



Digitalization

Industrie 4.0

Smart Production

E-Mobility

Smart Energy

Energy Efficiency

Smart Infrastructure

Smart Buildings

Renewables

Sebastian Wölk / Hochschule der Medien / 18.11.2021

Willkommen

**Prognose des Füllgrads
eines Distributionslagers
im automatischen Kleinteilelager der
Phoenix Contact GmbH & Co. KG**

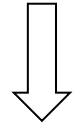
Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Problem

**Volatile & schwer planbare
Geschäftsentwicklung**

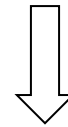
Global und mehrstufiges Vertriebs- und Distributionsnetzwerk

- Logistikstandort (am Headquarter) fungiert als zentrales HUB
- 75% der monetären Warenströme fließen durch das Distributionslager am zentralen HUB



Notwendigkeit der ausgewogenen Abstimmung dynamischer und statischer Kapazitäten

- Aufrechterhaltung und Stabilisierung der Supply Chain-Prozesse
- Termingereue Belieferung der globalen Kunden & Tochtergesellschaften innerhalb der Vertriebsstruktur




Füllgrad des Distributionslagers

Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Problem

- Überschreitung von kritischen Grenzwerten im Lagerfüllgrad
 - Überforderung der statischen Kapazitäten
 - Negative Auswirkungen auf die Performance der Maschinen & Anlagen

- 
- Unvorhergesehene Überschreitung der kritischen Grenzlinie

Ad-hoc-Maßnahmen als einziges Mittel:

- Kurzfristiges Aussetzen der Einlagerung
 - >2.000 Paletten im Rückhang
- Mehrfach bewegt sich der Füllgrad an seiner Kritischen Grenze und überschreitet diese sogar

- 
- Entspannung durch Erweiterung der max. möglichen Lagerkapazität

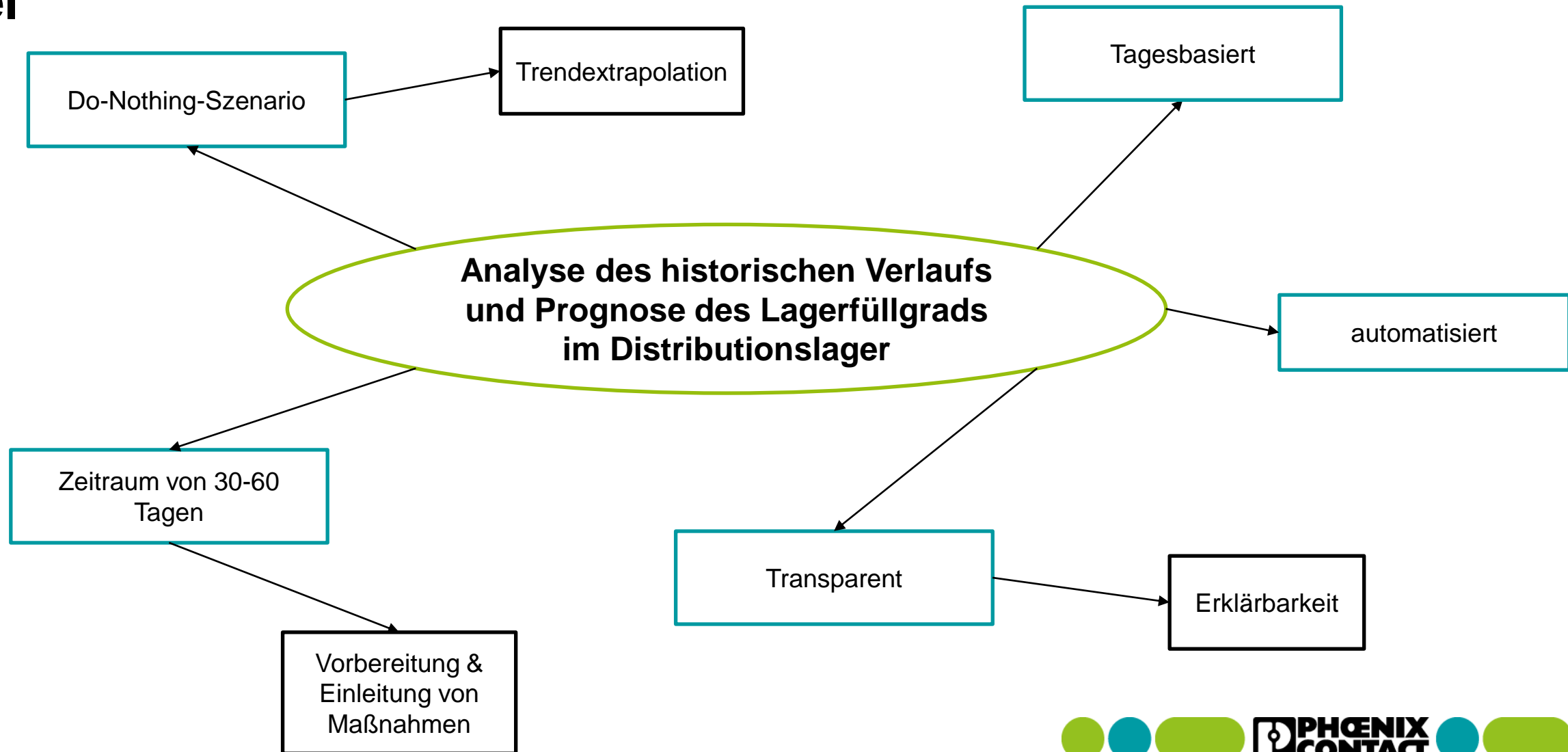
ABER:

