

ABSTRACT

Accurate forecasting of spot prices, especially in commodity markets, is of great interest because price predictions are crucial for business decisions across numerous industries and provide a solid foundation for risk management, investment planning, and cost control. Various approaches exist to improve spot price forecasting, including the analysis of historical data, expert opinions, and the use of futures prices. Futures prices are particularly promising as they can be interpreted as aggregated market expectations.

In this study, the models were applied to two perishable commodities, namely Skimmed Milk Powder ([SMP](#)) and butter. Spot prices are available at a monthly frequency, while futures prices are recorded daily. For modeling purposes, futures data were aggregated both temporally and structurally to create a suitable dataset.

Several models for spot price forecasting were examined: a random walk as a benchmark, Autoregressive Integrated Moving Average ([ARIMA](#)) as a classical statistical method based on historical data, the Kalman filter as a state-space model incorporating futures prices as exogenous variables, and eXtreme Gradient Boosting ([XGBoost](#)) as a modern machine learning approach also using futures data. The results show that [ARIMA](#) performs worse than a naive forecast on complex time series exhibiting properties such as non-stationarity, heteroskedasticity, structural instability, and considerable uncertainty. The Kalman filter and [XGBoost](#) proved to be more robust but only provided significantly better predictions than the naive forecast for futures with short maturities, where uncertainty and deviation from spot prices are low. While the Kalman filter is more sensitive and responsive over longer forecast horizons, [XGBoost](#) reacts more strongly to uncertainties in the training data.

Overall, the study demonstrates that reliable spot price forecasting requires a deep understanding of time series characteristics as well as the models and their underlying assumptions to select appropriate methods and thereby support well-founded decisions in commodity markets.

Keywords: Commodity markets, Perishable commodities, Spot price forecasting, Spot prices, Futures prices, Time series analysis, Price prediction, Risk management, Investment planning, [ARIMA](#), Kalman filter, State-space models, Machine learning, [XGBoost](#), [SMP](#), Butter

ZUSAMMENFASSUNG

Die genaue Vorhersage von Spot-Preisen, insbesondere in Commodity-Märkten, ist von großem Interesse, da die Preisprognose für unternehmerische Entscheidungen in zahlreichen Branchen entscheidend ist und zu einer fundierten Basis für Risikomanagement, Investitionsplanung und Kostenkontrolle beiträgt. Verschiedene Ansätze zur Verbesserung der Spot-Preisvorhersage existieren, darunter die Analyse historischer Daten, Expertenmeinungen sowie die Nutzung von Futures-Preisen. Futures-Preise sind besonders vielversprechend, da sie als aggregierte Markterwartungen interpretiert werden können.

In dieser Arbeit wurden die Modelle auf zwei verderbliche Rohstoffe, nämlich [SMP](#) und Butter, angewendet. Dabei liegen Spot-Preise in monatlichen und Futures-Preise in täglichen Frequenzen vor. Für die Modellierung wurden die Futures-Daten sowohl auf temporaler als auch auf struktureller Ebene aggregiert, um eine geeignete Datenbasis zu schaffen.

Untersucht wurden verschiedene Modelle zur Spot-Preisprognose: Ein Random Walk als Benchmark, [ARIMA](#) als klassisches statistisches Verfahren zur Zeitreihenanalyse basierend auf historischen Daten, der Kalman-Filter als State-Space-Modell unter Einbeziehung von Futures-Preisen als exogene Variable sowie [XGBoost](#) als moderner Machine-Learning-Ansatz mit Futures-Daten. Die Ergebnisse zeigen, dass [ARIMA](#) bei komplexen Zeitreihen mit Eigenschaften wie Nicht-Stationarität, Heteroskedastizität, struktureller Instabilität und hoher Unsicherheit schlechter abschneidet als eine naive Prognose. Kalman-Filter und [XGBoost](#) erweisen sich als robuster, liefern jedoch nur bei Futures mit kurzem Lieferzeitpunkt, bei denen geringe Unsicherheit und geringe Abweichung zum Spot-Preis vorliegen, signifikant bessere Vorhersagen. Während der Kalman-Filter bei längeren Prognosezeiträumen sensibler und reaktiver ist, reagiert [XGBoost](#) stärker auf Unsicherheiten in den Trainingsdaten.

Insgesamt zeigt die Untersuchung, dass für eine verlässliche Spot-Preisprognose ein tiefes Verständnis der Zeitreiheigenschaften sowie der Modelle und deren Annahmen notwendig ist, um geeignete Verfahren auszuwählen und somit fundierte Entscheidungen im Rohstoffmarkt zu unterstützen.

Keywords: Rohstoffmärkte, Verderbliche Rohstoffe, Spot-Preisprognose, Spot-Preise, Futures-Preise, Zeitreihenanalyse, Preisprognose, Risikomanagement, Investitionsplanung, [ARIMA](#), Kalman-Filter, State-Space-Modelle, Maschinelles Lernen, [XGBoost](#), [SMP](#), Butter