

Umweltforschungsplan des
Bundesministeriums für Umwelt,
Naturschutz, Bau und Reaktorsicherheit

Forschungskennzahl 3716 51 201 0
UBA-FB-00 [trägt die UBA-Bibliothek ein]

Sicherstellung der Ozonprognose

Schlussbericht

von

Dr. Lina Neunhäuserer

Volker Diegmann

Florian Tautz

Yvonne Breitenbach

IVU Umwelt GmbH
Emmy-Noether-Str. 2
D-79110 Freiburg

Telefon: +49 (0) 761 888 512 0

Fax: +49 (0) 761 888 512 12

Mail: info@ivu-umwelt.de

URL: www.ivu-umwelt.de

Im Auftrag des Umweltbundesamtes

Oktober 2018

Kurzbeschreibung

Das Umweltbundesamt (UBA) stellt seit Ende der 1990er Jahre u. a. Karten bundesweiter Ozonprognosen zur Information der Öffentlichkeit bereit. Diese Ozonprognosen beruhen auf Prognosen aus dem CAMS-Ensemble. Im Rahmen des hier beschriebenen Projekts wurde untersucht, inwieweit sich die Vorhersagegüte der CAMS-Ozonprognosen für Deutschland insbesondere hinsichtlich der Erfassung von Schwellenwertüberschreitungen durch eine nachgeschaltete Bearbeitung verbessern lässt. Eine entsprechende Untersuchung wurde auch für die CAMS-PM10-Prognosen durchgeführt.

Für die Nachbearbeitung werden im ersten Schritt die CAMS-Prognosen mit Hilfe von Messdaten an den Messorten korrigiert. Die dabei ermittelten Korrekturfaktoren an den Messorten werden dann im zweiten Schritt in die Fläche interpoliert und anschließend auf das Gitter der CAMS-Modellprognosen angewendet.

Es wurden drei verschiedene Korrekturverfahren (Hybrid-Vorhersage, Multiplikative Korrektur, Kalman-Filter) anhand der CAMS-Prognosedaten für das Jahr 2016 sowie zweier ausgewählter Schadstoffepisoden aus den Jahren 2015 und 2017 untersucht. Die Evaluierung sowohl der CAMS-Prognosedaten als auch der korrigierten Daten erfolgte anschließend auf der Basis statistischer Metriken und unter Verwendung des im Rahmen der FAIRMODE-Initiative entstandenen DELTA-Tools. Es zeigte sich, dass im Regelfall jedes der drei Korrekturverfahren die CAMS-Prognosen verbessert. Dabei führt der Kalman-Filter häufig zu den höchsten Korrelationskoeffizienten. Zudem erfüllen die Prognosedaten nach der Korrektur mit dem Kalman-Filter am besten die Qualitätskriterien des DELTA-Tools. Allerdings zeigt die Hybrid-Vorhersage häufig ein besseres Verhalten, wenn es um die Erfassung von Maximalwerten und die Überschreitung von Schwellenwerten geht.

Basierend auf den Ergebnissen der Evaluierung wurde ein Online-Tool zur Korrektur der CAMS-Prognosen im operationellen Betrieb entwickelt.

Abstract

Since the late 1990s, the Federal Environment Agency (UBA) has been publishing maps of nationwide ozone forecasts for information purposes. These ozone forecasts are based on forecasts from the CAMS ensemble. Aim of the project described here was to analyse the extent to which the predictive quality of the CAMS ozone forecasts for Germany can be improved by postprocessing, particularly with regard to the detection of limit value exceedances. A corresponding study was carried out for the CAMS-PM10 forecasts.

For postprocessing, the CAMS forecasts are corrected in a first step at measurement locations using measured data. In a second step, the correction factors determined at the measurement locations are interpolated and applied to the model grid of the CAMS forecasts.

Three different correction methods (hybrid forecast, multiplicative correction, Kalman filter) were analysed using the CAMS forecast data for 2016 and two selected pollutant episodes from 2015 and 2017. The evaluation of both the CAMS forecast data and the corrected data was then carried out on the basis of statistical metrics and using the DELTA tool developed as part of the FAIRMODE initiative. It turned out that generally each of the three correction methods improves the CAMS forecasts. The Kalman filter often leads to the highest correlation coefficients. In addition, after correction with the Kalman filter, the forecast data best meet the quality criteria of the DELTA tool. However, the hybrid forecast often shows better behaviour when it comes to the detection of maximum values and limit value exceedances.

Based on the results of the evaluation, an online tool for correcting the CAMS forecasts operationally was developed.

Inhaltsverzeichnis

Abbildungsverzeichnis	8
Tabellenverzeichnis	12
Abkürzungsverzeichnis	15
Zusammenfassung	16
Summary	28
1 Einführung	40
2 Vorgehensweise	42
3 Testumgebung	45
3.1 Aufbau	45
3.2 Eingangsdaten	47
3.2.1 Messdaten	47
3.2.2 Modelldaten	48
3.3 Implementierte Korrekturverfahren	48
3.3.1 Multiplikative Korrektur	48
3.3.2 Hybrid-Vorhersage	49
3.3.3 Kalman-Filter	49
3.4 Verwendete Metriken	54
3.4.1 Statistische Kennwerte	54
3.4.2 DELTA-Tool	57
4 Evaluierung der unkorrigierten Modelldaten	62
4.1 Ozon	62
4.1.1 Einfluss Mittelwertbildung der Modelldaten	62
4.1.2 Jahreszeitreihe 2016	63
4.1.3 Saisonale Auswertung 2016	73
4.2 PM10	79
4.2.1 Jahreszeitreihe 2016	79
4.2.2 Saisonale Auswertung 2016	86
5 Korrektur an den Messorten	91
5.1 Evaluierung der Ergebnisse	91
5.1.1 Ozon	91
5.1.1.1 Jahreszeitreihe 2016	91
5.1.1.2 Saisonale Auswertung 2016	105
5.1.1.3 Episode 2015	112
5.1.2 PM10	119

5.1.2.1	Jahreszeitreihe 2016	119
5.1.2.2	Saisonale Auswertung 2016	132
5.1.2.3	Episode 2017	139
5.2	Fazit.....	146
6	Korrektur in der Fläche	147
6.1	Systembeschreibung FLADIS.....	147
6.1.1	Interpolationsverfahren	148
6.1.2	Kreuzvalidierung.....	150
6.2	Ergebnisse der Interpolation	151
7	Online-Tool	154
7.1	Allgemeines.....	154
7.2	Aufbau.....	154
7.3	Implementierte Verfahren.....	156
7.3.1	Korrekturverfahren	156
7.3.2	Interpolationsverfahren	156
8	Schlussbemerkungen und Ausblick	157
9	Quellenverzeichnis.....	159

Abbildungsverzeichnis

Abbildung 3-1:	Ablaufschema Testumgebung	45
Abbildung 3-2:	Erforderliche CAMS-Prognosen der Vortage (violett) für die Korrektur der aktuellen CAMS-Prognose über die nächsten 72 h (grün) bei einem Vorlauf des Korrekturverfahrens von 3 Tagen. Der Zeitraum, in dem Messdaten verfügbar sind, ist orange hinterlegt. Die blauen Markierungen geben beispielhaft für die aktuellen Prognosestunden + 12 h und + 60 h an, welche Stundenwerte aus welchen Prognosen der Vortage zur Korrektur herangezogen werden.	46
Abbildung 3-3:	Beispielhafte Darstellung eines Soccer-Plots, der den Fractional Gross Error aufgetragen zeigt über den Mean Fractional Bias.....	56
Abbildung 3-4:	Beispielhafte Darstellung der vier Quadranten A – D zur Bestimmung der Metriken zur Bewertung von Schwellenwertüberschreitungen. Näheres siehe Text.	57
Abbildung 3-5:	Beispiel eines Benchmark-Reports des DELTA-Tools mit Assessment Target Plot (oben) und zusammenfassenden Statistiken (unten).....	59
Abbildung 3-6:	Beispiel eines Forecast-Sigmoid Target Plots des DELTA-Tools.....	61
Abbildung 4-1:	Box-Whisker-Plots ausgewählter Metriken für die unkorrigierten CAMS-Ozonprognosen, Bezugsjahr 2016	64
Abbildung 4-2:	Ozon-Zeitreihen von Mess- und Modelldaten an den Stationen mit dem niedrigsten (oben) bzw. dem höchsten (unten) Korrelationskoeffizienten, Bezugsjahr 2016.....	65
Abbildung 4-3:	Box-Whisker-Plots Schwellenwertüberschreitungen für die unkorrigierten CAMS-Ozonprognosen, Bezugsjahr 2016.....	67
Abbildung 4-4:	Assessment Target Plots (links) und Forecast-Sigmoid Target Plots (rechts) für CAMS-Rohdaten O ₃ , Bezugsjahr 2016, 1. bis 3. Prognosetag.....	72
Abbildung 4-5:	Soccer-Plots O ₃ für die Stundenmittelwerte 1SMW, 1. Prognosetag, Bezugsjahr 2016, unterteilt nach Jahreszeiten.....	76
Abbildung 4-6:	Soccer-Plots O ₃ für die maximalen gleitenden 8-Stundenmittelwerte eines Tages 8TMAX, 1. Prognosetag, Bezugsjahr 2016, unterteilt nach Jahreszeiten	77
Abbildung 4-7:	Soccer-Plots O ₃ für die maximalen Stundenmittelwerte eines Tages 1TMAX, 1. Prognosetag, Bezugsjahr 2016, unterteilt nach Jahreszeiten	78
Abbildung 4-8:	Box-Whisker-Plots ausgewählter Metriken für die unkorrigierten CAMS-PM10-Prognosen, Bezugsjahr 2016.....	80
Abbildung 4-9:	PM10-Zeitreihen von Mess- und Modelldaten an den Stationen mit dem niedrigsten (oben) bzw. dem höchsten (unten) Korrelationskoeffizienten, Bezugsjahr 2016.....	81
Abbildung 4-10:	Box-Whisker-Plots Schwellenwertüberschreitungen für die unkorrigierten CAMS-PM10-Prognosen, Bezugsjahr 2016.....	82

Abbildung 4-11:	Assessment Target Plots (links) und Forecast-Sigmoid Target Plots (rechts) für CAMS-Rohdaten PM10, Bezugsjahr 2016, 1. bis 3. Prognosetag.....	85
Abbildung 4-12:	Soccer-Plots PM10 für die Stundenmittelwerte 1SMW, 1. Prognosetag, Bezugsjahr 2016, unterteilt nach Jahreszeiten	89
Abbildung 4-13:	Soccer-Plots PM10 für die Tagesmittelwerte 1TMW, 1. Prognosetag, Bezugsjahr 2016, unterteilt nach Jahreszeiten.....	90
Abbildung 5-1:	Box-Whisker-Plots der Maxima an den Stationsorten für O ₃ , Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts), jeweils im Vergleich mit den Messwerten	94
Abbildung 5-2:	Box-Whisker-Plots der Standardabweichung an den Stationsorten für O ₃ , Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts), jeweils im Vergleich mit den Messwerten	95
Abbildung 5-3:	Soccer-Plots für O ₃ , Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	96
Abbildung 5-4:	Zeitreihen der 8TMAX für O ₃ , Station DEBW031, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	97
Abbildung 5-5:	Zeitreihen der 8TMAX für O ₃ , Station DENW030, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	98
Abbildung 5-6:	Assessment Target Plots O ₃ für CAMS-Rohdaten (links) und Multiplikative Korrektur (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag.....	100
Abbildung 5-7:	Assessment Target Plots O ₃ für Kalman-Filter (links) und Hybrid-Vorhersage (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag	101
Abbildung 5-8:	Forecast-Sigmoid Target Plots O ₃ für CAMS-Rohdaten (links) und Multiplikative Korrektur (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag.....	103
Abbildung 5-9:	Forecast-Sigmoid Target Plots O ₃ für Kalman-Filter (links) und Hybrid-Vorhersage (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag	104
Abbildung 5-10:	Box-Whisker-Plots der Hit Rate bezüglich der Informationsschwelle für O ₃ an den Stationsorten, Episode 2015, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts).....	116
Abbildung 5-11:	Zeitreihen der 1SMW für O ₃ , Station DEHE011, Episode 2015, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	117
Abbildung 5-12:	Zeitreihen der 8TMAX für O ₃ , Station DEHE011, Episode 2015, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	118

Abbildung 5-13:	Forecast-Sigmoid Target Plots für O ₃ , Episode 2015, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	119
Abbildung 5-14:	Box-Whisker-Plots der Maxima an den Stationsorten für PM10, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts), jeweils im Vergleich mit den Messwerten	122
Abbildung 5-15:	Box-Whisker-Plots der Standardabweichung an den Stationsorten für PM10, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts), jeweils im Vergleich mit den Messwerten	123
Abbildung 5-16:	Soccer-Plots für PM10, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	124
Abbildung 5-17:	Zeitreihen der 1TMW für PM10, Station DEHE024, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	125
Abbildung 5-18:	Zeitreihen der der 1TMW für PM10, Station DEBW022, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	126
Abbildung 5-19:	Assessment Target Plots PM10 für CAMS-Rohdaten (links) und Multiplikative Korrektur (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag	127
Abbildung 5-20:	Assessment Target Plots PM10 für Kalman-Filter (links) und Hybrid-Vorhersage (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag	128
Abbildung 5-21:	Forecast-Sigmoid Target Plots PM10 für CAMS-Rohdaten (links) und Multiplikative Korrektur (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag	130
Abbildung 5-22:	Forecast-Sigmoid Target Plots PM10 für Kalman-Filter (links) und Hybrid-Vorhersage (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag	131
Abbildung 5-23:	Box-Whisker-Plots der Hit Rate an den Stationsorten für PM10, Herbst (links) und Winter (rechts) 2016, Korrekturverfahren MK (oben), KF (Mitte) und HV (unten).....	138
Abbildung 5-24:	Zeitreihen der 1SMW für PM10, Station DETH060, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	141
Abbildung 5-25:	Zeitreihen der 1TMW für PM10, Station DETH060, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	142
Abbildung 5-26:	Zeitreihen der 1SMW für PM10, Station DEBY109, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	143

Abbildung 5-27:	Zeitreihen der 1TMW für PM10, Station DEBY109, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	144
Abbildung 5-28:	Forecast-Sigmoid Target Plots für PM10, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)	145
Abbildung 6-1:	Flussdiagramm des Programmsystems FLADIS	147
Abbildung 6-2:	Beispiel zur flächenhaften Korrektur: Ozon-Prognose 8TMAX für den 25.08.2016. CAMS-Rohdaten (oben links), nach Korrektur mit der Hybrid-Vorhersage, Interpolation der Korrekturfaktoren sowie Anwendung der flächenhaften Korrektur auf die CAMS-Rohdaten (oben rechts) sowie interpolierte Messdaten (unten links, vom Auftraggeber zur Verfügung gestellt)	153
Abbildung 7-1:	Ablaufschema Online-Tool	155

Tabellenverzeichnis

Tabelle 3-1:	Metriken O ₃ , Stundenmittelwerte 1SMW für den 1. Prognosetag, Korrekturverfahren KF mit $r = 0.06, 0.2, 0.3, 0.35, 0.4$ (jeweils Median über alle Stationen)	52
Tabelle 3-2:	Metriken O ₃ , maximale gleitende 8-Stundenmittelwerte eines Tages 8TMAX für den 1. Prognosetag, Korrekturverfahren KF mit $r = 0.06, 0.2, 0.3, 0.35, 0.4$ (jeweils Median über alle Stationen)	52
Tabelle 3-3:	Metriken PM ₁₀ , Stundenmittelwerte 1SMW für den 1. Prognosetag, Korrekturverfahren KF mit $r = 0.5, 0.8, 0.9, 1.0, 1.2$ (jeweils Median über alle Stationen)	53
Tabelle 3-4:	Metriken PM ₁₀ , Tagesmittelwerte 1TMW für den 1. Prognosetag, Korrekturverfahren KF mit $r = 0.5, 0.8, 0.9, 1.0, 1.2$ (jeweils Median über alle Stationen)	54
Tabelle 3-5:	Verwendete statistische Kennwerte / Metriken	55
Tabelle 4-1:	Metriken CAMS-Ozonprognose, basierend auf stundengemittelten Werten und auf Momentanwerten, 1. Prognosetag (jeweils Median über alle Stationen), Bezugsjahr 2016	62
Tabelle 4-2:	Metriken CAMS-Ozonprognose für die Stundenmittelwerte 1SMW, 1. bis 3. Prognosetag (jeweils Median über alle Stationen)	69
Tabelle 4-3:	Metriken CAMS-Ozonprognose für die maximalen gleitenden 8-Stundenmittelwerte eines Tages 8TMAX, 1. bis 3. Prognosetag (jeweils Median über alle Stationen)	69
Tabelle 4-4:	Metriken CAMS-Ozonprognose für die maximalen Stundenmittelwerte eines Tages 1TMAX, 1. bis 3. Prognosetag (jeweils Median über alle Stationen)	70
Tabelle 4-5:	Metriken O ₃ , Stundenmittelwerte 1SMW für den 1. Prognosetag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)	74
Tabelle 4-6:	Metriken O ₃ , maximale gleitende 8-Stundenmittelwerte eines Tages 8TMAX für den 1. Prognosetag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)	74
Tabelle 4-7:	Metriken O ₃ , maximale Stundenmittelwerte eines Tages 1TMAX für den 1. Prognosetag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)	75
Tabelle 4-8:	Metriken CAMS-PM ₁₀ -Prognose für die Stundenmittelwerte 1SMW, 1. bis 3. Prognosetag (jeweils Median über alle Stationen)	83
Tabelle 4-9:	Metriken CAMS-PM ₁₀ -Prognose für die Tagesmittelwerte 1TMW, 1. bis 3. Prognosetag (jeweils Median über alle Stationen)	84
Tabelle 4-10:	Metriken PM ₁₀ , Stundenmittelwerte 1SMW für den 1. Prognosetag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett	

	hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell).....	87
Tabelle 4-11:	Metriken PM ₁₀ , Tagesmittelwerte 1TMW für den 1. Prognosestag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell).....	88
Tabelle 5-1:	Metriken O ₃ , Stundenmittelwerte 1SMW für den 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell).....	92
Tabelle 5-2:	Metriken O ₃ , maximale gleitende 8-Stundenmittelwerte eines Tages 8TMAX für den 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell).....	92
Tabelle 5-3:	Metriken O ₃ , maximale Stundenmittelwerte eines Tages 1TMAX für den 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell).....	93
Tabelle 5-4:	Metriken O ₃ für 1SMW, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Frühling und Sommer 2016 (jeweils Median über alle Stationen).....	106
Tabelle 5-5:	Metriken O ₃ für 1SMW, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Herbst und Winter (jeweils Median über alle Stationen).....	107
Tabelle 5-6:	Metriken O ₃ für 8TMAX, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Frühling und Sommer 2016 (jeweils Median über alle Stationen).....	108
Tabelle 5-7:	Metriken O ₃ für 8TMAX, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Herbst und Winter 2016 (jeweils Median über alle Stationen).....	109
Tabelle 5-8:	Metriken O ₃ für 1TMAX, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Frühling und Sommer 2016 (jeweils Median über alle Stationen).....	110
Tabelle 5-9:	Metriken O ₃ für 1TMAX, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Herbst und Winter 2016 (jeweils Median über alle Stationen).....	111
Tabelle 5-10:	Metriken O ₃ , Stundenmittelwerte 1SMW Episode 2015 für den 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell).....	113
Tabelle 5-11:	Metriken O ₃ , maximale gleitende 8-Stundenmittelwerte eines Tages 8TMAX Episode 2015 für den 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell).....	114

Tabelle 5-12:	Metriken O ₃ , maximale Stundenmittelwerte eines Tages 1TMAX Episode 2015 für den 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell) 115
Tabelle 5-13:	Metriken PM ₁₀ , Stundenmittelwerte 1SMW für den 1. Prognosestag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell) 120
Tabelle 5-14:	Metriken PM ₁₀ , Tagesmittelwerte 1TMW für den 1. Prognosestag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell) 121
Tabelle 5-15:	Metriken PM ₁₀ für 1SMW, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Frühling und Sommer 2016 (jeweils Median über alle Stationen) 133
Tabelle 5-16:	Metriken PM ₁₀ für 1SMW, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Herbst und Winter (jeweils Median über alle Stationen) 134
Tabelle 5-17:	Metriken PM ₁₀ für 1TMW, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Frühling und Sommer 2016 (jeweils Median über alle Stationen) 135
Tabelle 5-18:	Metriken PM ₁₀ für 1TMW, 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Herbst und Winter 2016 (jeweils Median über alle Stationen) 136
Tabelle 5-19:	Metriken PM ₁₀ , Stundenmittelwerte 1SMW Episode 2017 für den 1. Prognosestag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell) 139
Tabelle 5-20:	Metriken PM ₁₀ , Tagesmittelwerte 1TMW für den 1. Prognosestag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell) 140
Tabelle 6-1:	O ₃ , Kenngrößen der Kreuzvalidierung für die Korrekturfaktoren der Hybrid-Vorhersage 151
Tabelle 6-2:	PM ₁₀ , Kenngrößen der Kreuzvalidierung für die Korrekturfaktoren der Hybrid-Vorhersage 151

Abkürzungsverzeichnis

1TMAX	maximaler Stundenmittelwert eines Tages
1TMW	Tagesmittelwert
8TMAX	maximaler gleitender 8-Stundenmittelwert eines Tages
CAMS	Copernicus Atmosphere Monitoring Service
CRMSE	zentrierter mittlerer quadratischer Fehler / Centralised Root Mean Square Error
CTM	Chemie-Transportmodell
FA	fälschlicherweise vorhergesagten Schwellenwertüberschreitungen / False Alarms
FAR	False Alarm Rate
FAIRMODE	Forum for air quality modelling in Europe
FGE	Fractional Gross Error
HV	Hybrid-Vorhersage
KF	Kalman-Filter
MA	nicht erfassten Schwellenwertüberschreitungen / Missed Alarms
MeasExc	Anzahl der gemessenen Überschreitungen
MEZ	mitteleuropäische Zeit
MFB	Mean Fractional Bias
MK	Multiplikative Korrektur
MPC	Modellierungsgüte-Kriterien / Modelling Performance Criteria
MPI	Modellierungsgüte-Indikator / Modelling Performance Indicator
MQI	Modellierungsqualitäts-Indikator / Modelling Quality Indicator
MQO	Modellierungsqualitäts-Ziel / Modelling Quality Objective
MRA	mittlere relative Abweichung
NMB	normierte Mittlere Abweichung / Normalised Mean Bias
NMSD	normierte mittlere Standardabweichung / Normalised Mean Standard Deviation
R	Korrelationskoeffizient
RSME	mittlerer quadratischer Fehler / Root Mean Square Error
SD	Standardabweichung / Standard Deviation
SMW	Stundenmittelwert
UBA	Umweltbundesamt
UTC	koordinierte Weltzeit / Coordinated Universal Time

Zusammenfassung

In Rahmen des europäischen Copernicus-Programms werden vom Copernicus Atmosphere Monitoring Service (CAMS) Daten zur atmosphärischen Schadstoffbelastung bereitgestellt. Dabei handelt es sich um die Ergebnisse von sieben Chemie-Transportmodellen (CTM) und deren Ensemble. Diese Modelle liefern täglich Prognosen für die nächsten 96 Stunden in stündlicher Auflösung für verschiedene Schadstoffe wie z. B. Ozon und PM10.

Die CTM-Simulationen werden in einer horizontalen Gitterweite von rund 10 km Kantenlänge durchgeführt und erfassen damit nur den regionalen Hintergrund. Im Vergleich mit Messwerten an Hintergrundstationen liefert das CAMS Ensemble in der Regel zu geringe maximale Ozonkonzentrationen. Dies hat negative Auswirkungen auf die korrekte Vorhersage von Schwellenwertüberschreitungen bei sehr hohen Ozonkonzentrationen zur Information und ggf. Warnung der Bevölkerung entsprechend der Luftqualitätsrichtlinie 2008/50/EG.

Das Umweltbundesamt (UBA) stellt seit Ende der 1990er Jahre u. a. Karten bundesweiter Ozonprognosen zur Information der Öffentlichkeit bereit. Die Ozonprognosen beruhen in der Vergangenheit auf einer statistischen Methode, die in der Folge zunächst durch die Prognosen aus dem CAMS-Ensemble ersetzt wurde. Im Rahmen des hier beschriebenen Projekts wurde untersucht, inwieweit sich die Vorhersagegüte der CAMS-Ozonprognosen für Deutschland insbesondere hinsichtlich der Erfassung von Schwellenwertüberschreitungen durch eine nachgeschaltete Bearbeitung verbessern lässt. Eine entsprechende Untersuchung wurde auch für die CAMS-PM10-Prognosen durchgeführt.

Die nachgeschaltete Bearbeitung der CAMS-Prognosen wurde in drei Schritten durchgeführt:

- a) Korrektur der CAMS-Prognosen mit Hilfe von Messdaten an den Messorten
- b) Interpolation der an den Messorten ermittelten Korrekturfaktoren in die Fläche, d. h. auf das Gitter, auf dem die CAMS-Prognosen vorliegen
- c) Anwendung der interpolierten Korrekturfaktoren auf die CAMS-Prognosen

Zur Identifikation geeigneter Korrekturverfahren für Schritt a) wurde eine Testumgebung in der Programmiersprache R implementiert mit drei verschiedenen Korrekturverfahren:

- ▶ Hybrid-Vorhersage (McKeen et al., 2005; Kang et al., 2008)
- ▶ Multiplikative Korrektur (McKeen et al., 2005; Borrego et al., 2011)
- ▶ Kalman-Filter (Delle Monache et al., 2006; Kang et al., 2008).

Abbildung I zeigt das Ablaufschema der Testumgebung. Zunächst werden die zu untersuchenden Daten eingelesen. Dies sind zum einen die Messdaten für den zu untersuchenden Schadstoff zusammen mit den Stationsmetadaten und zum anderen die zugehörigen CAMS-Prognosedaten. Die Messdaten werden für jede Station zu einer Jahreszeitreihe zusammengefügt und von MEZ nach UTC konvertiert, so dass sie mit den CAMS-Prognosedaten verglichen werden können. Mit Hilfe der Positionsdaten der Messstationen liest die Testumgebung dann die Prognosewerte an den Messorten aus den CAMS-Daten aus. Da die Messdaten als Stundenmittelwerte (SMW) und die CAMS-Prognosedaten als Momentanwerte zu jeder Stunde vorliegen, wurde in der Testumgebung die Möglichkeit vorgesehen, aus den Momentanwerten jeweils zu Anfang und zu Ende einer Stunde Stundenmittelwerte zu bilden.

Die CAMS-Prognosedaten liegen aus dem europäischen Copernicus-Programm täglich als Zeitreihen über die jeweils nächsten 96 Stunden ab Starttermin der Prognose vor. Im Rahmen dieses Projekts wurden davon jeweils die ersten 72 Stunden betrachtet. Um für die Testumgebung durchgehende Jahreszeitreihen der CAMS-Prognosen zu erhalten, werden diese aus den CAMS-Prognosezeitreihen in Abhängigkeit vom Prognosetag jeweils aus den Stunden des 1., 2. oder 3. Prognosetags gebildet.

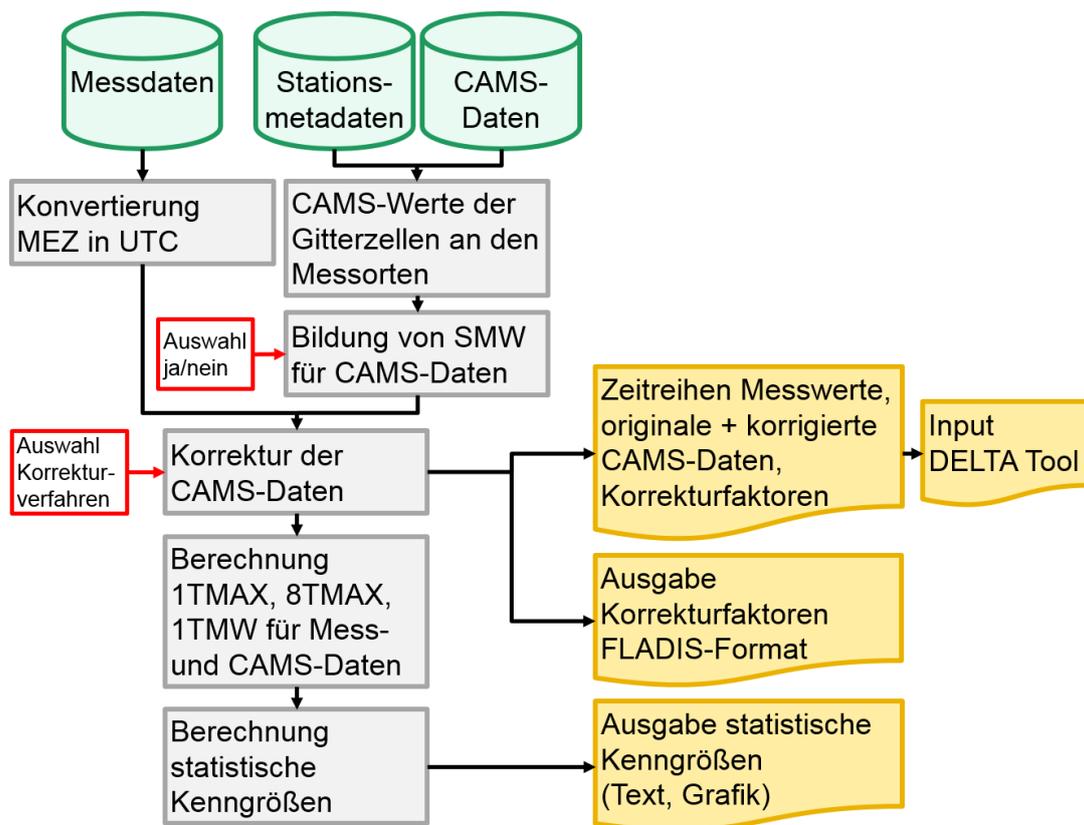
Anschließend werden die CAMS-Prognosedaten an den Messorten mit dem gewählten Korrekturverfahren für jede Stunde auf Basis der CAMS-Prognosedaten und der Messdaten der jeweils vorangegangenen Tage korrigiert. Die Anzahl der zur Korrektur herangezogenen vorangegangenen Tage richtet sich nach dem Korrekturverfahren.

Nach der Korrektur der CAMS-Daten werden für die Mess- und die korrigierten CAMS-Zeitreihen für die Schadstoffe Ozon und PM10 weitere nach 2008/50/EG relevante Bewertungsgrößen ermittelt. Für Ozon handelt es sich dabei um den höchsten Einstundenmittelwert pro Tag (1TMAX) und den höchsten Achtstundenmittelwert pro Tag aus stündlich gleitenden Achtstundenmittelwerten (8TMAX). Für PM10 wird tagesweise der Tagesmittelwert (1TMW) bestimmt.

Als Ausgangsbasis für die Evaluierung sowohl der unkorrigierten CAMS-Rohdaten als auch der implementierten Korrekturverfahren werden abschließend für die Stundenmittelwerte und die abgeleiteten Bewertungsgrößen 1TMAX und 8TMAX bzw. 1TMW die in Tabelle I aufgeführten Metriken bestimmt. Da die CAMS-Jahreszeitreihen in Abhängigkeit vom Prognosetag vorliegen, wurden auch die Metriken in Abhängigkeit vom Prognosetag bestimmt und darauf aufbauend die Evaluierung der Korrekturverfahren in Abhängigkeit vom Prognosetag durchgeführt.

Die berechneten Korrekturfaktoren und Metriken werden stationsweise als ASCII-Tabellen ausgegeben. Zusätzlich werden die Zeitreihen der betrachteten Messwerte und der originalen bzw. der korrigierten CAMS-Daten grafisch in Form von Zeitreihen- und Scatter-Plots aufbereitet.

Abbildung I: Ablaufschema Testumgebung



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Die in der Testumgebung implementierten Metriken zur Evaluierung der Korrekturverfahren nach Tabelle I werden für jede Station einzeln bestimmt und in Box-Whisker-Plots zusammengeführt, um die Spannweite der Metriken über alle Stationen darzustellen. Auf Basis des Mean Fractional Bias (MFB)

und des Fractional Gross Errors (FGE) werden von der Testumgebung zudem Soccer-Plots erstellt, die auf einen Blick die Charakterisierung der Modellfehler und die Einordnung der Vorhersagegüte des betrachteten Modells erlauben.

Tabelle I: Verwendete statistische Kennwerte / Metriken

Bezeichnung	Abkürzung
Minimum der Messwerte	Min Obs [$\mu\text{g}/\text{m}^3$]
Minimum der Modellwerte	Min Mod [$\mu\text{g}/\text{m}^3$]
Mittelwert der Messwerte	Mean Obs [$\mu\text{g}/\text{m}^3$]
Mittelwert der Modellwerte	Mean Mod [$\mu\text{g}/\text{m}^3$]
Maximum der Messwerte	Max Obs [$\mu\text{g}/\text{m}^3$]
Maximum der Modellwerte	Max Mod [$\mu\text{g}/\text{m}^3$]
Standardabweichung der Messwerte	SD Obs [$\mu\text{g}/\text{m}^3$]
Standardabweichung der Modellwerte	SD Mod [$\mu\text{g}/\text{m}^3$]
Korrelationskoeffizient nach Pearson	Korrelation
Bias	BIAS [$\mu\text{g}/\text{m}^3$]
Root Mean Square Error	RMSE [$\mu\text{g}/\text{m}^3$]
Mean Fractional Bias	MFB [%]
Fractional Gross Error	FGE [%]
Anzahl der gemessenen Überschreitungen	MeasExc
Hit Rate [%]	Hit Rate [%]
False Alarm Rate [%]	FAR [%]

In Erweiterung der statistischen Auswertung wurde das im Rahmen der FAIRMODE-Initiative entstandene DELTA-Tool für "Assessment & Planning" (A&P; Thunis & Cuvelier, 2017; FAIRMODE, 2017) auf die originalen und auf die korrigierten CAMS-Prognosen angewendet. Für die Analyse der Modellergebnisse wurden der „Assessment Target Plot“ im Benchmark-Modus und der „Forecast-Sigmoid Target Plot“ im Forecast-Modus verwendet. Die Target Plots fassen die im DELTA-Tool implementierten statistischen Parameter zur normierten Auswertung der Modellergebnisse in einem zentralen Diagramm zum MQI, dem Modellierungsqualitäts-Indikator, zusammen. Im Benchmark-Modus bewertet der MQI den Unterschied zwischen Modellergebnissen und Messungen im Vergleich zur Messunsicherheit der einzelnen Schadstoffkomponenten. Im Forecast-Modus trifft der MQI eine vergleichende Aussage zwischen zwei Modellen, in dem er für eine gegebene Vorhersagelänge die Modellprognose mit einem persistenten Modell vergleicht.

Die Evaluierung der originalen und der korrigierten CAMS-Prognosen erfolgte auf der Basis der in der Testumgebung implementierten statistischen Metriken und unter Verwendung des DELTA-Tools für das Jahr 2016 für Ozon und PM10 sowie für zwei ausgewählte Schadstoffepisoden aus den Jahren 2015 (Ozon) und 2017 (PM10). Die Untersuchungen für das Jahr 2016 wurden sowohl für die gesamten Jahreszeitreihen als auch getrennt nach Jahreszeiten durchgeführt.

Um den Einfluss der Bildung von Stundenmittelwerten für die CAMS-Daten zu betrachten, wurden die in Tabelle I aufgeführten Metriken testweise für O_3 für die Momentanwerte und für die stundengemittelten Werte bestimmt und miteinander verglichen. Im Ergebnis unterscheiden sich die Metriken für die stundengemittelten und die Momentanwerte nur marginal. Da die Verwendung stundengemittelter

CAMS-Daten den Stundenmittelwerten der Messdaten inhaltlich näher kommt, wurden im Weiteren nur noch stundengemittelte CAMS-Daten verwendet.

Die Evaluierung der Stundenmittelwerte 1SMW für Ozon hat gezeigt, dass die unkorrigierten CAMS-Prognosedaten die Messdaten im Mittel überschätzen, was sich in einem positiven Bias und einem positiven MFB niederschlägt. Tabelle II zeigt dazu exemplarisch die Mediane der Metriken nach Tabelle I für die Jahreszeitreihen 2016. Die gemessenen Maximalwerte der Jahreszeitreihen 2016 werden durch die CAMS-Rohdaten im Median über alle Stationen um rund 20 % unterschätzt, entsprechend ist auch die Standardabweichung der CAMS-Daten geringer als die der Messdaten. Der Median der Korrelationskoeffizienten liegt bei ca. 0.8. Bei saisonaler Betrachtung werden die Minimal-, Mittel- und Maximalwerte sowie die Standardabweichung der Messdaten im Median am besten im Winter durch die Modelldaten erfasst. Die größten Differenzen zwischen Mess- und Modelldaten zeigen sich für diese Metriken im Sommer. Die Korrelation zwischen Mess- und CAMS-Rohdaten hingegen ist im Winter sowie im Frühling mit 0.76 am niedrigsten, hier liegen Sommer und Herbst mit 0.81 bzw. 0.83 deutlich höher.

Grundsätzlich ergibt sich durch jedes Korrekturverfahren eine Verbesserung der CAMS-Rohdaten (Tabelle II, Spalten „MK“, „KF“, „HV“). Für die 1SMW erzeugt im Regelfall der Kalman-Filter die besten Ergebnisse. Ausnahme davon ist die Erfassung der Maximalwerte und der Standardabweichung, diese Größen werden am besten von der Hybrid-Vorhersage erreicht. Das deckt sich mit den Ergebnissen von Kang et al. (2008). Die Metriken der Multiplikativen Korrektur liegen überall im Mittelfeld, insbesondere werden damit höhere Maximalwerte erreicht als mit dem Kalman-Filter. Allerdings hat die Evaluierung gezeigt, dass die Multiplikative Korrektur ohne Beschränkung des Korrekturfaktors ein problematisches Korrekturverfahren darstellt, das in Einzelfällen extrem hohe Konzentrationen berechnet. Die Ergebnisse der Evaluierung der Korrekturverfahren gelten im Wesentlichen sowohl für die Betrachtung der Jahreszeitreihen 2016 als auch für die saisonale Betrachtung 2016 und für die Episode 2015. Dabei wird bei der Episode 2015 der Kalman-Filter hinsichtlich der Erfassung der Maximalwerte und der Standardabweichung deutlich von der Hybrid-Vorhersage übertroffen.

Tabelle II: Metriken O₃, Stundenmittelwerte 1SMW für den 1. Prognosestag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	0.52	0.52	0.52	0.52
Min Mod [$\mu\text{g}/\text{m}^3$]	0.82	0.23	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	44.72	45.54	44.87	44.95
Mean Mod [$\mu\text{g}/\text{m}^3$]	56.36	45.18	44.98	45.37
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.51	173.66	173.51
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	163.05	153.65	169.84
SD Obs [$\mu\text{g}/\text{m}^3$]	29.44	29.68	29.49	29.46
SD Mod [$\mu\text{g}/\text{m}^3$]	22.29	27.08	26.67	29.01
Korrelation	0.81	0.82	0.84	0.80
BIAS [$\mu\text{g}/\text{m}^3$]	10.17	-0.39	0.11	0.34
RMSE [$\mu\text{g}/\text{m}^3$]	19.82	17.21	15.85	18.08
MFB [%]	38.27	8.57	4.23	-2.23
FGE [%]	48.91	44.03	44.82	52.26

Bei den maximalen gleitenden Achtstundenmittelwerten eines Tages 8TMAX liegt der Median der Mittelwerte der CAMS-Rohdaten immer noch höher als derjenige der Messdaten, die Differenz ist jedoch geringer als bei den 1SMW. Entsprechend sinken auch Bias und MFB gegenüber den 1SMW. Tabelle III zeigt dazu exemplarisch die Mediane der Metriken nach Tabelle I für die Jahreszeitreihen 2016. Die Maximalwerte werden auch bei den 8TMAX durch die CAMS-Daten unterschätzt, ebenso die Standardabweichung. In der Folge liegt der Median der Hit Rate über alle Stationen für die Jahreszeitreihen 2016 bei nur ca. 13 %. Der Median der Korrelationskoeffizienten liegt bei knapp 0.9. Bei saisonaler Betrachtung ergibt sich für die 8TMAX ein etwas anderes Bild als für die 1SMW. Minimalwerte und Standardabweichung der Messdaten werden hier im Median am Besten im Frühling getroffen, die Mittelwerte im Sommer und die Maximalwerte wie bei den 1SMW im Winter. Die Korrelation zwischen Mess- und CAMS-Daten ist mit 0.91 im Herbst am höchsten und mit 0.79 im Winter am niedrigsten.

