

Stochastische Verbesserung von Eingangsvariablen und Modellprognosen

Basierend auf Arbeiten mit Oliver Grothe, Thomas Möbius und Felix Müsgens

Mira Watermeyer | 13. Mai 2022

Inhalt

1. Motivation & Ansatz

2. Verbessern von Eingangsvariablen (Preprocessing)

3. Verbessern von Modellprognosen (Postprocessing)

4. Zusammenfassung

Motivation & Ansatz

oooo

Verbessern von Eingangsvariablen (Preprocessing)

oooooo

Verbessern von Modellprognosen (Postprocessing)

ooooooo

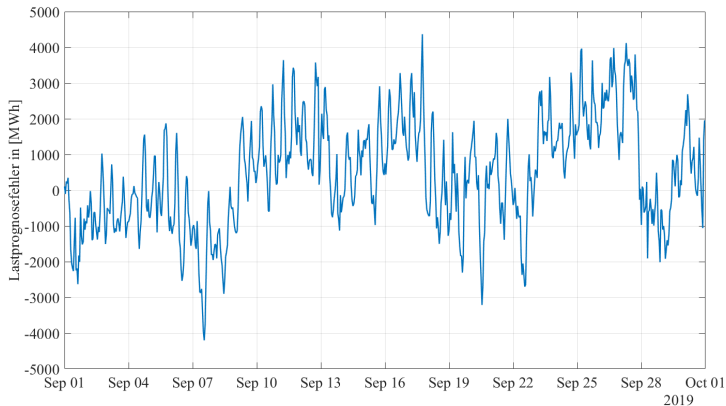
Zusammenfassung

o

Motivation

Modellierung von Fehlerzeitreihen

Fehler der Day-Ahead Lastprognose der Übertragungsnetzbetreiber (ÜNB) für Deutschland im September 2019



Motivation & Ansatz

●○○○

Verbessern von Eingangsvariablen (Preprocessing)

○○○○○

Verbessern von Modellprognosen (Postprocessing)

○○○○○○

Zusammenfassung

○

Motivation

Viele Prognosefehler

- besitzen saisonale Strukturen,
- sind autoregressiv, d.h. korrelieren mit vergangenen Werten ihrer selbst und
- korrelieren mit exogenen Größen.

Idee

Vorhersage des Prognosefehlers durch ein Zeitreihenmodell, welches die Muster und Strukturen erfasst

→ Kombination der Ausgangsdaten/Prognosen mit der Fehlervorhersage verkleinert den Fehler

Der Prognosefehler ϵ_t ergibt sich aus der Differenz der Echtdaten Y_t und der Prognosewerte \hat{Y}_t zu jedem Zeitpunkt t :

$$\epsilon_t = Y_t - \hat{Y}_t$$

Modellierung und Vorhersage dieses Prognosefehlers ϵ_t , so dass die Fehlervorhersage $\hat{\epsilon}_t = f(\cdot)$ durch das noch zu definierende Modell $f(\cdot)$ möglich ist.

Verbesserung der Prognose durch:

$$\hat{Y}_t^* = \hat{Y}_t + \hat{\epsilon}_t,$$

wobei

- $\hat{\epsilon}_t$: Vorhersage des Prognosefehlers
- \hat{Y}_t^* : Korrigierte Prognose für den Zeitpunkt t

- Reviews zeigen die Vielfältigkeit und Möglichkeiten im Bereich Forecasting (Bsp.: Rafal Weron 2014; Rafal Weron und Ziel 2019; Lago u. a. 2020; Petropoulos u. a. 2022)
- Fehlernachbearbeitung (Yang u. a. 2013)
- Empirische Studie und Kombination der Prognosen zweier Modellframeworks (Ziel und Rafal Weron 2018)
- Unterschiedliche Kalibrierungsfenster für ein Modell (Marcjasz, Serafin und Rafal Weron 2018)

Verbessern von Eingangsvariablen - Beispiel Last(prognose)

Motivation & Ansatz

○○○○

Verbessern von Eingangsvariablen (Preprocessing)

●○○○○○

Verbessern von Modellprognosen (Postprocessing)

○○○○○○○

Zusammenfassung

○

Verbessern von Eingangsvariablen (Preprocessing)

Daten

Der Datensatz umfasst die stündliche Lastprognose der Übertragungsnetzbetreiber sowie die stündliche reale Last für Deutschland und Luxemburg.

