + \beta_4 * Wind_{Actual} + \beta_5 * PV_{DAPrognose} + \beta_6 * PV_{Actual}$$

um

i. den Day-Ahead-Preis und

ii. den Intraday-Preis

zu beschreiben.

b) Interpretieren Sie alle Werte des Regressionsmodells (β_i , t-Statistics, p-Values, R^2). Welche Inputs würden Sie behalten, bzw. verwerfen (warum?).

Checkliste Strompreisbildung

- Wählen Sie vernünftige Modelle (Polynome siebten Grades mögen eine bessere Regression abgeben, sind aber meist nicht sinnvoll und weder realistisch noch interpretierbar!)
- Beachten Sie ob das Modell die tatsächliche Strompreisbildung wieder gibt.
- Unterschiede zwischen den Jahren 2012 und 2015 (siehe Graphik) sollten sich auch in den Modellen bzw. den Ergebnissen widerspiegeln.

Zusatzliteratur und Lehrmaterial

- Link zu einem Skriptum des IHS Wien in dem die wichtigsten Begriffe und Herleitungen beschrieben werden: <http://homepage.univie.ac.at/robert.kunst/emwi.pdf>
- Zudem finden Sie in zahlreichen Online-Tutorials weitere Hintergrundinformationen.
- Im Beispielcode „Regressionsbeispiel_Strompreise.m“ finden Sie hilfreiche Tipps zur Regressionsanalyse in Matlab. Auch ersichtlich als pdf-Report in „Regressionsbeispiel_Strompreise.pdf“
- Im File „Interpret Linear Regression Results.pdf“ finden Sie hilfreiche Informationen zur Interpretation der Ergebnisse und Teststatistiken der Regressionsanalyse in Matlab.

Abgabe

1. Protokoll

- Ergebnisse und Lösungsweg kommentieren (überflüssigen Text vermeiden!)
- Das Erscheinungsbild und die wissenschaftliche Gestaltung des Protokolls wird in die Beurteilung miteinbezogen.
 - <https://www.wissenschaftliches-arbeiten.org/>
 - Kriterien: Inhaltsverzeichnis, Abbildung- und Tabellenbeschriftung, Verweise, Modellbeschreibung, Lesbarkeit...
- Eine LaTeX Vorlage finden Sie im TUWEL (nicht verpflichtend).
- Abgabe des Protokolls als pdf Datei.

2. Matlab Code

Als Gruppenabgabe (Protokoll + Code gemeinsam als zip) ins TUWEL hochladen!

Deadline: 28.04.2019, 23:59 (keine spätere Abgabe möglich)



TECHNISCHE
UNIVERSITÄT
WIEN



Theresia Perger

TU Wien
Energy Economics Group – EEG
Gußhausstraße 25-29/E 370-3
1040 Vienna, Austria

+43 (1) 58801 370359
perger@eeg.tuwien.ac.at
www.eeg.tuwien.ac.at