Hinsichtlich der Korrekturverfahren erzeugen die Hybrid-Vorhersage und der Kalman-Filter die besten Ergebnisse für die 8TMAX, wobei die Hybrid-Vorhersage wie schon zuvor die Maximalwerte und die Standardabweichung, aber nun auch den Mittelwert am besten trifft, während der Kalman-Filter etwas bessere Werte bei Korrelation, RMSE, MFB und FGE aufweist (Tabelle III, Spalten „MK“, „KF“, „HV“). Die Ergebnisse der Evaluierung der Korrekturverfahren gelten im Wesentlichen sowohl für die Betrachtung der Jahreszeitreihen 2016 als auch für die saisonale Betrachtung 2016 und für die Episode 2015. Aufgrund der besseren Erfassung der Maximalwerte erreicht die Hybrid-Vorhersage die höchste Hit Rate, d. h. den höchsten Anteil korrekt vorhergesagter Schwellenwertüberschreitungen. Gleichzeitig ist damit die FAR, d. h. der höchste Anteil falsch vorhergesagter Schwellenwertüberschreitungen, für die Hybrid-Vorhersage am höchsten, allerdings auf niedrigem Niveau von wenigen Prozent.

Tabelle III: Metriken O₃, maximale gleitende 8-Stundenmittelwerte eines Tages 8TMAX für den 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	3.08	3.12	3.08	3.12
Min Mod [$\mu\text{g}/\text{m}^3$]	17.49	6.78	3.52	2.80
Mean Obs [$\mu\text{g}/\text{m}^3$]	64.59	64.76	64.87	64.75
Mean Mod [$\mu\text{g}/\text{m}^3$]	69.07	62.51	63.22	65.30
Max Obs [$\mu\text{g}/\text{m}^3$]	155.21	152.90	155.21	154.67
Max Mod [$\mu\text{g}/\text{m}^3$]	130.41	148.84	142.19	153.63
SD Obs [$\mu\text{g}/\text{m}^3$]	29.46	29.59	29.56	29.46
SD Mod [$\mu\text{g}/\text{m}^3$]	22.66	28.25	27.18	28.47
Korrelation	0.89	0.89	0.91	0.89
BIAS [$\mu\text{g}/\text{m}^3$]	3.63	-2.00	-1.57	0.88
RMSE [$\mu\text{g}/\text{m}^3$]	14.56	13.97	12.30	13.18
MFB [%]	12.68	-1.23	0.08	3.11
FGE [%]	22.90	22.16	19.61	20.60
gemessene Überschreitungen	15 ¹	14 ¹	15 ¹	15 ¹
Hit Rate [%]	13.33 ¹	53.33 ¹	50.00 ¹	65.45¹
FAR [%]	0.00¹	0.62 ¹	0.32 ¹	1.17 ¹

¹: Schwellenwert 120 $\mu\text{g}/\text{m}^3$

Die maximalen Stundenmittelwerte eines Tages 1TMAX liegen im Median der Mittelwerte nur noch wenig über den gemessenen Werten, entsprechend niedrig sind Bias und MFB für diese Größe. Tabelle IV zeigt dazu exemplarisch die Mediane der Metriken nach Tabelle I für die Jahreszeitreihen 2016. Die Unterschätzung der gemessenen Maximalwerte durch die CAMS-Daten entspricht derjenigen für die Stundenmittelwerte, entsprechend ist die Standardabweichung der CAMS-Daten auch bei den 1TMAX geringer als die der Messdaten. Die Anzahl der gemessenen Überschreitungen der Informationsschwelle von $180 \mu\text{g}/\text{m}^3$ war 2016 gering, keine dieser wenigen Überschreitungen wurde durch die CAMS-Daten erfasst. Der Median der Korrelationskoeffizienten liegt wie bei den 8TMAX bei knapp 0.9. Die saisonalen Ergebnisse entsprechen weitestgehend denjenigen für die 8TMAX.

Von den drei Korrekturverfahren liefern Kalman-Filter und Hybrid-Vorhersage die besten Ergebnisse für die 1TMAX, wobei wie zuvor die Hybrid-Vorhersage die Mittel- und Maximalwerte sowie die Standardabweichung besser trifft und der Kalman-Filter die besseren Werte für Korrelation, RMSE, MFB und FGE erzeugt (Tabelle IV, Spalten „MK“, „KF“, „HV“). Die Ergebnisse der Evaluierung der Korrekturverfahren gelten im Wesentlichen sowohl für die Betrachtung der Jahreszeitreihen 2016 als auch für die saisonale Betrachtung 2016. Für die Episode 2015 zeigt die Hybrid-Vorhersage in allen relevanten Größen die besten bzw. für Bias und MFB die zweitbesten Werte. Hinsichtlich der Erfassung von Schwellenwertüberschreitungen zeigen Auswertungen der Box-Whisker-Plots der Hit Rate für die Episode 2015, dass im oberen Quartil bis zu 67 % und im Maximum bis zu 100 % der Überschreitungen der Informationsschwelle durch die Korrekturverfahren erkannt werden. Den niedrigsten Wert im oberen Quartil weist hier mit 50 % der Kalman-Filter auf. Durch die unkorrigierten CAMS-Prognosen werden keine Überschreitungen der Informationsschwelle erkannt.

Tabelle IV: Metriken O₃, maximale Stundenmittelwerte eines Tages 1TMAX für den 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	4.79	4.86	4.81	4.86
Min Mod [$\mu\text{g}/\text{m}^3$]	20.36	9.03	6.64	4.61
Mean Obs [$\mu\text{g}/\text{m}^3$]	71.73	71.92	72.01	71.76
Mean Mod [$\mu\text{g}/\text{m}^3$]	72.56	69.44	69.58	73.48
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.06	173.66	173.66
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	162.92	153.57	169.73
SD Obs [$\mu\text{g}/\text{m}^3$]	30.63	30.62	30.62	30.64
SD Mod [$\mu\text{g}/\text{m}^3$]	23.19	29.27	28.10	29.69
Korrelation	0.88	0.87	0.90	0.87
BIAS [$\mu\text{g}/\text{m}^3$]	0.22	-2.27	-2.31	1.84
RMSE [$\mu\text{g}/\text{m}^3$]	15.23	15.58	13.73	15.12
MFB [%]	5.64	-1.88	-1.28	4.23
FGE [%]	20.13	20.95	18.64	19.87
gemessene Überschreitungen	0	0	0	0
Hit Rate [%]	0.00	0.00	0.00	0.00
FAR [%]	0.00	0.00	0.00	0.00

¹: Schwellenwert $180 \mu\text{g}/\text{m}^3$ (Informationsschwelle)

Die Evaluierung der 1SMW für PM10 hat gezeigt, dass die CAMS-Rohdaten die Messdaten im Mittel unterschätzen, was sich auch in einem negativen Bias und einem negativen MFB niederschlägt. Tabelle V zeigt dazu exemplarisch die Mediane der Metriken nach Tabelle I für die Jahreszeitreihen 2016. Die gemessenen Maximalwerte werden durch die CAMS-Daten im Median über alle Stationen um knapp zwei Drittel unterschätzt, entsprechend ist auch die Standardabweichung der CAMS-Daten deutlich geringer als die der Messdaten. Dies steht im Kontrast zu Ozon, wo die CAMS-Rohdaten die Messwerte im Mittel über- und nur in den Maximalwerten unterschätzen. Der Median der Korrelationskoeffizienten liegt bei 0.5, im Gegensatz zu ca. 0.8 bei Ozon. Bei saisonaler Betrachtung werden Mittel- und Maximalwerte sowie die Standardabweichung der Messdaten im Median am besten im Frühling erfasst. Die größten Differenzen zwischen Mess- und Modelldaten zeigen sich für diese Metriken im Winter, bis auf den Mittelwert, der im Sommer am schlechtesten vom Modell wiedergegeben wird. So werden z. B. die Maximalwerte im Winter im Median um ca. 67 % unterschätzt. Die Korrelation zwischen Mess- und CAMS-Daten ist im Herbst mit 0.58 am höchsten und im Sommer mit 0.43 am niedrigsten.

Für die Korrekturverfahren gilt wie bei Ozon, dass jedes Korrekturverfahren zu einer Verbesserung der CAMS-Prognosen führt. Der Kalman-Filter erzeugt für die Jahreszeitreihen 2016 insgesamt gute Metriken, wird allerdings hinsichtlich der Maximalwerte von der Hybrid-Vorhersage übertroffen, und die besten Metriken für die Standardabweichung zeigt hier die Multiplikative Korrektur (Tabelle V, Spalten „MK“, „KF“, „HV“). Bei saisonaler Betrachtung liegt das Hybrid-Verfahren vorn, das bis auf wenige Ausnahmen am besten oder gut die Mittel- und Maximalwerte sowie die Standardabweichung erfasst, sowie der Kalman-Filter, der außer im Frühling die besten Werte für Korrelation, RMSE und MFB erzielt. Für die Episode 2017 werden die besten Metriken bis auf wenige Ausnahmen von der Hybrid-Vorhersage erzielt. An zweiter Stelle stehen die Ergebnisse des Kalman-Filters.

Tabelle V: Metriken PM10, Stundenmittelwerte 1SMW für den 1. Prognosestag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	0.50	0.50	0.50	0.50
Min Mod [$\mu\text{g}/\text{m}^3$]	0.91	0.65	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.24	16.16	16.13	16.21
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.78	16.37	16.17	16.31
Max Obs [$\mu\text{g}/\text{m}^3$]	141.93	127.50	128.45	128.45
Max Mod [$\mu\text{g}/\text{m}^3$]	50.67	116.15	90.12	137.48
SD Obs [$\mu\text{g}/\text{m}^3$]	11.57	11.29	11.30	11.36
SD Mod [$\mu\text{g}/\text{m}^3$]	6.29	11.41	10.10	11.89
Korrelation	0.50	0.56	0.58	0.53
BIAS [$\mu\text{g}/\text{m}^3$]	-3.79	0.28	0.10	0.11
RMSE [$\mu\text{g}/\text{m}^3$]	10.87	10.55	9.68	11.09
MFB [%]	-16.34	3.78	2.23	-4.43
FGE [%]	48.65	44.02	45.95	52.05

Bei den Tagesmittelwerten 1TMW liegt der Median der Mittelwerte der modellierten Daten wie bei den 1SMW niedriger als derjenige der Messdaten. Tabelle VI zeigt dazu exemplarisch die Mediane der Metriken nach Tabelle I für die Jahreszeitreihen 2016. Die Maximalwerte der 1TMW werden durch die

CAMS-Rohdaten für die Jahreszeitreihen 2016 im Median um die Hälfte unterschätzt, die Standardabweichung der Modelldaten liegt rund 40 % niedriger als die der Messdaten. In der Folge liegt der Median der Hit Rate über alle Stationen bei 0 %. Der Median der Korrelationskoeffizienten liegt bei knapp 0.6. Bei saisonaler Betrachtung werden Mittelwerte und Standardabweichung der Messdaten wie bei den 1SMW im Median am Besten im Frühling getroffen, die Maximalwerte in Frühling und Herbst. Im Winter werden die Maximalwerte im Median um ca. 57 % unterschätzt. Die Korrelation zwischen Mess- und CAMS-Rohdaten ist mit 0.69 im Herbst am höchsten und mit 0.53 im Winter am niedrigsten.

Hinsichtlich der Korrekturverfahren zeigt der Kalman-Filter für die meisten Größen die besten Werte, die beste Erfassung der Maximalwerte und der Standardabweichung zeigt diesmal aber nicht die Hybrid-Vorhersage, sondern die Multiplikative Korrektur (Tabelle VI, Spalten „MK“, „KF“, „HV“). Die im Median gemessenen Überschreitungen des Schwellenwerts von $50 \mu\text{g}/\text{m}^3$ werden von keinem der Verfahren angezeigt. Auswertungen der Box-Whisker-Plots der Hit Rate zeigen aber, dass im oberen Quartil bis zu 50 % und im Maximum bis zu 100 % der Überschreitungen des Schwellenwerts durch die Korrekturverfahren erkannt werden. Die CAMS-Rohdaten erkennen im Maximum 25% der Überschreitungen. Bei saisonaler Betrachtung erzeugt der Kalman-Filter bis auf eine Ausnahme den höchsten Korrelationskoeffizienten und die niedrigsten Werte für RMSE und FGE. Mittel- und Maximalwerte sowie die Standardabweichung werden im Regelfall am besten durch die Hybrid-Vorhersage erfasst. In der überschreitungsrelevanten Jahreszeit Winter werden Maximalwerte und Standardabweichung am besten durch die Multiplikative Korrektur erfasst und am zweitbesten durch die Hybrid-Vorhersage. Im Gegensatz dazu werden für die Episode 2017 die besten Metriken bis auf wenige Ausnahmen von der Hybrid-Vorhersage erzielt, an zweiter Stelle stehen die Ergebnisse des Kalman-Filters.

Tabelle VI: Metriken PM10, Tagesmittelwerte 1TMW für den 1. Prognosetag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	2.02	2.05	2.02	2.03
Min Mod [$\mu\text{g}/\text{m}^3$]	3.10	2.79	1.97	1.41
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.27	16.18	16.15	16.22
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.77	16.38	16.16	16.34
Max Obs [$\mu\text{g}/\text{m}^3$]	72.84	66.99	67.31	67.31
Max Mod [$\mu\text{g}/\text{m}^3$]	35.88	65.94	58.33	71.18
SD Obs [$\mu\text{g}/\text{m}^3$]	9.20	9.05	8.93	9.06
SD Mod [$\mu\text{g}/\text{m}^3$]	5.48	9.19	8.23	9.24
Korrelation	0.58	0.69	0.72	0.72
BIAS [$\mu\text{g}/\text{m}^3$]	-3.80	0.29	0.09	0.11
RMSE [$\mu\text{g}/\text{m}^3$]	8.29	7.25	6.27	6.77
MFB [%]	-21.50	2.49	2.47	0.33
FGE [%]	39.01	29.64	27.62	28.81
gemessene Überschreitungen	2 ¹	2 ¹	2 ¹	2 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00¹	0.59 ¹	0.28 ¹	0.28 ¹

¹: Schwellenwert $50 \mu\text{g}/\text{m}^3$

Wird das DELTA-Tool im Benchmark-Modus auf die Jahreszeitreihen 2016 der Messwerte und der CAMS-Rohdaten angewendet, so zeigt sich, dass für O_3 bereits die CAMS-Rohdaten für alle Stationen einen $MQI < 1$ aufweisen und damit das vom DELTA-Tool definierte Modellqualitätsziel erfüllen. Die Auswertung bezieht sich dabei ausschließlich auf die 8TMAX. Ein Teil der Stationen weist sogar einen $MQI < 0.5$ auf, d. h. der RMSE zwischen Messung und Modell an diesen Stationen ist kleiner als die Messunsicherheit. Das 90 %-Perzentil der MQI aller Stationen beträgt für den 1. Prognosetag 0.346. Durch Anwendung der Korrekturverfahren wird das 90 %-Perzentil der MQI aller Stationen nur noch geringfügig verringert, der niedrigste Wert von 0.301 am 1. Prognosetag ergibt sich bei Anwendung des Kalman-Filters.

Im Forecast-Modus ist der MQI der Jahreszeitreihen 2016 der CAMS-Rohdaten deutlich schlechter als im Benchmark-Modus. Eine ganze Reihe von Stationen weist für den 1. Prognosetag einen $MQI > 1$ auf und wird damit schlechter modelliert als mit dem persistenten Modell. Das 90%-Perzentil der MQI aller Stationen liegt für den 1. Prognosetag bei 1.05 und ist damit ebenfalls > 1 . Für die Episode 2015 liegt das 90%-Perzentil der MQI aller Stationen am 1. Prognosetag für die CAMS-Rohdaten mit 1.37 sogar deutlich über 1. Auch für die Multiplikative Korrektur liegt das 90%-Perzentil der MQI aller Stationen mit 1.06 höher als 1 und ist damit immer noch schlechter als das persistente Modell. Das niedrigste 90%-Perzentil der MQI aller Stationen weist in der Episode das Hybrid-Verfahren mit 0.757 auf.

Bei PM10 beziehen sich die Auswertungen des DELTA-Tools auf die 1TMW. Für die Jahreszeitreihen 2016 der CAMS-Rohdaten ist der MQI einzelner Stationen größer 1. Das 90%-Perzentil der MQI aller Stationen liegt für die CAMS-Rohdaten mit 0.947 immer noch knapp unter 1. Damit bewertet das DELTA-Tool die CAMS-PM10-Prognosen wesentlich schlechter als die CAMS- O_3 -Prognosen. Die Anwendung der Korrekturverfahren führt in jedem Fall zu einer Verbesserung der Kennwerte für PM10. Der MQI aller Stationen liegt für die korrigierten CAMS-Daten bis auf wenige Ausnahmen (Multiplikative Korrektur) unter 1, und das 90%-Perzentil der MQI aller Stationen sinkt am 1. Prognosetag bis auf 0.703 (Kalman-Filter).

Im Forecast-Modus ist der MQI der Jahreszeitreihen 2016 der CAMS-Rohdaten wie schon bei O_3 deutlich schlechter als im Benchmark-Modus. Die meisten Stationen weisen einen $MQI > 1$ auf und werden damit schlechter modelliert als mit dem persistenten Modell. Das 90%-Perzentil der MQI aller Stationen liegt am 1. Prognosetag mit 1.29 deutlich über 1. Für die Episode 2017 liegt das 90%-Perzentil der MQI aller Stationen mit 2.08 deutlich über 1. Die Multiplikative Korrektur erscheint als Korrekturverfahren für diese Episode nicht tauglich zu sein, denn auch hier liegt das 90%-Perzentil der MQI mit 1.26 deutlich über 1. Sogar der Kalman-Filter erreicht für diese Episode nur ein 90%-Perzentil der MQI von 1.00 und ist damit gerade ebenso gut wie das persistente Modell. Das niedrigste 90%-Perzentil der MQI weist in der Episode die Hybrid-Vorhersage mit 0.942 auf.

Der Vergleich der Metriken und des 90%-Perzentils der MQI für die einzelnen Prognosetage hat gezeigt, dass der Einfluss der Vorhersagelänge auf die Metriken und auf den MQI im Benchmark-Modus sowohl bei O_3 als auch bei PM10 gering ist. Im Forecast-Modus sinkt der MQI mit steigender Vorhersagelänge, was bedeutet, dass die Vorhersagegüte der originalen und der korrigierten CAMS-Daten mit zunehmender Vorhersagelänge weniger stark abnimmt als die des persistenten Modells.

Es lässt sich zusammenfassen, dass von den untersuchten Korrekturverfahren der Kalman-Filter und die Hybrid-Vorhersage die besten Prognosedaten im Vergleich mit den Messdaten erzeugen. Dabei erzielt der Kalman-Filter häufiger bessere Ergebnisse im Sinne der betrachteten Metriken, d. h. für die damit korrigierten Prognosedaten ergeben sich im Regelfall höhere Korrelationskoeffizienten und niedrigere Werte für RMSE, MFB und/oder FGE. Auch das 90 %-Perzentil der MQI aller Stationen fällt sowohl im Benchmark-Modus als auch im Forecast-Modus für den Kalman-Filter am niedrigsten aus, d. h. nach der Korrektur mit dem Kalman-Filter erfüllen die Prognosedaten am besten die Qualitätskriterien des DELTA-Tools. Allerdings zeigt die Hybrid-Vorhersage häufig ein besseres Verhalten, wenn

es um die Erfassung von Maximalwerten und die Überschreitung von Schwellenwerten geht. Insbesondere bei den zusätzlich betrachteten Schadstoffepisoden in den Jahren 2015 und 2017 erzielt die Korrektur mit der Hybrid-Vorhersage bessere Ergebnisse als mit dem Kalman-Filter.

Die Multiplikative Korrektur liegt bei den Ergebnissen der statistischen Auswertung im Regelfall im Mittelfeld, in einigen Fällen erzielt sie auch bessere Ergebnisse als der Kalman-Filter und die Hybrid-Vorhersage. Allerdings ist der Einfluss der nach oben unbegrenzten Korrekturfaktoren ein Unsicherheitsfaktor, der in Einzelfällen zu extrem hohen Konzentrationen führen kann. Insofern sollte das Verfahren nicht ohne weitere Untersuchungen zur sinnvollen Beschränkung des Korrekturfaktors, z. B. in Abhängigkeit von Schadstoff, Standort und/oder Konzentrationsniveau, angewendet werden.

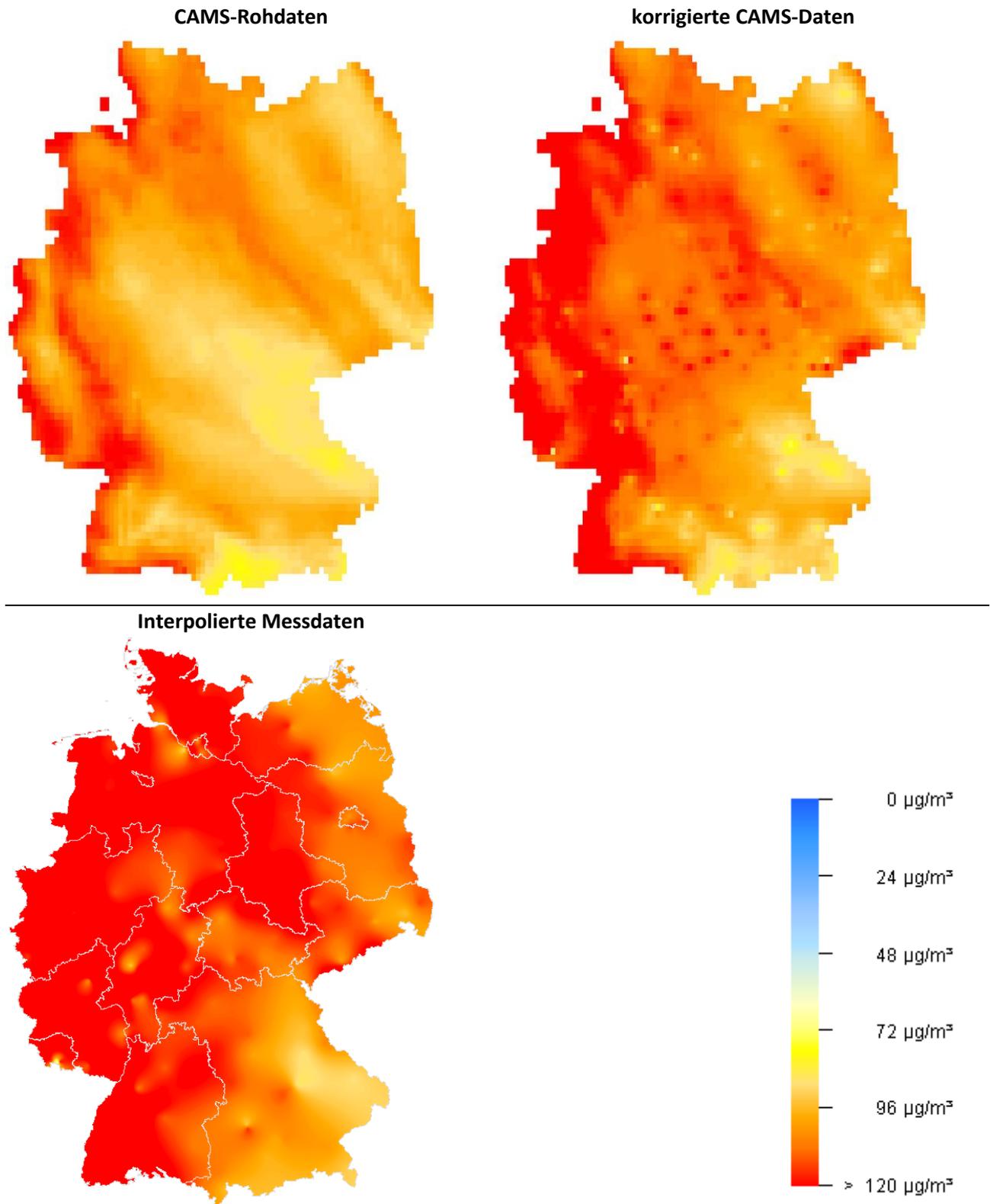
Zur Identifikation eines geeigneten Interpolationsverfahrens für Schritt b), d. h. zur Interpolation der mit dem Korrekturverfahren an den Messorten bestimmten Korrekturfaktoren auf das Gitter der CAMS-Daten, wurde das Programmsystem FLADIS eingesetzt. Dazu wurden die Korrekturfaktoren, die von der Testumgebung als stationsbezogene Zeitreihen im FLADIS-Format abgelegt wurden, als Stationsdaten in FLADIS eingelesen. Anschließend wurde die Interpolation der Korrekturfaktoren mit den verschiedenen in FLADIS implementierten Interpolationsverfahren vorgenommen. Abweichend vom üblichen Vorgehen in FLADIS war das Zielgitter der Interpolation dabei nicht das FLADIS-interne Gitter, sondern das ursprüngliche, durch die CAMS-Prognosedaten vorgegebene Gitter. Die Eignung der verschiedenen Interpolationsverfahren für die Interpolation der Korrekturfaktoren wurde dann mit der Kreuzvalidierung entsprechend VDI 4280 Blatt 5 (KRdL, 2009b) bewertet. Im Ergebnis wurde das globale Shepard-Verfahren mit Exponent = 2 zur Implementierung in das Online-Tool ausgewählt.

Abbildung II zeigt exemplarisch das Ergebnis der flächenhaften Korrektur einer CAMS-Ozonprognose am Beispiel der 8TMAX-Werte für den 25.08.2016. Abbildung II, oben links, zeigt das Raster der CAMS-Rohdaten für den 1. Prognosetag. Ein Vergleich mit den interpolierten 8TMAX der Messdaten für den 25.08.2016 (Abbildung II, unten links) zeigt, dass der Schwellenwert von $120 \mu\text{g}/\text{m}^3$ für die 8TMAX an diesem Tag in Deutschland großräumig überschritten wurde und dass die unkorrigierte CAMS-Prognose die gemessenen Konzentrationen deutlich unterschätzt und insbesondere die Bereiche mit Schwellenwertüberschreitungen nicht hinreichend erfasst. Die Korrektur der CAMS-Prognose durch Anwendung der Hybrid-Vorhersage an den Stationsorten, Interpolation der an den Stationsorten ermittelten Korrekturfaktoren auf das Raster der Modelldaten und Anwendung der flächenhaften Korrektur auf die CAMS-Rohdaten (Abbildung II, oben rechts) führt zu einem deutlichen Anstieg des Konzentrationsniveaus gegenüber den CAMS-Rohdaten und zu einer besseren Erfassung der Bereiche mit Überschreitungen des Schwellenwertes von $120 \mu\text{g}/\text{m}^3$.

Die Erkenntnisse aus der Evaluierung der Korrektur- und Interpolationsverfahren wurden in einem Online-Tool für die Korrektur der CAMS-Prognosedaten im operationellen Betrieb umgesetzt. Das Online-Tool wurde wie die Testumgebung in der Skriptsprache „R“ entwickelt. Es kann über eine Batchdatei oder über die Kommandozeile gestartet werden und korrigiert dann die CAMS-Prognosedaten für den heutigen oder einen vorgegebenen Tag auf Basis der Mess- und CAMS-Prognosedaten der vorangegangenen Tage. Erforderliche Einstellungen für das Online-Tool können über eine Konfigurationsdatei vorgenommen werden.

Das Ablaufschema des Online-Tools (Abbildung III) entspricht bis zur Korrektur der CAMS-Daten an den Messorten weitgehend dem Ablaufschema der Testumgebung (Abbildung I). Wesentlicher Unterschied ist, dass nicht die Zeitreihen eines ganzen Jahres eingelesen werden, sondern nur die zu korrigierende aktuelle CAMS-Prognose sowie die Messdaten und die CAMS-Prognosen der zur Korrektur herangezogenen vorangegangenen Tage.

Abbildung II: Beispiel zur flächenhaften Korrektur: Ozon-Prognose 8TMAX für den 25.08.2016. CAMS-Rohdaten (oben links), nach Korrektur mit der Hybrid-Vorhersage, Interpolation der Korrekturfaktoren sowie Anwendung der flächenhaften Korrektur auf die CAMS-Rohdaten (oben rechts) sowie interpolierte Messdaten (unten links, vom Auftraggeber zur Verfügung gestellt)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Mit der Hybrid-Vorhersage wurde dasjenige Korrekturverfahren in das Online-Tool implementiert, das häufig ein besseres Verhalten zeigt, wenn es um die Erfassung von Maximalwerten und die Überschreitung von Schwellenwerten geht. Dieses Verfahren wird für die standardmäßige Anwendung des Online-Tools empfohlen.

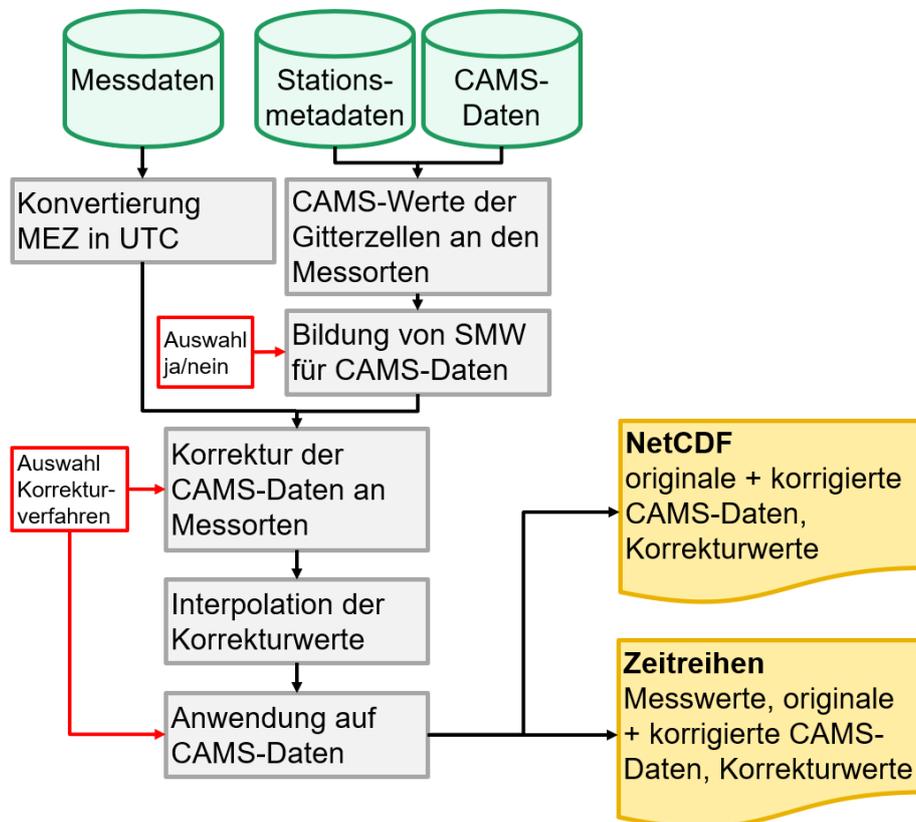
Die Multiplikative Korrektur wurde aufgrund guter Ergebnisse in der Literatur und als in der Leistungsbeschreibung vorgegebenes Verfahren zu einem frühen Zeitpunkt im Projektablauf in das Online-Tool implementiert. Im Laufe der systematischen Evaluierung der Korrekturverfahren wurde wie bereits beschrieben deutlich, dass die Multiplikative Korrektur in einigen Fällen durchaus gute Ergebnisse erzielt, dass der Einfluss der nach oben unbegrenzten Korrekturfaktoren jedoch ein Unsicherheitsfaktor ist, der in Einzelfällen zu extrem hohen Konzentrationen führen kann. Insofern sollte das Verfahren nicht ohne weitere Untersuchungen angewendet werden.

Nach der Korrektur werden für jede Stunde die an den Messorten bestimmten Korrekturfaktoren mit dem globalen Shepard-Verfahren auf das Raster der CAMS-Prognosen interpoliert. Das Korrekturraster wird dann nach den Vorgaben des verwendeten Korrekturverfahrens auf die ggf. stundengemittelten CAMS-Prognosedaten angewendet, um den korrigierten Prognosedatensatz zu erhalten.

Als letzter Schritt werden die Ausgabedateien geschrieben. Für jede Prognosestunde werden als NetCDF-Rasterdateien die CAMS-Prognosedaten, das Korrekturraster und der korrigierte Prognosedatensatz ausgegeben sowie optional für jede Messstation Zeitreihen im ASCII-Format mit u. a. dem CAMS-Prognosewert, dem korrigierter Prognosewert und dem Korrekturfaktor zur jeder Stunde.

Das Online-Tool wurde nach Übergabe an das UBA in den operationellen Betrieb eingebunden. Die vom UBA veröffentlichten Karten der bundesweiten Ozonprognose basieren nun auf den korrigierten CAMS-Daten.

Abbildung III: Ablaufschema Online-Tool



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Summary

The Copernicus Atmosphere Monitoring Service (CAMS) provides data on atmospheric pollution as part of the European Copernicus programme. These are the results of seven chemical transport models (CTM) and their ensemble. These models provide daily forecasts for the next 96 hours in hourly resolution for various pollutants such as ozone and PM10.

The CTM simulations are being carried out in a horizontal grid of about 10 km grid size and thus only cover the regional background. In comparison with measurements at background stations, the CAMS Ensemble usually provides too low maximum ozone concentrations. This has a negative impact on the correct prediction of threshold value exceedances at very high ozone concentrations for information and, if necessary, warning of the population in accordance with the Air Quality Directive 2008/50/EC.

Since the late 1990s, the Federal Environment Agency (UBA) has been publishing maps of nationwide ozone forecasts for information purposes. In the past, ozone forecasts were based on a statistical method, initially replaced by forecasts from the CAMS ensemble. The project described here explored to what extent the predictive quality of the CAMS ozone forecasts for Germany can be improved by post-processing, particularly with regard to the detection of limit value exceedances. A corresponding study was carried out for the CAMS-PM10 forecasts.

The post-processing of the CAMS forecasts was carried out in three steps:

- a) correction of CAMS forecasts at measurement locations using measurement data
- b) interpolation of the correction factors determined at the measurement locations onto the grid which the CAMS prognoses are available on
- c) application of the interpolated correction factors to the CAMS forecasts

To identify suitable correction techniques for step a), a test environment in the programming language R was implemented with three different correction techniques:

- ▶ hybrid forecast (McKeen et al., 2005; Kang et al., 2008)
- ▶ multiplicative correction (McKeen et al., 2005; Borrego et al., 2011)
- ▶ Kalman filter (Delle Monache et al., 2006; Kang et al., 2008).

Figure I shows the flowchart of the test environment. First, the data to be examined is read in. These are on the one hand the measurement data for the pollutant to be investigated together with the station metadata and on the other hand the associated CAMS forecast data. The measurement data for each station are merged into annual time series and converted from CET to UTC so that they can be compared with the CAMS forecast data. The test environment then reads the forecast values at the measuring points from the CAMS data. Since the measurement data are available as hourly mean values (SMW) and the CAMS forecast data as instantaneous values at each hour, the test environment provides the possibility of calculating hourly mean values from the instantaneous values at the beginning and end of each hour.

The CAMS forecast data from the European Copernicus program are available daily as time series over the next 96 hours from the start date of the forecast. Within the framework of this project, the first 72 hours were considered. In order to obtain continuous annual time series of CAMS forecasts for the test environment, these are formed from the CAMS forecast time series depending on the forecast day, using the hours of the 1st, 2nd or 3rd forecast day, respectively.

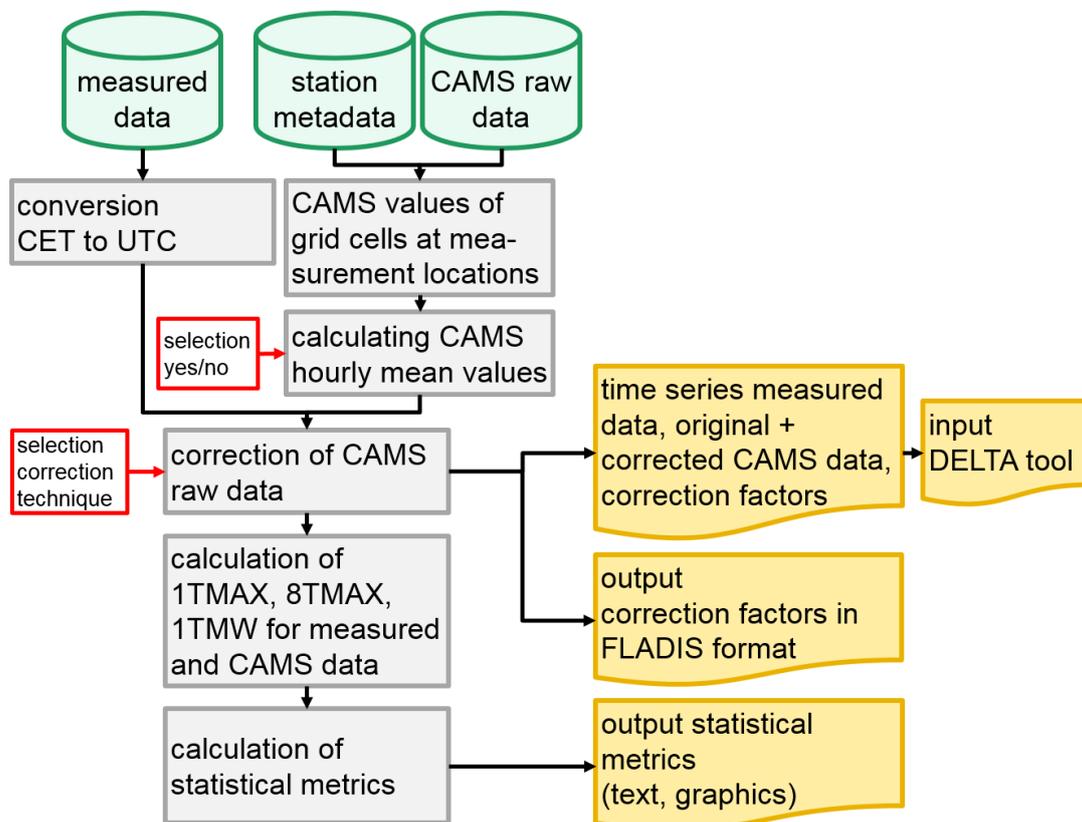
Subsequently, the CAMS forecast data at the measurement locations are corrected with the selected correction technique for each hour based on the CAMS forecast data and the measurement data of the respective previous days. The number of previous days used for the correction depends on the correction technique.

After correcting the CAMS data, further relevant quantifying parameters are determined according to 2008/50/EC for the measured and for the corrected CAMS time series for ozone and PM10. For ozone, this is the maximum one hour mean value per day (1TMAX) and the maximum eight-hour mean value per day from hourly running eight-hour mean values (8TMAX). For PM10, the daily mean value (1TMW) is determined.

As a basis for the evaluation of both the uncorrected CAMS raw data and the implemented correction techniques, the metrics listed in Table I are determined for the hourly mean values and the derived quantifying parameters 1TMAX, 8TMAX and 1TMW. Since the CAMS annual time series are available depending on the forecast day, the metrics were also determined depending on the forecast day and the evaluation of the correction techniques were then carried out depending on the forecast day as well.

The calculated correction factors and metrics are written station-wise to ASCII tables. In addition, graphics of the time series of the observed values and the original or corrected CAMS data are produced as well as scatter plots.

Figure I: Flowchart test environment



Source: Author's illustration (IVU Umwelt GmbH)

The metrics implemented in the test environment to evaluate the correction techniques according to Table I are determined individually for each station and merged into box whisker plots to represent the range of metrics across all stations. Based on Mean Fractional Bias (MFB) and Fractional Gross Errors (FGE), the test environment also creates soccer plots that allow the characterization of model errors and the classification of the predictive quality of the model under consideration at a single glance.