- Zeitraum: 01.01.2016 - 31.12.2019
- Quelle: ENTSO-E Transparency Platform 2021
- Einheit: MWh

Aus der Literatur ist bekannt, dass der Datensatz **hohe Vorhersagefehler** aufweist (Hirth, Mühlentpfordt und Bulkeley 2018; Maciejowska, Nitka und T. Weron 2021).

Verbessern von Eingangsvariablen (Preprocessing)

Modell

Zerlegung des Prognosefehlers ϵ_t in zwei Komponenten:

$$\epsilon_t = SC_t + RC_t$$

1. Saisonale Komponente SC_t :

$$SC_t = \sum_{d=1}^7 \sum_{h=1}^{24} HoW_t^{h,d} \cdot HS^{h,d}$$

$HoW_t^{h,d}$: Dummy-Variable für die Wochenstunde

l_w : Länge des Kalibrierungsfensters

$$HS^{h,d} = \frac{\sum_{s=t-l_w-1}^{t-l_w-1} \epsilon_s \cdot HoW_s^{h,d}}{\sum_{s=t-l_w-1}^{t-l_w-1} HoW_s^{h,d}}$$

$HS^{h,d}$: Wochenstunden-spezifischer Mittelwert über den Zeitraum des Kalibrierungsfensters

Verbessern von Eingangsvariablen (Preprocessing)

Modell

2. Residualkomponente RC_t :

Die Residualkomponente wird als stochastische Komponente über ein SARMA(1, 1)x(1, 1)₂₄ Modell abgebildet:

$$\begin{aligned} RC_t = & \phi_0 + \phi_1 \cdot RC_{t-1} + \phi_{24} \cdot RC_{t-24} - \phi_1 \phi_{24} \cdot RC_{t-25} \\ & + \omega_1 \cdot \psi_{t-1} + \omega_{24} \cdot \psi_{t-24} + \omega_1 \omega_{24} \cdot \psi_{t-25} \\ & + \psi_t \end{aligned}$$

ϕ_j : Koeffizient AR-Lag

ω_j : Koeffizient MA-Lag

ψ_t : Fehler des geschätzten SARMA-Modells zum Zeitpunkt t

Ergebnis

Jährliche Verbesserung

Mittelwerte, Standardabweichungen und Fehlermaße für die **Fehlerzeitreihe** der ÜNB Day-Ahead Lastprognose und die der verbesserten Lastprognose; RMSE und MAE in [MWh], Verbesserung in [%].

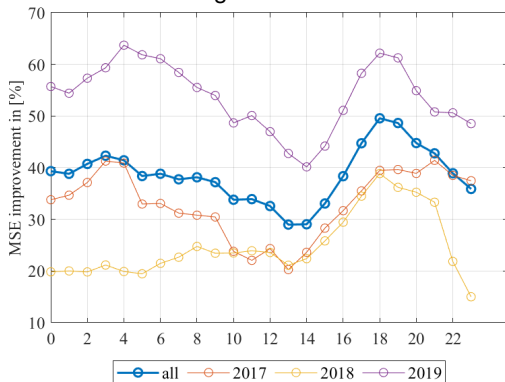
		all	2017	2018	2019
RMSE	TSO	2,224.63	1,802.61	2,360.51	2,454.70
	Preprocessed	1,746.69	1,483.45	2,043.87	1,665.76
	% Verbesserung	21.48	17.71	13.41	32.14
MAE	TSO	1,691.37	1,396.45	1,726.67	1,951.00
	Preprocessed	1,253.95	1,106.12	1,372.55	1,283.17
	% Verbesserung	25.86	20.79	20.51	34.23

Maciejowska, Nitka und T. Weron 2021 erreichen durch ein neues Prognosemodell für die Last Verbesserungen der Lastprognose im RMSE zwischen 19% und 34%, im MAE zwischen 27% und 38%.

Ergebnis

Verbesserung des MSE über den Tages- und Wochenverlauf

Stündliche Verbesserung des MSE in % über den Tagesverlauf



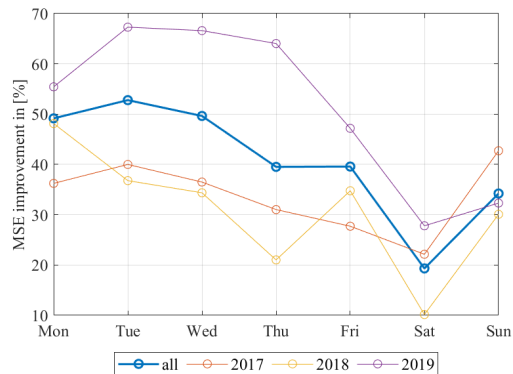
Motivation & Ansatz
○○○○

Verbessern von Eingangsvariablen (Preprocessing)
○○○○●

Verbessern von Modellprognosen (Postprocessing)
○○○○○○

Zusammenfassung
○

Wochentägliche Verbesserung des MSE in %



Verbessern von Modellprognosen - Beispiel Preisprognose

Motivation & Ansatz

○○○○

Verbessern von Eingangsvariablen (Preprocessing)