Solche Entwicklungen können jederzeit wieder eintreten



Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Ziel



Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Lösungsansatz

Entwicklung verschiedener potentieller Prognosemodelle & Vorgehensweisen zur Ableitung einer Vorhersage des tagesbasierten, zukünftigen Lagerfüllgrads

Auswahl des besten Modells zur Ableitung einer Vorhersage

Erweiterung, **Optimierung** & **Deployment** des Modells

Folgende Modelle werden betrachtet:

ARIMA

XGBoost

**Monte Carlo
Simulation**

Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Daten & Tools

- **Datengrundlage:** Zeitreihendaten des Lagerfüllgrads im Distributionslager

- Zeitraum: 01.01.2019 – 31.10.2021

Trainingsdaten
vom 01.01.2019 – 31.07.2021

Testdaten
vom 01.08.2021 – 31.10.2021

Datensatz	Füllgrad	File_Date	Anzahl Stellplätze	Freie Plätze	Belegte Plätze
1	86.6	2019-01-02	196636	26349	170287
2	86	2019-01-03	196636	27529	169107
3	84.9	2019-01-04	196636	29692	166944
4	83.92	2019-01-07	196636	31619	165017
5	82.54	2019-01-08	196636	34333	162303

- **Tool:** Alteryx (Low-Code Data Science Tool)

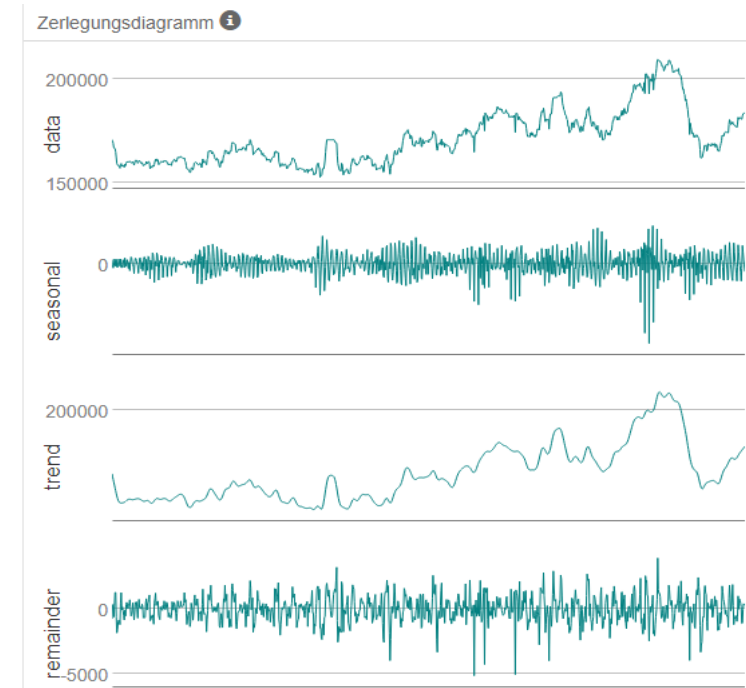
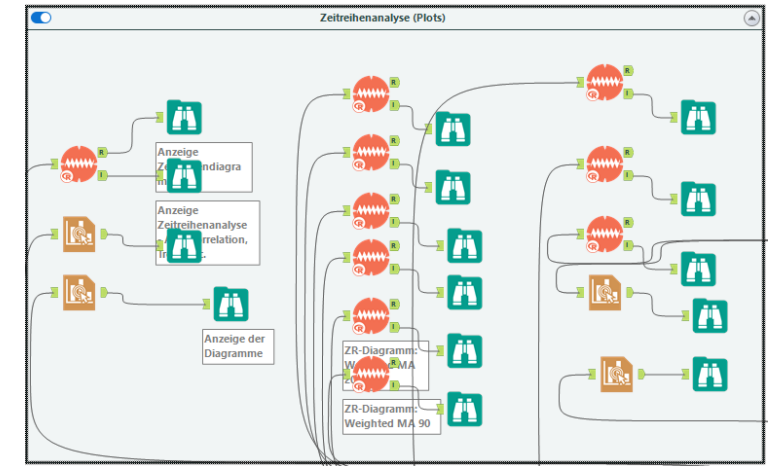
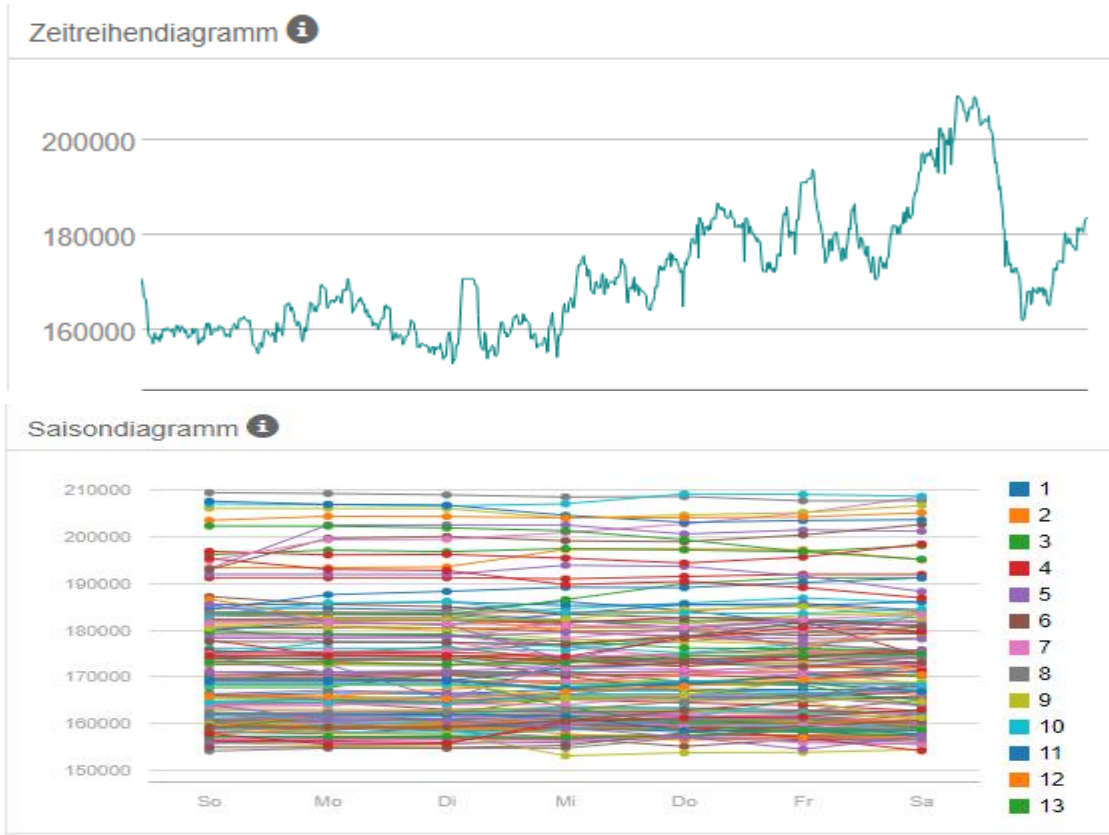
- Low-Code-Tool für den gesamten Data Science Prozess
- Bereits im Unternehmen im Einsatz
- Bisher keinen geeigneten Use-Case zur Verprobung
- Lizenzen sehr teuer (5.000 EUR p.a.)
- Verwendung der Studenten-Lizenz