Table I: Statistical parameters / metrics

description	abbreviation
minimum observation	Min Obs [$\mu\text{g}/\text{m}^3$]
minimum model	Min Mod [$\mu\text{g}/\text{m}^3$]
mean observation	Mean Obs [$\mu\text{g}/\text{m}^3$]
mean model	Mean Mod [$\mu\text{g}/\text{m}^3$]
maximum observation	Max Obs [$\mu\text{g}/\text{m}^3$]
maximum model	Max Mod [$\mu\text{g}/\text{m}^3$]
standard deviation observation	SD Obs [$\mu\text{g}/\text{m}^3$]
standard deviation model	SD Mod [$\mu\text{g}/\text{m}^3$]
Pearson correlation coefficient	Korrelation
bias	BIAS [$\mu\text{g}/\text{m}^3$]
root mean square error	RMSE [$\mu\text{g}/\text{m}^3$]
mean fractional bias	MFB [%]
fractional gross error	FGE [%]
number of observed exceedances	MeasExc
hit rate [%]	Hit Rate [%]
false alarm rate [%]	FAR [%]

In addition to the statistical evaluation, the DELTA tool for "Assessment & Planning" (A&P; Thunis & Cuvelier, 2017; FAIRMODE, 2017), which was developed as part of the FAIRMODE initiative, was applied to the original and to the corrected CAMS forecasts. The "Assessment Target Plot" in benchmark mode and the "Forecast-Sigmoid Target Plot" in forecast mode were used to analyse the model results. The Target Plots summarise the statistical parameters implemented in the DELTA tool for standardised evaluation of the model results in a central diagram for the MQI, the modelling quality indicator. In benchmark mode, the MQI evaluates the difference between model results and measurements compared to the measurement uncertainty of the individual pollutant components. In forecast mode, the MQI makes a comparative statement between two models by comparing for a given forecast length the model forecast with a persistent model.

The evaluation of the original and of the corrected CAMS forecasts was based on the statistical metrics implemented in the test environment and on the DELTA tool results for the year 2016 for ozone and PM10 and for two selected pollutant episodes from the years 2015 (ozone) and 2017 (PM10). The studies for the year 2016 were carried out both for the entire annual time series and separately by seasons.

In order to assess the influence of using hourly mean values for the CAMS data, the metrics listed in Table I were determined for O_3 for the original CAMS instantaneous values and for the CAMS hourly mean values calculated by the test environment and compared with each other as a test. As a result, the metrics for the hourly mean and the instantaneous values differ only marginally. Since the use of hourly mean CAMS data is closer to the hourly mean values of the measured data, only hourly mean CAMS data were used in the following.

The evaluation of the hourly mean values 1SMW for ozone has shown that the uncorrected CAMS forecast data overestimate the measurement data on average, which is reflected in a positive bias and a

positive MFB. Table II shows the medians of the metrics according to Table I for the 2016 annual time series as an example. The maximum values observed for the 2016 annual time series are underestimated by the CAMS raw data by 20 % in the median for all stations; accordingly, the standard deviation of the CAMS data is also smaller than that of the measurement data. The median of the correlation coefficients is about 0.8. On seasonal inspection, the minimum, mean and maximum values as well as the standard deviation of the measurement data in the median are best captured by the model data in winter. The biggest differences between measured and modelled data can be found for these metrics in the summer. The correlation between measurement and CAMS raw data however is lowest in winter and spring at 0.76, with summer and autumn being significantly higher at 0.81 and 0.83 respectively.

Basically, each correction technique improves the CAMS raw data (Table II, columns "MK": multiplicative correction, "KF": Kalman filter, "HV": hybrid forecast). For the 1SMW, the Kalman filter usually produces the best results. The exception to this is the reproduction of the maximum values and the standard deviation, these values are best achieved by the hybrid forecast. This is in line with the results of Kang et al (2008). The metrics of the multiplicative correction are consistently average, especially higher maximum values are attained compared to the Kalman filter. However, the evaluation has shown that multiplicative correction without limiting the correction factor is calculating extremely high concentrations in individual cases. The evaluation's results of the correction techniques essentially apply to both the 2016 annual time series, the 2016 seasonal time series and the 2015 episode, where the Kalman filter is significantly outperformed by the hybrid forecast in terms of reproducing maximum values and standard deviation.

Table II: Metrics O₃, hourly mean values 1SMW for the 1st forecast day, CAMS raw data and correction techniques MK, KF and HV (in bold: values with the greatest correspondence between measurement and model)

Metrics	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	0.52	0.52	0.52	0.52
Min Mod [$\mu\text{g}/\text{m}^3$]	0.82	0.23	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	44.72	45.54	44.87	44.95
Mean Mod [$\mu\text{g}/\text{m}^3$]	56.36	45.18	44.98	45.37
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.51	173.66	173.51
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	163.05	153.65	169.84
SD Obs [$\mu\text{g}/\text{m}^3$]	29.44	29.68	29.49	29.46
SD Mod [$\mu\text{g}/\text{m}^3$]	22.29	27.08	26.67	29.01
Korrelation	0.81	0.82	0.84	0.80
BIAS [$\mu\text{g}/\text{m}^3$]	10.17	-0.39	0.11	0.34
RMSE [$\mu\text{g}/\text{m}^3$]	19.82	17.21	15.85	18.08
MFB [%]	38.27	8.57	4.23	-2.23
FGE [%]	48.91	44.03	44.82	52.26

For the maximum daily eight hour mean values 8TMAX, the median of the mean values of the CAMS raw data is still higher than that of the measured data, but the difference is smaller than for the 1SMW. Accordingly, bias and MFB also decrease compared to the 1SMW. Table III shows the medians of the

metrics according to Table I for the 2016 annual time series as an example. Maximum values are underestimated by the CAMS data for the 8TMAX as well, as is the standard deviation. As a result, the median hit rate across all stations for the 2016 annual time series is only around 13%. The median of the correlation coefficients is just under 0.9. Seasonally, the picture for the 8TMAX is slightly different from that for the 1SMW. The median values and standard deviation of the measurement data are best in spring, the mean values in summer and the maximum values as for the 1SMW in winter. The correlation between measurement and CAMS data is highest at 0.91 in autumn and lowest at 0.79 in winter.

With regard to the correction techniques, the hybrid forecast and the Kalman filter produce the best results for the 8TMAX, whereby the hybrid forecast, as before, best meets the maximum values and the standard deviation, but now also the mean value, while the Kalman filter has slightly better values for correlation, RMSE, MFB and FGE (Table III, columns "MK": multiplicative correction, "KF": Kalman filter, "HV": hybrid forecast). The evaluation's results of the correction techniques essentially apply to both the 2016 annual time series, the 2016 seasonal time series, and the 2015 episode, with the hybrid forecast achieving the highest hit rate, i.e. the highest proportion of correctly predicted threshold value exceedances, due to better reproducing the maximum values. At the same time, the FAR, i.e. the highest proportion of incorrectly predicted threshold value exceedances, is the highest for hybrid prediction, albeit at the low level of a few percent.

Table III: Metrics O₃, maximum daily eight hour mean values 8TMAX for the 1st forecast day, CAMS raw data and correction techniques MK, KF and HV (in bold: values with the greatest correspondence between measurement and model)

Metrics	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	3.08	3.12	3.08	3.12
Min Mod [$\mu\text{g}/\text{m}^3$]	17.49	6.78	3.52	2.80
Mean Obs [$\mu\text{g}/\text{m}^3$]	64.59	64.76	64.87	64.75
Mean Mod [$\mu\text{g}/\text{m}^3$]	69.07	62.51	63.22	65.30
Max Obs [$\mu\text{g}/\text{m}^3$]	155.21	152.90	155.21	154.67
Max Mod [$\mu\text{g}/\text{m}^3$]	130.41	148.84	142.19	153.63
SD Obs [$\mu\text{g}/\text{m}^3$]	29.46	29.59	29.56	29.46
SD Mod [$\mu\text{g}/\text{m}^3$]	22.66	28.25	27.18	28.47
Korrelation	0.89	0.89	0.91	0.89
BIAS [$\mu\text{g}/\text{m}^3$]	3.63	-2.00	-1.57	0.88
RMSE [$\mu\text{g}/\text{m}^3$]	14.56	13.97	12.30	13.18
MFB [%]	12.68	-1.23	0.08	3.11
FGE [%]	22.90	22.16	19.61	20.60
MeasExc	15 ¹	14 ¹	15 ¹	15 ¹
Hit Rate [%]	13.33 ¹	53.33 ¹	50.00 ¹	65.45¹
FAR [%]	0.00¹	0.62 ¹	0.32 ¹	1.17 ¹

¹: threshold value 120 $\mu\text{g}/\text{m}^3$

For the maximum daily one hour mean values 1TMAX, the median of the mean values of the CAMS raw data is only slightly above the measured data, Bias and MFB are correspondingly low for this parameter. Table IV shows the medians of the metrics according to Table I for the 2016 annual time series as

an example. The underestimation of the measured maximum values by the CAMS data corresponds to that for the 1SMW; accordingly, the standard deviation of the CAMS data is smaller for the 1TMAX than that of the measurement data. The number of measured exceedances of the information threshold of $180 \mu\text{g}/\text{m}^3$ was low in 2016, none of these few exceedances was recorded by the CAMS data. The median of the correlation coefficient is close to 0.9 as for the 8TMAX. The seasonal results largely correspond to those for the 8TMAX.

Of the three correction techniques, Kalman filter and hybrid prediction provide the best results for the 1TMAX, with the mean and maximum values as well as the standard deviation best met by the hybrid forecast and the Kalman filter generating the better values for correlation, RMSE, MFB and FGE (Table IV, columns "MK": multiplicative correction, "KF": Kalman filter, "HV": hybrid forecast). The evaluation's results of the correction techniques are essentially valid for both the 2016 annual time series and the 2016 seasonal time series; for episode 2015 the hybrid forecast shows the best values in all relevant quantities and the second-best values for bias and MFB. With regard to the reproduction of threshold value exceedances, evaluation of the box whisker plots of the hit rate for episode 2015 shows that in the upper quartile up to 67 % and in the maximum up to 100 % of the exceedances of the information threshold are detected by the correction techniques. The lowest hit rate in the upper quartile is 50 %, provided by the Kalman filters. The uncorrected CAMS forecasts do not detect any exceedance of the information threshold.

Table IV: Metrics O_3 , maximum daily one hour mean values 1TMAX for the 1st forecast day, CAMS raw data and correction techniques MK, KF and HV (in bold: values with the greatest correspondence between measurement and model)

Metrics	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	4.79	4.86	4.81	4.86
Min Mod [$\mu\text{g}/\text{m}^3$]	20.36	9.03	6.64	4.61
Mean Obs [$\mu\text{g}/\text{m}^3$]	71.73	71.92	72.01	71.76
Mean Mod [$\mu\text{g}/\text{m}^3$]	72.56	69.44	69.58	73.48
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.06	173.66	173.66
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	162.92	153.57	169.73
SD Obs [$\mu\text{g}/\text{m}^3$]	30.63	30.62	30.62	30.64
SD Mod [$\mu\text{g}/\text{m}^3$]	23.19	29.27	28.10	29.69
Korrelation	0.88	0.87	0.90	0.87
BIAS [$\mu\text{g}/\text{m}^3$]	0.22	-2.27	-2.31	1.84
RMSE [$\mu\text{g}/\text{m}^3$]	15.23	15.58	13.73	15.12
MFB [%]	5.64	-1.88	-1.28	4.23
FGE [%]	20.13	20.95	18.64	19.87
MeasExc	0	0	0	0
Hit Rate [%]	0.00	0.00	0.00	0.00
FAR [%]	0.00	0.00	0.00	0.00

¹: threshold value $180 \mu\text{g}/\text{m}^3$ (information threshold)

The evaluation of the 1SMW for PM10 has shown that the CAMS raw data underestimate the measured data on average, which is also reflected in a negative bias and a negative MFB. Table V shows the medians of the metrics according to Table I for the 2016 annual time series as an example. The CAMS data underestimate the measured maximum values the median for all stations by almost two thirds, correspondingly the standard deviation of the CAMS data is also significantly smaller than that of the measured data. This contrasts with ozone, where the CAMS raw data overestimate the measured values on average and only underestimate the maximum values. The median of the correlation coefficients is 0.5, in contrast to approx. 0.8 for ozone. Seasonally, mean and maximum values and the standard deviation of the measured data in the median are best reproduced in spring. The biggest differences between measured and modelled data can be seen for these metrics in winter, except for the mean value which is worst represented by the model in summer. For example, the maximum values in winter are underestimated in the median by approx. 67 %. The correlation between measured and CAMS data is highest in autumn at 0.58 and lowest in summer at 0.43.

As with ozone, each correction technique leads to an improvement of the CAMS forecasts for PM10. The Kalman filter produces good metrics for the 2016 annual time series in general, but is surpassed by the hybrid forecast for the maximum values, and the best metrics for the standard deviation are obtained with the multiplicative correction (Table V, columns "MK": multiplicative correction, "KF": Kalman filter, "HV": hybrid forecast). On seasonal inspection, the hybrid forecast is the best technique, with a few exceptions, for reproducing mean and maximum values and standard deviation, as well as the Kalman filter, which achieves the best values for correlation, RMSE and MFB except in spring. For the 2017 episode, the best metrics are achieved with a few exceptions from the hybrid forecast. In second place are the results of the Kalman filter.

Table V: Metrics PM10, hourly mean values 1SMW for the 1st forecast day, CAMS raw data and correction techniques MK, KF and HV (in bold: values with the greatest correspondence between measurement and model)

Metrics	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	0.50	0.50	0.50	0.50
Min Mod [$\mu\text{g}/\text{m}^3$]	0.91	0.65	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.24	16.16	16.13	16.21
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.78	16.37	16.17	16.31
Max Obs [$\mu\text{g}/\text{m}^3$]	141.93	127.50	128.45	128.45
Max Mod [$\mu\text{g}/\text{m}^3$]	50.67	116.15	90.12	137.48
SD Obs [$\mu\text{g}/\text{m}^3$]	11.57	11.29	11.30	11.36
SD Mod [$\mu\text{g}/\text{m}^3$]	6.29	11.41	10.10	11.89
Korrelation	0.50	0.56	0.58	0.53
BIAS [$\mu\text{g}/\text{m}^3$]	-3.79	0.28	0.10	0.11
RMSE [$\mu\text{g}/\text{m}^3$]	10.87	10.55	9.68	11.09
MFB [%]	-16.34	3.78	2.23	-4.43
FGE [%]	48.65	44.02	45.95	52.05

For the daily mean values 1TMW, the median of the mean values of the modelled data is lower than that of the measured data, as it is for the 1SMW. Table VI shows the medians of the metrics according to Table I for the 2016 annual time series as an example. 1TMW maximum values are underestimated

by half in the median by the CAMS raw data for the 2016 annual time series, the standard deviation of the model data is about 40 % less than that of the measurement data. As a result, the median hit rate across all stations is 0%. The median of the correlation coefficients is just under 0.6. On seasonal inspection, mean values and standard deviation of the measurement data are in the median best achieved in spring, as with the 1SMW, the maximum values in spring and autumn. In winter, the maximum values are underestimated in the median by approx. 57 %. The correlation between measured and CAMS raw data is highest at 0.69 in autumn and lowest at 0.53 in winter.

With regard to the correction techniques, the Kalman filter shows the best values for most metrics, but this time the best reproduction of the maximum values and the standard deviation is not obtained with the hybrid forecast but with the multiplicative correction (Table VI, columns "MK": multiplicative correction, "KF": Kalman filter, "HV": hybrid forecast). The measured exceedances of the limit value of $50 \mu\text{g}/\text{m}^3$ are not indicated in the median by any of the techniques. However, evaluation of the box whisker plots of the hit rate shows that in the upper quartile up to 50 % and in the maximum up to 100 % of the exceedances of the limit value are detected by the correction techniques. The CAMS raw data detects a maximum of 25% of the exceedances. On seasonal inspection, the Kalman filter produces the highest correlation coefficient and the lowest values for RMSE and FGE with one exception. Average and maximum values as well as the standard deviation are generally best captured by the hybrid forecast. In the winter time relevant to limit value exceedances, maximum values and standard deviation are best reproduced by the multiplicative correction and second-best by the hybrid forecast. In contrast, the best metrics for the 2017 episode are achieved with a few exceptions from the hybrid forecast, followed by the results of the Kalman filter.

Table VI: Metrics PM10, daily mean values 1TMW for the 1st forecast day, CAMS raw data and correction techniques MK, KF and HV (in bold: values with the greatest correspondence between measurement and model)

Metrics	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	2.02	2.05	2.02	2.03
Min Mod [$\mu\text{g}/\text{m}^3$]	3.10	2.79	1.97	1.41
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.27	16.18	16.15	16.22
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.77	16.38	16.16	16.34
Max Obs [$\mu\text{g}/\text{m}^3$]	72.84	66.99	67.31	67.31
Max Mod [$\mu\text{g}/\text{m}^3$]	35.88	65.94	58.33	71.18
SD Obs [$\mu\text{g}/\text{m}^3$]	9.20	9.05	8.93	9.06
SD Mod [$\mu\text{g}/\text{m}^3$]	5.48	9.19	8.23	9.24
Korrelation	0.58	0.69	0.72	0.72
BIAS [$\mu\text{g}/\text{m}^3$]	-3.80	0.29	0.09	0.11
RMSE [$\mu\text{g}/\text{m}^3$]	8.29	7.25	6.27	6.77
MFB [%]	-21.50	2.49	2.47	0.33
FGE [%]	39.01	29.64	27.62	28.81
MeasExc	2 ¹	2 ¹	2 ¹	2 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00 ¹	0.59 ¹	0.28 ¹	0.28 ¹

¹: limit value $50 \mu\text{g}/\text{m}^3$

If the DELTA tool is applied in benchmark mode to the 2016 annual time series of the measured values and the CAMS raw data, it becomes apparent that the CAMS raw data for O₃ already have an MQI < 1 for all stations and therefore meet the model quality target defined by the DELTA tool. The evaluation of the DELTA tool refers exclusively to the 8TMAX. Some of the stations even have an MQI < 0.5, i.e. the RMSE between measurement and model at these stations is smaller than the measurement uncertainty. The 90% percentile of the MQI of all stations for the 1st forecast day is 0.346. By applying the correction techniques, the 90% percentile of the MQI of all stations is only slightly reduced, the lowest value of 0.301 on the 1st forecast day is obtained by using the Kalman filter.

In forecast mode, the 2016 annual time series MQI of the CAMS raw data is significantly worse than in benchmark mode. A whole number of stations have an MQI > 1 for the first forecast day and are therefore modelled worse than with the persistent model. The 90% percentile of the MQI of all stations is 1.05 for the first forecast day and thus also > 1. For the episode 2015, the 90% percentile of the MQI of all stations on the first forecast day for the CAMS raw data amounts to 1.37, being even significantly higher than 1. For the multiplicative correction, the 90% percentile of the MQI of all stations is 1.06 and therefore still worse than the persistent model. The lowest 90% percentile of the MQI of all stations obtained for the episode is provided by the hybrid method with 0.757.

For PM₁₀, the evaluation of the DELTA tool refers to the 1TMW. For the 2016 annual time series of CAMS PM₁₀ raw data of the 1st forecast day, the MQI of individual stations is larger than 1. The 90% percentile of the MQI of all stations for the CAMS raw data results to 0.947, being still just below 1. Thus, the DELTA tool evaluates the CAMS PM₁₀ forecasts much worse than the CAMS ozone forecasts. Applying the correction techniques leads in any case to an improvement of MQI results for PM₁₀. The MQI of all stations, with a few exceptions (multiplicative correction), is below 1, and the 90% percentile of the MQI of all stations drops to 0.703 (Kalman filter).

In forecast mode, the MQI of the 2016 annual time series of the CAMS PM₁₀ raw data is significantly worse than in benchmark mode, as it is with O₃. Most stations have an MQI > 1 and are therefore modelled worse than with the persistent model. The 90% percentile of the MQI of all stations amounts to 1.29 on the 1st forecast day, being significantly higher than 1. For the episode 2017, the 90% percentile of the MQI of all stations amounts to 2.08 for the CAMS raw data. The multiplicative correction does not seem to be a suitable correction technique for this episode as the 90% percentile of the MQI is at 1.26 still larger than 1. Even the Kalman filter only achieves a 90% percentile of the MQI of 1.00 for this episode, therefore performing just as good as the persistent model. The lowest 90% percentile of the MQI for the 2017 episode is obtained at 0.942 with the hybrid forecast.

Comparison of the metrics and 90% percentile of the MQI for the individual forecast days has shown that the influence of the forecast length on the metrics and on the MQI in benchmark mode is small for both O₃ and PM₁₀. In forecast mode, the MQI decreases as the prediction length increases, meaning that the predictive quality of the original and the corrected CAMS data decreases less than that of the persistent model as the prediction length increases.

In summary, of the correction methods investigated the Kalman filter and the hybrid forecast produce the best forecasts in comparison with the measured data. The Kalman filter often achieves better results in terms of the metrics considered, i.e. higher correlation coefficients and lower values for RMSE, MFB and/or FGE are usually obtained for the corrected forecast data. The 90% percentile of the MQI of all stations is also the lowest in both benchmark mode and forecast mode for the Kalman filter, i.e. after correction with the Kalman filter, the forecast data best meet the quality criteria of the DELTA tool. However, the hybrid forecast often shows better results when it comes to reproducing maximum values and exceeding threshold or limit values. Especially for the additionally considered pollutant episodes in 2015 and 2017, the correction with the hybrid prediction achieves better results than with the Kalman filter.

The multiplicative correction obtains usually medium quality results in terms of the statistical evaluation. In some cases, it also achieves better results than the Kalman filter and the hybrid forecast. However, the influence of the unlimited correction factors is an uncertainty factor leading to extremely high concentrations in individual cases. Due to this, the multiplicative correction should not be used without further studies to limit the correction factor, e.g. depending on pollutant, location and/or concentration level.

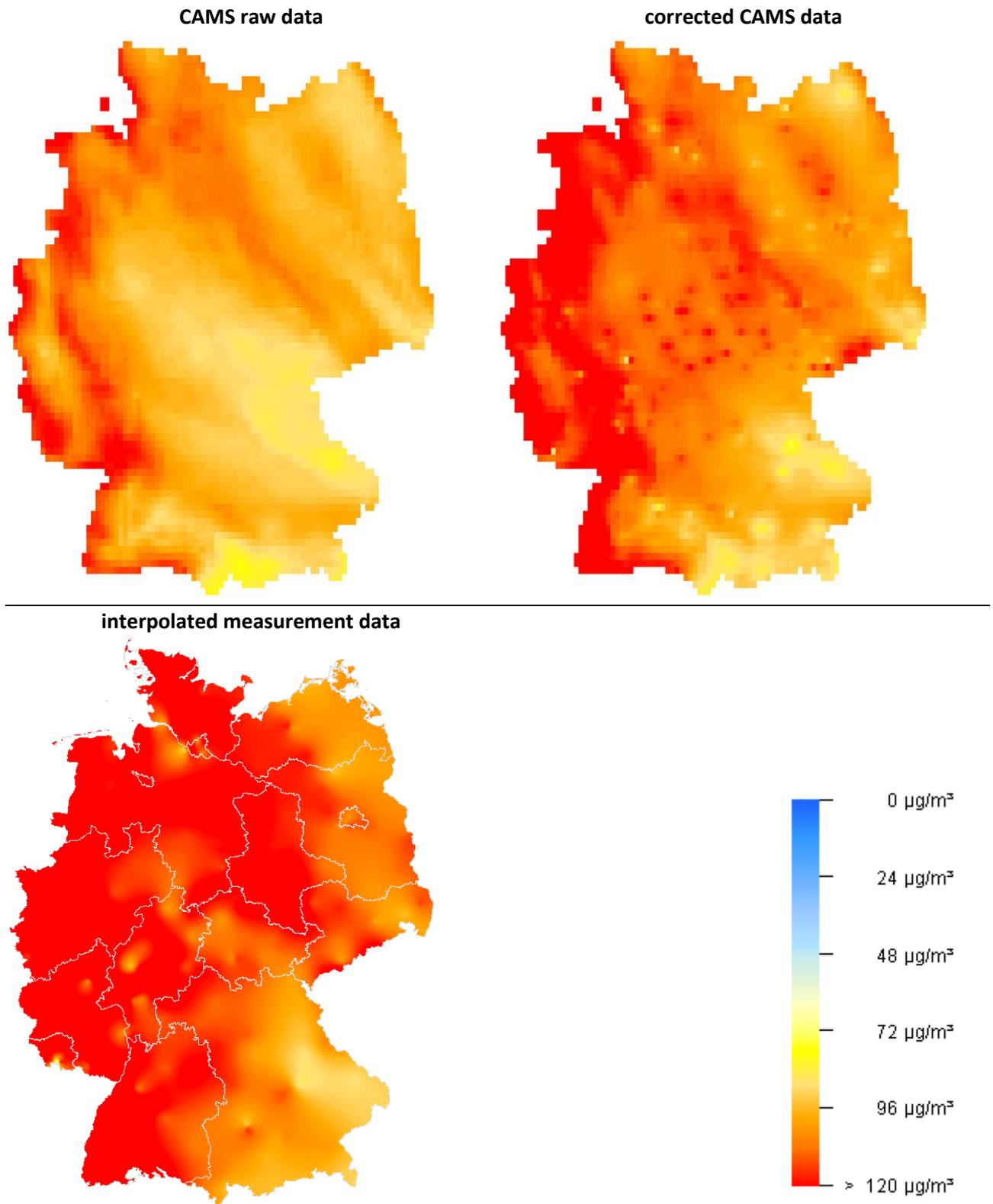
The FLADIS program system was used to identify a suitable interpolation method for step b), i.e. to interpolate the correction factors determined with the chosen correction technique at the measurement locations onto the grid of the CAMS data. The correction factors, stored by the test environment as station-related time series in FLADIS format, were imported into FLADIS as station data. The correction factors were then interpolated with the various interpolation methods implemented in FLADIS. In contrast to the usual procedure in FLADIS, the target grid of the interpolation was not the FLADIS internal grid, but the original grid specified by the CAMS forecast data. The suitability of the different interpolation methods for the interpolation of the correction factors was then evaluated by cross-validation according to VDI 4280 Part 5 (KRdL, 2009b). As a result, the global Shepard interpolation with exponent = 2 was selected for implementation in the online tool.

Figure II shows the result of the spatial correction of a CAMS ozone forecast using the example of the 8TMAX values for August 25, 2016. Figure II, top left, shows the grid of CAMS raw data for the 1st forecast day. A comparison with the interpolated 8TMAX of the measurement data for August 25, 2016 (Figure II, bottom left) shows that the threshold value of $120 \mu\text{g}/\text{m}^3$ for the 8TMAX was exceeded on a large scale in Germany on this day and that the original CAMS forecast clearly underestimates the measured concentrations and particularly does not sufficiently capture the areas with threshold exceedances. The correction of the CAMS forecast by applying the hybrid forecast at the station locations, interpolation of the correction factors determined at the station locations to the grid of the model data and application of the interpolated correction factors to the CAMS raw data (Figure II, top right) leads to a significant increase in the concentration level compared to the CAMS raw data and to a better representation of the areas where the threshold value of $120 \mu\text{g}/\text{m}^3$ is exceeded.

The findings from the evaluation of the correction techniques and interpolation methods were implemented in an online tool for the correction of the CAMS forecast data in operational mode. Like the test environment, the online tool was developed in the script language "R". It can be started from a batch file or from the command line and then corrects the CAMS forecast data for the current day or a specified day based on the measurement and CAMS forecast data of the previous days. Necessary settings for the online tool can be made via a configuration file.

The flowchart of the online tool (Figure III) largely corresponds to the flowchart of the test environment (Figure I) until the CAMS data at the measurement locations are corrected. The main difference is that not the time series of an entire year are read in, but only the current CAMS forecast to be corrected as well as the measurement data and the CAMS forecasts of the previous days used for correction.

Figure II: Example for spatial correction: Ozone forecast 8TMAX for August 25, 2016. CAMS raw data (top left), after correction with the hybrid prediction, interpolation of the correction factors and application of the interpolated correction factors to the CAMS raw data (top right) and interpolated measurement data (bottom left, provided by the client).



Source: Author's illustration (IVU Umwelt GmbH)

With the hybrid forecast, the correction technique was implemented in the online tool usually showing better behaviour when it comes to capturing maximum values and exceeding threshold values. This procedure is recommended for the standard use of the online tool.

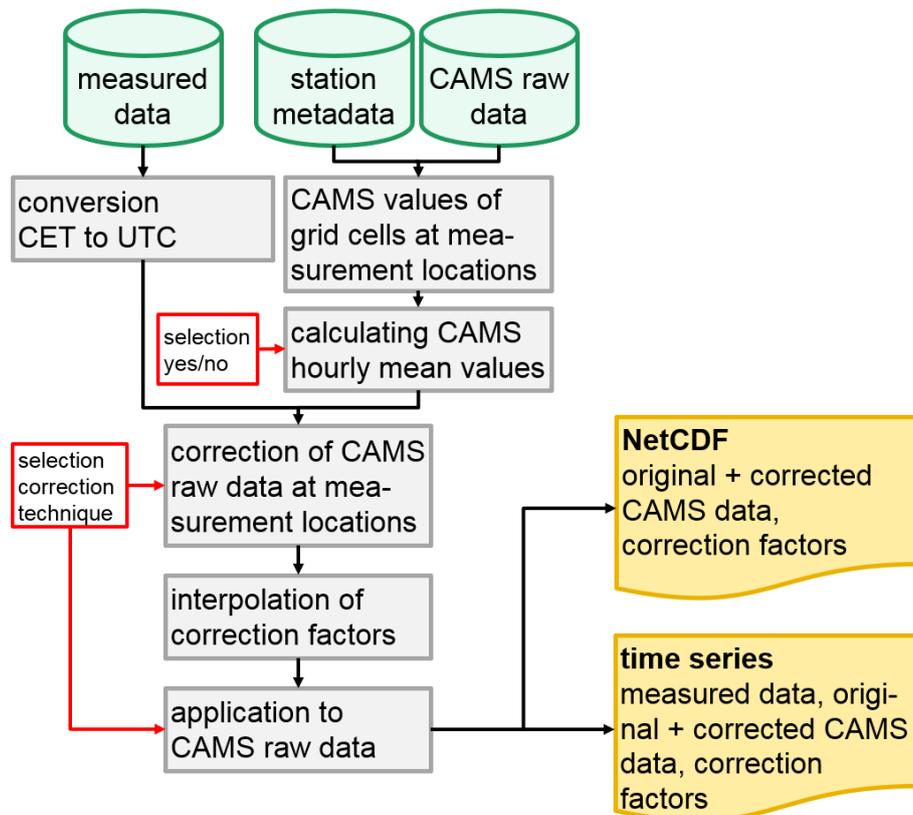
The multiplicative correction was implemented in the online tool at an early stage of the project due to good results in the literature and as a procedure specified in the terms of reference. In the course of the systematic evaluation of the correction techniques it became apparent, as already described, that the multiplicative correction achieves good results in some cases, but that the influence of the unlimited correction factors is an uncertainty factor leading to extremely high concentrations in individual cases. In this respect, the multiplicative correction should not be used without further investigation.

After correction, the correction factors determined at the measurement locations are interpolated to the grid of the CAMS forecasts for each hour using the global Shepard interpolation. The grid of interpolated correction factors is then applied to the (optionally hourly averaged) CAMS forecast data according to the specifications of the correction method used to obtain the corrected forecast data set.

As a last step, the output files are written. For each forecast hour, the CAMS forecast data, the correction grid and the corrected forecast data set are provided as NetCDF grid files. Optionally, for each measurement station time series in ASCII format with the CAMS forecast value, the corrected forecast value and the correction factor for each hour are produced.

The online tool was integrated into operational mode after being delivered to UBA. The maps of the nationwide ozone forecast published by UBA are now based on the corrected CAMS data.

Figure III: Flowchart online tool



Source: Author's illustration (IVU Umwelt GmbH)

1 Einführung

Aufgrund der negativen gesundheitlichen Wirkung des Schadstoffs Ozon stellt das Umweltbundesamt (UBA) seit Ende der 1990er Jahre Karten der aktuellen Ozonsituation sowie bundesweiter Ozonprognosen zur Information der Öffentlichkeit bereit. Die Ozonprognosen beruhen in der Vergangenheit auf einer statistischen Methode, die mittlerweile durch Chemie-Transportmodell-Rechnungen (CTM) ersetzt wurde.

CTM berechnen die Konzentration von Luftschadstoffen an einem oder mehreren Orten in Abhängigkeit der Emissionen, der physikalischen und chemischen Prozesse während des Transportvorgangs in der Atmosphäre sowie der meteorologischen und topographischen Einflüsse. Dazu wird der relevante Ausbreitungsraum, der ganz Europa oder auch nur eine Stadt oder ein Emittentengebiet umfassen kann, mit einem dreidimensionalen Gitter überspannt, in dessen Gitterzellen die Berechnung der Konzentrationen erfolgt. Basierend auf meteorologischen Prognosedaten und Annahmen über die zukünftige Verteilung der Emissionen lassen sich mit Hilfe von CTM Prognosen der Luftschadstoffbelastung erstellen, die gegenüber statistischen Methoden den Vorteil haben, dass sie räumlich-zeitlich-physikalisch-chemische Zusammenhänge während des Transportvorgangs im betrachteten Ausbreitungsraum stärker berücksichtigen als dies bei rein statistischen Verfahren der Fall ist.

Die Qualität der CTM-Prognosen der Luftschadstoffbelastung hat sich in den letzten Jahren kontinuierlich verbessert. Gründe dafür sind z. B. stetig steigende Rechnerleistungen, die eine höhere Auflösung und bessere Erfassung der physikalischen und chemischen Prozesse durch die CTM ermöglichen, sowie höher aufgelöste und verbesserte Emissionsinventare. Trotz dieser Fortschritte können CTM-Prognosen durchaus erhebliche Fehler aufweisen, sei es aufgrund der Unsicherheiten der meteorologischen Eingangsdaten, möglicher Fehler in den Anfangs- und Randbedingungen, fehlerhafter Annahmen bei der Prognose der räumlichen und zeitlichen Verteilung der Emissionen oder einer unzureichenden Beschreibung der den Transportvorgängen in der Atmosphäre zu Grunde liegenden physikalischen und chemischen Prozesse (Neal et al., 2014).

Vor diesem Hintergrund gibt es verschiedene Ansätze, CTM-Prognosen zu verbessern (Neal et al., 2014). Eine Möglichkeit ist die Nachbearbeitung der CTM-Prognosen durch eine fachlich qualifizierte Person, die die Prognosen begutachtet, auf Grund ihrer Erfahrung und der jüngsten Messdaten beurteilt und ggf. anpasst. Alternativ kann die nachgeschaltete Bearbeitung der CTM-Prognosen mit Hilfe automatischer Verfahren verschiedener Komplexitätsgrade durchgeführt werden. Der fortschrittlichste, aber auch rechenaufwendigste und komplexeste Ansatz zur Verbesserung der CTM-Prognosen ist der Einsatz von Datenassimilationsverfahren (z. B. Inness et al., 2013)

In Rahmen des europäischen Copernicus-Programms werden vom Copernicus Atmosphärendienst (Copernicus Atmosphere Monitoring Service, CAMS) Daten zur Entwicklung der atmosphärischen Schadstoffbelastung bereitgestellt. Dabei handelt es sich um die Ergebnisse von sieben CTM, deren Output zum Teil durch Assimilation von Satellitendaten optimiert wird, und deren Ensemble. Diese Modelle liefern täglich Prognosen für die nächsten 96 Stunden in stündlicher Auflösung und für verschiedene Schadstoffe wie z. B. Ozon.

Die Simulationen mit den sieben CTM werden in einer Gitterweite von rund 10 km Kantenlänge durchgeführt und erfassen damit nur den regionalen Hintergrund. Im Vergleich mit Messwerten an Hintergrundstationen liefert das CAMS Ensemble in der Regel zu geringe maximale Ozonkonzentrationen. Dies hat negative Auswirkungen auf die korrekte Vorhersage von Schwellenwertüberschreitungen bei sehr hohen Ozonkonzentrationen zur Information und ggf. Warnung der Bevölkerung entsprechend der Luftqualitätsrichtlinie 2008/50/EG (EU, 2008).

Die Ozonvorhersage aus dem CAMS-Ensemble wird am UBA täglich für das Teilgebiet Deutschland aufbereitet und veröffentlicht. Zusätzlich sollen auch die PM10-Prognosen aus dem CAMS-Ensemble für

das Teilgebiet Deutschland aufbereitet werden. Im vorliegenden Projekt wurde für das Jahr 2016 untersucht, inwieweit sich die Vorhersagegüte der CAMS-Prognosen für die beiden Schadstoffe durch eine nachgeschaltete Bearbeitung verbessern lässt. Dazu wurden im ersten Schritt die CAMS-Prognosen mit Hilfe von Messdaten an den Messorten korrigiert. Die dabei ermittelten Korrekturfaktoren an den Messorten wurden dann im zweiten Schritt in die Fläche interpoliert und anschließend auf das Gitter der CAMS-Modellprognosen angewendet.

Im Rahmen einer Testumgebung wurden verschiedene Korrektur- und Interpolationsverfahren untersucht. Für das Umweltbundesamt (UBA) wurde ein Online-Tool zur Korrektur der CAMS-Prognosen entwickelt, das die Korrektur an den Messorten, die Interpolation der Korrekturfaktoren in die Fläche und die abschließende Korrektur auf dem Gitter der CAMS-Modellprognosen im operationellen Betrieb ermöglicht.

Anmerkung

Im Text wird die wissenschaftliche Zahlendarstellung mit dem Punkt als Dezimaltrennzeichen verwendet.

2 Vorgehensweise

Die Bearbeitung der in Abschnitt 1 beschriebenen Aufgabenstellung gliedert sich in zwei Arbeitspakete. Inhalt des ersten Arbeitspaketes sind der Aufbau einer Testumgebung, in die insgesamt drei Verfahren zur Korrektur der CAMS-Prognosen an den Messorten implementiert wurden, die Evaluierung der Prognosen ohne und mit Korrektur für das Jahr 2016 sowie die Auswahl eines geeigneten Verfahrens zur Interpolation der an den Messorten ermittelten Korrekturfaktoren in die Fläche. Inhalt des zweiten Arbeitspaketes ist die Implementierung und der Test des Online-Tools zur Korrektur der CAMS-Prognose im operationellen Betrieb.

Die Implementierung der Testumgebung (Abschnitt 3.1) erfolgte in R. R ist eine freie Programmiersprache und -umgebung für statistische Berechnungen und Grafiken. Für R existieren zahlreiche, ebenfalls frei abrufbare Pakete mit zusätzlichen Funktionen für spezifische Fragestellungen. Damit lässt sich eine flexible Testumgebung aufbauen, in der zum einen verschiedene Korrekturverfahren für die CAMS-Prognosen implementiert werden können und zum anderen ein umfangreiches Angebot an statistischen Funktionen zur Bewertung der Verbesserung der Vorhersagegüte durch die Korrekturen genutzt werden kann.

Mögliche Korrekturverfahren für die Ozonprognose sind z. B. die multiplikative Korrektur, die Hybrid-Vorhersage und die Kalman-Filter Technik (Borrego et al., 2011). Diese Verfahren sind z. B. in McKeen et al. (2005), Delle Monache et al. (2006) oder Kang et al. (2008) beschrieben. McKeen et al. (2005) verwenden als weiteres Verfahren die Subtraktion des mittleren Bias (mean subtraction oder subtraction of mean bias). Dieses unterscheidet sich inhaltlich jedoch nicht von der Hybrid-Vorhersage. McKeen et al. (2005) wenden die Subtraktion des mittleren Bias und die multiplikative Korrektur auf Ozonvorhersagen verschiedener Modelle und des zugehörigen Ensembles an. Sie finden eine grundsätzlich bessere Korrekturleistung der multiplikativen Korrektur, was die Vorhersagegüte betrifft, allerdings eine bessere Leistung der Subtraktion des mittleren Bias, was die Trefferquote bei Schwellenwertüberschreitungen betrifft. Borrego et al. (2011) wenden die multiplikative Korrektur und einen Kalman-Filter auf drei CTM und die Stoffe Ozon, NO₂, SO₂, PM₁₀ und PM_{2.5} an und sehen hinsichtlich der statistischen Kennwerte eine leichte Überlegenheit der multiplikativen Korrektur gegenüber dem Kalman Filter. Insbesondere werden durch den Kalman-Filter große Bias-Werte, die von einer Überschätzung durch das Modell herrühren, nicht korrigiert. Bei der Korrektur von Ozon-Stundenwerten zeigen aber beide Verfahren gute Ergebnisse.