○○○○○○

Verbessern von Modellprognosen (Postprocessing)

●○○○○○○

Zusammenfassung

○

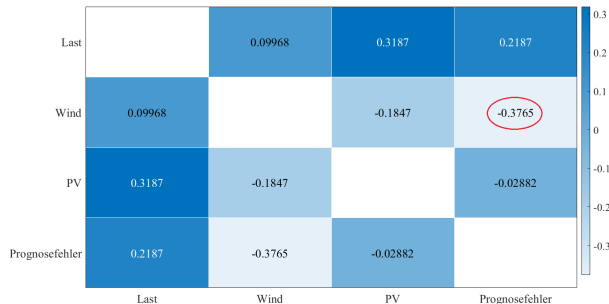
Verbessern von Modellprognosen (Postprocessing)

Daten

Preisprognosen als Output aus dem im ersten Vortrag vorgestellten Fundamentalmarktmodell
(Zeitraum: 01.01.2015 - 31.12.2020)

- Saisonale Abhängigkeiten und Autokorrelation im Prognosefehler
- Neu: Korrelation des Prognosefehlers mit exogenen Variablen, insbesondere mit Wind
- Modellwahl: Verwendung von Modellen mit Wind als exogenem Faktor

Korrelationsmatrix



Verbessern von Modellprognosen (Postprocessing)

Modell

Für die Verbesserung von Modellprognosen werden ein **univariater** und ein **multivariater** Modellansatz verwendet.

1. Univariater Ansatz

Modellierung der Fehlerzeitreihe als einen gesamten hochfrequenten Datensatz (über ARMAX Modell)

$$\begin{aligned}\epsilon_t = & \phi_0 + \phi_1 \cdot \epsilon_{t-1} + \phi_2 \cdot \epsilon_{t-2} + \phi_3 \cdot \epsilon_{t-24} + \phi_4 \cdot \epsilon_{t-168} + \phi_5 \cdot \psi_{t-1} \\ & + \omega_1 \cdot \epsilon_{min,d-1} + \omega_2 \cdot \epsilon_{max,d-1} + \omega_3 \cdot W_t + \omega_4 \cdot D_t \\ & + \psi_t\end{aligned}$$

ϕ_j : Koeffizient AR- und MA-Lag

ω_i : Koeffizient exogener Größe

ψ_t : Fehler des geschätzten ARMAX-Modells zum Zeitpunkt t

$\epsilon_{min,d-1}, \epsilon_{max,d-1}$: Minimaler/maximaler Prognosefehler des Vortages

W_t : Windprognose für den Zeitpunkt t

D_t : Feiertag-Dummyvariable für den Zeitpunkt t

Motivation & Ansatz

○○○○

Verbessern von Eingangsvariablen (Preprocessing)

○○○○○

Verbessern von Modellprognosen (Postprocessing)

○○●○○○

Zusammenfassung

○

Verbessern von Modellprognosen (Postprocessing)

Modell

2. Multivariater Ansatz

Aufspaltung der Fehlerzeitreihe in 24 Zeitreihen für jeweils eine Stunde des Tages (über ARX Modell)

$$\begin{aligned}\epsilon_{h,d} = & \phi_0 + \phi_1 \cdot \epsilon_{h,d-1} + \phi_2 \cdot \epsilon_{h,d-7} \\ & + \omega_1 \cdot \epsilon_{min,d-1} + \omega_2 \cdot \epsilon_{max,d-1} + \omega_3 \cdot W_{h,d} + \omega_4 \cdot D_{h,d} \\ & + \psi_{h,d}\end{aligned}$$

ϕ_j : Koeffizient AR-Lag

ω_j : Koeffizient exogener Größe

$\psi_{h,d}$: Fehler des geschätzten ARX-Modells zum Zeitpunkt (h, d)

$\epsilon_{min,d-1}, \epsilon_{max,d-1}$: Minimaler/maximaler Prognosefehler des Vortages