alteryx

Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Zeitreihenanalyse

- Identifikation wichtiger Eigenschaften der Zeitreihe
 - Saisonalität
 - Trend



Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Zeitreihenanalyse

- Identifikation wichtiger Eigenschaften der Zeitreihe
 - Korrelationen innerhalb der Zeitreihe
 - Autokorrelation
 - Partielle Autokorrelation

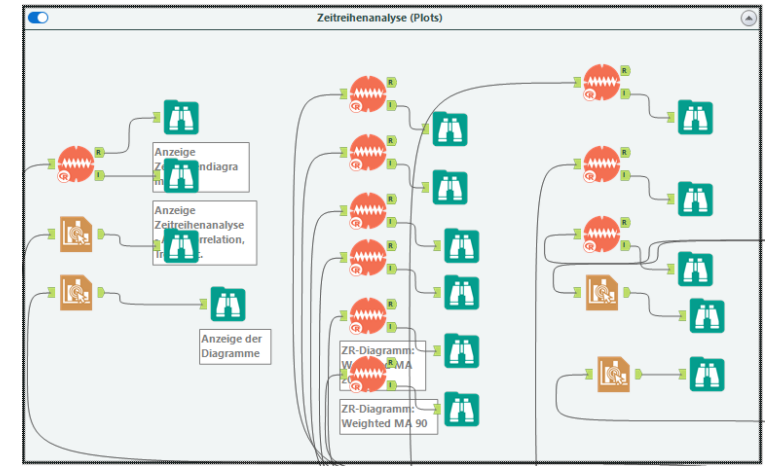


Diagramm Autokorrelationsfunktion ⓘ

ACF

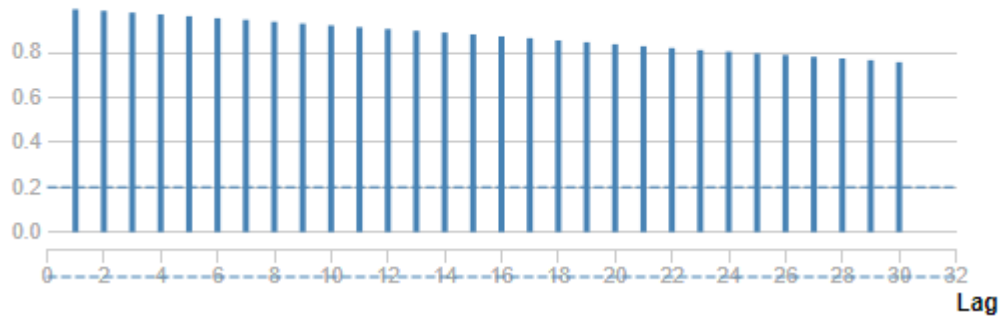
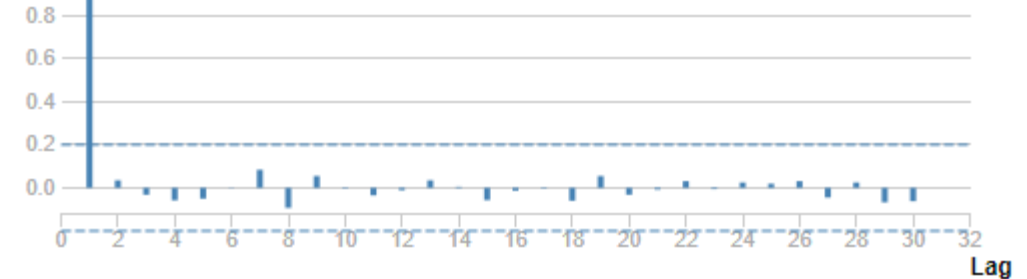


Diagramm partielle Autokorrelationsfunktion ⓘ

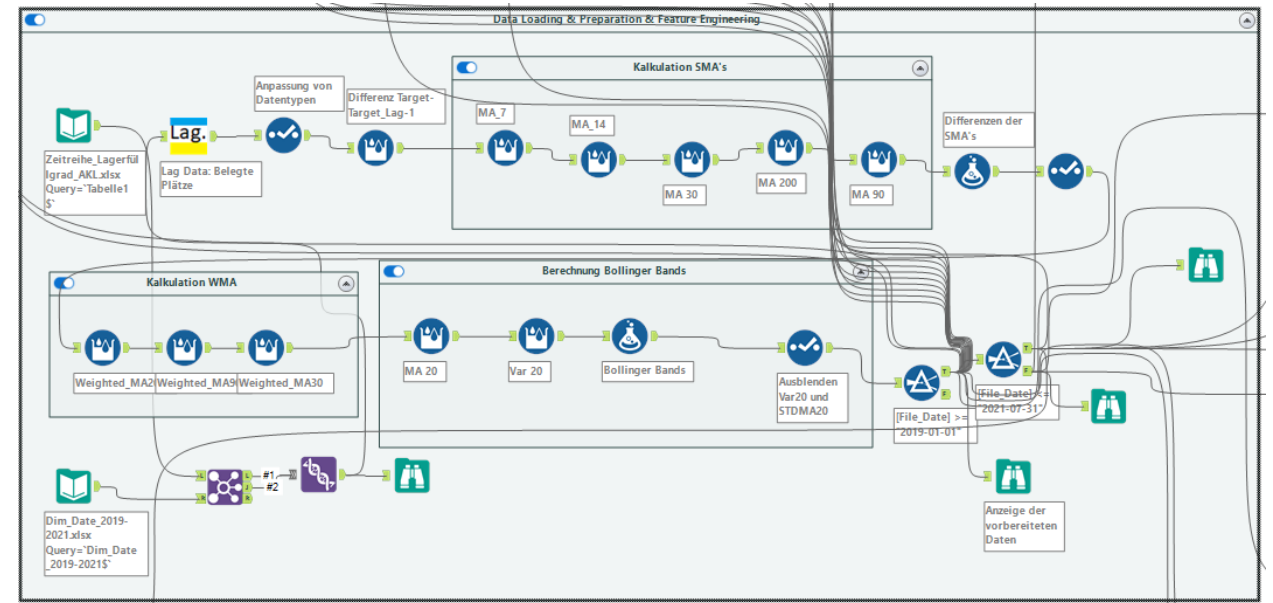
PACF



Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Feature Engineering

- **Simple Moving Average**
 - Verschiedene Ausprägungen
 - Ermöglicht das Erkennen von Trends im Datensatz
 - Glättet den Verlauf der Zeitreihe
- **Weighted Moving Averages**
 - Verschiedene Ausprägungen
 - Ermöglicht das Erkennen von Trends im Datensatz
 - Höhere zeitliche Gewichtung der jüngsten Vergangenheit
- **Zeitverzögerte Variablen**
 - Verzögerung der Zielvariable
- **Bollinger Bänder**
 - Bekannt aus der Analyse von Aktienkursen
 - Können u.U. auf Trendveränderungen in der Zeitreihe hinweisen