Kang et al. (2008) wenden die Hybridvorhersage und die Kalman-Filter Technik auf Ozonprognosen eines CTM an und sehen eine Überlegenheit der Kalman-Filter Technik gegenüber der Hybridvorhersage, allerdings werden Schwellenwertüberschreitungen durch den Kalman Filter nicht so gut erfasst. Neal et al. (2014) schließlich wenden eine Form der Hybridvorhersage auf ein CTM und die Stoffe Ozon, NO₂, PM₁₀ und PM_{2.5} an und erzielen damit eine deutliche Verbesserung der Vorhersagegüte gegenüber den Originalprognosen, insbesondere für Ozon und PM₁₀. Alle Literaturstellen weisen aber auf die Abhängigkeit der Verbesserung durch das Korrekturverfahren vom Verhalten des Modells im Vergleich zu den Messwerten hin (Über- oder Unterschätzung durch das Modell, ggf. auch gegenläufiges Verhalten für verschiedenen Konzentrationsbereiche). Grundsätzlich erscheinen daher neben der in der Leistungsbeschreibung explizit geforderten multiplikativen Korrektur sowohl die Hybridvorhersage als auch die Kalman-Filter Technik testenswert, insbesondere im Hinblick auf die Verbesserung der Trefferquote bei Schwellenwertüberschreitung. In Abstimmung mit dem Auftraggeber wurden diese drei Korrekturverfahren in die R-Testumgebung implementiert (Abschnitt 3.3).

Die CAMS-Prognosedaten wurden zunächst ohne Korrektur evaluiert, um das Modellverhalten erfassen zu können. Anschließend wurde die Evaluierung für die mit den ausgewählten Korrekturverfahren (multiplikative Korrektur, Hybrid-Vorhersage, Kalman-Filter) korrigierten CAMS-Daten durchgeführt.

Die Evaluierung erfolgte an den Messorten, da die zu implementierenden Korrekturverfahren Messdaten nutzen und damit nur an den Messorten anwendbar sind. Dazu wurden im ersten Schritt die erforderlichen Daten, d.h. die Prognosen für Ozon und PM10 aus dem CAMS-Ensemble und die stündlichen Zeitreihen der Ozon- und PM10-Messdaten in Deutschland für das Jahr 2016, übernommen und auf Vollständigkeit geprüft (Abschnitt 3.2). Die Metriken zur Evaluierung der CAMS-Prognosedaten (u. a. Korrelationskoeffizient, RMSE, Bias, Hit Rate, False Alarm Rate) wurden in die R-Testumgebung implementiert (Abschnitt 3.4). Anschließend wurde die Evaluierung für die unkorrigierten Rohdaten der CAMS-Prognose (Abschnitt 4) und die zu untersuchenden Korrekturverfahren (Abschnitt 5) durchgeführt.

In Erweiterung der statistischen Auswertung wurde zudem das im Rahmen der FAIRMODE-Initiative (FAIRMODE, 2016) entstandene DELTA-Tool auf die originalen und auf die korrigierten CAMS-Prognosen angewendet. Das DELTA-Tool (Thunis & Cuvelier, 2017) bietet umfangreiche statistische Auswertungen, so z. B. Target Plots und die Bestimmung des MQI (Modelling Quality Indicator) (Abschnitt 3.4.2). Das DELTA-Tool beinhaltet in seiner aktuellen Version 5.6 einen Benchmark- und einen Forecast-Modus, wobei insbesondere letzterer sich noch in der Testphase befindet und damit möglicherweise auch in den folgenden Versionen noch inhaltlichen Änderungen unterworfen sein wird. Im Benchmark-Modus liefert ein Vergleich der mit dem DELTA-Tool erzielten statistischen Parameter für die originalen und die korrigierten Ozonprognosen Aussagen zur Bewertung der Veränderung der Vorhersagegüte durch die Korrektur. Der Forecast-Modus hingegen vergleicht die eingelesenen Modellprognosen mit den Ergebnissen eines internen persistenten Modells. Damit lassen sich Aussagen über die Güte der originalen und der korrigierten CAMS-Prognosen im Vergleich mit dem persistenten Modell und in Abhängigkeit der Vorhersagelänge treffen.

Das im zweiten Arbeitspaket zu entwickelnde Online-Tool soll nicht nur die CAMS-Prognose an den Messorten korrigieren, sondern auch in der Lage sein, eine Interpolation der an den Messorten berechneten Korrekturfaktoren auf das Modellgitter vorzunehmen und die interpolierten Korrekturfaktoren anschließend auf das Gitter der Modellprognosen anzuwenden. Für die Interpolation der Korrekturfaktoren wird ein geeignetes Verfahren benötigt. Daher wurde in das Programmsystem FLADIS (IVU Umwelt GmbH, 2016) eine Testumgebung implementiert, in der verschiedene Interpolationsverfahren getestet und bewertet wurden (Abschnitt 6). FLADIS ist ein Programmsystem zur flächenhaften Darstellung von Immissionssituationen auf der Basis von Punktmessdaten und/oder Modelldaten, das seit Anfang der 1990er Jahre von IVU Umwelt entwickelt wird (IVU GmbH, 1992; IVU GmbH, 1996; Diegmann & Wiegand, 2000) und das vom UBA zur Erstellung der Kartendarstellung der aktuellen Luftdaten genutzt wird.

In FLADIS existieren bereits eine Schnittstelle zum NetCDF-Format¹, in dem die CAMS-Prognosen vorliegen, sowie eine Funktionalität zum Einlesen von Stationsdaten und den zugehörigen Metadaten (Stationscode, Lage, Höhe, Stationstyp). Die von der Testumgebung in R erzeugten stationsbezogenen Zeitreihen der Korrekturfaktoren konnten daher als Stationsdaten in FLADIS eingelesen und mit den in FLADIS implementierten Interpolationsverfahren interpoliert sowie auf das NetCDF-Modellgitter angewendet werden. In FLADIS sind insgesamt sieben Interpolationsverfahren implementiert, u. a. inverse Abstands-Wichtungs-Methoden, radiale Basisfunktions-Methoden sowie mit der Optimalen Interpolation ein geostatistisches Verfahren. Die Eignung der verschiedenen Interpolationsverfahren für die Interpolation der Korrekturfaktoren wurde mit der in FLADIS implementierten Kreuzvalidierung entsprechend VDI 4280 Blatt 5 (KRdL, 2009b) bewertet und ein geeignetes Interpolationsverfahren ausgewählt.

¹ <http://www.unidata.ucar.edu/software/netcdf/>

Inhalt des zweiten Arbeitspakets ist, wie eingangs aufgeführt, die Implementierung und der Test des Online-Tools zur Korrektur der Prognosedaten aus dem CAMS-Ensemble im operationellen Betrieb (Abschnitt 7). Das Online-Tool wurde ebenfalls in R implementiert und nutzt einzelne Elemente aus der im ersten Arbeitspaket implementierten Testumgebung, insbesondere die dort implementierten Korrekturverfahren. Im Gegensatz zur Testumgebung, die eine zusammenhängende Zeitreihe über einen vorgegebenen Zeitraum bearbeitet, ist das Online-Tool darauf ausgerichtet, täglich eine Korrektur der aktuellen Prognosedaten aus dem CAMS-Ensemble für die nächsten 72 Stunden auf Basis der Mess- und Prognosedaten der vorangegangenen Tage (je nach Korrekturverfahren bis zu drei Tage) durchzuführen. Anschließend nimmt das Online-Tool die Interpolation der an den Messorten berechneten Korrekturfaktoren auf das Modellgitter vor und wendet die interpolierten Korrekturfaktoren auf das Gitter der Modellprognosen an. Als Interpolationsverfahren wird dabei das im ersten Arbeitspaket als geeignet ausgewählte Verfahren eingesetzt. Als Ergebnis gibt das Online-Tool die korrigierten CAMS-Prognosedaten analog zum Eingangsformat der CAMS-Daten im NetCDF-Format aus. Metriken zur Evaluierung der korrigierten Modelldaten sind im Online-Tool nicht implementiert. Der reibungslose Ablauf des Online-Tools im operationellen Betrieb wurde über einen Zeitraum von mehreren Monaten erfolgreich getestet.

3 Testumgebung

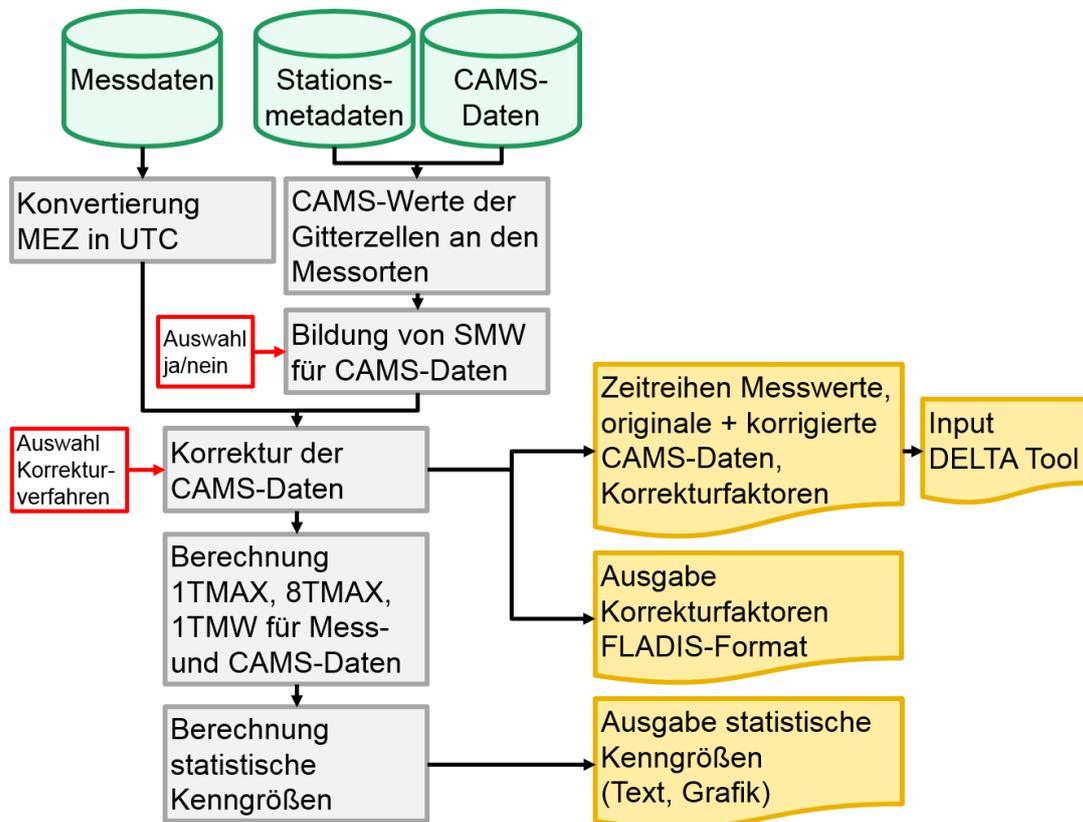
3.1 Aufbau

Um die Wirkung der ausgewählten Korrekturverfahren evaluieren zu können, wurde eine flexible Testumgebung implementiert. Diese wurde gemäß Abschnitt 2 in der Skriptsprache „R“ entwickelt.

Abbildung 3-1 zeigt das Ablaufschema der Testumgebung. Zunächst werden die zu untersuchenden Daten eingelesen. Dies sind zum einen die Messdaten für den zu untersuchenden Schadstoff zusammen mit den Stationsmetadaten (Abschnitt 3.2.1) und zum anderen die zugehörigen CAMS-Prognosedaten (Abschnitt 3.2.2). Die Messdaten werden für jede Station zu einer Jahreszeitreihe zusammengefügt und von MEZ nach UTC konvertiert, so dass sie mit den CAMS-Prognosedaten (Abschnitt 3.2.2) verglichen werden können. Mit Hilfe der Positionsdaten der Messstationen liest die Testumgebung dann die CAMS-Prognosewerte an den Messorten aus den NetCDF-Eingangsdateien aus. Dabei wird pro Station der Konzentrationswert derjenigen Gitterzelle des CAMS-Modells verwendet, in der die Messstation verortet ist; eine Interpolation auf den Messort innerhalb der Gitterzelle wird nicht vorgenommen.

Wie in Abschnitt 3.2 erläutert, liegen die Messdaten als Stundenmittelwerte (SMW) und die CAMS-Prognosedaten als Momentanwerte zu jeder Stunde vor. In der Testumgebung kann gewählt werden, ob die weiteren Auswertungen auf Basis der Momentanwerte der CAMS-Daten durchgeführt werden sollen, oder ob zunächst aus den Momentanwerten Stundenmittel gebildet werden sollen, indem aus dem Momentanwerten jeweils zu Anfang und zu Ende einer Stunde der Mittelwert bestimmt wird. Inhaltlich kommt die Verwendung von stundengemittelten CAMS-Daten den Stundenmittelwerten der Messdaten näher.

Abbildung 3-1: Ablaufschema Testumgebung



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

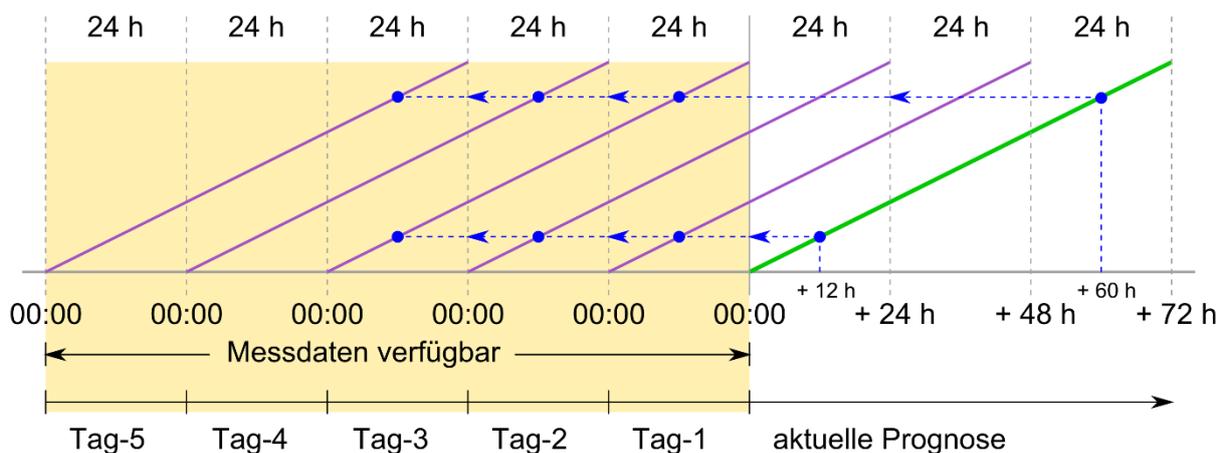
Anschließend werden die CAMS-Prognosedaten an den Messorten für jede Stunde auf Basis der CAMS-Prognosedaten und der Messdaten der jeweils vorangegangenen Tage korrigiert. In die Testumgebung wurden 3 Korrekturverfahren implementiert:

- ▶ Multiplikative Korrektur (Abschnitt 3.3.1)
- ▶ Hybrid-Vorhersage (Abschnitt 3.3.2)
- ▶ Kalman-Filter (Abschnitt 3.3.3)

Das zu verwendende Korrekturverfahren kann in der Testumgebung ausgewählt werden. Alternativ können die Original-Modelldaten auch ohne Korrektur evaluiert werden. Die Anzahl der zur Korrektur herangezogenen vorangegangenen Tage richtet sich nach dem Korrekturverfahren. Abbildung 3-2 zeigt die erforderlichen CAMS-Prognosen der Vortage (violett) für die Korrektur der aktuellen CAMS-Prognose (grün) bei einem Vorlauf des Korrekturverfahrens von 3 Tagen. Wesentlich ist dabei, dass für jeden zur Korrektur herangezogenen Modellwert ein entsprechender Messwert zur Verfügung steht (zumindest theoretisch, Fehlwerte dürfen auftreten; die Behandlung der Fehlwerte wird in Abschnitt 3.3 erläutert). Daher sind für die Korrektur aktueller Prognosestunden > + 24 h CAMS-Prognosen erforderlich, die länger als drei Tage zurück liegen.

Die CAMS-Prognosedaten liegen aus dem europäischen Copernicus-Programm täglich als Zeitreihen über die jeweils nächsten 96 Stunden ab Starttermin der Prognose vor (Abschnitt 1). Im Rahmen dieses Projekts werden davon jeweils die ersten 72 Stunden betrachtet (Abschnitt 3.2.2). Um für die Testumgebung durchgehende Jahreszeitreihen der CAMS-Prognosen zu erhalten, werden diese in Abhängigkeit vom Prognosetag gebildet. Beispielsweise werden für den Prognosetag 1 die Jahreszeitreihen durch Aneinanderreihung der täglichen Prognosestunden + 01 h bis + 24 h ab Initialisierung gebildet, für den Prognosetag 2 durch die Aneinanderreihung der täglichen Prognosestunden + 25 h bis + 48 h etc. Auf diese Weise erhält man durchgehende Jahreszeitreihen der CAMS-Prognosen, die eine bestimmte Vorhersagelänge repräsentieren. Führt man die Evaluierung gemäß Abschnitt 3.4 für alle 3 Prognosetage durch und vergleicht man deren Ergebnisse untereinander, so kann damit die Änderung der Vorhersagegüte mit der Vorhersagelänge bewertet werden. Für die hier untersuchten CAMS-Daten werden die Ergebnisse dieser Evaluierung im Wesentlichen in Abschnitt 4, in Einzelfällen auch in Abschnitt 5.1 vorgestellt.

Abbildung 3-2: Erforderliche CAMS-Prognosen der Vortage (violett) für die Korrektur der aktuellen CAMS-Prognose über die nächsten 72 h (grün) bei einem Vorlauf des Korrekturverfahrens von 3 Tagen. Der Zeitraum, in dem Messdaten verfügbar sind, ist orange hinterlegt. Die blauen Markierungen geben beispielhaft für die aktuellen Prognosestunden + 12 h und + 60 h an, welche Stundenwerte aus welchen Prognosen der Vortage zur Korrektur herangezogen werden.



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Nach der Korrektur der CAMS-Daten werden für die Mess- und die korrigierten CAMS-Zeitreihen für die Schadstoffe Ozon und PM10 weitere nach 2008/50/EG relevante Bewertungsgrößen ermittelt. Für Ozon handelt es sich dabei um den höchsten Einstundenmittelwert pro Tag (1TMAX) und den höchsten Achtstundenmittelwert pro Tag aus stündlich gleitenden Achtstundenmittelwerten (8TMAX). Zur Bestimmung des 1TMAX müssen dabei mindestens 18 gültige Stundenmittelwerte eines Tages vorliegen. Entsprechend werden gleitende Achtstundenmittelwerte nur bestimmt, wenn in dem jeweiligen Achtstunden-Zeitraum mindestens 6 gültige Stundenmittelwerte vorliegen, und 8TMAX werden nur bestimmt, wenn mindestens 18 gültige Achtstundenmittelwerte eines Tages vorliegen. Andernfalls wird der entsprechende 1TMAX bzw. 8TMAX mit einem Fehlwert belegt. Für PM10 wird tagesweise der Tagesmittelwert (1TMW) bestimmt, sofern mindestens 18 gültige Stundenmittelwerte für den jeweiligen Tag vorliegen, andernfalls wird der entsprechende 1TMW mit einem Fehlwert belegt.

Abschließend werden für die Stundenmittelwerte und die abgeleiteten Bewertungsgrößen 1TMAX und 8TMAX bzw. 1TMW die in Abschnitt 3.4.1 aufgeführten Metriken bestimmt.

Die Testumgebung erzeugt folgende Ausgaben:

- ▶ Tabellen im ASCII-Format mit den stündlichen Zeitreihen der Messwerte, der originalen und der korrigierten CAMS-Daten und der Korrekturfaktoren für alle Stationen
- ▶ Tabellen im ASCII-Format mit den Zeitreihen der abgeleiteten Bewertungsgrößen für die Messwerte und die korrigierten CAMS-Daten für alle Stationen
- ▶ Zeitreihen der Korrekturfaktoren pro Station im FLADIS-Input-Format zur Verwendung in Abschnitt 6
- ▶ Tabellen im ASCII-Format mit den nach Abschnitt 3.4.1 berechneten Metriken für Stundenmittelwerte und abgeleitete Bewertungsgrößen für alle Stationen
- ▶ Zeitreihenplots und Scatterplots pro Station mit den Mess- und den korrigierten Modelldaten, für die Stundenmittelwerte und für die abgeleiteten Bewertungsgrößen
- ▶ Box-Whisker-Plots für die einzelnen Metriken aus Abschnitt 3.4.1, die die Ergebnisse über alle Stationen zusammenfassen, für die Stundenmittelwerte und für die abgeleiteten Bewertungsgrößen
- ▶ Soccerplots (Abschnitt 3.4.1) für die Stundenmittelwerte und für die abgeleiteten Bewertungsgrößen

Zur Erzeugung der Daten, die als Input zur Auswertung der Performance von Modell und Korrekturverfahren im DELTA-Tool (Abschnitt 3.4.2) erforderlich sind, wird ein der Testumgebung nachgeschaltetes R-Skript eingesetzt.

3.2 Eingangsdaten

3.2.1 Messdaten

Vom Auftraggeber wurden Messdaten für Ozon und PM10 für das Bezugsjahr 2016 in stündlicher Auflösung als Stundenmittelwerte in MEZ (ohne Berücksichtigung der Sommerzeit) bereitgestellt. Der angegebene Zeitstempel bezieht sich dabei jeweils auf das Ende des Mittelungszeitraums. Die Messdaten wurden vom UBA bereits geprüft übergeben, sie enthalten Werte bis zur negativen Nachweisgrenze.

Die zu den Messstationen benötigten Metadaten, wie z.B. Stationscode, Position der Station als Rechts- und Hochwert sowie Typ der Station („Spot“ oder „Area“), liegen in einer separaten Datei vor, die ebenfalls vom Auftraggeber bereitgestellt wurde. Für die Korrektur der CAMS-Prognosedaten wurden ausschließlich Stationen vom Typ „Area“ verwendet. Damit fließen für Ozon die Daten von 260 Messstationen in die Korrektur ein und für PM10 von 216 Messstationen.

3.2.2 Modelldaten

Vom Auftraggeber wurden Prognosedaten für Ozon und PM10 aus dem CAMS-Ensemble (Abschnitt 1) im NetCDF-Format für das Bezugsjahr 2016 bereitgestellt. In der Regel liegen damit für jeden Tag des Jahres CAMS-Prognosedaten für Ozon und PM10 für jeweils 72 Prognosestunden (drei Prognosestage) vor. Für jede Prognosestunde gibt es eine Datei im NetCDF-Format, deren Name sich aus dem Starttermin der Prognose sowie der Prognosestunde zusammensetzt, z B.:

HRES_ENS_2016010100+072.nc

für den Starttermin 01.01.2016 00:00 Uhr und die Prognosestunde 72. Für Starttermine im Zeitraum 07.06.2016 bis einschließlich 11.06.2016 sind keine CAMS-Ozonprognosen vorhanden. Für den Starttermin 20.12.2016 liegen weder Ozon- noch PM10-Prognosen aus CAMS vor.

Die CAMS-Daten liegen in UTC als Momentanwerte (statt als Stundenmittelwerte wie die Messdaten) vor.

3.3 Implementierte Korrekturverfahren

3.3.1 Multiplikative Korrektur

Die Multiplikative Korrektur ist ein einfaches Verfahren zur Korrektur von Luftschadstoffvorhersagen, die mit CTM erzeugt wurden. Es wurde z. B. von McKeen et al. (2005) zur Verbesserung der Ozonvorhersage und von Borrego et al. (2011) zur Verbesserung der Vorhersage von Ozon, NO₂, SO₂, PM10 und PM2.5 eingesetzt. Dabei wird ein für einen Messort für den zukünftigen Zeitpunkt $t+\Delta t$ prognostizierter Modellwert C_{mod} mit einem Faktor multipliziert, der sich als das Verhältnis von Messwert C_{obs} zu Modellwert C_{mod} an diesem Ort zu n Vorlauf früheren Zeitpunkten berechnet:

$$C_{mod,t+\Delta t}^{korr} = \frac{\sum_{k=1}^{nVorlauf} C_{obs,k}}{\sum_{k=1}^{nVorlauf} C_{mod,k}} \times C_{mod,t+\Delta t} \quad (1)$$

Δt beträgt dabei 24 h. n Vorlauf wurde in Anlehnung an Borrego et al. (2011) auf 3 Tage gesetzt, ein Zeitraum, durch den z. B. auch die synoptischen Bedingungen charakterisiert sind, durch die die Luftschadstoffverteilung geprägt wird. Demnach wird ein für heute, 11:00 Uhr, prognostizierter Modellwert korrigiert mit den Mess- und Modellwerten von gestern, vorgestern und vorgestern jeweils um 11:00 Uhr. Bei der Wahl von n Vorlauf ist zu berücksichtigen, dass der zeitliche Versatz, mit dem gemessene Spitzenkonzentrationen in die Korrektur eingespeist werden, umso länger ist, je größer n Vorlauf ist.

Im Gegensatz zu den anderen beiden im Rahmen dieses Projektes eingesetzten Verfahren (Abschnitt 3.3.2, Abschnitt 3.3.3) ist das Ergebnis einer Multiplikativen Korrektur implizit größer Null, d. h. nicht negativ, wenn sowohl Mess- als auch Modellwerte größer Null sind. Im vorliegenden Projekt wurden Messzeitreihen bereit gestellt, in denen Werte bis zur negativen Nachweisgrenze enthalten sind (Abschnitt 3.2.1). Diese negativen Werte wurden in Abstimmung mit dem Auftraggeber vor der Anwendung der multiplikativen Korrektur zu Null gesetzt.

Bei einem ungünstigen Verhältnis der Mess- zu den Modellwerten an den Vorlauf-Tagen kann der zur Korrektur des prognostizierten Modellwerts bestimmte Faktor sehr große Werte annehmen und damit zu unrealistischen korrigierten Modellwerten führen (Abschnitt 5.1.1). Grundsätzlich sollte der Korrekturfaktor daher nach oben begrenzt werden. Zu der Frage, auf welchen Wert der Korrekturfaktor sinnvollerweise zu begrenzen ist, ggf. in Abhängigkeit von Schadstoff, Standort und/oder Konzentrationsniveau, liegen jedoch keine Untersuchungen vor und wurden im Rahmen dieses Projekts auch keine Untersuchungen durchgeführt. Die Multiplikative Korrektur wurde daher sowohl in der Testumgebung als auch im Online-Tool für den operationellen Betrieb ohne Begrenzung des Korrekturfaktors implementiert.

3.3.2 Hybrid-Vorhersage

Ein weiteres einfaches Verfahren zur Korrektur von Luftschadstoffvorhersagen, die mit Hilfe von Chemie-Transportmodell-Rechnungen (CTM) erzeugt wurden, ist die Hybrid-Vorhersage. Dabei wird ein für einen Messort für den zukünftigen Zeitpunkt $t+\Delta t$ prognostizierter Modellwert C_{mod} mit dem Mittelwert der Differenzen zwischen Messwert C_{obs} und Modellwert C_{mod} an diesem Ort zu $nVorlauf$ früheren Zeitpunkten korrigiert:

$$C_{mod,t+\Delta t}^{korr} = C_{mod,t+\Delta t} + \frac{1}{nVorlauf} \sum_{k=1}^{nVorlauf} (C_{obs,k} - C_{mod,k}) \quad (2)$$

Δt beträgt dabei wie schon bei der Multiplikativen Korrektur in Abschnitt 3.3.1 24 h. Kang et al. (2008) setzen $nVorlauf$ auf einen Tag, mithin wird der prognostizierte Modellwert immer mit der Differenz zwischen Mess- und Modellwert vom Vortag zur selben Stunde korrigiert. McKeen et al. (2005) hingegen betrachten eine 56-tägige Episode und verwenden für $nVorlauf$ alle an der jeweiligen Station vorliegenden Tage mit Messwerten. Das Verfahren wird dort als „mean subtraction“ oder „subtraction of mean bias“ bezeichnet. Neal et al. (2014) setzen $nVorlauf$ wie Kang et al. (2008) auf einen Tag. Zur Verringerung des Einflusses von Messfehlern auf den Korrekturterm setzen Neal et al. (2014) den Korrekturterm zu 50 % aus der Differenz zwischen Mess- und Modellwert vom Vortag am Stationsort und zu 50 % aus dem Median der Differenzen zwischen Mess- und Modellwert vom Vortag an allen anderen Stationen gleichen Typs (städtisch, vorstädtisch, ländlich) zusammen. Zudem lassen Neal et al. (2014) den Korrekturterm linear über die Vorhersagelänge (bei Neal et al. (2014) 5 Tage) gegen den 30-Tage-Mittelwert der Korrekturterme für die jeweilige Station und Stunde konvergieren.

Im hier beschriebenen Projekt wurde die Hybrid-Vorhersage entsprechend Gleichung (2) implementiert. Testläufe mit statistischer Auswertung gemäß Abschnitt 3.4.1 für $nVorlauf = 3$ und $nVorlauf = 1$ haben gezeigt, dass ein Vorlauf von 3 Tagen für die Stundenmittelwerte zu einer geringfügig höheren Korrelation sowie für 1TMAX und 8TMAX zu einem etwas geringeren Mean Fractional Bias (MFB, Abschnitt 3.4.1) führt als ein Vorlauf von nur einem Tag. Allerdings werden die Maxima sowie die Standardabweichung bei $nVorlauf = 3$ deutlich gedämpft, und wie in Abschnitt 3.3.1 ist der zeitliche Versatz, mit dem gemessene Spitzenkonzentrationen in die Korrektur eingespeist werden, für $nVorlauf = 3$ größer als für $nVorlauf = 1$. Aus diesem Grund, und weil mit den korrigierten Modellwerten insbesondere auch Schwellenwertüberschreitungen besser erfasst werden sollen, wird im weiteren Verlauf für das Hybrid-Verfahren $nVorlauf = 1$ gesetzt.

Die Hybrid-Vorhersage kann im Gegensatz zur Multiplikativen Korrektur in Abschnitt 3.3.1 auch bei ausschließlich positiven Mess- und Modellwerten zu negativen korrigierten Modellwerten führen. Um dies zu vermeiden, wurde für die korrigierten Modellwerte eine untere Schranke von $0 \mu\text{g}/\text{m}^3$ angesetzt.

3.3.3 Kalman-Filter

Als drittes Korrekturverfahren wurde ein Kalman-Filter nach Delle Monache et al. (2006), Kang et al. (2008) und Borrego et al. (2011) auf die vom UBA zur Verfügung gestellten CAMS-Prognosen angewendet. Berechnet wird $\hat{x}_{t+\Delta t}$ als Schätzwert der Differenz zwischen Modellvorhersage und Messung an einem Messort für den zukünftigen Zeitpunkt $t+\Delta t$:

$$\hat{x}_{t+\Delta t} = \hat{x}_t + \beta_t \times (y_t - \hat{x}_t) \quad (3)$$

Δt beträgt dabei wie schon bei der Multiplikativen Korrektur in Abschnitt 3.3.1 24 h. \hat{x}_t stellt den Schätzwert der Differenz zwischen Modellvorhersage und Messung für den bereits bekannten Zeitpunkt t dar und y_t die beobachtete Differenz zwischen Modellvorhersage und Messung für den Zeitpunkt t . Der Schätzwert $\hat{x}_{t+\Delta t}$ wird demnach auf Basis des Schätzwertes \hat{x}_t gewonnen, der wiederum auf Basis des Schätzwertes $\hat{x}_{t-\Delta t}$ gewonnen wurde etc. Für y_t wird angenommen, dass es aufgrund von Messfehlern, Modellungenauigkeiten etc. von der wahren Differenz zwischen Modellvorhersage und Messung x_t um den Fehler ε_t („Messrauschen“) abweicht. Die wahre Differenz x_t zum Zeitpunkt t wird

wiederum angesetzt als Summe aus der wahren Differenz $x_{t-\Delta t}$ zum Zeitpunkt $t-\Delta t$ und einem Term $\eta_{t-\Delta t}$, der das Prozessrauschen beschreibt („white noise term“). Damit ergibt sich:

$$y_t = x_t + \varepsilon_t = x_{t-\Delta t} + \eta_{t-\Delta t} + \varepsilon_t \quad (4)$$

Sowohl für das Messrauschen ε_t als auch für das Prozessrauschen $\eta_{t-\Delta t}$, wird angenommen, dass es zeitlich nicht korreliert ist und normal verteilt mit Mittelwert Null und Varianz σ_ε^2 bzw. σ_η^2 .

Die Differenz zwischen beobachteter Differenz y_t und geschätzter Differenz \hat{x}_t zum Zeitpunkt t in Gleichung (3) wird mit dem Kalman-Faktor β_t („kalman gain factor“) gewichtet. β_t berechnet sich mit Hilfe der beschriebenen Fehlervarianzen σ_ε^2 und σ_η^2 wie folgt:

$$\beta_t = \frac{p_{t-\Delta t} + \sigma_\eta^2}{p_{t-\Delta t} + \sigma_\eta^2 + \sigma_\varepsilon^2} \quad (5)$$

p_t ist der Erwartungswert des mittleren quadratischen Fehlers zwischen wahrer und geschätzter Differenz zwischen Modellvorhersage und Messung ($E[(x_t - \hat{x}_t)^2]$) und lässt sich berechnen wie folgt:

$$p_t = (p_{t-\Delta t} + \sigma_\eta^2) \times (1 - \beta_t) \quad (6)$$

Um einen prognostizierten Modellwert $C_{mod,t+\Delta t}$ an einem Messort zu korrigieren, wird also zunächst der Kalman-Faktor β_t nach Gleichung (5) berechnet, und anschließend werden der Schätzwert $\hat{x}_{t+\Delta t}$ nach Gleichung (1) und (als Basis für den nächsten Zeitschritt) der Erwartungswert des mittleren quadratischen Fehlers p_t nach Gleichung (6) bestimmt. Der korrigierte Modellwert für den Zeitpunkt $t+\Delta t$ ergibt sich dann zu:

$$C_{mod,t+\Delta t}^{korr} = C_{mod,t+\Delta t} - \hat{x}_{t+\Delta t} \quad (7)$$

Um zu vermeiden, dass sich aus Gleichung (7) negative korrigierte Modellwerte ergeben, wurde analog zur Hybrid-Vorhersage (Abschnitt 3.3.2) eine untere Schranke von $0 \mu\text{g}/\text{m}^3$ angesetzt (Delle Monache et al., 2006). Als Startwerte zu Beginn der Zeitreihe wurden $p_{t-\Delta t} = 10\,000$ und $\hat{x}_t = 0$ vorgegeben, und die Kalman-Faktoren wurden einer „Trainingsphase“ von 2 Tagen unterzogen. Entsprechend dem Δt von 24 h wurde der Kalman-Filter für jede Stunde des Tages getrennt eingesetzt, d. h. ein für heute, 11:00 Uhr, prognostizierter Modellwert wurde ausschließlich auf Basis der Mess- und Modellwerte der vorangegangenen Tage für 11:00 Uhr korrigiert. Weist eine Messstation zu einer bestimmten Stunde einen Fehlwert auf, so wird die Korrektur der Modellprognose mit dem letzten gültigen Schätzwert für diese Stunde durchgeführt. Liegt jedoch an dieser Messstation zu dieser Stunde länger als 3 Tage ein Fehlwert vor, so werden der Schätzwert und damit auch der korrigierte Modellwert für alle folgenden Zeitschritte ebenfalls auf Fehlwert gesetzt. Der letzte gültige Schätzwert und der letzte gültige Erwartungswert des mittleren quadratischen Fehlers werden jedoch beibehalten und dienen als Startwert, falls die Messstation zu einem späteren Zeitpunkt wieder Messwerte liefert.

Die Fehlervarianzen σ_ε^2 und σ_η^2 wurden nach Delle Monache et al. (2006) als zeitabhängige Größen implementiert. Die Bestimmung von σ_ε^2 erfolgt ebenfalls mit Hilfe des Kalman-Filters. Die Gleichungen (1), (5) und (6) werden dazu entsprechend umformuliert:

$$\sigma_{\varepsilon,t+\Delta t}^2 = \sigma_{\varepsilon,t}^2 + \beta_t \times \left(\frac{(y_t - y_{t-\Delta t})^2}{2+r} - \sigma_{\varepsilon,t}^2 \right) \quad (8)$$

$$\beta_t^\sigma = \frac{p_{t-\Delta t}^\sigma + \sigma_\eta^2}{p_{t-\Delta t}^\sigma + \sigma_\eta^2 + \sigma_\varepsilon^2} \quad (9)$$

$$p_t^\sigma = (p_{t-\Delta t}^\sigma + \sigma_\eta^2) \times (1 - \beta_t^\sigma) \quad (10)$$

Als Startwerte zu Beginn der Zeitreihe wurden analog zur Bestimmung des Schätzwerts $p_{t-\Delta t}^\sigma = 10\,000$ und $\sigma_{\varepsilon,t}^2 = 0$ vorgegeben. $\sigma_{\sigma_\varepsilon}^2$ und $\sigma_{\sigma_\eta}^2$ wurden entsprechend Delle Monache et al. (2006) auf konstant 1 bzw. 0.0005 gesetzt. Zur Bestimmung von $\sigma_{\varepsilon,t+\Delta t}^2$ nach Gleichung (8) wird nicht nur die beobachtete

Differenz zwischen Modellvorhersage und Messung für den Zeitpunkt t , y_t , herangezogen, sondern auch $y_{t-\Delta t}$ aus dem vorangegangenen Zeitpunkt $t-\Delta t$.

Mit r wird das Verhältnis der Fehlervarianzen bezeichnet:

$$r = \sigma_{\eta}^2 / \sigma_{\varepsilon}^2 \quad \Rightarrow \quad \sigma_{\eta}^2 = r \sigma_{\varepsilon}^2 \quad (11)$$

Die Performanz des Kalman-Filters hängt stark von der Wahl von r ab (Delle Monache et al., 2006; Kang et al., 2008). Ist σ_{η}^2 sehr viel größer als σ_{ε}^2 , dann konvergiert β_t gegen 1, und das Ergebnis des Kalman-Filters konvergiert gegen das Ergebnis der Hybrid-Vorhersage.

Delle Monache et al. (2006) setzen zur Korrektur von Ozonprognosen $r = 0.01$ auf Basis von Literaturwerten. In einer weiteren Veröffentlichung untersuchen Delle Monache et al. (2008) die Sensitivität des Kalman-Filters bezüglich der Wahl von r und kommen auf einen optimalen Wert von $r=0.4$, schränken aber die Allgemeingültigkeit dieses Wertes mit Hinweis auf die untypischen Verhältnisse der zu Grunde liegenden Daten (kühler Sommer, wenige Luftschadstoffepisoden) ein. Ausführliche stationsbezogenen Sensitivitätstests von Kang et al. (2008) für Ozonprognosen ergaben optimale Werte für r zwischen 0.001 und 10. Höhere Werte für r ergaben sich vor allem dann, wenn die Modellergebnisse die Messungen eher schlecht prognostiziert haben. Der „optimale Wert“ für r wurde dabei sowohl bei Delle Monache et al. (2008) als auch bei Kang et al. (2008) als derjenige Wert bestimmt, für den sich der minimale RMSE zwischen Messung und Modellvorhersage ergab. Kang et al. (2008) zeigten dann, dass der RMSE nur unwesentlich schlechter wurde, wenn statt des stationspezifisch optimalen Wertes für r ein einheitlicher Wert von $r=0.06$ für alle Stationen gewählt wurde.

Um den Einfluss von r auf das Ergebnis des Kalman-Filters bei der vorliegenden Datengrundlage abschätzen zu können, wurden einige Sensitivitätstest unter Verwendung der Messdaten und der CAMS-Prognosen für das Jahr 2016 durchgeführt. Dazu wurden die CAMS-Prognosen mit dem Kalman-Filter und unterschiedlichen r korrigiert und die in Abschnitt 3.4.1 in Tabelle 3-5 beschriebenen Metriken bestimmt.