$W_{h,d}$: Windprognose für den Zeitpunkt (h, d)

$D_{h,d}$: Feiertag-Dummyvariable für den Zeitpunkt (h, d)

Motivation & Ansatz

○○○○

Verbessern von Eingangsvariablen (Preprocessing)

○○○○○

Verbessern von Modellprognosen (Postprocessing)

○○●○○

Zusammenfassung

○

Verbessern von Modellprognosen (Postprocessing)

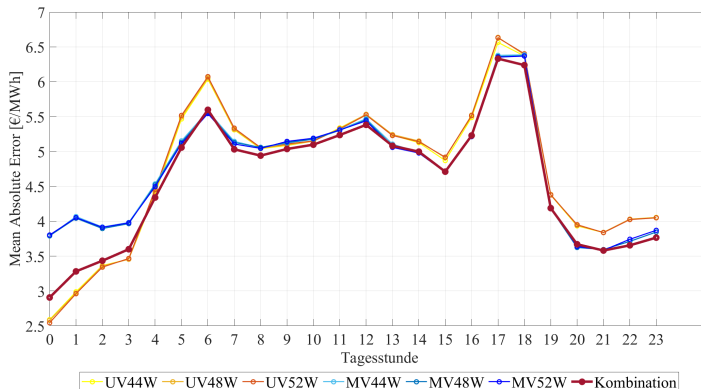
Variation und Kombination

- Modellschätzung: Wie viele historische Daten bilden die Grundlage?
 - Verwenden von drei unterschiedlichen Fensterlängen
- Zwischenergebnis: sechs individuelle Punktvorhersagen für den Fehler
- Arithmetisches Mittel aller Fehlervorhersagen kombiniert diese zu einer neuen Fehlerpunktvorhersage
- Referenzliteratur zu der Kombination verschiedener Kalibrierungsfenster bei Marcjasz, Serafin und Rafał Weron 2018, zu univariatem und multivariatem Modellframework bei Ziel und Rafał Weron 2018

Verbessern von Modellprognosen (Postprocessing)

Kombination der Teilmodelle

Stündlicher MAE der Vorhersage der Fehlerzeitreihe jedes Teilmodells



- Univariate Vorhersagen insbesondere für die ersten Tagesstunden besser
- Alle multivariaten MAE ab der 12. Tagesstunde unterhalb aller MAE des univariaten Modells
- Kombination verläuft nahe dem minimalen stündlichen MAE

Ergebnisse

Initiales Modell ESM mit Postprocessing

Fehlermaße der Fehlerzeitreihe (alle Werte in [€/MWh]):

		Initial	Postprocessing							
		ESM	UV 44 W.	UV 48 W.	UV 52 W.	MV 44 W.	MV 48 W.	MV 52 W.	Komb.	(Verb.)
Alle	RMSE	9.50	7.57	7.59	7.59	7.51	7.50	7.51	7.38	(22 %)
	MAE	6.00	4.73	4.75	4.75	4.75	4.74	4.74	4.60	(23 %)
2016	RMSE	7.60	6.07	6.13	6.09	5.95	5.94	5.92	5.82	(23 %)
	MAE	4.83	3.69	3.71	3.69	3.64	3.62	3.61	3.48	(28 %)
2017	RMSE	12.01	9.22	9.20	9.13	8.92	8.86	8.86	8.79	(27 %)
	MAE	7.09	5.55	5.55	5.51	5.37	5.31	5.28	5.25	(26 %)
2018	RMSE	8.61	7.49	7.47	7.54	7.40	7.38	7.40	7.28	(15 %)
	MAE	5.91	5.20	5.18	5.21	5.24	5.23	5.25	5.07	(14 %)
2019	RMSE	8.09	7.11	7.13	7.14	7.21	7.22	7.21	7.05	(13 %)
	MAE	5.07	4.46	4.49	4.49	4.61	4.61	4.59	4.43	(13 %)
2020	RMSE	10.49	7.65	7.69	7.71	7.77	7.79	7.85	7.63	(27 %)
	MAE	7.13	4.77	4.80	4.83	4.91	4.94	4.99	4.77	(33 %)

Referenzwerte: RMSE 7.62 €/MWh, MAE 4.24 €/MWh (für 2017) Lago u. a. 2020; RMSE 8.434 €/MWh, MAE 5.92 €/MWh (für 2016-2019)

Maciejowska, Nitka und T. Weron 2021; MAE 5.007 €/MWh (für 2012-2016) Ziel und Rafal Weron 2018