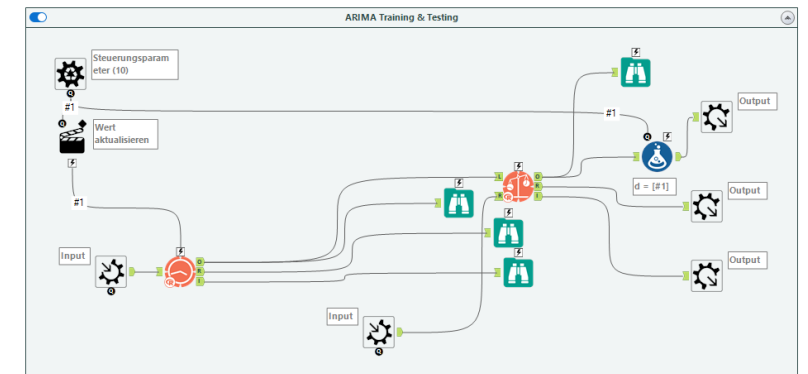
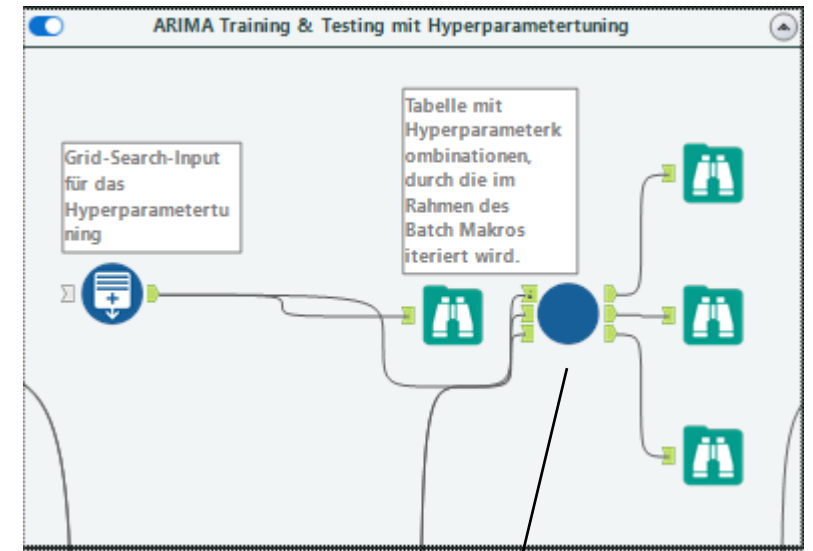
Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Modellierung I - ARIMA

- Ein für Zeitreihenprobleme optimiertes Modell

„Auto-Regressive Integrated Moving Average“

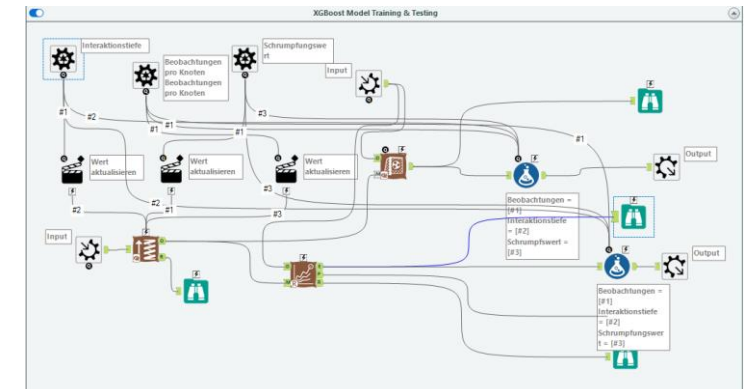
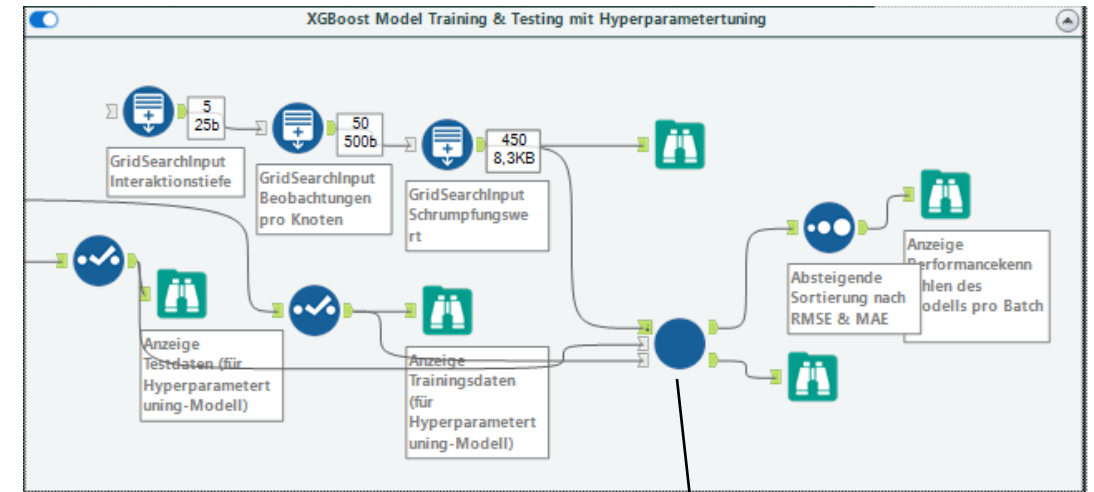
- AR: autoregressiver Teil
- I: Trendbeseitigung in der Zeitreihe (Herstellung einer Stationarität innerhalb der Zeitreihe)
- MA: gleitender Mittelwertbeitrag
- Hyperparametertuning mit Batch Makro:
 - P: Autoregressive Komponente (AR)
 - D: Grad der ersten Differenzierung (I)
 - Q: Komponente des gleitenden Mittelwerts (MA)