Für O_3 sind die Ergebnisse des Sensitivitätstests in Tabelle 3-1 und in Tabelle 3-2 aufgeführt. Dargestellt ist jeweils der Median der Metriken über alle Stationen. Darin zeigt sich ein deutlicher Einfluss von r auf die Maximalwerte und auf die Standardabweichung, sowohl für die Stundenmittelwerte (Tabelle 3-1) als auch für die gleitenden 8-Stundenmittelwerte (Tabelle 3-2). Beide Größen steigen mit steigendem r an. Korrelationskoeffizient, Bias und RMSE zeigen bei den Stundenmittelwerten für kleinere r bessere Werte als für größere r , der Einfluss von r auf diese Größen ist aber nicht besonders stark, wie auch Kang et al. (2008) bereits für den RMSE beschrieben haben. Bei den gleitenden 8-Stundenmittelwerten sind Bias und RMSE für größere r etwas besser als für kleinere r , der Korrelationskoeffizient ändert sich nicht. Im Hinblick auf die korrekte Erfassung von Maximalwerten und Schwellenwertüberschreitungen wäre demnach ein eher höherer Wert für r zu wählen. Allerdings wird für die 8-Stundenmittelwerte mit $r = 0.35$ die False Alarm Rate deutlich schlechter. Insofern wird für die weiteren Auswertungen in diesem Projekt für O_3 $r = 0.3$ angesetzt.

Tabelle 3-1: Metriken O₃, Stundenmittelwerte 1SMW für den 1. Prognosetag, Korrekturverfahren KF mit $r = 0.06, 0.2, 0.3, 0.35, 0.4$ (jeweils Median über alle Stationen)

Metrik	r = 0.06	r = 0.2	r = 0.3	r = 0.35	r = 0.4
Min Obs [$\mu\text{g}/\text{m}^3$]	0.52	0.52	0.52	0.52	0.52
Min Mod [$\mu\text{g}/\text{m}^3$]	0.00	0.00	0.00	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	44.87	44.87	44.87	44.87	44.87
Mean Mod [$\mu\text{g}/\text{m}^3$]	44.96	44.97	44.98	44.98	44.99
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.66	173.66	173.66	173.66
Max Mod [$\mu\text{g}/\text{m}^3$]	145.35	151.91	153.65	154.26	154.88
SD Obs [$\mu\text{g}/\text{m}^3$]	29.49	29.49	29.49	29.49	29.49
SD Mod [$\mu\text{g}/\text{m}^3$]	26.04	26.46	26.67	26.76	26.84
Korrelation	0.85	0.85	0.84	0.84	0.84
BIAS [$\mu\text{g}/\text{m}^3$]	0.08	0.10	0.11	0.12	0.12
RMSE [$\mu\text{g}/\text{m}^3$]	15.54	15.71	15.85	15.89	15.92
MFB [%]	5.33	4.63	4.23	3.98	3.81
FGE [%]	43.80	44.28	44.82	45.02	45.18

Tabelle 3-2: Metriken O₃, maximale gleitende 8-Stundenmittelwerte eines Tages 8TMAX für den 1. Prognosetag, Korrekturverfahren KF mit $r = 0.06, 0.2, 0.3, 0.35, 0.4$ (jeweils Median über alle Stationen)

Metrik	r = 0.06	r = 0.2	r = 0.3	r = 0.35	r = 0.4
Min Obs [$\mu\text{g}/\text{m}^3$]	3.08	3.08	3.08	3.08	3.08
Min Mod [$\mu\text{g}/\text{m}^3$]	5.30	4.16	3.52	3.24	3.15
Mean Obs [$\mu\text{g}/\text{m}^3$]	64.87	64.87	64.87	64.87	64.87
Mean Mod [$\mu\text{g}/\text{m}^3$]	62.87	63.11	63.22	63.27	63.32
Max Obs [$\mu\text{g}/\text{m}^3$]	155.21	155.21	155.21	155.21	155.21
Max Mod [$\mu\text{g}/\text{m}^3$]	135.58	140.45	142.19	142.96	143.46
SD Obs [$\mu\text{g}/\text{m}^3$]	29.56	29.56	29.56	29.56	29.56
SD Mod [$\mu\text{g}/\text{m}^3$]	26.66	27.01	27.18	27.25	27.31
Korrelation	0.91	0.91	0.91	0.91	0.91
BIAS [$\mu\text{g}/\text{m}^3$]	-2.01	-1.72	-1.57	-1.50	-1.44
RMSE [$\mu\text{g}/\text{m}^3$]	12.53	12.35	12.30	12.29	12.28
MFB [%]	-0.48	-0.05	0.08	0.11	0.12
FGE [%]	19.83	19.60	19.61	19.64	19.68
gemessene Überschreitungen	15 ¹				
Hit Rate [%]	36.36 ¹	46.67 ¹	50.00 ¹	50.00 ¹	50.00 ¹
FAR [%]	0.29 ¹	0.31 ¹	0.32 ¹	0.58 ¹	0.58 ¹

¹: Schwellenwert 120 $\mu\text{g}/\text{m}^3$

Für PM10 sind die Ergebnisse des Sensitivitätstests in Tabelle 3-3 und in Tabelle 3-4 aufgeführt. Dargestellt ist wieder der Median der Metriken über alle Stationen. Auch hier hat die Variation von r einen deutlichen Einfluss auf Maximalwerte und Standardabweichung, sowohl für die Stundenmittelwerte (Tabelle 3-3) als auch für die Tagesmittelwerte (Tabelle 3-4). Wie bei O₃ steigen beide Größen mit steigendem r an. Korrelationskoeffizient, Bias und RMSE zeigen bei den Stundenmittelwerten für kleinere r bessere Werte als für größere r , der Einfluss von r auf diese Größen ist aber nicht besonders stark. Dies entspricht den Ergebnissen für O₃ in Tabelle 3-1. Bei den Tagesmittelwerten sind Bias und RMSE im Gegensatz zu O₃ für kleinere r etwas besser als für größere r , der Korrelationskoeffizient hingegen steigt mit r leicht an. Insgesamt ist der Einfluss von r auf diese Größen aber auch bei PM10 eher gering. Im Hinblick auf die korrekte Erfassung von Maximalwerten und Schwellenwertüberschreitungen wird für die weiteren Auswertungen in diesem Projekt für PM10 $r = 1.0$ angesetzt.

Insgesamt liegt r für PM10 deutlich höher als für O₃. Wie in Abschnitt 4 gezeigt wird, wird PM10 durch die CAMS-Prognosen auch deutlich schlechter erfasst als O₃. Dies entspricht der Beobachtung von Kang et al. (2008), dass höhere optimale Werte für r im Regelfall in Bereichen gefunden wurden, in denen die Modellergebnisse die Messungen nicht so gut erfassen.

Tabelle 3-3: Metriken PM10, Stundenmittelwerte 1SMW für den 1. Prognosetag, Korrekturverfahren KF mit $r = 0.5, 0.8, 0.9, 1.0, 1.2$ (jeweils Median über alle Stationen)

Metrik	$r = 0.5$	$r = 0.8$	$r = 0.9$	$r = 1.0$	$r = 1.2$
Min Obs [$\mu\text{g}/\text{m}^3$]	0.50	0.50	0.50	0.50	0.50
Min Mod [$\mu\text{g}/\text{m}^3$]	0.00	0.00	0.00	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.13	16.13	16.13	16.13	16.13
Mean Mod [$\mu\text{g}/\text{m}^3$]	16.17	16.17	16.17	16.17	16.17
Max Obs [$\mu\text{g}/\text{m}^3$]	128.45	128.45	128.45	128.45	128.45
Max Mod [$\mu\text{g}/\text{m}^3$]	80.78	88.13	89.30	90.12	92.63
SD Obs [$\mu\text{g}/\text{m}^3$]	11.30	11.30	11.30	11.30	11.30
SD Mod [$\mu\text{g}/\text{m}^3$]	9.63	9.92	10.02	10.10	10.21
Korrelation	0.59	0.58	0.58	0.58	0.58
BIAS [$\mu\text{g}/\text{m}^3$]	0.09	0.09	0.10	0.10	0.10
RMSE [$\mu\text{g}/\text{m}^3$]	9.43	9.61	9.63	9.68	9.74
MFB [%]	4.04	2.86	2.51	2.23	1.78
FGE [%]	44.65	45.49	45.75	45.95	46.31

Tabelle 3-4: Metriken PM10, Tagesmittelwerte 1TMW für den 1. Prognosestag, Korrekturverfahren KF mit $r = 0.5, 0.8, 0.9, 1.0, 1.2$ (jeweils Median über alle Stationen)

Metrik	$r = 0.5$	$r = 0.8$	$r = 0.9$	$r = 1.0$	$r = 1.2$
Min Obs [$\mu\text{g}/\text{m}^3$]	2.02	2.02	2.02	2.02	2.02
Min Mod [$\mu\text{g}/\text{m}^3$]	2.30	2.07	2.02	1.97	1.87
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.15	16.15	16.15	16.15	16.15
Mean Mod [$\mu\text{g}/\text{m}^3$]	16.17	16.16	16.16	16.16	16.16
Max Obs [$\mu\text{g}/\text{m}^3$]	67.31	67.31	67.31	67.31	67.31
Max Mod [$\mu\text{g}/\text{m}^3$]	55.10	56.97	57.54	58.33	59.11
SD Obs [$\mu\text{g}/\text{m}^3$]	8.93	8.93	8.93	8.93	8.93
SD Mod [$\mu\text{g}/\text{m}^3$]	8.01	8.18	8.21	8.23	8.28
Korrelation	0.71	0.72	0.72	0.72	0.73
BIAS [$\mu\text{g}/\text{m}^3$]	0.09	0.09	0.09	0.09	0.10
RMSE [$\mu\text{g}/\text{m}^3$]	6.23	6.27	6.27	6.27	6.29
MFB [%]	3.24	2.71	2.60	2.47	2.25
FGE [%]	27.86	27.66	27.64	27.62	27.62
gemessene Überschreitungen	2 ¹				
Hit Rate [%]	0.00 ¹				
FAR [%]	0.28 ¹				

¹: Schwellenwert 50 $\mu\text{g}/\text{m}^3$

3.4 Verwendete Metriken

3.4.1 Statistische Kennwerte

Für die Evaluierung der unkorrigierten und der korrigierten Prognosedaten an den Messorten wurden die in Tabelle 3-5 aufgeführten statistischen Kennwerte (Metriken) verwendet. Die Metriken wurden für jede Station einzeln bestimmt und in Box-Whisker-Plots zusammengeführt, um die Spannbreite der Metriken darzustellen.

Vergleich von Minimum, Mittelwert, Maximum und Standardabweichung zwischen Messung und Modell verschaffen einen ersten Überblick über die Güte der Modellergebnisse. Der Korrelationskoeffizient nach Pearson zeigt an, inwieweit der zeitliche Verlauf der Messdaten von den Modelldaten erfasst wird. Der Bias beschreibt den systematischen Fehler des Modells, d. h. ob das Modell die Messwerte über- oder unterschätzt. Der Root Mean Square Error (RMSE) erfasst sowohl den systematischen als auch unsystematische (zufällige) Fehler des Modells. Er repräsentiert die Standardabweichung der Differenzen zwischen Messung und Modell und wird i. d. R. verwendet, um den Fehler verschiedener Modelle für einen bestimmten Datensatz zu vergleichen. Für den Vergleich der Modellperformance zwischen verschiedenen Datensätzen (z. B. verschiedene Schadstoffe) ist der RMSE nicht geeignet, da er skalenabhängig ist. Das gilt auch für den Bias. Weiterhin ist der RMSE sensitiv bezüglich Ausreißern, weil jede Abweichung zwischen Messung und Modell im Quadrat in den RMSE eingeht und damit größere Fehler einen überproportionalen Einfluss auf den RMSE haben.

Um eine skalenunabhängige Bewertung der Modellergebnisse zu erhalten, wurden zwei weitere statistische Größen betrachtet, der Mean Fractional Bias (MFB):

$$MFB = \frac{2}{N} \sum_{i=1}^N \frac{(C_m - C_o)}{(C_m + C_o)} \quad (12)$$

und der Fractional Gross Error (FGE):

$$FGE = \frac{2}{N} \sum_{i=1}^N \frac{|C_m - C_o|}{(C_m + C_o)} \quad (13).$$

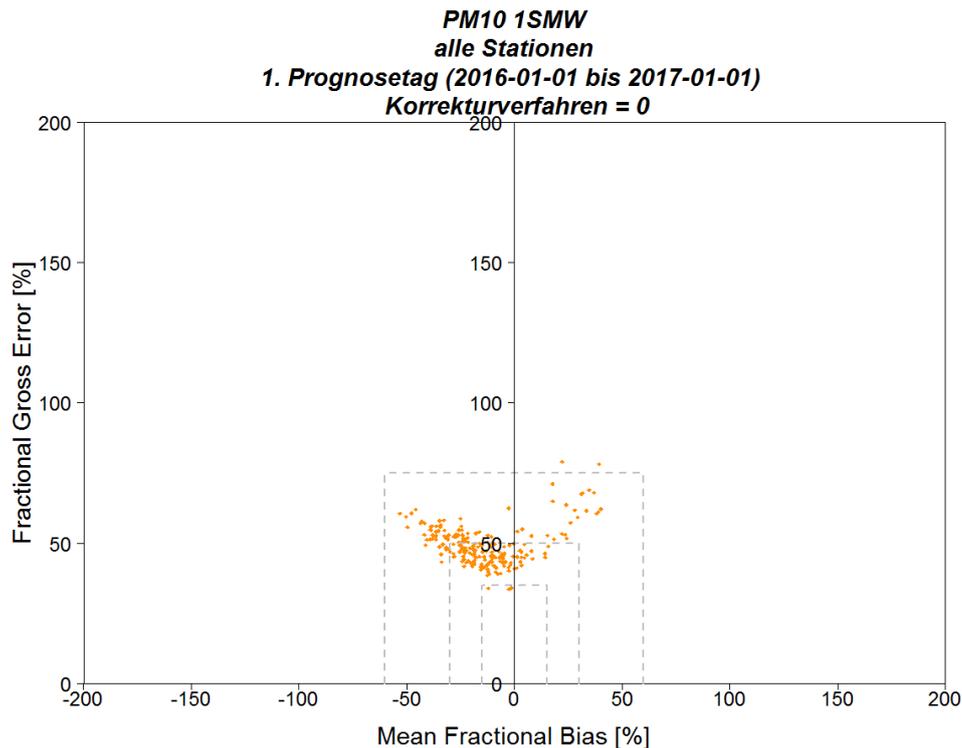
Beide Größen werden mit Hilfe des Mittelwerts aus Messung und Modell normalisiert und sind damit unabhängig von der Skala des betrachteten Stoffs. Sie verhalten sich symmetrisch hinsichtlich Unter- und Überschätzung. Der MFB kann Werte zwischen -200 % und 200 % annehmen, optimal ist ein Wert von 0 %. Der FGE kann Werte zwischen 0 und 200 % annehmen, optimal ist auch hier ein Wert von 0 %.

Tabelle 3-5: Verwendete statistische Kennwerte / Metriken

Bezeichnung	Abkürzung
Minimum der Messwerte	Min Obs [$\mu\text{g}/\text{m}^3$]
Minimum der Modellwerte	Min Mod [$\mu\text{g}/\text{m}^3$]
Mittelwert der Messwerte	Mean Obs [$\mu\text{g}/\text{m}^3$]
Mittelwert der Modellwerte	Mean Mod [$\mu\text{g}/\text{m}^3$]
Maximum der Messwerte	Max Obs [$\mu\text{g}/\text{m}^3$]
Maximum der Modellwerte	Max Mod [$\mu\text{g}/\text{m}^3$]
Standardabweichung der Messwerte	SD Obs [$\mu\text{g}/\text{m}^3$]
Standardabweichung der Modellwerte	SD Mod [$\mu\text{g}/\text{m}^3$]
Korrelationskoeffizient nach Pearson	Korrelation
Bias	BIAS [$\mu\text{g}/\text{m}^3$]
Root Mean Square Error	RMSE [$\mu\text{g}/\text{m}^3$]
Mean Fractional Bias	MFB [%]
Fractional Gross Error	FGE [%]
Anzahl der gemessenen Überschreitungen	MeasExc
Hit Rate [%]	Hit Rate [%]
False Alarm Rate [%]	FAR [%]

Wird in einer Grafik der FGE über dem MFB aufgetragen, so erhält man eine Darstellung, die als „Soccer-Plot“ bezeichnet wird (Savage et al., 2013). Abbildung 3-3 zeigt beispielhaft einen Soccer-Plot für die unkorrigierten CAMS-Prognosen 2016 für PM10. Jeder Punkt gibt den FGE und den MFB für eine Station wieder. Für eine perfekte Vorhersage läge der Punkt der Station im Ursprung der Grafik. Die drei eingezeichneten Boxen dienen der visuellen Orientierung, sie liegen analog zu Savage et al. (2013) bei 15 % / 35 %, 30 % / 50 % und 60 % / 75 %. Ein systematischer Fehler erscheint als eine lineare Gruppierung von Punkten. Eine Streuung der Punkte deutet hingegen darauf hin, dass zufällige Fehler dominieren. Diese Art der Darstellung erlaubt auf einen Blick die Charakterisierung der Modellfehler und die Einordnung der Vorhersagegüte des betrachteten Modells. Im vorliegenden Beispiel zeigen die meisten Messorte einen negativen Bias, eine geringere Anzahl Messorte weist auch einen positiven Bias auf. Die nur schwach lineare Gruppierung der Punkte der einzelnen Stationen deutet darauf hin, dass bei PM10 eher zufällige Fehler in den CAMS-Prognosedaten dominieren.

Abbildung 3-3: Beispielhafte Darstellung eines Soccer-Plots, der den Fractional Gross Error aufgetragen zeigt über den Mean Fractional Bias



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Die bisher beschriebenen Metriken dienen der Bewertung der Vorhersagegüte eines Modells auf allen Konzentrationsniveaus. Diese Metriken sind jedoch nur bedingt geeignet, um die Vorsagegüte eines Modells hinsichtlich der Erfassung von Schwellenwertüberschreitungen zu beschreiben. Um dafür geeignete Metriken zu definieren, können Mess- und Modellwerte entsprechend Abbildung 3-4 in vier Bereiche eingeteilt werden:

- ▶ A: korrekte Vorhersage einer Schwellenwertüberschreitung
- ▶ B: Schwellenwertüberschreitung vom Modell falsch vorhergesagt
- ▶ C: Schwellenwertüberschreitung vom Modell nicht erkannt
- ▶ D: korrekt erkannt, dass keine Schwellenwertüberschreitung stattgefunden hat

Daraus wurde zur Orientierung zunächst die Anzahl der gemessenen Überschreitungen („measured exceedances“) bestimmt:

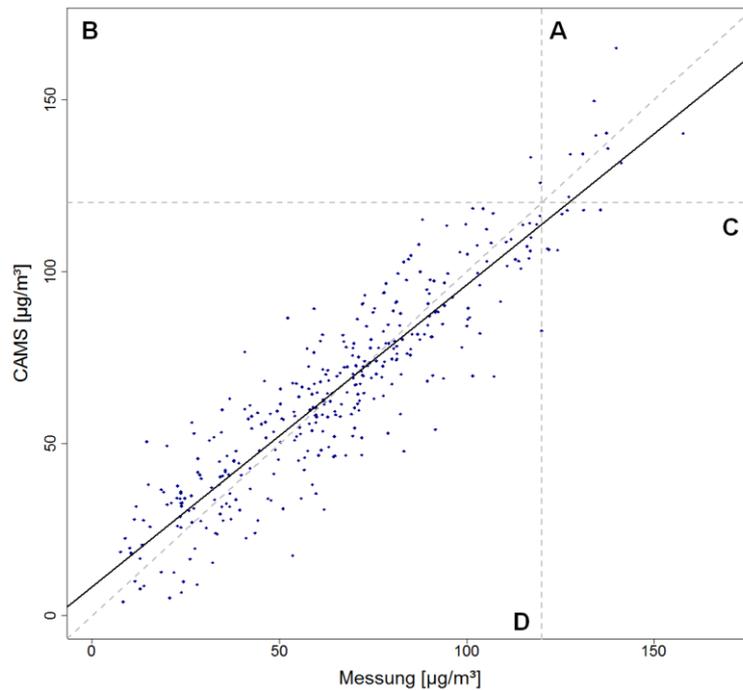
$$MeasExc = A + C \quad (14)$$

Des Weiteren wurden die Hit Rate und die False Alarm Rate (FAR) ermittelt (Savage et al., 2013). Die Hit Rate gibt den Anteil korrekter Vorhersagen einer Schwellenwertüberschreitung an der Anzahl der gemessenen Schwellenwertüberschreitungen an, während die FAR den Anteil der vom Modell fälschlicherweise vorhergesagten Schwellenwertüberschreitungen an der Anzahl der gemessenen Werte unterhalb des Schwellenwerts angibt:

$$Hit\ Rate = 100 * A / (A + C) \quad [\%] \quad (15)$$

$$FAR = 100 * B / (B + D) \quad [\%] \quad (16)$$

Abbildung 3-4: Beispielhafte Darstellung der vier Quadranten A – D zur Bestimmung der Metriken zur Bewertung von Schwellenwertüberschreitungen. Näheres siehe Text.



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

3.4.2 DELTA-Tool

In Erweiterung der statistischen Auswertung wurde das im Rahmen der FAIRMODE-Initiative entstandene DELTA-Tool für "Assessment & Planning" (A&P; Thunis & Cuvelier, 2017; FAIRMODE, 2017) auf die originalen und auf die korrigierten CAMS-Prognosen angewendet. Das DELTA-Tool ist ein in der Entwicklung befindliches Analysetool, das Modellergebnisse im Vergleich mit Messungen statistisch bewertet. Entwickelt wird das DELTA-Tool im Rahmen der Arbeitsgruppe 1 (WG1 Assessment) des Europäischen Forums für Luftqualitätsmodellierung (FAIRMODE, Forum for air quality modelling in Europe, <http://fairmode.jrc.ec.europa.eu>), mit dem Ziel, Modellevaluierungen zu harmonisieren. Es werden Modellqualitätsziele festgelegt und in Form des DELTA-Tools in einem Analysewerkzeug bereitgestellt, mit dem Modellierungen im Vergleich mit Messungen nach den implementierten Methoden statistisch bewertet werden können.

Im sogenannten Benchmark-Modus werden zusammenfassende Bewertungen anhand von Gütekriterien für verschiedene statistische Indikatoren erstellt. Diese Gütekriterien sind fest im Tool implementiert und können damit vom Benutzer nicht verändert werden (Thunis & Cuvelier, 2017; FAIRMODE, 2017). Die wichtigsten im Benchmark-Mode verwendeten statistischen Indikatoren sind:

- ▶ Mittlerer quadratischer Fehler (RMSE),
- ▶ Korrelationskoeffizient (R),
- ▶ Normierte Mittlere Abweichung (NMB),
- ▶ Normierte mittlere Standardabweichung (NMSD) und
- ▶ Zentrierter mittlerer quadratischer Fehler (CRMSE).

Die mathematischen Formulierungen dieser Größen können Thunis et al. (2012), FAIRMODE (2017) und Thunis & Cuvelier (2017) entnommen werden. Die Qualität von Modellergebnissen wird abgeschätzt mit Hilfe eines aggregierenden Modellierungsqualitäts-Indikators („modelling quality indicator“, MQI), der mit einem festgelegten Modellierungsqualitäts-Ziel („modelling quality objective“,

MQO) verglichen wird. Zusätzlich werden weitere Modellierungsgüte-Indikatoren („modelling performance indicator“, MPI) und entsprechende Vergleichskriterien („modelling performance criteria“, MPC) definiert und damit die an die Modelle zu stellenden Genauigkeitsanforderungen bezüglich ihres Einsatzes für Fragestellungen in der Luftreinhaltung für die einzelnen statistischen Indikatoren festgelegt. Die MQI und MPI werden in Abhängigkeit von den Messunsicherheiten der einzelnen Schadstoffkomponenten bestimmt, d. h. den Modellergebnissen wird ein Toleranzbereich zugeordnet, der vom Unsicherheitsbereich der Messungen abhängt.

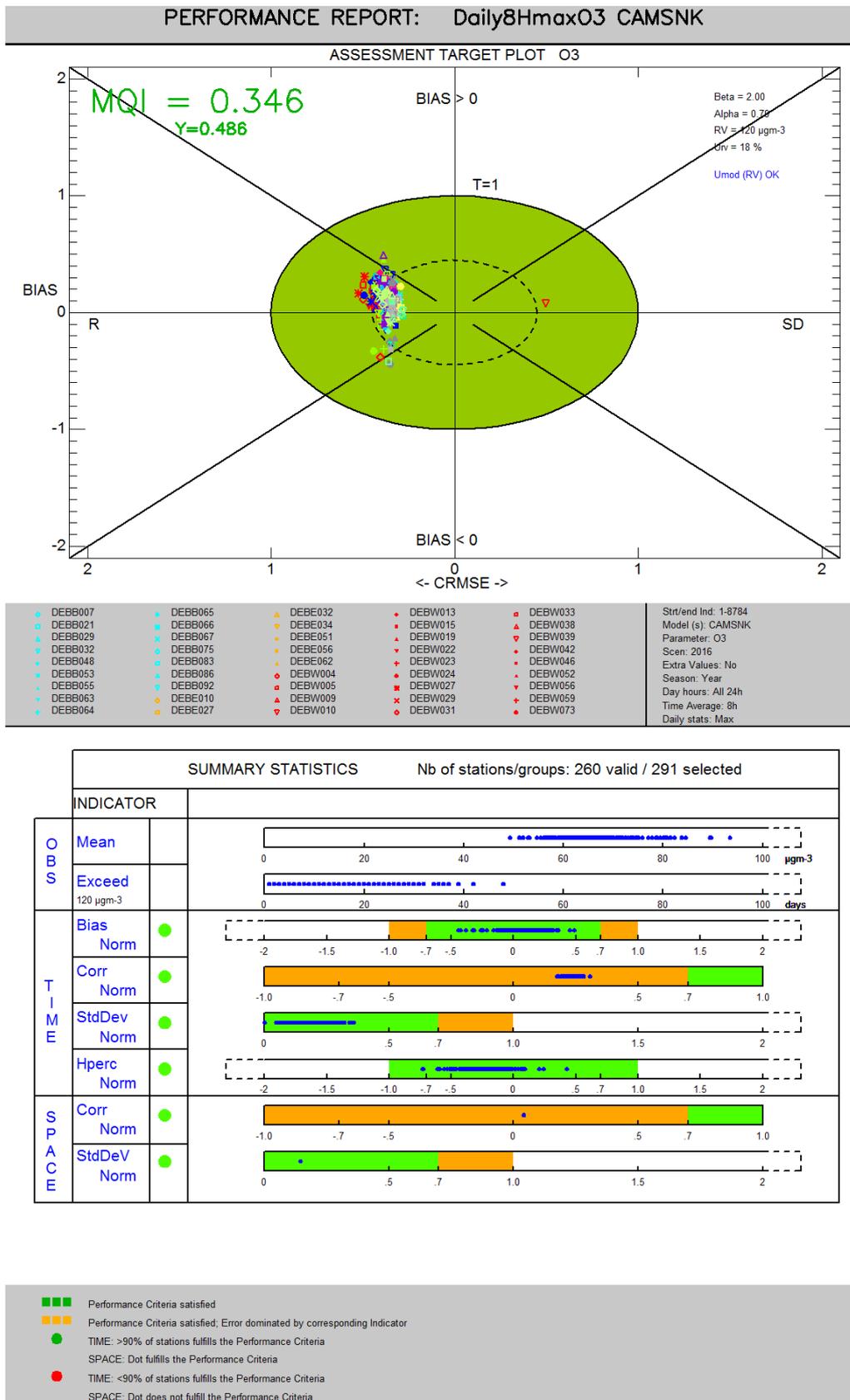
Die Beurteilung der Güte der Modellierung über den MQI wird anhand dreier Kategorien vorgenommen:

- ▶ $MQI \leq 0.5$. Der Unterschied zwischen Modellergebnissen und Messungen an einer Station ist kleiner als die Messunsicherheit. Die Modellergebnisse liegen damit im Mittel im Bereich der Messunsicherheit und erfüllen das Modellqualitätsziel. Eine Verbesserung der Modellgüte im Sinne einer „besseren“ Abbildung der Messungen ist nicht möglich.
- ▶ $0.5 < MQI \leq 1 = MQO$. Der Unterschied zwischen Modellergebnissen und Messungen an einer Station ist größer als die Messunsicherheit, die Unsicherheitsbereiche von Modellergebnissen und Beobachtungen überdecken sich. Damit ist das Modellergebnis noch in dem geforderten Bereich und erfüllt das Modellqualitätsziel.
- ▶ $MQI > 1$. Die Unterschiede zwischen Modellergebnissen und Beobachtungen sind signifikant. Das Modellqualitätsziel ist nicht erreicht.

Der MQI gibt einen allgemeinen Überblick zur Güte einer Modellierung. Die zusätzlichen Parameter ermöglichen eine detailliertere Beurteilung der Güte und zeigen damit mögliche Aspekte einer Verbesserung der Modellierung auf. Die entscheidende Rolle bei allen statistischen Parametern spielt die Definition der Normierungsgröße, also des Unsicherheitsbereichs der Messungen. Für die Ableitung dieses Unsicherheitsbereiches wird auf die o. g. Quellen verwiesen. Die normierte Auswertung der statistischen Parameter wird in den Benchmark-Reports des DELTA-Tools in einem zentralen Diagramm zum MQI, dem sogenannten „Assessment Target Plot“, und diversen zusammenfassenden Statistiken zu den einzelnen MPI dargestellt (Abbildung 3-5). In den Target Plots erfüllen Stationspunkte, die innerhalb der grünen Ellipse liegen, die Bedingung „ $MQI < 1$ “, und Stationspunkte, die innerhalb der gestrichelten grünen Ellipse liegen, die Bedingung „ $MQI < 0.5$ “. Für Stationspunkte oberhalb der x-Achse bei $y > 0$ liegt ein positiver und für Stationspunkte unterhalb ein negativer Bias der Modelldaten gegenüber den Messwerten vor. Die Aufteilung der Stationspunkte auf die linke und die rechte Hälfte des Diagramms erfolgt in Abhängigkeit von Korrelationskoeffizient (R) und Standardabweichung (SD) (FAIRMODE, 2017)

Das Modellierungsqualitäts-Ziel („modelling quality objective“, MQO) gilt, entsprechend der Luftqualitätsrichtlinie 2008/50/EG (EU, 2008), als erfüllt, wenn es für 90 % der Stationen erfüllt ist. In den Benchmark-Reports des DELTA-Tools wird dementsprechend der MQI-Wert in grün dargestellt, wenn das 90-Perzentil der MQI aller Stationen ≤ 1 ist. Im vorliegenden Projekt wurden die Auswertungen mit dem DELTA-Tool in der Version 5.6 durchgeführt.

Abbildung 3-5: Beispiel eines Benchmark-Reports des DELTA-Tools mit Assessment Target Plot (oben) und zusammenfassenden Statistiken (unten)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Die Version 5.6 des DELTA-Tools lag zum Zeitpunkt der Arbeiten als Beta-Version vor, die neben dem Benchmark-Modus eine Bewertung von Modellprognosen im Forecast-Modus erlaubt. Der Forecast-Modus vergleicht die betrachteten Prognosedaten mit einem persistenten Modell, welches davon ausgeht, dass der Konzentrationswert, der zum Zeitpunkt i (z. B. morgen 11:00 Uhr) gemessen werden wird, dem Konzentrationswert entspricht, der zum Zeitpunkt j (z. B. gestern um 11:00 Uhr) gemessen wurde. Die Vorhersagelänge, d. h. die Differenz zwischen den Zeitpunkten i und j , ist dabei entsprechend dem Vorhersagezeitraum der betrachteten Prognosedaten vorzugeben.

Der MQI wird im Forecast-Modus nicht wie im Benchmark-Modus in Abhängigkeit von den Messunsicherheiten der einzelnen Schadstoffkomponenten bestimmt, sondern auf eine Größe bezogen, die die „Intra-Day-Variation“ der Messdaten, mithin die Differenz der Messwerte zu den Zeitpunkten i und j , repräsentiert. Für die mathematische Formulierung dieser Größe wird auf FAIRMODE (2017) verwiesen. Die Güte der Vorhersage wird dann wie folgt beurteilt:

- ▶ $MQI < 1$. Die betrachteten Modellprognosen sind besser als das persistente Modell.
- ▶ $MQI \geq 1$. Die betrachteten Modellprognosen sind schlechter als das persistente Modell.

Wichtig für das Verständnis des MQI im Forecast-Modus ist, dass er im Gegensatz zum MQI im Benchmark-Modus, der sich mit den Messwerten auf eine feste Referenzgröße bezieht, eine vergleichende Aussage zwischen zwei Modellen trifft, in dem er für eine gegebene Vorhersagelänge die Modellprognose mit dem persistenten Modell vergleicht. Hat man beispielsweise ein Modell, dessen Vorhersagegüte mit zunehmender Vorhersagelänge weniger stark abnimmt als die des persistenten Modells, dann führt dies dazu, dass der MQI des Forecast-Modus für das Modell mit steigender Vorhersagelänge kleiner wird, auch wenn die Vorhersagegüte des Modells mit steigender Vorhersagelänge abnimmt (was durch den steigenden MQI im Benchmark-Modus angezeigt wird).

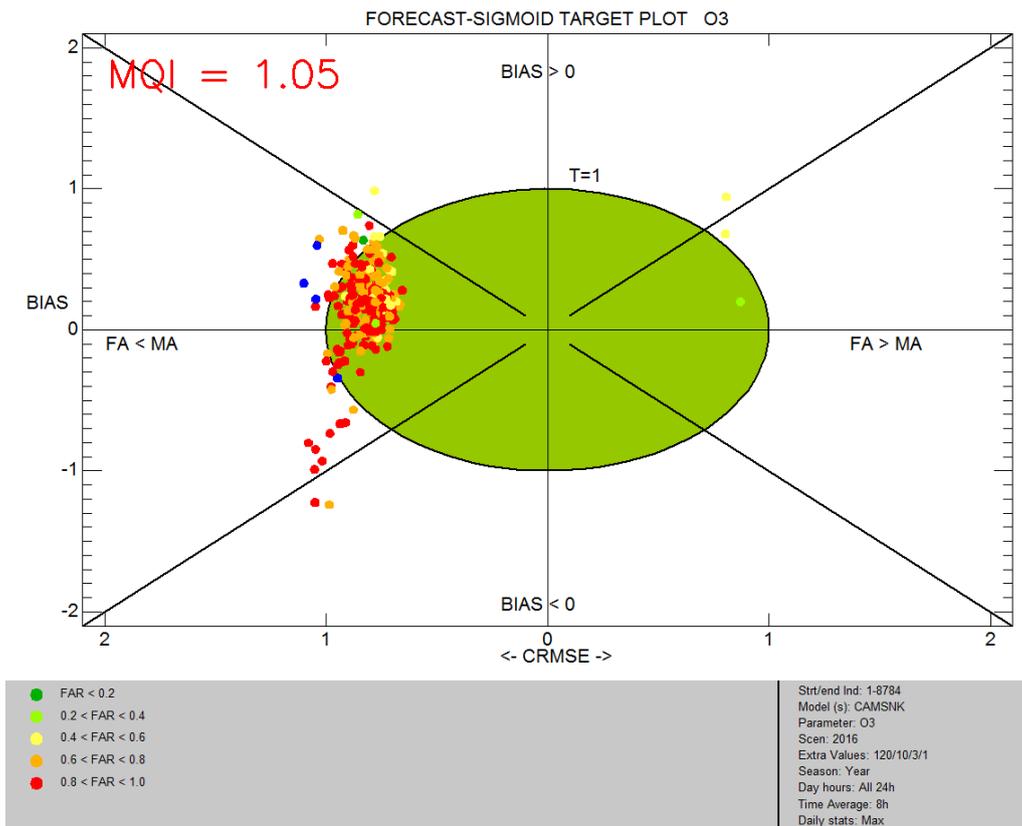
Die normierte Auswertung der statistischen Parameter wird vom DELTA-Tool im Forecast-Modus analog zum „Assessment Target Plot“ in einem „Forecast-Sigmoid Target Plot“ zusammengefasst (Abbildung 3-6). An Stationspunkten, die innerhalb der grünen Ellipse liegen, ist die Bedingung „ $MQI < 1$ “ erfüllt, d. h. die betrachteten Prognosedaten sind für die betrachtete Vorhersagelänge besser als das persistente Modell. Die Position der Stationspunkte ober- und unterhalb der x-Achse ($y=0$) zeigt analog zum Assessment Target Plot den Bias der Prognosedaten an, und die Aufteilung der Stationspunkte auf die linke und die rechte Hälfte des Target Plots wird bestimmt in Abhängigkeit der fälschlicherweise vorhergesagten Schwellenwertüberschreitungen (False Alarms, FA) und der nicht erfassten Schwellenwertüberschreitungen (Missed Alarms, MA). Die Festlegung, ob eine Schwellenwertüberschreitung gemessen oder modelliert wurde, erfolgt im DELTA-Tool unter Berücksichtigung der Messunsicherheiten der einzelnen Schadstoffkomponenten (FAIRMODE, 2017). Die Einfärbung der Stationspunkte erfolgt nach FAIRMODE (2017) in Klassen in Abhängigkeit der hier vom DELTA-Tool ebenfalls als FAR bezeichnete False Alarm Ratio. Die False Alarm Ratio berechnet sich im Gegensatz zur False Alarm Rate nach Gleichung (16) in Abschnitt 3.4.1 zu

$$FAR = 100 * B / (B + A) [-] \quad (17)$$

Die Größen A und B sind in Abschnitt 3.4.1 definiert.

Da es sich bei der verwendeten Version 5.6 des DELTA-Tools um eine Beta-Version handelt, sind die in den folgenden Kapiteln dargestellten Ergebnisse, insbesondere bezüglich des noch stärker in der Entwicklung befindlichen Forecast-Modus, als vorläufig zu betrachten.

Abbildung 3-6: Beispiel eines Forecast-Sigmoid Target Plots des DELTA-Tools



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

4 Evaluierung der unkorrigierten Modelldaten

4.1 Ozon

4.1.1 Einfluss Mittelwertbildung der Modelldaten

Wie in Abschnitt 3.2 beschrieben, liegen die CAMS-Prognosedaten als Momentanwerte vor, die Messdaten hingegen als Stundenmittelwerte. In der Testumgebung kann daher gewählt werden, ob die weiteren Auswertungen auf Basis der Momentanwerte der CAMS-Daten durchgeführt werden sollen, oder ob zunächst aus den Momentanwerten Stundenmittel gebildet werden sollen, indem aus dem Momentanwerten jeweils zu Anfang und zu Ende einer Stunde der Mittelwert bestimmt wird (Abschnitt 3.1).

Um den Einfluss der Bildung von Stundenmittelwerten für die CAMS-Daten zu betrachten, wurden die in Tabelle 3-5 aufgeführten Metriken testweise für O₃ für die Momentanwerte und für die stundengemittelten Werte bestimmt und miteinander verglichen. Die Ergebnisse sind in Tabelle 4-1 dargestellt, es ist jeweils der Median der betrachteten Metrik über alle Stationen angegeben. Neben den Stundenmittelwerten 1SMW wurden auch die Metriken der abgeleiteten Größen 1TMAX (höchster Einstundenmittelwert pro Tag) und 8TMAX (höchster gleitender Achtstundenmittelwert pro Tag) berechnet. Im Ergebnis unterscheiden sich die Metriken für die stundengemittelten und die Momentanwerte nur marginal. Da die Verwendung stundengemittelter CAMS-Daten den Stundenmittelwerten der Messdaten inhaltlich näher kommt, wurden im Weiteren nur noch stundengemittelte CAMS-Daten verwendet.