Motivation & Ansatz
○○○○

Verbessern von Eingangsvariablen (Preprocessing)
○○○○○○

Verbessern von Modellprognosen (Postprocessing)
○○○○○○●

Zusammenfassung
○

Zusammenfassung und Ausblick

- Vorhersagefehler haben **Strukturen**, z.B. ÜNB-Lastprognose
- **Zeitreihenansätze zur Abbildung der Vorhersagefehler** erlauben, diese zu verbessern
- Verbesserung ist **unabhängig vom Problem** möglich
- Zur **Verbesserung** genügen bereits **parameterarme Modelle**, z.B. Verbesserung einer Preisprognose um 22 % RMSE

- Zeitreihenansätze zu dem Vorhersagefehler erlauben, auch die Dichte des Vorhersagefehlers zu betrachten und Vorhersagedichten zu erzeugen

Arbeiten des Vortrages

Grothe u. a. 2022b, Paper under review

Grothe u. a. 2022a, Working paper

Literatur I

- [1] ENTSO-E Transparency Platform. *Total Load - Day Ahead / Actual*. 2021. URL: <https://transparency.entsoe.eu/>.
- [2] Oliver Grothe u. a. „Developing an Hybrid Model for day Ahead Electricity Price Forecasting“. Working Paper. 2022.
- [3] Oliver Grothe u. a. „Dynamic Preprocessing of Load Data for Energy System Models“. Paper under review. 2022.
- [4] Lion Hirth, Jonathan Mühlenpfordt und Marisa Bulkeley. „The ENTSO-E Transparency Platform – A review of Europe’s most ambitious electricity data platform“. In: *Applied Energy* 225 (2018), S. 1054–1067. ISSN: 0306-2619. DOI: <https://doi.org/10.1016/j.apenergy.2018.04.048>. URL: <https://www.sciencedirect.com/science/article/pii/S0306261918306068>.
- [5] Jesus Lago u. a. *Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark*. 2020. arXiv: 2008.08004 [stat.AP].

Literatur II

- [6] Katarzyna Maciejowska, Weronika Nitka und Tomasz Weron. „Enhancing load, wind and solar generation for day-ahead forecasting of electricity prices“. In: *Energy Economics* 99 (Juli 2021), S. 105273. ISSN: 01409883. DOI: 10.1016/j.eneco.2021.105273. URL: <https://linkinghub.elsevier.com/retrieve/pii/S014098832100178X>.
- [7] Grzegorz Marcjasz, Tomasz Serafin und Rafał Weron. „Selection of Calibration Windows for Day-Ahead Electricity Price Forecasting“. In: *Energies* 2018, Vol. 11, Page 2364 11 (9 Sep. 2018), S. 2364. ISSN: 1996-1073. DOI: 10.3390/EN11092364. URL: <https://www.mdpi.com/1996-1073/11/9/2364/htm%20https://www.mdpi.com/1996-1073/11/9/2364>.
- [8] Fotios Petropoulos u. a. „Forecasting: theory and practice“. In: *International Journal of Forecasting* (2022). ISSN: 0169-2070. DOI: <https://doi.org/10.1016/j.ijforecast.2021.11.001>. URL: <https://www.sciencedirect.com/science/article/pii/S0169207021001758>.

Literatur III

- [9] Rafał Weron. „Electricity price forecasting: A review of the state-of-the-art with a look into the future“. In: *International Journal of Forecasting* 30.4 (2014), S. 1030–1081. ISSN: 0169-2070. DOI: <https://doi.org/10.1016/j.ijforecast.2014.08.008>. URL: <https://www.sciencedirect.com/science/article/pii/S0169207014001083>.
- [10] Rafał Weron und Florian Ziel. „Electricity price forecasting“. In: *Routledge Handbook of Energy Economics*. Routledge, Sep. 2019, S. 506–521. DOI: 10.4324/9781315459653-36.
- [11] Yi Yang u. a. „A New Strategy for Short-Term Load Forecasting“. In: *Abstract and Applied Analysis* 2013 (2013). DOI: 10.1155/2013/208964. URL: <https://www.hindawi.com/journals/aaa/2013/208964/>.
- [12] Florian Ziel und Rafał Weron. „Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks“. In: *Energy Economics* 70 (2018), S. 396–420. ISSN: 0140-9883. DOI: <https://doi.org/10.1016/j.eneco.2017.12.016>. URL: <https://www.sciencedirect.com/science/article/pii/S014098831730436X>.