Datensatz	Model	ME	RMSE	MAE	MPE	MAPE	MASE	d
1	ARIMA	-2452.8765	6914.9652	6079.4437	-1.5384	3.5453	6.5017	0
2	ARIMA	-7639.9176	9528.4052	8021.3796	-4.5297	4.7382	8.5785	1
3	ARIMA	140750.799	167227.7094	140804.8228	80.3573	80.3885	150.5836	2
4	ARIMA	619973.373	810312.1444	620029.1354	351.5886	351.6207	663.0897	3
5	ARIMA	12854205.5792	19053940.9303	12854205.5792	7231.4293	7231.4293	13746.921	4
6	ARIMA	2782232097.0187	4590633625.0523	2782232097.0187	1555855.4076	1555855.4076	2975456.1265	5
7	ARIMA	62206635197.1921	110544088153.221	62206635197.1921	34662681.2698	34662681.2698	66526841.5258	6

Modellierung II – XG Boost

- Entscheidungsbaummodell mit Gradient Boosting
- Kombination einfacher und komplexer Modelle
 - Erzeugung mehrerer Entscheidungsbäume
 - Unterschiedliche Bäume mit unterschiedlicher Komplexität
- Verwendung der selbst berechneten Features
- Hyperparametertuning mit Batch Makro:
 - Schrumpfungswert
 - Interaktionstiefe
 - Minimal erforderliche Anzahl Objekte pro Baumknoten



Datensatz	Model	Correlation	RMSE	MAE	MPE	MAPE	Beobachtungen	Interaktionstiefe	Schrumpfungswert
1	XGBoost	0.95843	1666.015064	1147.159455	0.204068	0.663364	5	4	0.004
2	XGBoost	0.958024	1666.517468	1138.775669	0.187981	0.658474	6	8	0.005
3	XGBoost	0.958276	1668.064763	1141.947871	0.203965	0.660511	5	8	0.005
4	XGBoost	0.958619	1671.886391	1160.886826	0.225389	0.670798	4	8	0.004
5	XGBoost	0.958934	1675.068609	1170.267077	0.250671	0.676838	4	6	0.004
6	XGBoost	0.956991	1677.489601	1143.354778	0.153965	0.661565	6	3	0.004
7	XGBoost	0.958975	1678.221421	1182.652997	0.256815	0.683758	4	5	0.004

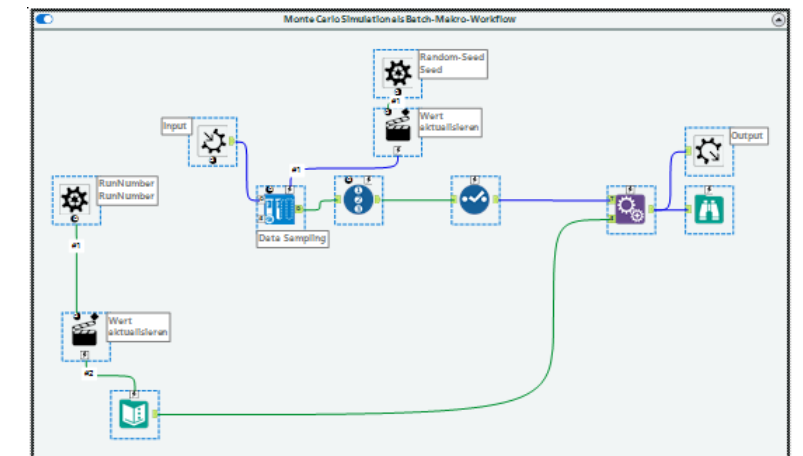
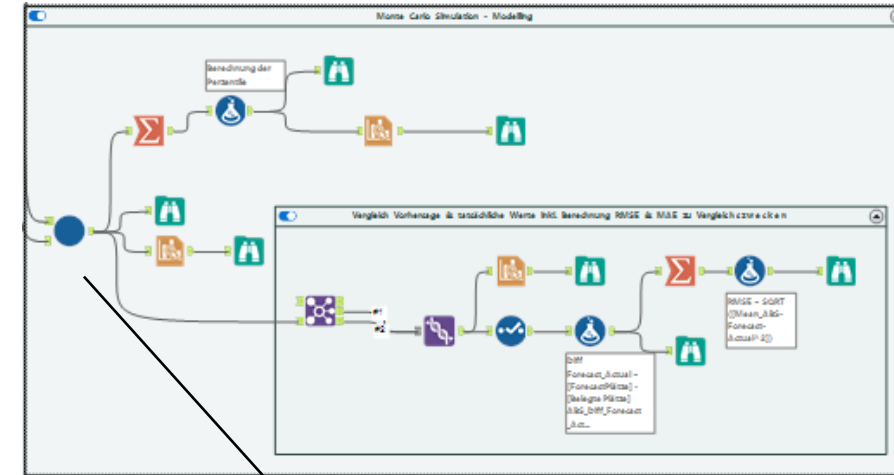
Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Modellierung III – Monte Carlo Simulation