Tabelle 4-1: Metriken CAMS-Ozonprognose, basierend auf stundengemittelten Werten und auf Momentanwerten, 1. Prognosestag (jeweils Median über alle Stationen), Bezugsjahr 2016

Metrik	stundengemittelte Werte O ₃			Momentanwerte O ₃		
	1SMW	1TMAX	8TMAX	1SMW	1TMAX	8TMAX
Min Obs [µg/m ³]	0.52	4.79	3.08	0.52	4.79	3.08
Min Mod [µg/m ³]	0.82	20.36	17.49	0.55	20.87	17.59
Mean Obs [µg/m ³]	44.72	71.73	64.59	44.71	71.73	64.59
Mean Mod [µg/m ³]	56.36	72.56	69.07	56.36	73.13	69.17
Max Obs [µg/m ³]	173.66	173.66	155.21	173.66	173.66	155.21
Max Mod [µg/m ³]	139.15	139.15	130.41	139.74	139.74	130.81
SD Obs [µg/m ³]	29.44	30.63	29.46	29.44	30.63	29.46
SD Mod [µg/m ³]	22.29	23.19	22.66	22.41	23.22	22.69
Korrelation	0.81	0.88	0.89	0.81	0.88	0.89
BIAS [µg/m ³]	10.17	0.22	3.63	10.17	0.74	3.76
RMSE [µg/m ³]	19.82	15.23	14.56	19.90	15.25	14.56
MFB [%]	38.27	5.64	12.68	37.68	6.51	12.80
FGE [%]	48.91	20.13	22.90	48.64	20.19	22.90
gemessene Überschreitungen	--	0	15 ²	--	0	15 ²
Hit Rate [%]	--	0.00 ¹	13.33 ²	--	0.00 ¹	14.29 ²
FAR [%]	--	0.00 ¹	0.00 ²	--	0.00 ¹	0.00 ²

¹: Schwellenwert 180 µg/m³ (Informationsschwelle) ²: Schwellenwert 120 µg/m³

4.1.2 Jahreszeitreihe 2016

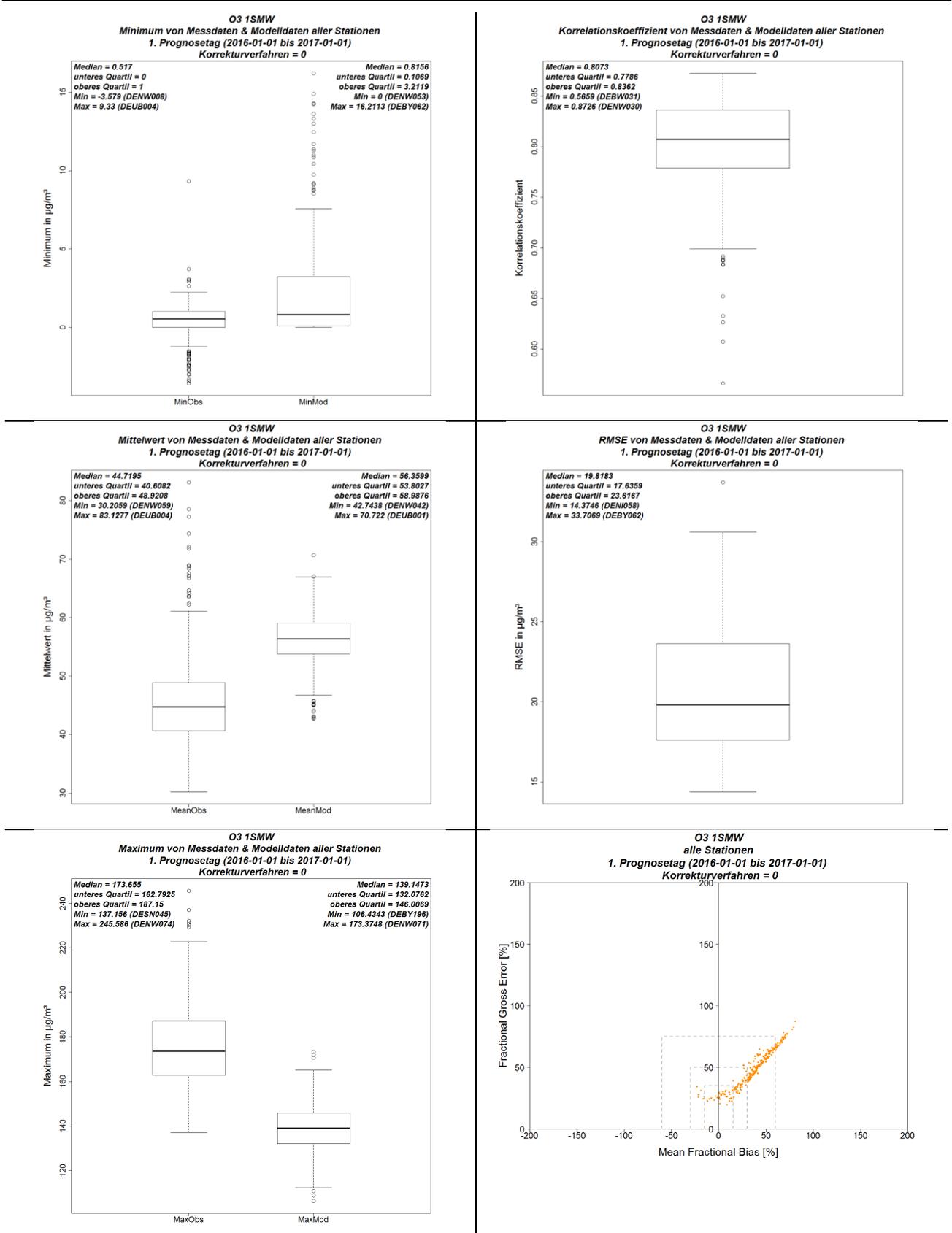
Die CAMS-Prognosedaten wurden zunächst ohne Korrektur evaluiert, um das Modellverhalten erfassen zu können. Bezugsjahr der Evaluierung ist 2016. Um für die Testumgebung durchgehende Jahreszeitreihen der CAMS-Prognosen zu erhalten, wurden diese in Abhängigkeit vom Prognosetag gebildet (Abschnitt 3.1). Die statistische Auswertung (Metriken nach Tabelle 3-5, DELTA-Tool) erfolgte dann in Abhängigkeit vom Prognosetag.

Abbildung 4-1 zeigt für den 1. Prognosetag die Box-Whisker-Plots der Minima, Mittelwerte, Maxima, Korrelationskoeffizienten und RMSE an allen Messorten sowie den zugehörigen Soccer-Plot für die Stundenmittelwerte. Die Mediane der Minima liegen für Messwerte und Modelldaten in der gleichen Größenordnung. Auffällig sind bei den Minima die vielen negativen Messwerte. Diese lassen sich im Wesentlichen auf Messstandorte in Nordrhein-Westfalen zurückverfolgen und entsprechen etwa der dort festgelegten negativen Nachweisgrenze. Negative Messwerte sind sowohl für einzelne Metriken (MFB, FGE) als auch für die Korrekturverfahren problematisch. In Abstimmung mit dem Auftraggeber wurden diese negativen Messwerte für die Korrekturverfahren zu Null gesetzt (Abschnitt 3.3).

Der Box-Whisker-Plot der Mittelwerte zeigt, dass die CAMS-Prognosedaten die gemessenen O_3 -Werte im Mittel deutlich überschätzen. Die gemessenen Maximalwerte hingegen werden von den Modellprognosen unterschätzt. Der Korrelationskoeffizient zwischen Messung und Modell liegt für O_3 zwischen 0.57 (Station DEBW031: Schwarzwald-Süd) und 0.87 (Station DENW030: Wesel-Feldmark) mit dem Median bei 0.81. Der RMSE beträgt im Median knapp $20 \mu\text{g}/\text{m}^3$ und damit fast die Hälfte des Medians der Mittelwerte der Messdaten von $44.7 \mu\text{g}/\text{m}^3$. Im Soccer-Plot ist ein deutlich positiver Bias zu erkennen, der MFB zeigt Werte bis knapp über 80 %. Die lineare Gruppierung der Punkte der einzelnen Stationen deutet auf einen überwiegend systematischen Fehler in den CAMS-Prognosedaten hin.

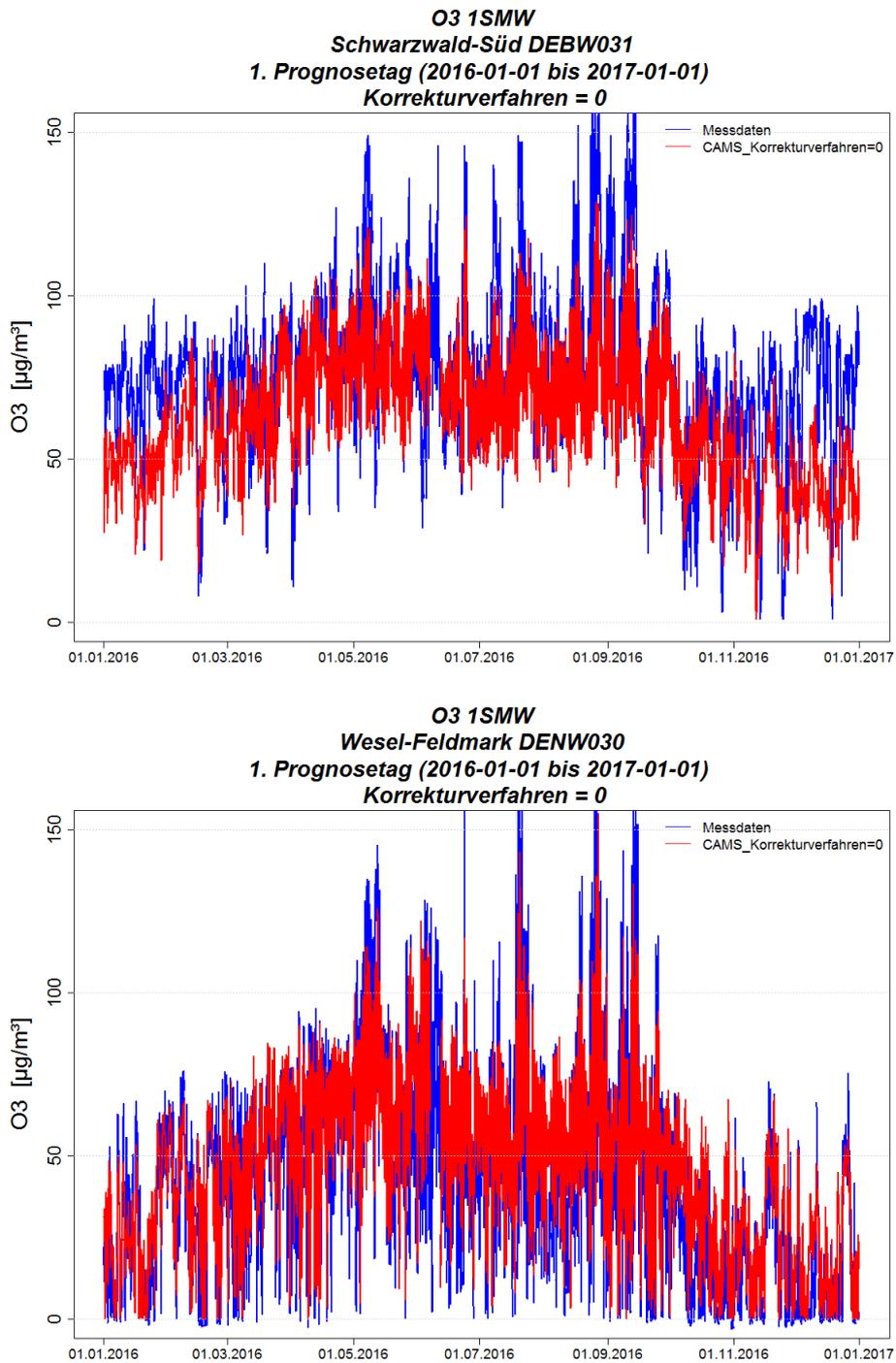
In Abbildung 4-2 sind die Zeitreihen der Stundenmittelwerte für die Station mit dem niedrigsten (DEBW031) und mit dem höchsten (DENW030) Korrelationskoeffizienten dargestellt. Der Korrelationskoeffizient nach Pearson zeigt an, inwieweit der zeitliche Verlauf der Messdaten von den Modelldaten erfasst wird. An der Station DEBW031 unterschätzen die CAMS-Prognosen vor allem die gemessenen Maximalwerte, in den Wintermonaten (Januar, Februar, November, Dezember) wird auch der grundsätzliche Verlauf der gemessenen Daten nicht vom Modell wiedergegeben. An der Station DENW030 hingegen werden sowohl die Minimal- und Maximalwerte als auch der grundsätzliche Verlauf der Messdaten recht gut zu allen Jahreszeiten erfasst.

Abbildung 4-1: Box-Whisker-Plots ausgewählter Metriken für die unkorrigierten CAMS-Ozonprognosen, Bezugsjahr 2016



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 4-2: Ozon-Zeitreihen von Mess- und Modelldaten an den Stationen mit dem niedrigsten (oben) bzw. dem höchsten (unten) Korrelationskoeffizienten, Bezugsjahr 2016



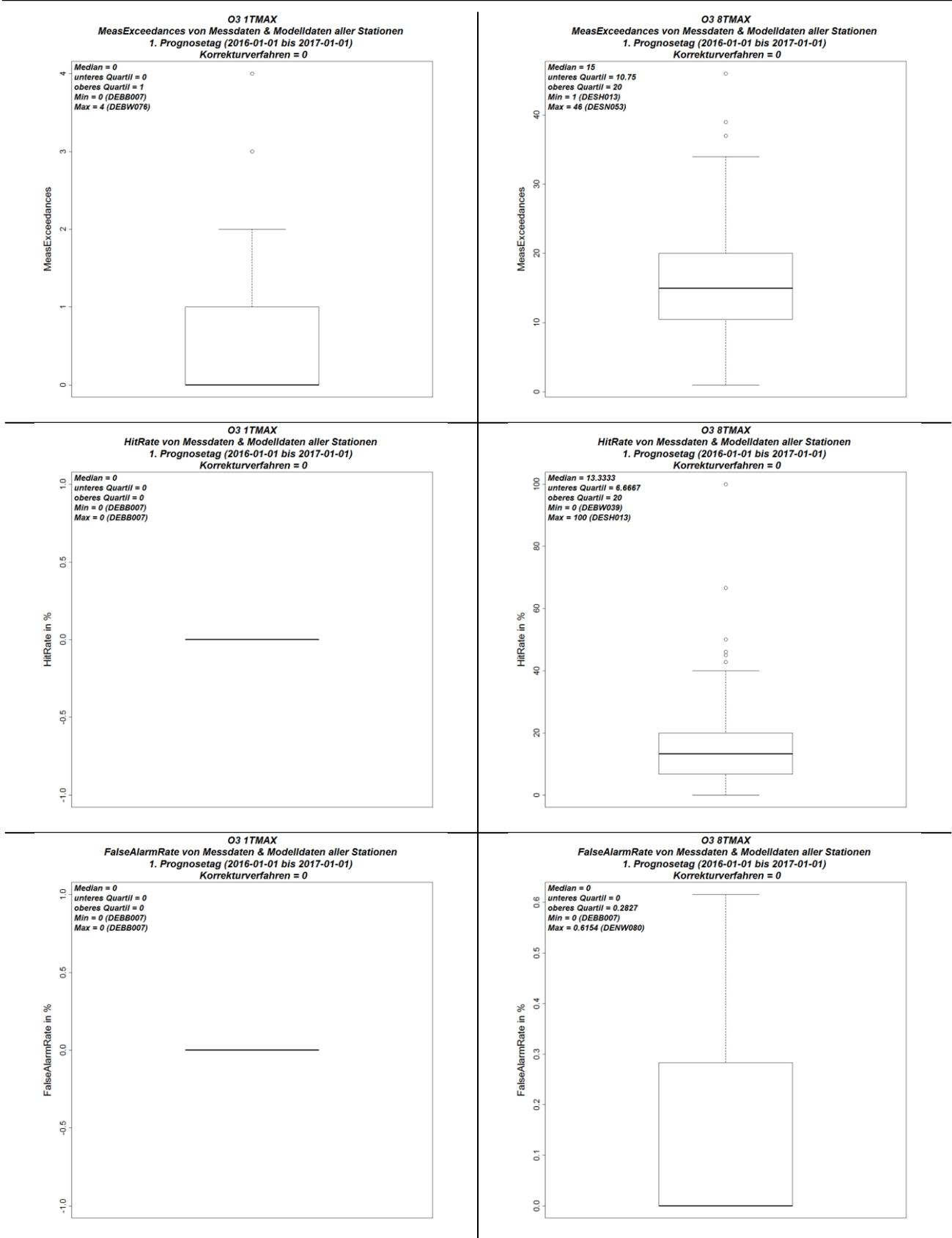
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 4-3 zeigt für den 1. Prognosetag die Box-Whisker-Plots der Anzahl der gemessenen Überschreitungen, der Hit Rate und der False Alarm Rate an allen Messorten. Als Schwellenwerte wurden gemäß 39. BImSchV die Informationsschwelle von $180 \mu\text{g}/\text{m}^3$ für die Einstundenmittelwerte (Abbildung 4-3, links) bzw. der Zielwert zum Schutz der menschlichen Gesundheit von $120 \mu\text{g}/\text{m}^3$ für die höchsten Achtstundenmittelwerte eines Tages (Abbildung 4-3, rechts) angesetzt.

Wie Abbildung 4-3 zu entnehmen ist, war die Anzahl der gemessenen Schwellenwertüberschreitungen im Jahr 2016 gering. Die Informationsschwelle von $180 \mu\text{g}/\text{m}^3$ wurde an einer Station (DEBW076) viermal überschritten, das obere Quartil über alle Stationen liegt bei einer Überschreitung und der Median bei keiner Überschreitung. Sowohl die Hit Rate als auch die False Alarm Rate für diese Überschreitungen beträgt an allen Stationen 0 %, d. h. keine der wenigen Überschreitungen wurde von den CAMS-Ensembledaten prognostiziert, allerdings auch keine Überschreitung fälschlicherweise angezeigt.

Der Zielwert zum Schutz der menschlichen Gesundheit von $120 \mu\text{g}/\text{m}^3$ bezogen auf den höchsten Achtstundenmittelwert eines Tages wurde im Jahr 2016 etwas häufiger überschritten als die Informationsschwelle. Im Maximum wurde er 46mal an einer Station (DESN053) überschritten, das obere Quartil über alle Stationen liegt bei 20 Überschreitungen und der Median bei 15 Überschreitungen. Es gab keine Station ohne Überschreitungen, das Minimum liegt bei einer Überschreitung (Station DESH013). Diese Überschreitung wurde vom CAMS-Ensemble auch korrekt erkannt (Hit Rate von 100 % für diese Station). Das obere Quartil der Hit Rate liegt über alle Stationen allerdings nur bei 20 % und der Median bei 13 %, so dass sich auch hier wie schon bei den Überschreitungen der Informationsschwelle bemerkbar macht, dass die CAMS-Prognosedaten für O_3 die gemessenen Maximalwerte unterschätzen. Trotz alledem gibt es eine geringe Anzahl fälschlicherweise prognostizierter Überschreitungen, die False Alarm Rate beträgt im Maximum ca. 0.6 % (DENW080) und im oberen Quartil knapp 0.3 %.

Abbildung 4-3: Box-Whisker-Plots Schwellenwertüberschreitungen für die unkorrigierten CAMS-Ozonprognosen, Bezugsjahr 2016



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

In Tabelle 4-2 bis Tabelle 4-4 sind die Mediane der für alle Stationen berechneten Metriken aus Tabelle 3-5 in Abhängigkeit der Vorhersagelänge (1., 2., 3. Prognosetag) aufgeführt. Tabelle 4-2 zeigt die Metriken für die Stundenmittelwerte. Hier wird noch einmal deutlich, dass die unkorrigierten CAMS-Prognosedaten die Messdaten im Mittel überschätzen, was sich auch in einem positiven Bias und einem positiven MFB niederschlägt. Die gemessenen Maximalwerte werden durch die CAMS-Daten im Median über alle Stationen um rund 20 % unterschätzt, entsprechend ist auch die Standardabweichung der CAMS-Daten geringer als die der Messdaten. Der Median der Korrelationskoeffizienten liegt bei ca. 0.8.

Bei den maximalen gleitenden Achtstundenmittelwerten eines Tages 8TMAX (Tabelle 4-3) liegt der Median der Mittelwerte der modellierten Daten immer noch höher als derjenige der Messdaten, die Differenz ist jedoch geringer als bei den 1SMW. Entsprechend sinken auch Bias und MFB gegenüber den Stundenmittelwerten. Die Maximalwerte werden auch bei den 8TMAX durch die CAMS-Daten unterschätzt, ebenso die Standardabweichung. In der Folge liegt der Median der Hit Rate über alle Stationen bei nur ca. 13 %. Der Median der Korrelationskoeffizienten liegt bei knapp 0.9.

Die maximalen Stundenmittelwerte eines Tages 1TMAX (Tabelle 4-4) liegen im Median der Mittelwerte nur noch wenig über den gemessenen Werten, entsprechend niedrig sind Bias und MFB für diese Größe. Die Unterschätzung der gemessenen Maximalwerte durch die CAMS-Daten entspricht derjenigen für die Stundenmittelwerte, entsprechend ist die Standardabweichung der CAMS-Daten auch bei den 1TMAX geringer als die der Messdaten. Wie bereits im Zusammenhang mit Abbildung 4-3 diskutiert, wird keine der wenigen gemessenen Überschreitungen der Informationsschwelle von $180 \mu\text{g}/\text{m}^3$ durch die CAMS-Daten erfasst. Der Median der Korrelationskoeffizienten liegt wie bei den 8TMAX bei knapp 0.9.

Im Vergleich der Metriken für die einzelnen Prognosetage zeigt sich, dass der Einfluss der Vorhersagelänge auf die Metriken bei O_3 eher gering ist. Der Mittelwert der Modelldaten bleibt für die drei Prognosetage fast konstant. Der Maximalwert der Modelldaten nimmt geringfügig ab, ebenso die Standardabweichung und der Korrelationskoeffizient. RMSE und FGE nehmen geringfügig zu. Das gilt sowohl für die Stundenmittelwerte (Tabelle 4-2) als auch für die abgeleiteten Größen 8TMAX (Tabelle 4-3) und 1TMAX (Tabelle 4-4). Durch die Abnahme der Maximalwerte der Modelldaten über die Vorhersagelänge sinkt auch der Median der Hit Rate für die 8TMAX geringfügig von ca. 13 % am 1. Prognosetag auf 10 % am 3. Prognosetag.

Tabelle 4-2: Metriken CAMS-Ozonprognose für die Stundenmittelwerte 1SMW, 1. bis 3. Prognosestag (jeweils Median über alle Stationen)

Metrik	1. Prognosestag	2. Prognosestag	3. Prognosestag
Min Obs [$\mu\text{g}/\text{m}^3$]	0.52	0.52	0.52
Min Mod [$\mu\text{g}/\text{m}^3$]	0.82	1.00	0.89
Mean Obs [$\mu\text{g}/\text{m}^3$]	44.72	44.80	45.00
Mean Mod [$\mu\text{g}/\text{m}^3$]	56.36	56.54	56.56
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.66	173.66
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	135.87	134.34
SD Obs [$\mu\text{g}/\text{m}^3$]	29.44	29.49	29.51
SD Mod [$\mu\text{g}/\text{m}^3$]	22.29	22.14	21.96
Korrelation	0.81	0.80	0.78
BIAS [$\mu\text{g}/\text{m}^3$]	10.17	10.16	10.00
RMSE [$\mu\text{g}/\text{m}^3$]	19.82	20.10	20.62
MFB [%]	38.27	38.78	37.74
FGE [%]	48.91	49.30	50.22

Tabelle 4-3: Metriken CAMS-Ozonprognose für die maximalen gleitenden 8-Stundenmittelwerte eines Tages 8TMAX, 1. bis 3. Prognosestag (jeweils Median über alle Stationen)

Metrik	1. Prognosestag	2. Prognosestag	3. Prognosestag
Min Obs [$\mu\text{g}/\text{m}^3$]	3.08	3.08	3.08
Min Mod [$\mu\text{g}/\text{m}^3$]	17.49	18.56	19.60
Mean Obs [$\mu\text{g}/\text{m}^3$]	64.59	64.80	65.06
Mean Mod [$\mu\text{g}/\text{m}^3$]	69.07	69.07	68.98
Max Obs [$\mu\text{g}/\text{m}^3$]	155.21	155.21	155.21
Max Mod [$\mu\text{g}/\text{m}^3$]	130.41	127.51	126.32
SD Obs [$\mu\text{g}/\text{m}^3$]	29.46	29.51	29.49
SD Mod [$\mu\text{g}/\text{m}^3$]	22.66	22.40	22.29
Korrelation	0.89	0.88	0.87
BIAS [$\mu\text{g}/\text{m}^3$]	3.63	3.36	3.10
RMSE [$\mu\text{g}/\text{m}^3$]	14.56	15.19	15.83
MFB [%]	12.68	12.31	11.79
FGE [%]	22.90	23.26	23.53
gemessene Überschreitungen	15 ¹	15 ¹	15 ¹
Hit Rate [%]	13.33 ¹	12.50 ¹	10.00 ¹
FAR [%]	0.00 ¹	0.00 ¹	0.00 ¹

¹: Schwellenwert 120 $\mu\text{g}/\text{m}^3$

Tabelle 4-4: Metriken CAMS-Ozonprognose für die maximalen Stundenmittelwerte eines Tages 1TMAX, 1. bis 3. Prognosetag (jeweils Median über alle Stationen)

Metrik	1. Prognosetag	2. Prognosetag	3. Prognosetag
Min Obs [$\mu\text{g}/\text{m}^3$]	4.79	4.79	4.79
Min Mod [$\mu\text{g}/\text{m}^3$]	20.36	21.55	20.94
Mean Obs [$\mu\text{g}/\text{m}^3$]	71.73	71.90	72.25
Mean Mod [$\mu\text{g}/\text{m}^3$]	72.56	72.58	72.62
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.66	173.66
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	135.87	134.34
SD Obs [$\mu\text{g}/\text{m}^3$]	30.63	30.71	30.62
SD Mod [$\mu\text{g}/\text{m}^3$]	23.19	23.03	22.84
Korrelation	0.88	0.88	0.86
BIAS [$\mu\text{g}/\text{m}^3$]	0.22	-0.04	-0.26
RMSE [$\mu\text{g}/\text{m}^3$]	15.23	15.71	16.51
MFB [%]	5.64	5.31	4.88
FGE [%]	20.13	20.51	20.79
gemessene Überschreitungen	0 ¹	0 ¹	0 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00 ¹	0.00 ¹	0.00 ¹

¹: Schwellenwert 180 $\mu\text{g}/\text{m}^3$ (Informationsschwelle)

Wird das DELTA-Tool im Benchmark-Modus auf die Messwerte und die CAMS-Rohdaten angewendet, so zeigt sich, dass für O_3 bereits die CAMS-Rohdaten für alle Stationen einen $\text{MQI} < 1$ aufweisen und damit das vom DELTA-Tool definierte Modellqualitätsziel erfüllen (Abbildung 4-4, links). Die Auswertung bezieht sich dabei ausschließlich auf die 8TMAX. Ein Teil der Stationen weist sogar einen $\text{MQI} < 0.5$ auf, d. h. der RMSE zwischen Messung und Modell an diesen Stationen ist kleiner als die Messunsicherheit. Das 90 %-Perzentil der MQI aller Stationen beträgt für den 1. Prognosetag 0.346.

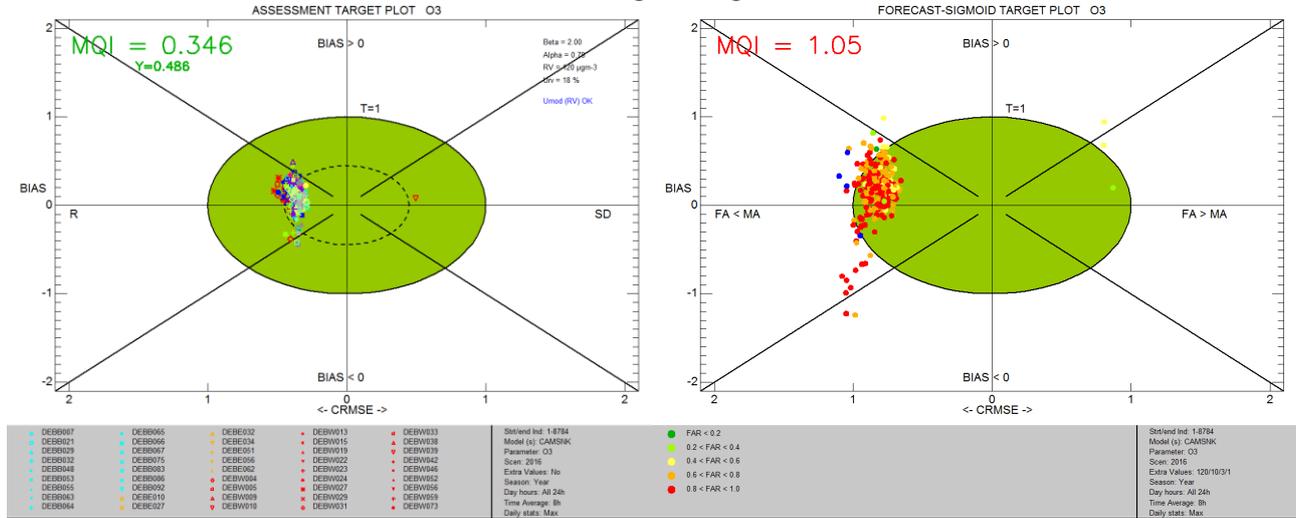
Im Forecast-Modus des DELTA-Tools zeigen die Target Plots die Güte der Modellprognosen im Vergleich zur Prognosegüte eines persistenten Modells. Die Auswertung bezieht sich wie im Benchmark-Modus auf die 8TMAX. Im Forecast-Modus ist der MQI der CAMS-Rohdaten (Abbildung 4-4, rechts) deutlich schlechter als im Benchmark-Modus. Eine ganze Reihe von Stationen weist einen $\text{MQI} > 1$ auf und wird damit schlechter modelliert als mit dem persistenten Modell. Das 90%-Perzentil der MQI aller Stationen liegt für den 1. Prognosetag bei 1.05 und ist damit ebenfalls > 1 . Die Aufteilung der Stationenpunkte entlang der x-Achse weist darauf hin, dass im Regelfall die Anzahl der falsch vorhergesagten Schwellenwertüberschreitungen (FA, false alarms) geringer ist als die Anzahl der nicht erkannten gemessenen Schwellenwertüberschreitungen (MA, missed alarms).

Werden der Benchmark- und der Forecast-Modus auf den 2. und den 3. Prognosetag der CAMS-Prognosen und die Messwerte angewendet, so zeigt sich in den Ergebnissen des Benchmark-Modus (Abbildung 4-4, links) ein ähnliches Bild wie zuvor bei der Diskussion der Metriken in Abhängigkeit der Vorhersagelänge. Der MQI steigt mit der Vorhersagelänge nur geringfügig von 0.346 auf 0.359, der Einfluss der Vorhersagelänge auf den MQI ist damit vernachlässigbar. Im Forecast-Modus zeigen die CAMS-Rohdaten eine umso bessere Performanz im Vergleich mit dem persistenten Modell, je länger die Vorhersage ist; der MQI sinkt von 1.05 am 1. Prognosetag auf 0.746 am 3. Prognosetag (Abbildung

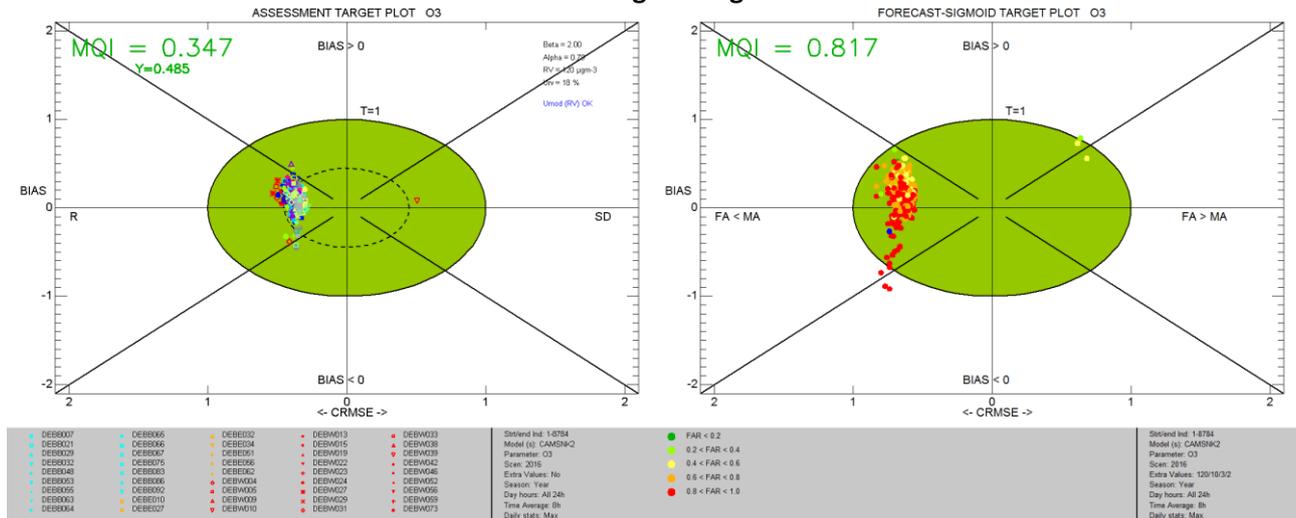
4-4, rechts). Demnach nimmt die Vorhersagegüte der CAMS-Rohdaten mit zunehmender Vorhersage-
länge weniger stark ab als die des persistenten Modells.

Abbildung 4-4: Assessment Target Plots (links) und Forecast-Sigmoid Target Plots (rechts) für CAMS-Rohdaten O₃, Bezugsjahr 2016, 1. bis 3. Prognostag

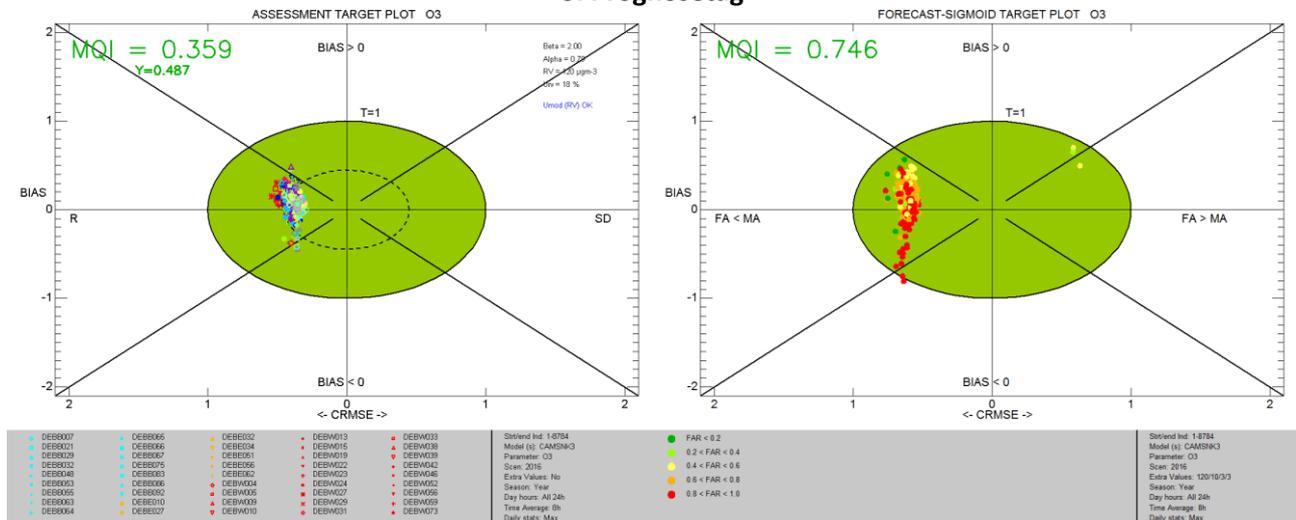
1. Prognostag



2. Prognostag



3. Prognostag



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

4.1.3 Saisonale Auswertung 2016

Um das Modellverhalten des CAMS-Ensembles auch saisonal erfassen zu können, wurden die Jahreszeitreihen der Mess- und der CAMS-Prognosedaten 2016 nach Jahreszeiten unterteilt, und anschließend wurden für jede Jahreszeit getrennt die Metriken nach Tabelle 3-5 bestimmt. Die zeitliche Aufteilung wurde wie folgt vorgenommen:

Frühling:	01.03.2016 01:00 UTC bis 01.06.2016 00:00 UTC
Sommer:	01.06.2016 01:00 UTC bis 01.09.2016 00:00 UTC
Herbst:	01.09.2016 01:00 UTC bis 01.12.2016 00:00 UTC
Winter:	01.01.2016 01:00 UTC bis 01.03.2016 00:00 UTC und 01.12.2016 01:00 UTC bis 01.01.2017 00:00 UTC

In Tabelle 4-5 bis Tabelle 4-7 sind die Mediane der für alle Stationen berechneten Metriken für das Jahr 2016 und für die vier Jahreszeiten aufgeführt. Fett hervorgehobene Zahlen markieren jeweils den Wert, der die größte Übereinstimmung zwischen Messung und Modell für diese Größe anzeigt. Dabei werden Bias und RMSE nicht markiert, da sie aufgrund ihrer Skalenabhängigkeit für den Vergleich der Modellperformanz zwischen verschiedenen Datensätzen nicht geeignet sind.

Tabelle 4-5 zeigt die Metriken für die Stundenmittelwerte. Minimal-, Mittel- und Maximalwerte sowie die Standardabweichung der Messdaten werden demnach im Median am besten im Winter durch die Modelldaten erfasst. Die größten Differenzen zwischen Mess- und Modelldaten zeigen sich für diese Metriken im Sommer, bis auf den Mittelwert, der im Herbst am schlechtesten vom Modell wiedergegeben wird. Die Korrelation zwischen Mess- und CAMS-Daten hingegen ist im Winter sowie im Frühling mit 0.76 am niedrigsten, hier liegen Sommer und Herbst mit 0.81 bzw. 0.83 deutlich höher. MFB und FGE sind im Median im Frühling am niedrigsten und im Herbst am höchsten.

Für die höchsten gleitenden 8-Stundenmittelwerte (Tabelle 4-6) ergibt sich ein etwas anderes Bild. Minimalwerte und Standardabweichung der Messdaten werden hier im Median am Besten im Frühling getroffen, die Mittelwerte im Sommer und die Maximalwerte wie bei den Stundenmittelwerten im Winter. Die Korrelation zwischen Mess- und CAMS-Daten ist mit 0.91 im Herbst am höchsten und mit 0.79 im Winter am niedrigsten. Die niedrigsten Werte für MFB und FGE ergeben sich wie bei den Stundenmittelwerten im Frühling; der MFB ist im Herbst am höchsten und der FGE im Winter. Ähnlich sieht es bei den maximalen Stundenmittelwerten eines Tages 1TMAX (Tabelle 4-7) aus, nur dass hier der Mittelwert der Messwerte im Median am besten im Frühling und nicht im Sommer durch das Modell erfasst wird.

Zusammenfassend lässt sich sagen, dass sowohl für die Stundenmittelwerte 1SMW als auch für die abgeleiteten Größen 8TMAX und 1TMAX die zur Erfassung der Überschreitungen relevanten Maximalwerte im Median insbesondere im Sommer durch das Modell unterschätzt werden. Die Korrelation, d. h. die Erfassung des zeitlichen Verlaufs, ist im Herbst am besten, gefolgt vom Sommer. MFB und FGE sind im Frühling am niedrigsten, wiederum gefolgt vom Sommer.

Tabelle 4-5: Metriken O₃, Stundenmittelwerte 1SMW für den 1. Prognosetag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	Jahr	Frühling	Sommer	Herbst	Winter
Min Obs [$\mu\text{g}/\text{m}^3$]	0.52	1.25	1.19	0.60	0.60
Min Mod [$\mu\text{g}/\text{m}^3$]	0.82	11.65	31.79	1.88	1.11
Mean Obs [$\mu\text{g}/\text{m}^3$]	44.72	57.60	55.96	33.40	31.75
Mean Mod [$\mu\text{g}/\text{m}^3$]	56.36	68.62	70.48	48.82	38.92
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	135.97	173.66	155.00	83.28
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	121.34	139.11	124.94	74.75
SD Obs [$\mu\text{g}/\text{m}^3$]	29.44	26.67	29.45	28.73	21.96
SD Mod [$\mu\text{g}/\text{m}^3$]	22.29	18.39	16.72	21.02	15.15
Korrelation	0.81	0.76	0.81	0.83	0.76
BIAS [$\mu\text{g}/\text{m}^3$]	10.17	8.73	13.36	13.50	5.20
RMSE [$\mu\text{g}/\text{m}^3$]	19.82	19.19	23.05	20.63	15.77
MFB [%]	38.27	23.11	34.41	57.70	36.78
FGE [%]	48.91	32.43	39.76	64.89	55.79

Tabelle 4-6: Metriken O₃, maximale gleitende 8-Stundenmittelwerte eines Tages 8TMAX für den 1. Prognosetag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	Jahr	Frühling	Sommer	Herbst	Winter
Min Obs [$\mu\text{g}/\text{m}^3$]	3.08	37.03	39.02	4.42	3.83
Min Mod [$\mu\text{g}/\text{m}^3$]	17.49	44.40	65.36	22.35	18.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	64.59	79.18	83.36	52.94	43.45
Mean Mod [$\mu\text{g}/\text{m}^3$]	69.07	82.28	85.76	61.03	47.31
Max Obs [$\mu\text{g}/\text{m}^3$]	155.21	130.17	154.62	140.54	78.44
Max Mod [$\mu\text{g}/\text{m}^3$]	130.41	115.83	130.41	115.72	71.12
SD Obs [$\mu\text{g}/\text{m}^3$]	29.46	19.57	22.29	31.54	20.13
SD Mod [$\mu\text{g}/\text{m}^3$]	22.66	15.20	13.07	21.84	12.69
Korrelation	0.89	0.81	0.81	0.91	0.79
BIAS [$\mu\text{g}/\text{m}^3$]	3.63	2.13	2.27	7.22	2.55
RMSE [$\mu\text{g}/\text{m}^3$]	14.56	12.22	15.68	16.52	13.57
MFB [%]	12.68	4.04	5.23	24.71	16.43
FGE [%]	22.90	12.15	14.90	31.80	32.22
gemessene Überschreitungen	15 ¹	4 ¹	7 ¹	4 ¹	0 ¹
Hit Rate [%]	13.33 ¹	0.00 ¹	28.57¹	0.00 ¹	0.00 ¹
FAR [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹

¹: Schwellenwert 120 $\mu\text{g}/\text{m}^3$

Tabelle 4-7: Metriken O₃, maximale Stundenmittelwerte eines Tages 1TMAX für den 1. Prognostag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	Jahr	Frühling	Sommer	Herbst	Winter
Min Obs [$\mu\text{g}/\text{m}^3$]	4.79	41.59	40.60	6.79	5.82
Min Mod [$\mu\text{g}/\text{m}^3$]	20.36	48.39	66.91	25.02	21.10
Mean Obs [$\mu\text{g}/\text{m}^3$]	71.73	85.89	91.12	61.51	49.47
Mean Mod [$\mu\text{g}/\text{m}^3$]	72.56	86.31	88.85	64.73	50.41
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	135.94	173.66	155.00	83.24
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	121.41	139.11	124.87	74.75
SD Obs [$\mu\text{g}/\text{m}^3$]	30.63	19.63	25.05	33.97	20.26
SD Mod [$\mu\text{g}/\text{m}^3$]	23.19	15.62	14.29	22.72	12.44
Korrelation	0.88	0.80	0.83	0.91	0.75
BIAS [$\mu\text{g}/\text{m}^3$]	0.22	-0.19	-1.72	2.95	-0.19
RMSE [$\mu\text{g}/\text{m}^3$]	15.23	12.18	16.42	16.52	14.12
MFB [%]	5.64	0.59	0.71	13.45	8.01
FGE [%]	20.13	11.16	14.16	25.75	28.30
gemessene Überschreitungen	0 ¹				
Hit Rate [%]	0.00 ¹				
FAR [%]	0.00 ¹				

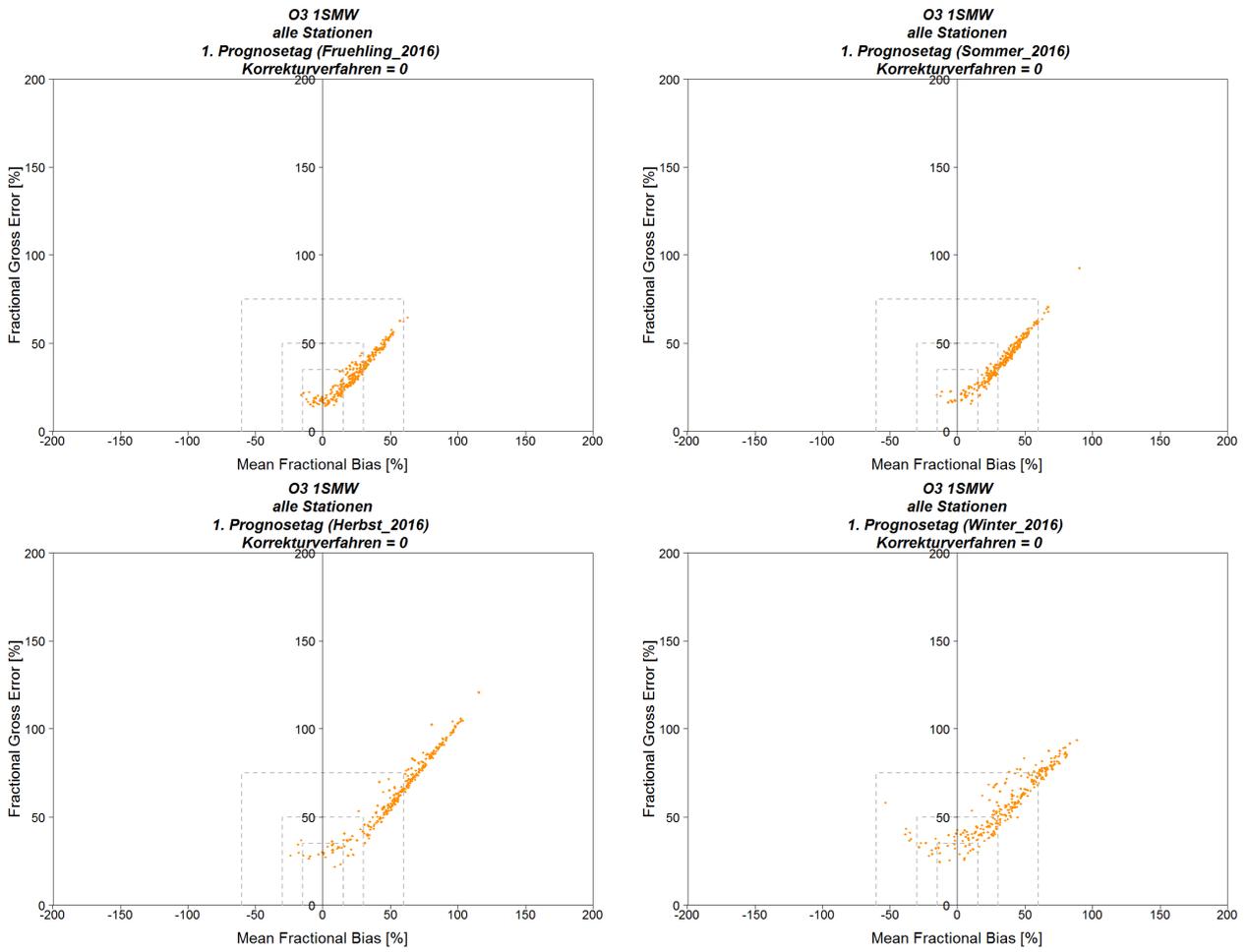
¹: Schwellenwert 180 $\mu\text{g}/\text{m}^3$ (Informationsschwelle)

MFB und FGE sind in Abbildung 4-5 bis Abbildung 4-7 für die Stundenmittelwerte 1SMW und für die abgeleiteten Größen 8TMAX und 1TMAX als Soccer-Plots über alle Stationen für jede der vier Jahreszeiten dargestellt. Die 1SMW zeigen einen überwiegend positiven Bias. Die höchsten Werte für MFB und FGE ergeben sich, wie schon in Tabelle 4-5 dargestellt, im Herbst und im Winter. Dabei fällt auf, dass die Stationspunkte im Sommer eine sehr deutliche lineare Gruppierung zeigen, die auf einen systematischen Fehler in den CAMS-Prognosedaten hinweist. Diese lineare Gruppierung ist im Frühling und Herbst immer noch zu erkennen, löst sich aber bereits teilweise auf. Im Winter deutet dir nur noch schwach lineare Gruppierung darauf hin, dass zu dieser Jahreszeit eher zufällige Fehler in den CAMS-Prognosedaten dominieren.

Für die 8TMAX (Abbildung 4-6) sind MFB und FGE in Frühling und Sommer relativ gering, im Herbst und Winter deutlich höher, aber geringer als für die 1SMW. Es zeigt sich weiterhin ein überwiegend positiver Bias. Hinsichtlich des Einflusses systematischer oder zufälliger Fehler lässt sich bei den 8TMAX aufgrund der Dichte der Punktwolke in Frühling und Sommer keine Aussage treffen. Im Herbst dominieren eher systematische und im Winter wie bei den 1SMW eher zufällige Fehler die CAMS-Prognosen.

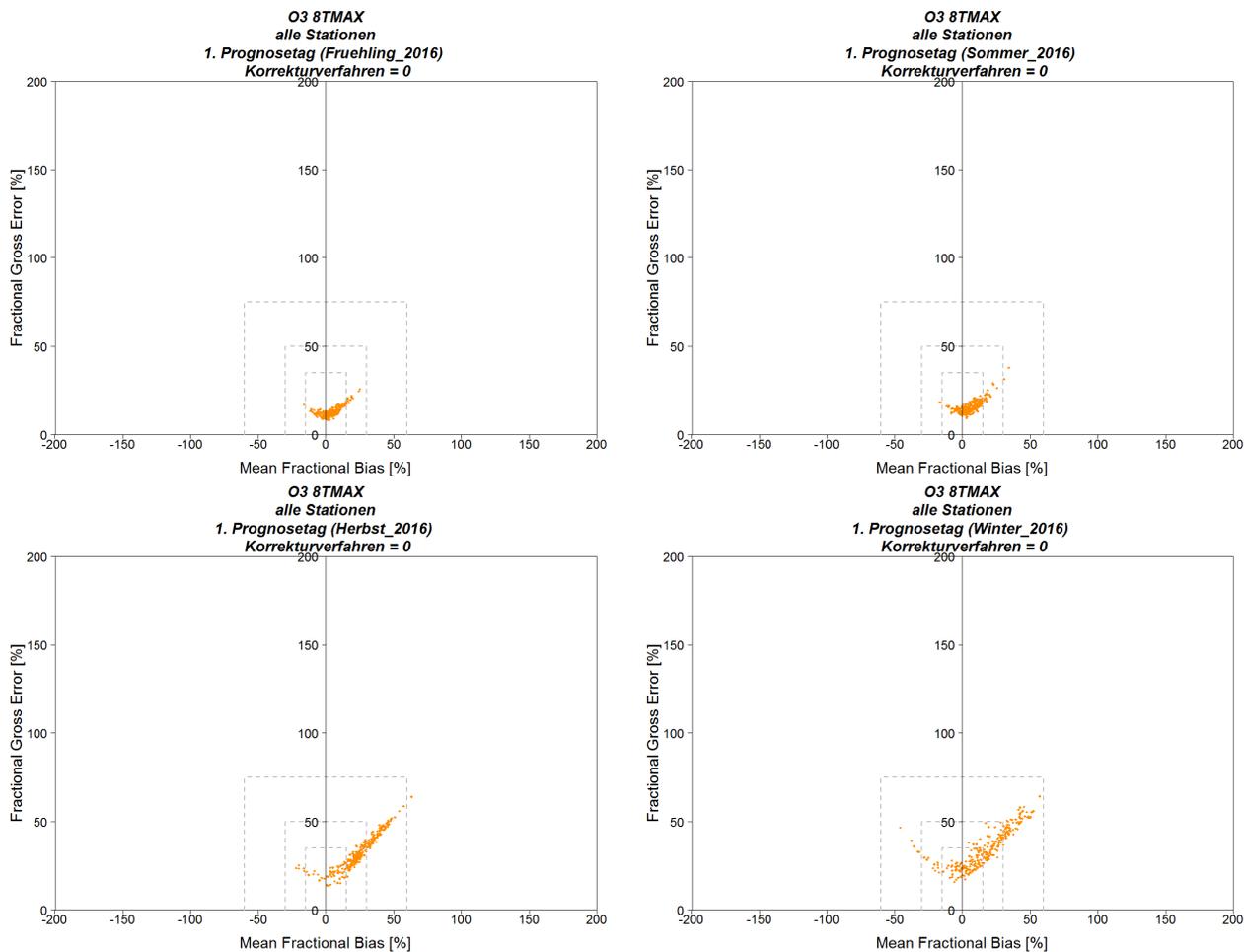
Für die 1TMAX (Abbildung 4-7) gilt im Grundsatz das gleiche wie für die 8TMAX. Allerdings treten hier deutlich mehr Stationen mit negativem Bias auf, so dass im Frühling und Sommer der Median des MFB nahe bei 0 % liegt. Zudem erscheint der Einfluss zufälliger Fehler auf die CAMS-Prognosen gegenüber den 8TMAX im Herbst zuzunehmen.

Abbildung 4-5: Soccer-Plots O₃ für die Stundenmittelwerte 1SMW, 1. Prognosestag, Bezugsjahr 2016, unterteilt nach Jahreszeiten



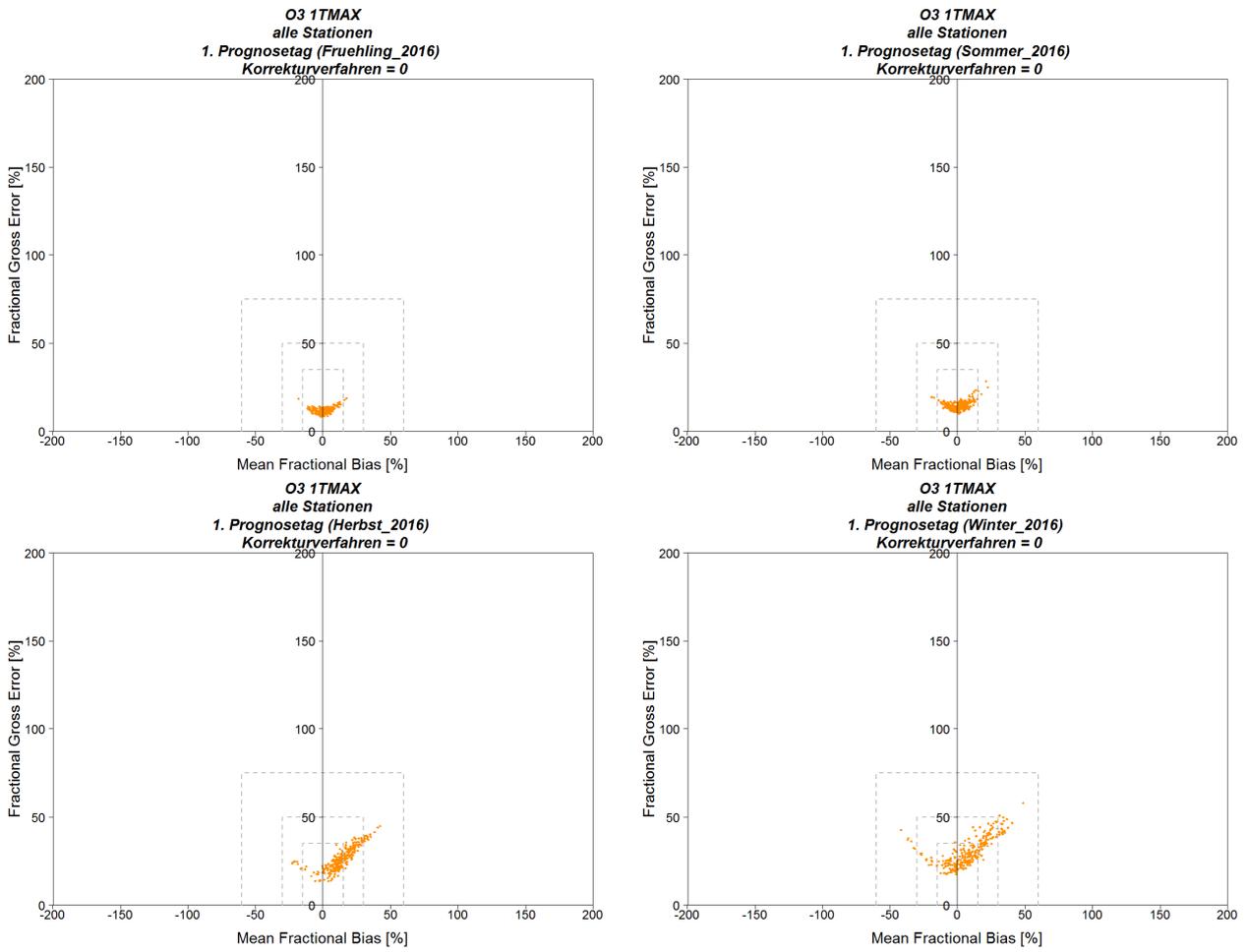
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 4-6: Soccer-Plots O₃ für die maximalen gleitenden 8-Stundenmittelwerte eines Tages 8TMAX, 1. Prognostetag, Bezugsjahr 2016, unterteilt nach Jahreszeiten



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 4-7: Soccer-Plots O₃ für die maximalen Stundenmittelwerte eines Tages 1TMAX, 1. Prognose- tag, Bezugsjahr 2016, unterteilt nach Jahreszeiten



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

4.2 PM10

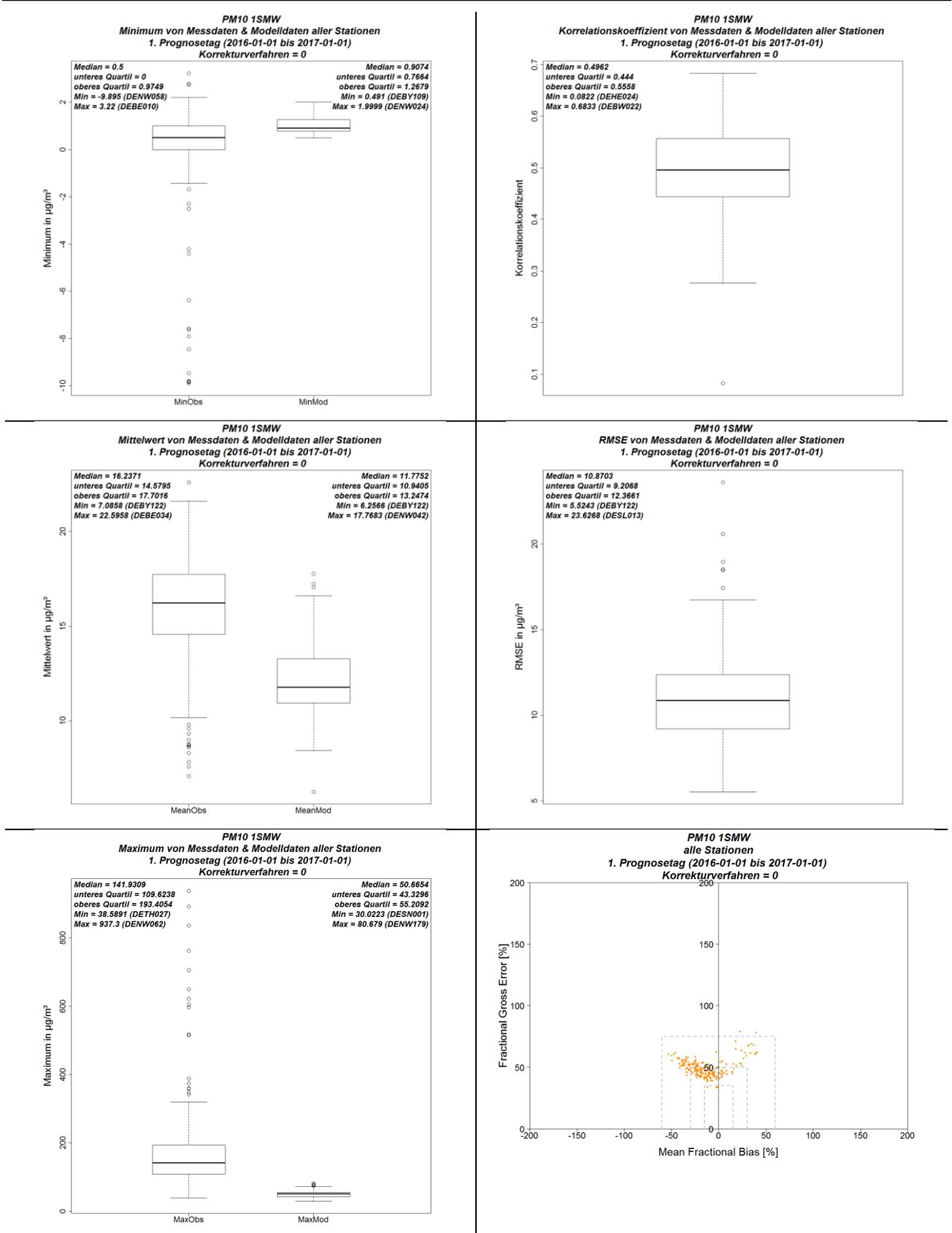
4.2.1 Jahreszeitreihe 2016

Für PM10 wurden die Jahreszeitreihen 2016 der CAMS-Prognosedaten analog zu O₃ (Abschnitt 4.1.2) in Abhängigkeit vom Prognosetag gebildet und ausgewertet. Abbildung 4-8 zeigt für den 1. Prognosetag die Box-Whisker-Plots der Minima, Mittelwerte, Maxima, Korrelationskoeffizienten und RMSE an allen Messorten sowie den zugehörigen Soccer-Plot für die Stundenmittelwerte. Die Mediane der Minima liegen für Messwerte und Modelldaten wie bei O₃ in der gleichen Größenordnung, auffällig sind auch hier die vielen negativen Messwerte. Diese lassen sich analog zu O₃ im Wesentlichen auf Messstandorte in Nordrhein-Westfalen zurückverfolgen und entsprechen etwa der dort festgelegten negativen Nachweisgrenze. Auch diese negativen Messwerte wurden für die Korrekturverfahren zu Null gesetzt (Abschnitt 3.3).

Der Box-Whisker-Plot der Mittelwerte zeigt, dass die CAMS-Prognosedaten die gemessenen PM10-Werte im Gegensatz zu O₃ bereits im Median deutlich unterschätzen. Die gemessenen Maximalwerte werden von den Modellprognosen bei PM10 noch stärker unterschätzt als bei O₃, die modellierten Maximalwerte erreichen im Median nur ein gutes Drittel der gemessenen Maximalwerte. Der Korrelationskoeffizient zwischen Messung und Modell liegt für PM10 zwischen 0.08 (Station DEHE024: Witzenhäuser/Wald) und 0.68 (Station DEBW022: Kehl) mit dem Median bei knapp 0.5. Der Korrelationskoeffizient ist damit für PM10 deutlich niedriger als für O₃. Der RMSE beträgt im Median knapp 11 µg/m³ und damit ca. zwei Drittel des Medians der Mittelwerte der Messdaten von 16.2 µg/m³. Im Soccer-Plot zeigen die meisten Messorte einen negativen Bias, eine geringere Anzahl Messorte weist auch einen positiven Bias auf. Die nur schwach lineare Gruppierung der Punkte der einzelnen Stationen deutet darauf hin, dass bei PM10 eher zufällige Fehler in den CAMS-Prognosedaten dominieren.

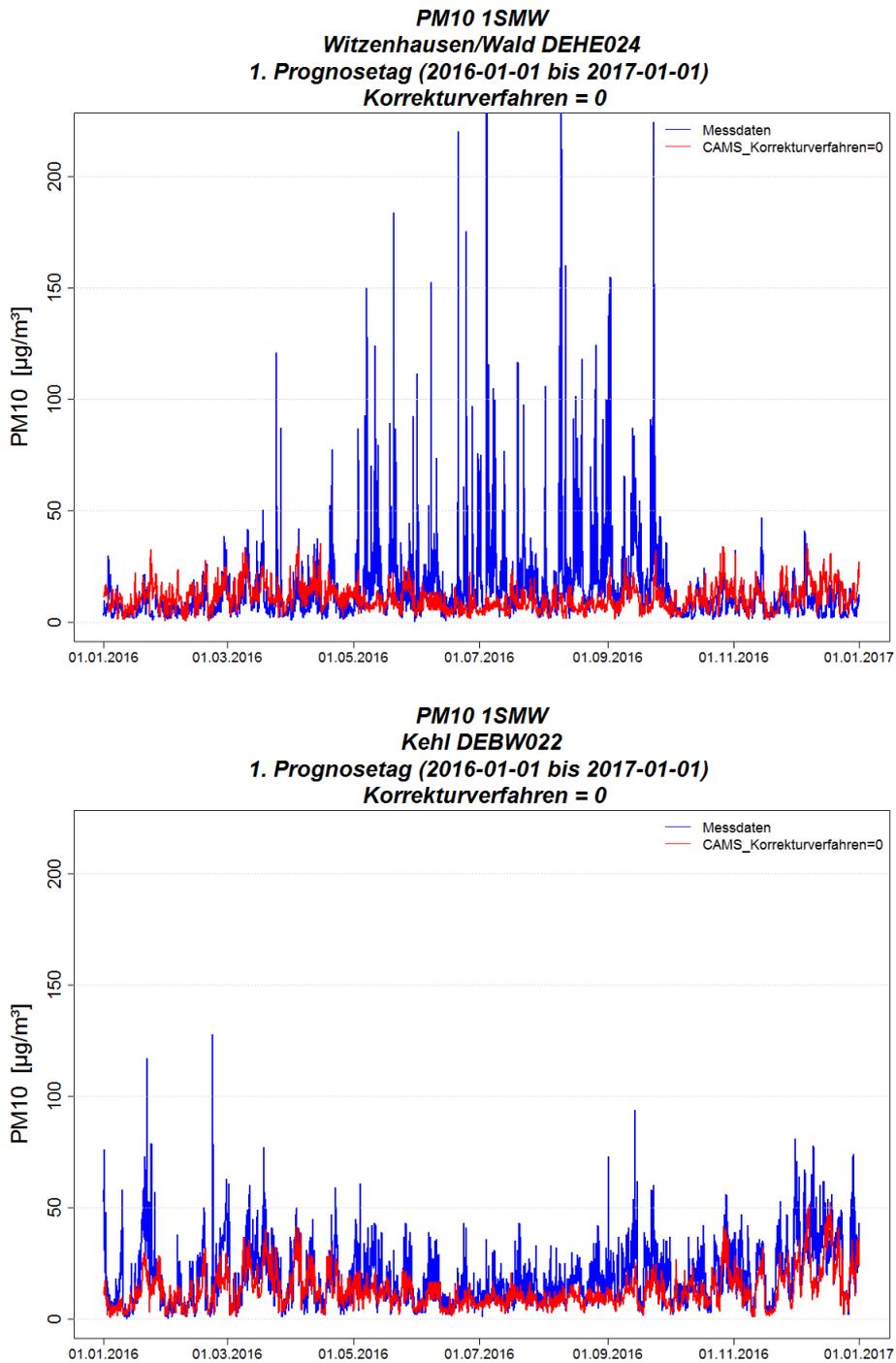
In Abbildung 4-9 sind die Zeitreihen der Stundenmittelwerte für die Station mit dem niedrigsten (DEHE024) und mit dem höchsten (DEBW022) Korrelationskoeffizienten dargestellt. An der Station DEHE024 unterschätzen die CAMS-Prognosen die gemessenen PM10-Werte bei Weitem, es wurden, insbesondere in dem Zeitraum von Ende März 2016 bis Ende September 2016, Werte von über 200 µg/m³ PM10 gemessen. Laut Stationsdatenbank des Umweltbundesamtes wurde neben der Messstation ab Mitte Dezember 2015 bis Januar 2017 ein Windpark gebaut, möglicherweise ist hier der Grund für die ungewöhnlich hohen PM10-Werte (DEHE024 ist als ländliche Hintergrundstation klassifiziert) zu finden. Zudem wurden die Emissionen der Baustelle in den CAMS-Modellläufen sicher nicht berücksichtigt, was die große Diskrepanz zwischen Mess- und Modellwerten erklären könnte. An der Station DEBW022 werden die gemessenen Maximalwerte im Regelfall ebenfalls nicht durch die CAMS-Prognosen erreicht, allerdings sind die Differenzen zwischen Messung und Modell deutlich geringer als an der Station DEHE024. Der zeitliche Verlauf der Messungen wird von den CAMS-Prognosen zumindest teilweise erfasst, wenn auch die Maxima nicht getroffen werden. So wurden beispielsweise Episoden im Januar, im März und April sowie im November und Dezember grundsätzlich wiedergegeben.

Abbildung 4-8: Box-Whisker-Plots ausgewählter Metriken für die unkorrigierten CAMS-PM10-Prognosen, Bezugsjahr 2016



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 4-9: PM10-Zeitreihen von Mess- und Modelldaten an den Stationen mit dem niedrigsten (oben) bzw. dem höchsten (unten) Korrelationskoeffizienten, Bezugsjahr 2016

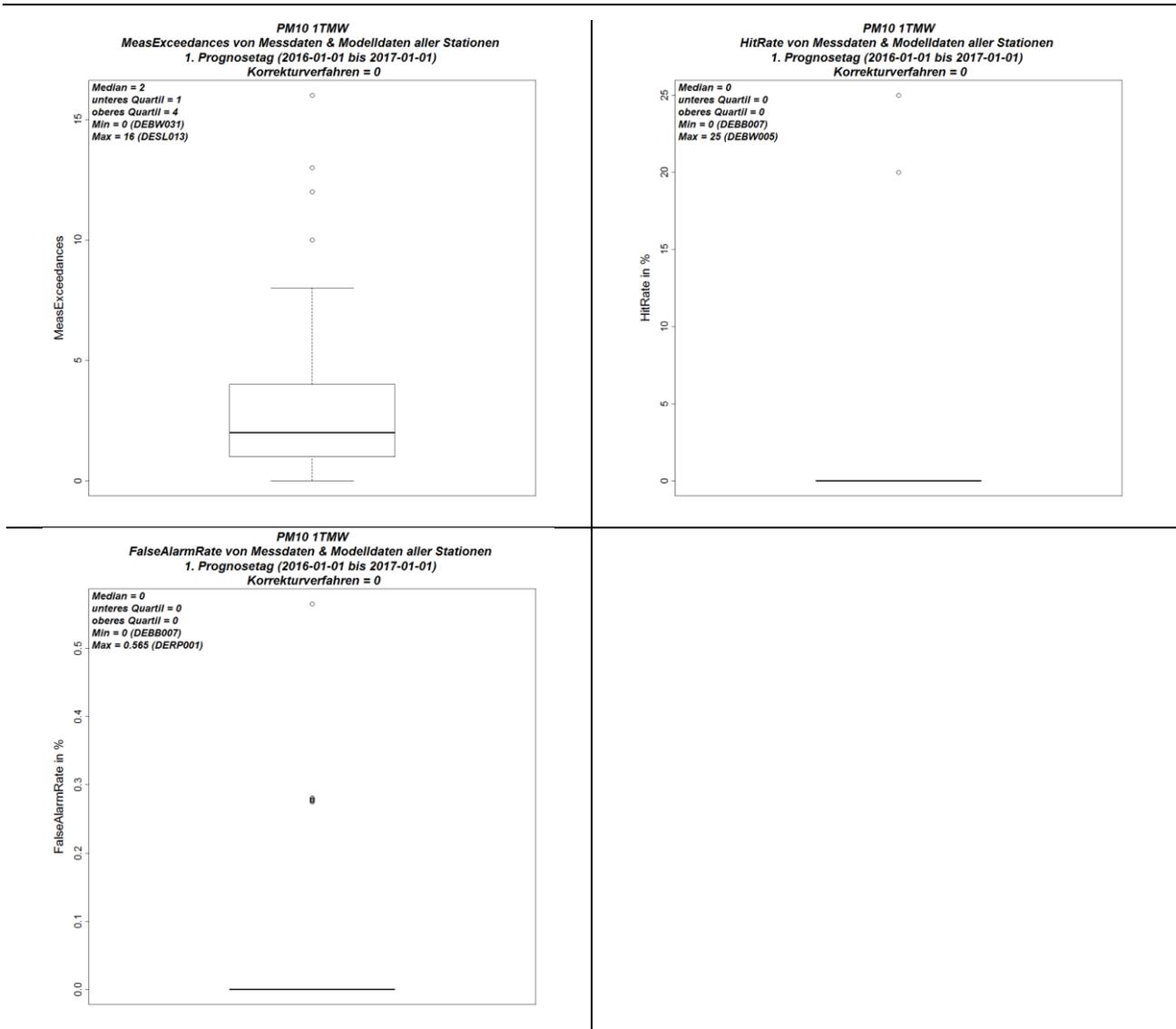


Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 4-10 zeigt für den 1. Prognosezeitraum die Box-Whisker-Plots der Anzahl der gemessenen Überschreitungen, der Hit Rate und der False Alarm Rate an allen Messorten. Als Schwellenwert wurde gemäß 39. BImSchV der zum Schutz der menschlichen Gesundheit definierte, über den Tag gemittelte Immissionsgrenzwert von $50 \mu\text{g}/\text{m}^3$ angesetzt.

Wie Abbildung 4-10 zu entnehmen ist, war die Anzahl der gemessenen Schwellenwertüberschreitungen im Jahr 2016 ebenso wie für O₃ gering. Der Tagesmittelwert von 50 µg/m³ wurde an einer Station (DESL013) 16mal überschritten, das obere Quartil über alle Stationen liegt bei vier Überschreitungen und der Median bei zwei Überschreitungen. Grund für die geringe Anzahl der Überschreitungen ist auch, dass hier keine Verkehrsstationen betrachtet werden. Die Hit Rate beträgt im Maximum 25% (DEBW005), an einer weiteren Station 20 % (DERP003) und an allen anderen Stationen 0 %. Damit wurde an der Station DEBW005 eine von vier und an der Station DERP003 eine von fünf Überschreitungen durch die CAMS-Daten prognostiziert, alle anderen Überschreitungen wurden nicht erkannt. Grund für die geringe Anzahl der korrekt prognostizierten Überschreitungen ist in noch stärkerem Maße als bei O₃ die deutliche Unterschätzung der Messwerte durch die CAMS-Daten. Trotzdem gibt es auch bei PM₁₀ eine geringe Anzahl fälschlicherweise prognostizierter Überschreitungen. Die False Alarm Rate beträgt im Maximum knapp 0.6 % (DERP001), aber bereits im oberen Quartil 0 %.

Abbildung 4-10: Box-Whisker-Plots Schwellenwertüberschreitungen für die unkorrigierten CAMS-PM₁₀-Prognosen, Bezugsjahr 2016



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

In Tabelle 4-8 und Tabelle 4-9 sind die Mediane der für alle Stationen berechneten Metriken aus Tabelle 3-5 in Abhängigkeit der Vorhersagelänge (1., 2., 3. Prognosetag) aufgeführt. Tabelle 4-8 zeigt die

Metriken für die Stundenmittelwerte. Hier wird noch einmal deutlich, dass die CAMS-Rohdaten die Messdaten im Mittel unterschätzen, was sich auch in einem negativen Bias und einem negativen MFB niederschlägt. Die gemessenen Maximalwerte werden durch die CAMS-Daten im Median über alle Stationen um knapp zwei Drittel unterschätzt, entsprechend ist auch die Standardabweichung der CAMS-Daten deutlich geringer als die der Messdaten. Der Median der Korrelationskoeffizienten liegt bei 0,5, im Gegensatz zu ca. 0,8 bei O₃.

Bei den Tagesmittelwerten 1TMW (Tabelle 4-9) liegt der Median der Mittelwerte der modellierten Daten ebenfalls niedriger als derjenige der Messdaten. Der Bias ist ähnlich hoch und der MFB ist höher als bei den Stundenmittelwerten. Die Maximalwerte der 1TMW werden durch die CAMS-Daten im Median um die Hälfte unterschätzt, die Standardabweichung der Modelldaten liegt rund 40 % niedriger als die der Messdaten. In der Folge liegt der Median der Hit Rate über alle Stationen bei 0 %. Der Median der Korrelationskoeffizienten liegt bei knapp 0,6.

Im Vergleich der Metriken für die einzelnen Prognosestage zeigt sich, dass der Einfluss der Vorhersagelänge auf die Metriken wie bei O₃ eher gering ist. Der Mittelwert der Modelldaten bleibt für die drei Prognosestage fast konstant. Der Median von Standardabweichung und Korrelationskoeffizient nimmt geringfügig ab. RMSE und FGE nehmen geringfügig zu. Das gilt sowohl für die Stundenmittelwerte (Tabelle 4-8) als auch für die abgeleiteten Tagesmittelwerte (Tabelle 4-9). Der Median der modellierten Maximalwerte nimmt bei den Stundenmittelwerten relativ betrachtet etwas stärker ab als bei O₃, nämlich knapp 11 %, bezogen auf den 1. Prognosestag. Bei den Tagesmittelwerten nehmen die modellierten Maximalwerte über die Vorhersagelänge um rund 7 % ab. Der Median von Hit Rate und False Alarm Rate liegt für alle drei Prognosestage konstant bei 0 %.

Tabelle 4-8: Metriken CAMS-PM10-Prognose für die Stundenmittelwerte 1SMW, 1. bis 3. Prognosestag (jeweils Median über alle Stationen)

Metrik	1. Prognosestag	2. Prognosestag	3. Prognosestag
Min Obs [$\mu\text{g}/\text{m}^3$]	0.50	0.50	0.50
Min Mod [$\mu\text{g}/\text{m}^3$]	0.91	1.01	1.06
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.24	16.20	16.13
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.78	11.64	11.62
Max Obs [$\mu\text{g}/\text{m}^3$]	141.93	128.45	128.45
Max Mod [$\mu\text{g}/\text{m}^3$]	50.67	48.41	45.11
SD Obs [$\mu\text{g}/\text{m}^3$]	11.57	11.37	11.36
SD Mod [$\mu\text{g}/\text{m}^3$]	6.29	6.13	5.95
Korrelation	0.50	0.49	0.46
BIAS [$\mu\text{g}/\text{m}^3$]	-3.79	-3.92	-3.82
RMSE [$\mu\text{g}/\text{m}^3$]	10.87	10.88	10.99
MFB [%]	-16.34	-16.94	-16.04
FGE [%]	48.65	49.07	49.73

Tabelle 4-9: Metriken CAMS-PM10-Prognose für die Tagesmittelwerte 1TMW, 1. bis 3. Prognosetag (jeweils Median über alle Stationen)

Metrik	1. Prognosetag	2. Prognosetag	3. Prognosetag
Min Obs [$\mu\text{g}/\text{m}^3$]	2.02	2.02	2.02
Min Mod [$\mu\text{g}/\text{m}^3$]	3.10	2.92	2.83
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.27	16.21	16.13
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.77	11.64	11.63
Max Obs [$\mu\text{g}/\text{m}^3$]	72.84	67.31	67.31
Max Mod [$\mu\text{g}/\text{m}^3$]	35.88	34.57	33.18
SD Obs [$\mu\text{g}/\text{m}^3$]	9.20	9.09	9.10
SD Mod [$\mu\text{g}/\text{m}^3$]	5.48	5.29	5.15
Korrelation	0.58	0.57	0.55
BIAS [$\mu\text{g}/\text{m}^3$]	-3.80	-3.93	-3.80
RMSE [$\mu\text{g}/\text{m}^3$]	8.29	8.30	8.35
MFB [%]	-21.50	-22.51	-21.40
FGE [%]	39.01	39.22	39.46
gemessene Überschreitungen	2 ¹	2 ¹	2 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00 ¹	0.00 ¹	0.00 ¹

¹: Schwellenwert 50 $\mu\text{g}/\text{m}^3$

Wird das DELTA-Tool im Benchmark-Modus auf die Messwerte und die CAMS-Rohdaten angewendet, so zeigt sich, dass für PM10 der MQI einzelner Stationen größer 1 ist, die Modellprognosen also dort nicht die Qualitätskriterien des DELTA-Tools erfüllen (Abbildung 4-11, links). Die Auswertung bezieht sich dabei ausschließlich auf die 1TMW. Das 90%-Perzentil der MQI aller Stationen liegt für die CAMS-Rohdaten mit 0.947 immer noch knapp unter 1. Damit bewertet das DELTA-Tool die CAMS-PM10-Prognosen wesentlich schlechter als die CAMS-O₃-Prognosen (Abbildung 4-4, links).

Im Forecast-Modus ist der MQI der CAMS-Rohdaten wie schon bei O₃ deutlich schlechter als im Benchmark-Modus (Abbildung 4-11, rechts). Die meisten Stationen weisen einen MQI > 1 auf und werden damit schlechter modelliert als mit dem persistenten Modell. Das 90%-Perzentil der MQI aller Stationen liegt mit 1.29 deutlich über 1. Die Aufteilung der Stationspunkte entlang der x-Achse weist darauf hin, dass im Regelfall die Anzahl der falsch vorhergesagten Schwellenwertüberschreitungen (FA, false alarms) geringer ist als die Anzahl der nicht erkannten gemessenen Schwellenwertüberschreitungen (MA, missed alarms).