- Monte Carlo Simulation zur Prognose der belegten Lagerplätze
 - Simulation von 30 Durchläufen pro Ausführung (30 Tage)
- Stichprobenziehung aus benannter Variable der zugrundeliegenden Daten
 - Variable „PercentChange“ (Belegte Plätze (t) / Belegte Plätze (t-1))
 - Relative Veränderung der Zielvariable zum Vortag (annähernd normalverteilt)
- Parameter, die das Batch-Makro beeinflussen:
 - Anzahl „Runs“ (Ausführungshäufigkeit des Batch Makros)
 - Je „Run“ werden 30 Durchläufe der Stichprobenziehung getätigt
 - Random Seed
 - Jeder „Run“ erzeugt andere Vorhersagen/Stichprobenziehungen
 - Jeder „Run“ hat einen anderen Wert als „Seed“

Beispiel:

Anzahl „Runs“ = 50
Anzahl **Durchläufe** je Run = 30 (Tage)
= **1500 Vorhersageergebnisse** (50 Vorhersagen je Tag)



Evaluation der Modelle

1. ARIMA-Modell

- Starkes Rauschen in der Zeitreihe
- Geringe Datenmenge
- Fehlende Saisonalität in der Zeitreihe



Datensatz	Model	ME	RMSE	MAE	MPE	MAPE	MASE	d
1	ARIMA	-2452.8765	6914.9652	6079.4437	-1.5384	3.5453	6.5017	0

RMSE

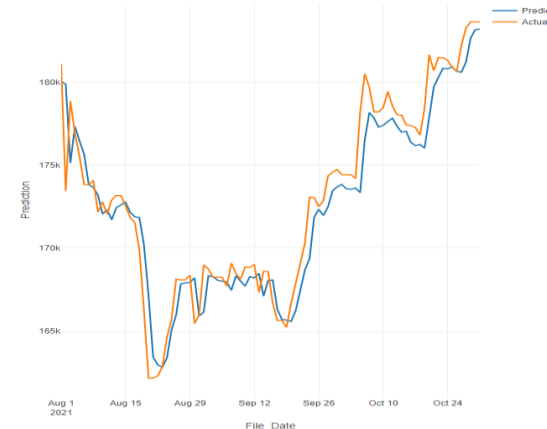
6914

MAE

6079

2. XGBoost

- Vollständigkeit aller Features im Datensatz notwendig
- Overfitting wird im XGBoost über Parameter vorgebeugt
- Validierung im praktischen Einsatz notwendig
 - Immer noch möglicher BIAS in den Testdaten
(z.B. bedingt durch Lag 1, Lag 2 etc. der Zielvariable)



Datensatz	RecordID	Model	Correlation	RMSE	MAE	MPE	MAPE	Beobachtungen	Interaktionstiefe	Schrumpfungswert
1	1	XGBoost	0.95843	1666.015064	1147.159455	0.204068	0.663364	5	4	0.004

RMSE

1666

MAE

1147

Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Evaluation der Modelle

3. Monte-Carlo-Simulation

- Ermöglicht den Ausdruck von Wahrscheinlichkeiten über die Perzentile der Simulationsergebnisse
- Relativ lange Laufzeit bei hoher Anzahl an Ausführungen (i.d.R. mind. 10.000)
- Zufällige Stichprobenziehung vernachlässigt die in jüngster Vergangenheit eingetretenen Trendbewegungen
- Vorhersageleistung der Simulationsergebnisse nicht brauchbar
 - Genauigkeit nimmt über einen größeren Vorhersagezeitraum stark ab

Datensatz	Mean_Forecast-Actual	Mean_ABS_Forecast-Actual (MAE)	Mean_ABS-Forecast-Actual^2	RMSE
1	11371.792	11557.558667	194829950.525333	13958



XGBoost

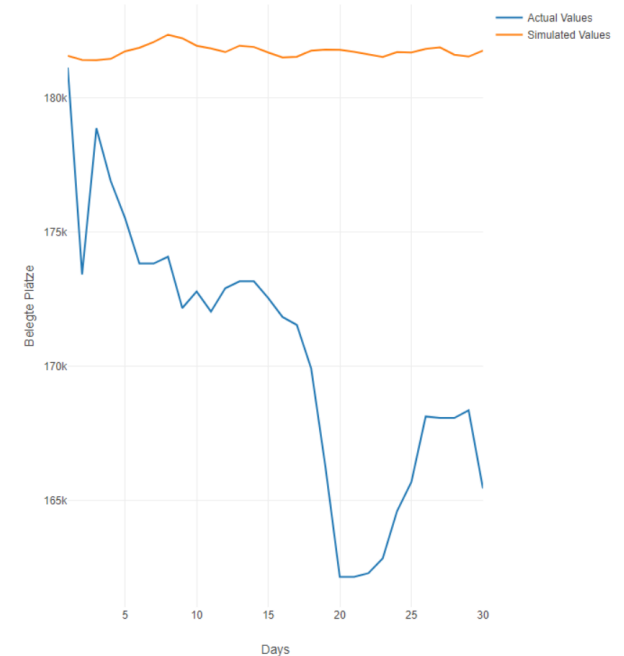
Monte Carlo
Simulation

RMSE

MAE

13958

11558

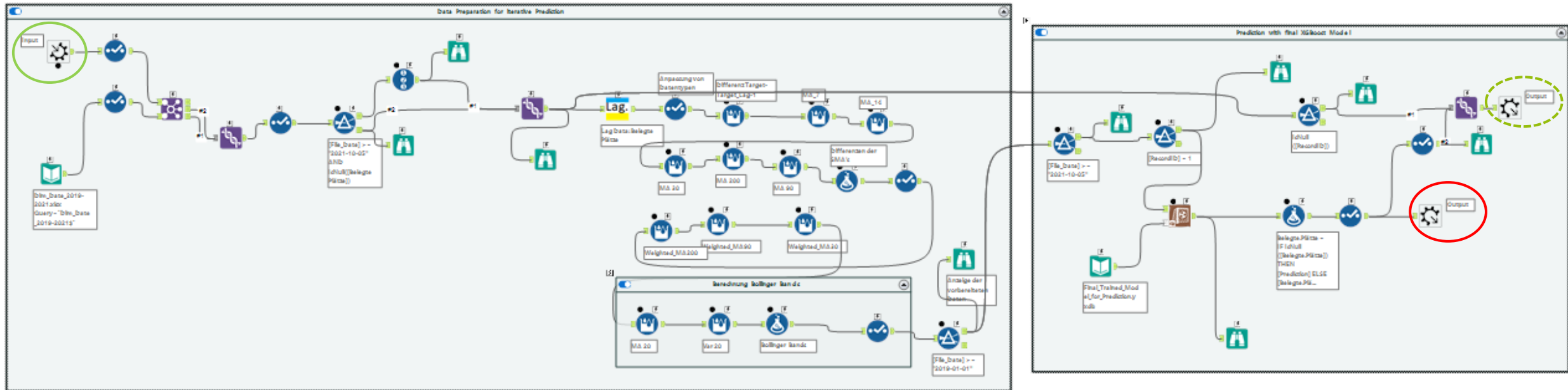
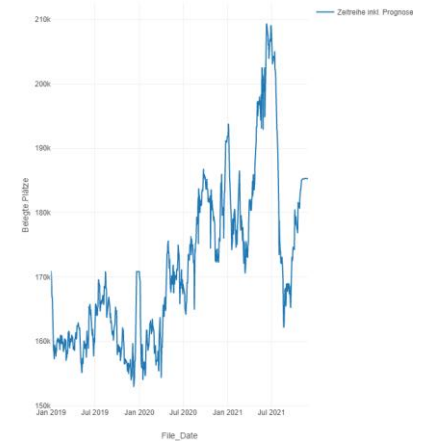


Entwicklung eines Prognosemodells zur Vorhersage des Lagerfüllgrads

Das beste Modell - Iterativer Prognoseprozess

Ziel: Bei jeder Vorhersage liegt die komplette Zeitreihe inkl. aller berechneten Feature aktualisiert vor

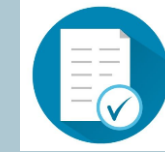
1. Input = Zeitreihendaten bestehend aus Datum & Belegte Plätze
2. Durchführung der Data Preparation, Feature Engineering & Vorhersage für den Zeitpunkt $t+1$ (t = heute)
3. Loop-Back des Vorhersageergebnisses in den Input-Datensatz & erneute Ausführung ab Punkt 1 (max. Iterationen begrenzt)



Zusammenfassung



- Zeitreihenanalyse und Feature Engineering
- Trainieren und Testen verschiedener Modellvarianten
- Auswahl des besten Modells
- Aufbau eines iterativen Prognoseprozesses unter Verwendung der optimalen Hyperparameterkombination für das beste Modell
 - Anzahl „x“ Tage werden iterativ prognostiziert
 - Vorhersagen von $t+1$ werden zur Neuberechnung aller Features verwendet, um $t+2$ vorherzusagen



- Deployment (Inbetriebnahme + Operationalisierung)
 - Automatisierung der Alteryx-Workflows
 - Training mit Hyperparameter-tuning
 - Final Training (mit besten Hyperparametern)
 - Iterative Prognose mit final trainiertem Modell
 - Definition geeigneter Retraining-Zyklen
 - Verarbeitung und Einbindung des Outputs in einem Reportingsystem
- Validierung des besten Modells im praktischen Einsatz
 - Vergleich der Vorhersage mit tats. Werten
- Weiterentwicklung des Modells
 - Einbindung weiterer domänenspezifischer Features



Danke

**Prognose des Füllgrads eines
Distributionslagers
im automatischen Kleinteilelager der
Phoenix Contact GmbH & Co. KG**