Werden der Benchmark- und der Forecast-Modus auf den 2. und den 3. Prognosetag der CAMS-Prognosen und die Messwerte angewendet, so zeigt sich in den Ergebnissen des Benchmark-Modus (Abbildung 4-11, links) ein ähnliches Bild wie zuvor bei der Diskussion der Metriken in Abhängigkeit der Vorhersagelänge. Der MQI steigt mit der Vorhersagelänge nur geringfügig von 0.947 auf 0.972, der Einfluss der Vorhersagelänge auf den MQI ist damit auch bei PM10 vernachlässigbar. Im Forecast-Modus zeigen die CAMS-Rohdaten eine umso bessere Performanz im Vergleich mit dem persistenten Modell, je länger die Vorhersage ist; der MQI sinkt von 1.29 am 1. Prognosetag auf 0.907 am 3. Prognosetag (Abbildung 4-11, rechts).

4.2.2 Saisonale Auswertung 2016

Analog zu O₃ wurden auch für PM₁₀ die Jahreszeitreihen 2016 der CAMS-Prognosen nach Jahreszeiten getrennt untersucht. In Tabelle 4-10 und Tabelle 4-11 sind die Mediane der für alle Stationen berechneten Metriken für das Jahr 2016 und für die vier Jahreszeiten aufgeführt.

Tabelle 4-10 zeigt die Metriken für die Stundenmittelwerte. Mittel- und Maximalwerte sowie die Standardabweichung der Messdaten werden demnach im Median am besten im Frühling erfasst. Die größten Differenzen zwischen Mess- und Modelldaten zeigen sich für diese Metriken im Winter, bis auf den Mittelwert, der im Sommer am schlechtesten vom Modell wiedergegeben wird. Die Korrelation zwischen Mess- und CAMS-Daten ist im Herbst mit 0.58 am höchsten und im Sommer mit 0.43 am niedrigsten. MFB und FGE sind im Median im Frühling am niedrigsten und im Sommer am höchsten.

Zum besseren Verständnis der in Tabelle 4-10 angegebenen Werte wird darauf hingewiesen, dass es sich um die Mediane der Metriken über alle Stationen bezogen auf den jeweils zu Grunde liegenden Zeitraum handelt. Daher findet sich z. B. der Median der gemessenen Maximalwerte an den Stationen (Max Obs) von 141.93 µg/m³ für das Jahr 2016 nicht in der jahreszeitlichen Auswertung wieder. Bei Betrachtung des gesamten Jahres liegt für die Messzeitreihe jeder Station genau ein Maximum vor, und der Wert von 141.93 µg/m³ stellt hier den Median dieser Maximalwerte über alle Stationen dar. Bei jahreszeitlicher Betrachtung liegen für die Messzeitreihe jeder Station vier Maxima vor, eins zu jeder Jahreszeit, und nur eines davon entspricht dem Maximum des gesamten Jahres. Der Median der Maxima aller Stationen für eine Jahreszeit wird daher im Regelfall niedriger liegen als der Median der Maxima für das gesamte Jahr, weil bei jahreszeitlicher Betrachtung Werte als Maximalwerte berücksichtigt werden, die bei Betrachtung des gesamten Jahres bestenfalls Nebenmaxima darstellen.

Analoge Überlegungen gelten für die Mediane der Minimalwerte (Min Obs, Min Mod). Eine Abweichung findet sich bei Ozon (z. B. Tabelle 4-5). Hier ist der Median der gemessenen Maximalwerte (Max Obs) mit 173.66 µg/m³ identisch für das Jahr 2016 und für den Sommer 2016. Grund dafür ist, dass die Maximalwerte für Ozon für jede der Stationen in den Sommermonaten liegen, d. h. der Datenpool der Maximalwerte über alle Stationen ist für den Sommer weitgehend identisch mit dem für das ganze Jahr, und damit sind auch die zugehörigen Mediane identisch.

Tabelle 4-10: Metriken PM10, Stundenmittelwerte 1SMW für den 1. Prognosetag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	Jahr	Frühling	Sommer	Herbst	Winter
Min Obs [$\mu\text{g}/\text{m}^3$]	0.50	1.00	1.04	1.00	0.50
Min Mod [$\mu\text{g}/\text{m}^3$]	0.91	2.00	1.68	1.40	0.98
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.24	15.95	13.44	16.78	18.52
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.78	13.80	8.10	11.80	13.56
Max Obs [$\mu\text{g}/\text{m}^3$]	141.93	79.01	71.89	74.21	122.70
Max Mod [$\mu\text{g}/\text{m}^3$]	50.67	44.86	24.55	39.58	39.88
SD Obs [$\mu\text{g}/\text{m}^3$]	11.57	10.55	7.96	10.29	14.66
SD Mod [$\mu\text{g}/\text{m}^3$]	6.29	6.76	3.32	6.16	6.55
Korrelation	0.50	0.56	0.43	0.58	0.45
BIAS [$\mu\text{g}/\text{m}^3$]	-3.79	-1.51	-4.78	-4.63	-4.50
RMSE [$\mu\text{g}/\text{m}^3$]	10.87	8.86	8.69	9.77	13.74
MFB [%]	-16.34	0.22	-34.70	-22.63	-9.35
FGE [%]	48.65	43.35	51.44	48.31	49.43

Für die Tagesmittelwerte (Tabelle 4-11) ergibt sich ein ähnliches Bild wie für die Stundenmittelwerte in Tabelle 4-10. Mittelwerte und Standardabweichung der Messdaten werden wie bei den 1SMW im Median am Besten im Frühling getroffen, die Maximalwerte in Frühling und Herbst. Die Korrelation zwischen Mess- und CAMS-Daten ist mit 0.69 im Herbst am höchsten und mit 0.53 im Winter am niedrigsten. Die niedrigsten Werte für MFB und FGE ergeben sich wie bei den Stundenmittelwerten im Frühling, die höchsten Werte im Sommer.

Zusammenfassend lässt sich sagen, dass sowohl für die Stundenmittelwerte 1SMW als auch für die Tagesmittelwerte 1TMW die zur Erfassung der Überschreitungen relevanten Maximalwerte im Median insbesondere im Winter durch das Modell unterschätzt werden, und zwar bei den 1SMW im Median um ca. 67 % und bei den 1TMW um ca. 57 %. Die Korrelation, d. h. die Erfassung des zeitlichen Verlaufs, ist im Herbst am höchsten, gefolgt vom Frühling, während sie in Sommer und Winter deutlich niedriger ist. MFB und FGE sind im Median sowohl für 1SMW als auch für 1TMW im Frühling am niedrigsten und im Sommer am höchsten.

Tabelle 4-11: Metriken PM10, Tagesmittelwerte 1TMW für den 1. Prognosetag, bezogen auf das Jahr 2016 und auf die Jahreszeiten (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

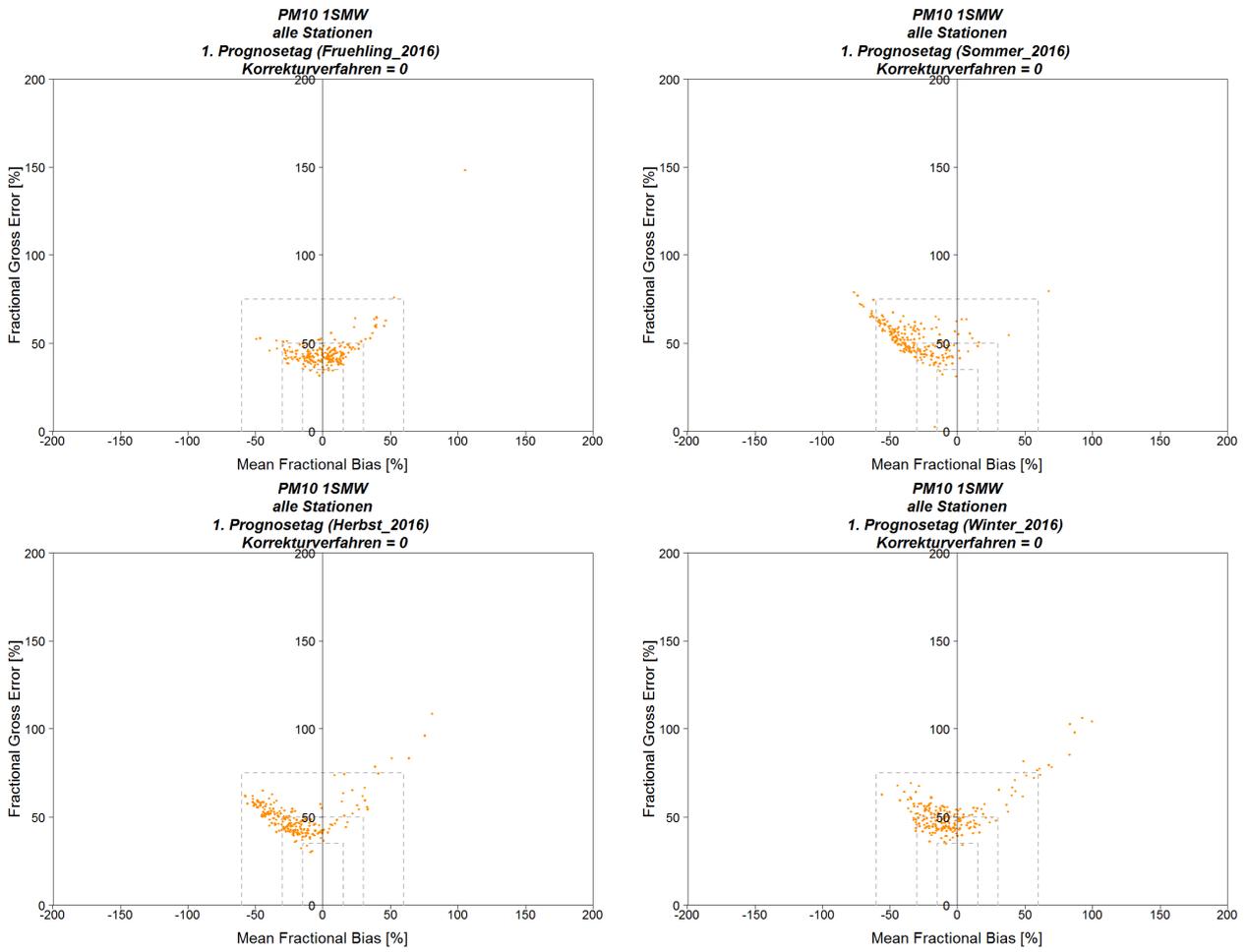
Metrik	Jahr	Frühling	Sommer	Herbst	Winter
Min Obs [$\mu\text{g}/\text{m}^3$]	2.02	3.81	4.95	3.41	2.28
Min Mod [$\mu\text{g}/\text{m}^3$]	3.10	4.52	3.53	3.20	4.02
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.27	15.95	13.43	16.80	18.53
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.77	13.80	8.09	11.80	13.56
Max Obs [$\mu\text{g}/\text{m}^3$]	72.84	43.47	29.71	38.78	71.40
Max Mod [$\mu\text{g}/\text{m}^3$]	35.88	32.81	15.60	28.23	30.50
SD Obs [$\mu\text{g}/\text{m}^3$]	9.20	7.73	5.06	7.92	12.57
SD Mod [$\mu\text{g}/\text{m}^3$]	5.48	5.60	2.48	5.31	5.58
Korrelation	0.58	0.66	0.54	0.69	0.53
BIAS [$\mu\text{g}/\text{m}^3$]	-3.80	-1.47	-4.78	-4.62	-4.60
RMSE [$\mu\text{g}/\text{m}^3$]	8.29	6.14	6.32	7.51	11.73
MFB [%]	-21.50	-5.22	-40.16	-28.21	-13.03
FGE [%]	39.01	31.74	44.44	38.27	38.64
gemessene Überschreitungen	2 ¹	0 ¹	0 ¹	0 ¹	2 ¹
Hit Rate [%]	0.00 ¹				
FAR [%]	0.00 ¹				

¹: Schwellenwert 50 $\mu\text{g}/\text{m}^3$

MFB und FGE sind in Abbildung 4-12 und Abbildung 4-13 für die Stundenmittelwerte 1SMW und für die abgeleitete Größe 1TMW als Soccer-Plots über alle Stationen für jede der vier Jahreszeiten dargestellt. Die 1SMW zeigen im Sommer und im Herbst einen überwiegend negativen Bias. Im Winter verschiebt sich die Wolke der Stationspunkte in Richtung des positiven Bias. Sowohl im Herbst als auch im Winter gibt es eine Reihe Stationen mit einem hohen positiven Bias. Im Frühling halten sich Stationen mit negativem und positivem Bias in etwa die Waage. Dies spiegelt sich auch in dem Median des MFB nahe Null für den Frühling wider (Tabelle 4-10). Die eher schwach lineare Gruppierung der Stationspunkte im Sommer und im Herbst weist auf einen Einfluss sowohl systematischer als auch zufälliger Fehler auf die Prognosedaten zu diesen beiden Jahreszeiten hin. Im Winter und im Frühling dominieren eher zufällige Fehler die Prognosedaten.

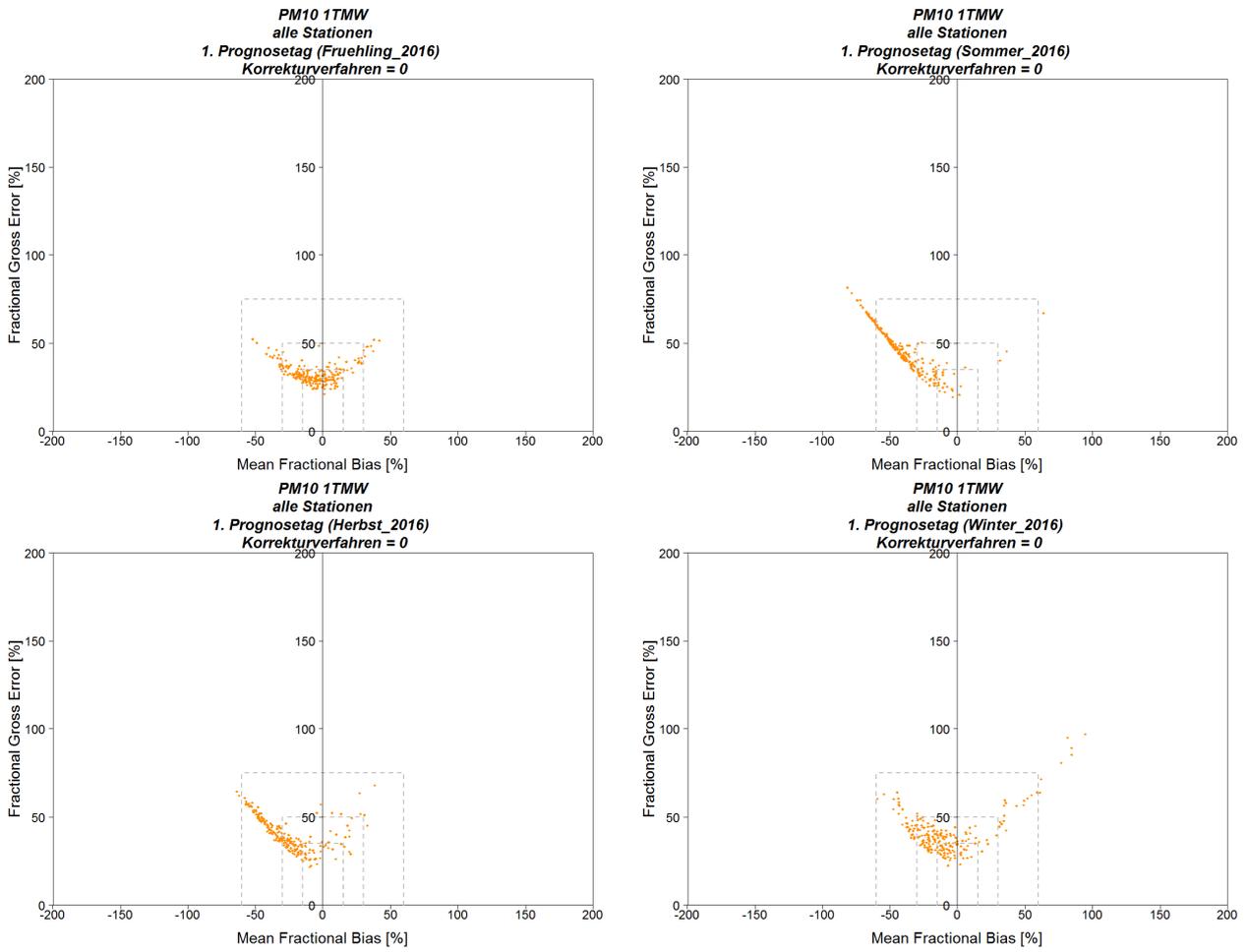
Die Soccer-Plots der 1TMW (Abbildung 4-13) zeigen eine ähnliche Charakteristik wie die der 1SMW. Allerdings ergibt sich zu allen Jahreszeiten eine stärkere lineare Gruppierung der Stationspunkte als bei den 1SMW, und der Median des MFB liegt weiter im negativen Bereich als bei den 1SMW.

Abbildung 4-12: Soccer-Plots PM10 für die Stundenmittelwerte 1SMW, 1. Prognosestag, Bezugsjahr 2016, unterteilt nach Jahreszeiten



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 4-13: Soccer-Plots PM10 für die Tagesmittelwerte 1TMW, 1. Prognosestag, Bezugsjahr 2016, unterteilt nach Jahreszeiten



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

5 Korrektur an den Messorten

5.1 Evaluierung der Ergebnisse

5.1.1 Ozon

5.1.1.1 Jahreszeitreihe 2016

Die Jahreszeitreihen der CAMS-Ozonprognosen für 2016 an den Messorten wurden für die drei betrachteten Prognosetage einer Korrektur durch die in Abschnitt 3.3 beschriebenen Korrekturverfahren unterzogen. Die Verfahren werden im Folgenden mit MK (Multiplikative Korrektur, Abschnitt 3.3.1), KF (Kalman-Filter, Abschnitt 3.3.3) und HV (Hybrid-Vorhersage, Abschnitt 3.3.2) abgekürzt. Für die korrigierten CAMS-Ozonprognosen wurde ebenso wie für die CAMS-Rohdaten in Abschnitt 4.1 eine statistische Auswertung mit Hilfe der in Abschnitt 3.4.1 aufgeführten Metriken und auf Basis der in Abschnitt 3.4.2 erläuterten Target-Plots des DELTA-Tools durchgeführt.

Tabelle 5-1 enthält die für den 1. Prognosetag für die Einstundenmittelwerte (1SMW) berechneten Metriken für die CAMS-Rohdaten und im Vergleich dazu für die korrigierten CAMS-Daten. Angegeben ist jeweils der Median über alle betrachteten Stationen. Fett hervorgehobene Zahlen markieren jeweils den Wert, der die größte Übereinstimmung zwischen Messung und Modell für diese Größe anzeigt. Die reinen Messwert-Metriken (Min Obs, Mean Obs, Max Obs, SD Obs) weisen zwischen den Verfahren leichte Unterschiede auf. Grund dafür ist, dass die einzelnen Korrekturverfahren unterschiedliche Vorlauf-Längen haben, so dass einzelne Messwerte zu Beginn des Jahres je nach Korrekturverfahren mit in der Metrik berücksichtigt werden oder eben nicht.

Grundsätzlich ergibt sich durch jedes Korrekturverfahren eine Verbesserung der CAMS-Rohdaten. Für die Einstundenmittelwerte erzeugt im Regelfall der Kalman-Filter die besten Ergebnisse. Ausnahme davon ist die Erfassung der Maximalwerte und der Standardabweichung, diese Größen werden am besten von der Hybrid-Vorhersage erreicht. Das deckt sich mit den Ergebnissen von Kang et al. (2008). Die Hybrid-Vorhersage weist gegenüber dem Kalman-Filter eine schlechtere Korrelation zu den Messwerten und einen höheren FGE auf. Die Metriken der Multiplikativen Korrektur liegen überall im Mittelfeld, insbesondere werden damit höhere Maximalwerte erreicht als mit dem Kalman-Filter. Allerdings ist die Multiplikative Korrektur, wie in Abschnitt 3.3.1 bereits angesprochen, ohne Beschränkung des Korrekturfaktors ein problematisches Korrekturverfahren. Dies wird im Folgenden noch verdeutlicht.

Für die maximalen gleitenden 8-Stundenmittelwerte eines Tages (8TMAX) zeigt sich ein etwas anderes Bild (Tabelle 5-2). Hier erzeugen die Hybrid-Vorhersage und der Kalman-Filter die besten Ergebnisse, wobei die Hybrid-Vorhersage wie schon zuvor die Maximalwerte und die Standardabweichung, aber nun auch den Mittelwert am besten trifft, während der Kalman-Filter etwas bessere Werte bei Korrelation, RMSE, MFB und FGE aufweist. Aufgrund der besseren Erfassung der Maximalwerte erreicht die Hybrid-Vorhersage auch die höchste Hit Rate, d. h. den höchsten Anteil korrekt vorhergesagter Schwellenwertüberschreitungen. Gleichzeitig ist damit auch die FAR, d. h. der höchste Anteil falsch vorhergesagter Schwellenwertüberschreitungen, für die Hybrid-Vorhersage am höchsten, allerdings liegt der Wert mit knapp 1.2 % auf niedrigem Niveau.

Für die maximalen Einstundenmittelwerte eines Tages (1TMAX) in Tabelle 5-3 werden die Mittelwerte der gemessenen Zeitreihen am besten durch die unkorrigierten CAMS-Rohdaten getroffen, entsprechend ist hier auch der Bias am niedrigsten. Von den drei Korrekturverfahren liefern Kalman-Filter und Hybrid-Vorhersage die besten Ergebnisse, wobei wie zuvor die Hybrid-Vorhersage die Mittel- und Maximalwerte sowie die Standardabweichung besser trifft und der Kalman-Filter die besseren Werte für Korrelation, RMSE, MFB und FGE erzeugt. Schwellenwertüberschreitungen können nicht bewertet werden, da 2016 im Median über alle Stationen weder Überschreitungen der Alarmschwelle ($240 \mu\text{g}/\text{m}^3$) noch der Informationsschwelle ($180 \mu\text{g}/\text{m}^3$) erscheinen.

Tabelle 5-1: Metriken O₃, Stundenmittelwerte 1SMW für den 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	0.52	0.52	0.52	0.52
Min Mod [$\mu\text{g}/\text{m}^3$]	0.82	0.23	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	44.72	45.54	44.87	44.95
Mean Mod [$\mu\text{g}/\text{m}^3$]	56.36	45.18	44.98	45.37
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.51	173.66	173.51
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	163.05	153.65	169.84
SD Obs [$\mu\text{g}/\text{m}^3$]	29.44	29.68	29.49	29.46
SD Mod [$\mu\text{g}/\text{m}^3$]	22.29	27.08	26.67	29.01
Korrelation	0.81	0.82	0.84	0.80
BIAS [$\mu\text{g}/\text{m}^3$]	10.17	-0.39	0.11	0.34
RMSE [$\mu\text{g}/\text{m}^3$]	19.82	17.21	15.85	18.08
MFB [%]	38.27	8.57	4.23	-2.23
FGE [%]	48.91	44.03	44.82	52.26

Tabelle 5-2: Metriken O₃, maximale gleitende 8-Stundenmittelwerte eines Tages 8TMAX für den 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	3.08	3.12	3.08	3.12
Min Mod [$\mu\text{g}/\text{m}^3$]	17.49	6.78	3.52	2.80
Mean Obs [$\mu\text{g}/\text{m}^3$]	64.59	64.76	64.87	64.75
Mean Mod [$\mu\text{g}/\text{m}^3$]	69.07	62.51	63.22	65.30
Max Obs [$\mu\text{g}/\text{m}^3$]	155.21	152.90	155.21	154.67
Max Mod [$\mu\text{g}/\text{m}^3$]	130.41	148.84	142.19	153.63
SD Obs [$\mu\text{g}/\text{m}^3$]	29.46	29.59	29.56	29.46
SD Mod [$\mu\text{g}/\text{m}^3$]	22.66	28.25	27.18	28.47
Korrelation	0.89	0.89	0.91	0.89
BIAS [$\mu\text{g}/\text{m}^3$]	3.63	-2.00	-1.57	0.88
RMSE [$\mu\text{g}/\text{m}^3$]	14.56	13.97	12.30	13.18
MFB [%]	12.68	-1.23	0.08	3.11
FGE [%]	22.90	22.16	19.61	20.60
gemessene Überschreitungen	15 ¹	14 ¹	15 ¹	15 ¹
Hit Rate [%]	13.33 ¹	53.33 ¹	50.00 ¹	65.45¹
FAR [%]	0.00¹	0.62 ¹	0.32 ¹	1.17 ¹

¹: Schwellenwert 120 $\mu\text{g}/\text{m}^3$

Tabelle 5-3: Metriken O₃, maximale Stundenmittelwerte eines Tages 1TMAX für den 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	4.79	4.86	4.81	4.86
Min Mod [$\mu\text{g}/\text{m}^3$]	20.36	9.03	6.64	4.61
Mean Obs [$\mu\text{g}/\text{m}^3$]	71.73	71.92	72.01	71.76
Mean Mod [$\mu\text{g}/\text{m}^3$]	72.56	69.44	69.58	73.48
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.06	173.66	173.66
Max Mod [$\mu\text{g}/\text{m}^3$]	139.15	162.92	153.57	169.73
SD Obs [$\mu\text{g}/\text{m}^3$]	30.63	30.62	30.62	30.64
SD Mod [$\mu\text{g}/\text{m}^3$]	23.19	29.27	28.10	29.69
Korrelation	0.88	0.87	0.90	0.87
BIAS [$\mu\text{g}/\text{m}^3$]	0.22	-2.27	-2.31	1.84
RMSE [$\mu\text{g}/\text{m}^3$]	15.23	15.58	13.73	15.12
MFB [%]	5.64	-1.88	-1.28	4.23
FGE [%]	20.13	20.95	18.64	19.87
gemessene Überschreitungen	0	0	0	0
Hit Rate [%]	0.00	0.00	0.00	0.00
FAR [%]	0.00	0.00	0.00	0.00

¹: Schwellenwert 180 $\mu\text{g}/\text{m}^3$ (Informationsschwelle)

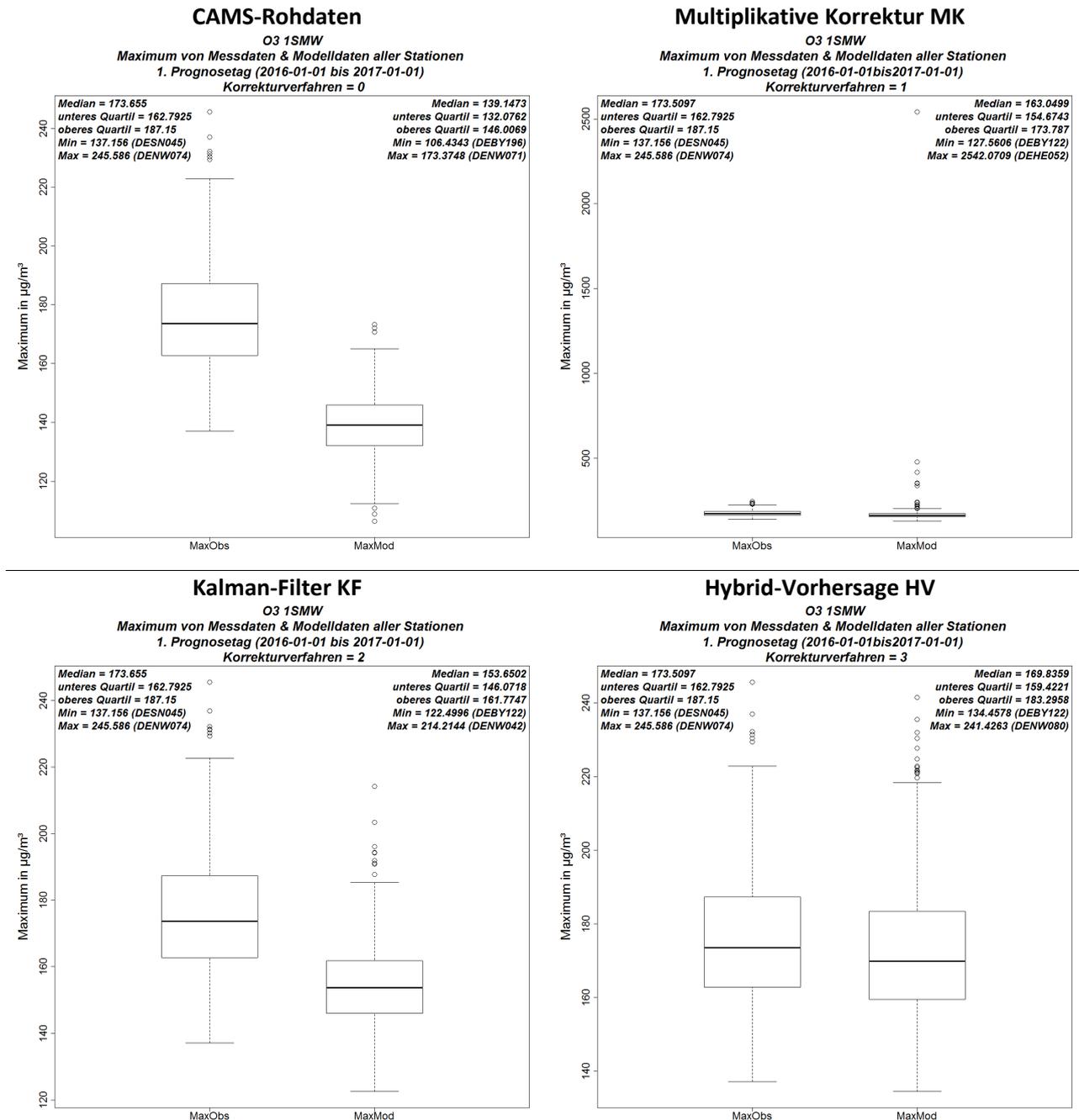
Abbildung 5-1 zeigt für Ozon beispielhaft für den 1. Prognosetag Box-Whisker-Plots der Maxima der 1SMW an den Stationsorten, für die CAMS-Rohdaten und für die drei Korrekturverfahren, jeweils im Vergleich mit den Messwerten. Gleichzeitig sind die Werte für Median, oberes und unteres Quartil sowie Minimum und Maximum in den Grafiken angegeben. Es wird deutlich, dass der Kalman-Filter und noch mehr die Hybrid-Vorhersage zu einer verbesserten Erfassung der Maximalwerte führen. Die Multiplikative Korrektur weist zwar im Median und in den Quartilen gute Werte im Vergleich zu den Messdaten auf. Allerdings ergeben sich für einige Stationen zu bestimmten Zeitpunkten sehr ungünstige Verhältnisse der Mess- zu den Modellwerten und damit einige extrem hohen Korrekturfaktoren, die entsprechend unrealistisch hohe Konzentrationswerte in den korrigierten Modellprognosen erzeugen. Im Maximum werden hier Konzentrationen von über 2500 $\mu\text{g}/\text{m}^3$ erreicht (Station DEHE052).

Abbildung 5-2 zeigt die zugehörigen Box-Whisker-Plots der Standardabweichungen. Es ergibt sich das gleiche Bild wie bei den Maxima in Abbildung 5-1. Der Kalman-Filter und noch mehr die Hybrid-Vorhersage führen zu einer Angleichung der modellierten Standardabweichung an die der Messwerte. Die Multiplikative Korrektur weist im Median und in den Quartilen gute Werte im Vergleich zu den Messdaten auf, wenn auch niedrigere und damit ungünstigere Werte als die Hybrid-Vorhersage. Sie führt jedoch an der Station DEHE052, für die der Konzentrationswert von über 2500 $\mu\text{g}/\text{m}^3$ berechnet wurde, auch zu einer im Vergleich zu den anderen Stationen sehr hohen Standardabweichung.

In Abbildung 5-3 sind für den 1. Prognosetag die Soccer-Plots der 1SMW für die CAMS-Rohdaten und die drei Korrekturverfahren dargestellt. Demnach weisen die CAMS-Rohdaten MFB-Werte bis 80%

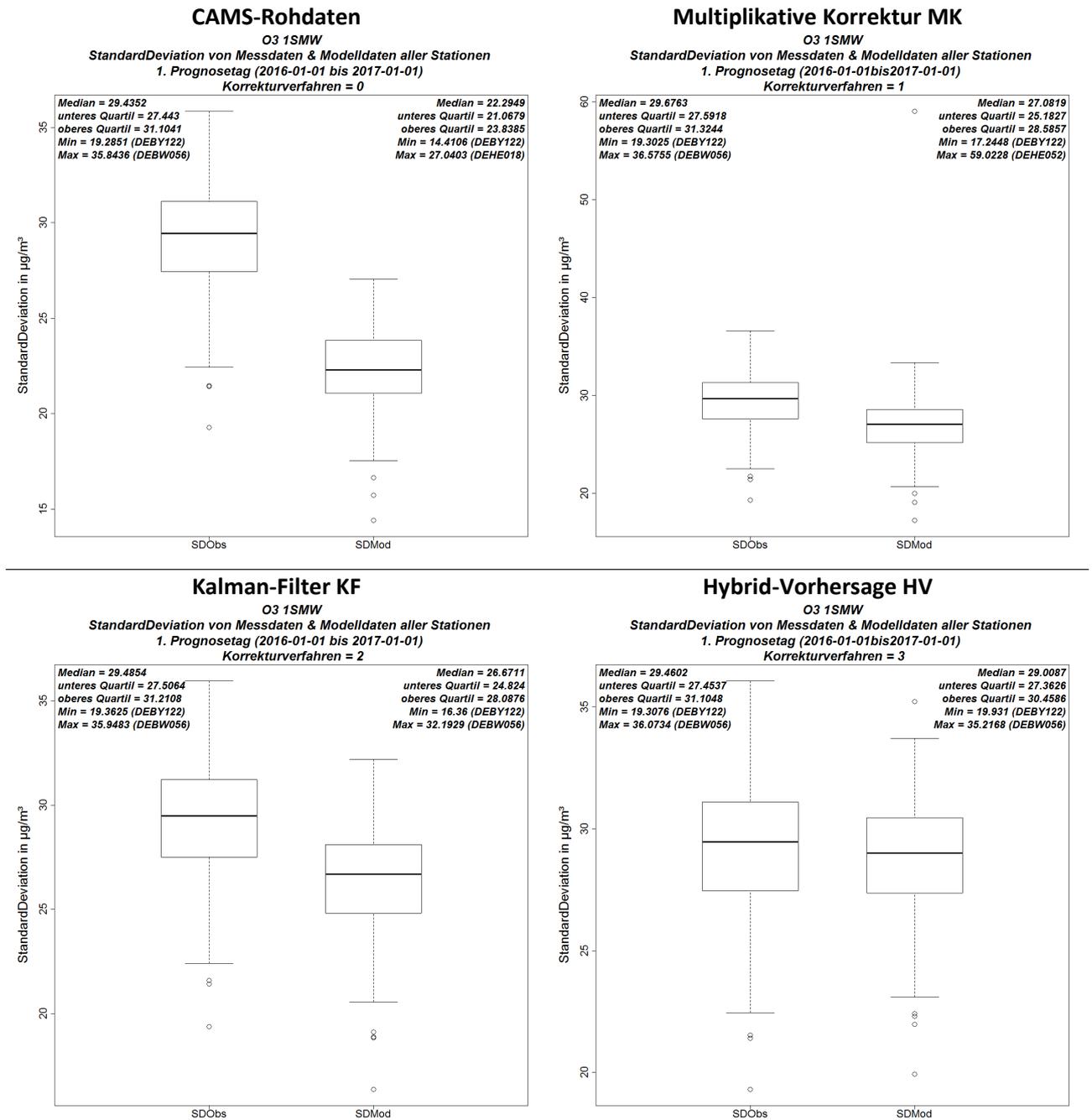
und FGE-Werte bis zu knapp 90 % auf, die lineare Gruppierung der Werte lässt auf einen systematischen Bias in den Modelldaten schließen. Die Anwendung der Korrekturverfahren verringert in jedem Fall den MFB und in geringerem Maße auch den FGE.

Abbildung 5-1: Box-Whisker-Plots der Maxima an den Stationsorten für O₃, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts), jeweils im Vergleich mit den Messwerten



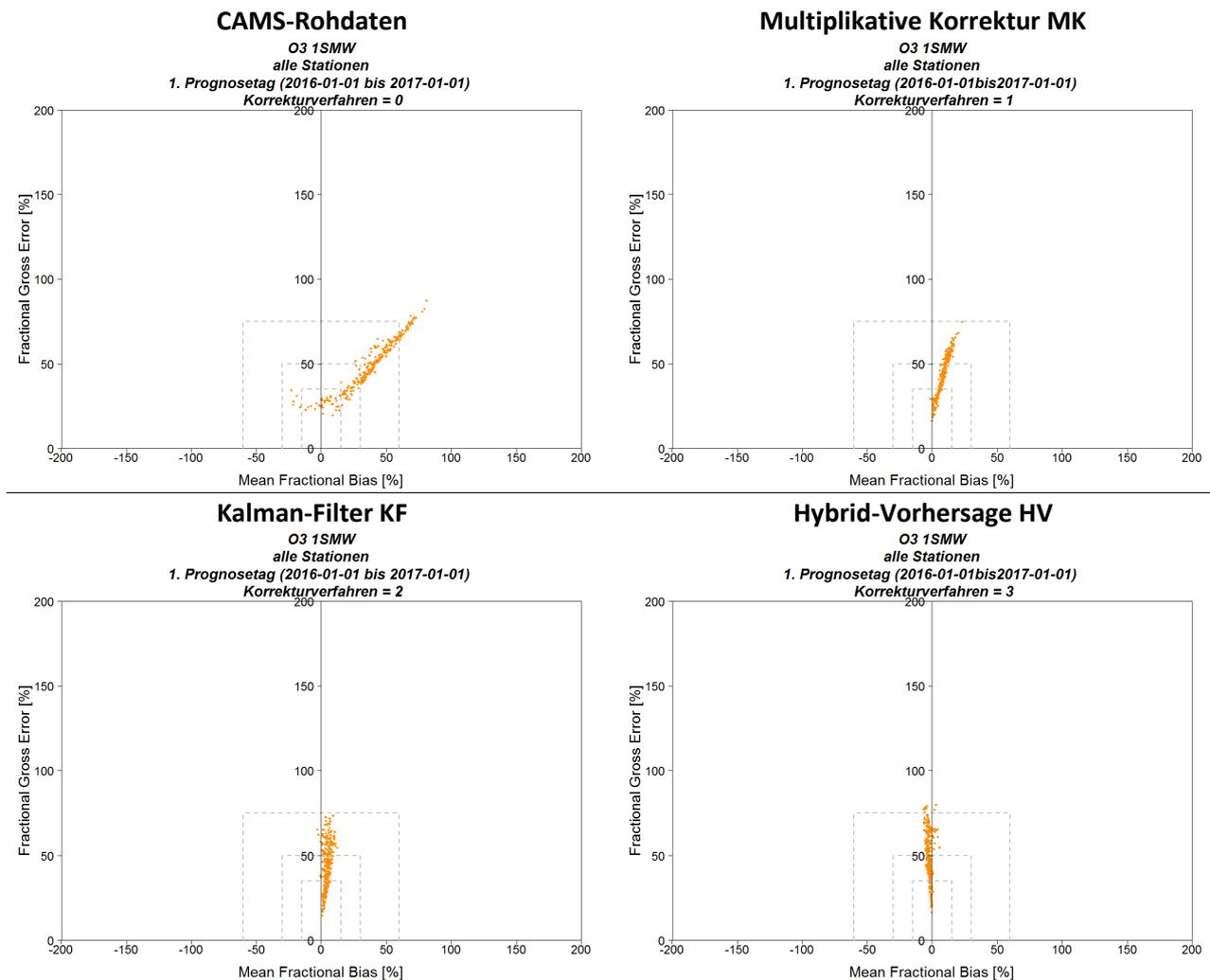
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-2: Box-Whisker-Plots der Standardabweichung an den Stationsorten für O₃, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts), jeweils im Vergleich mit den Messwerten



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-3: Soccer-Plots für O₃, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

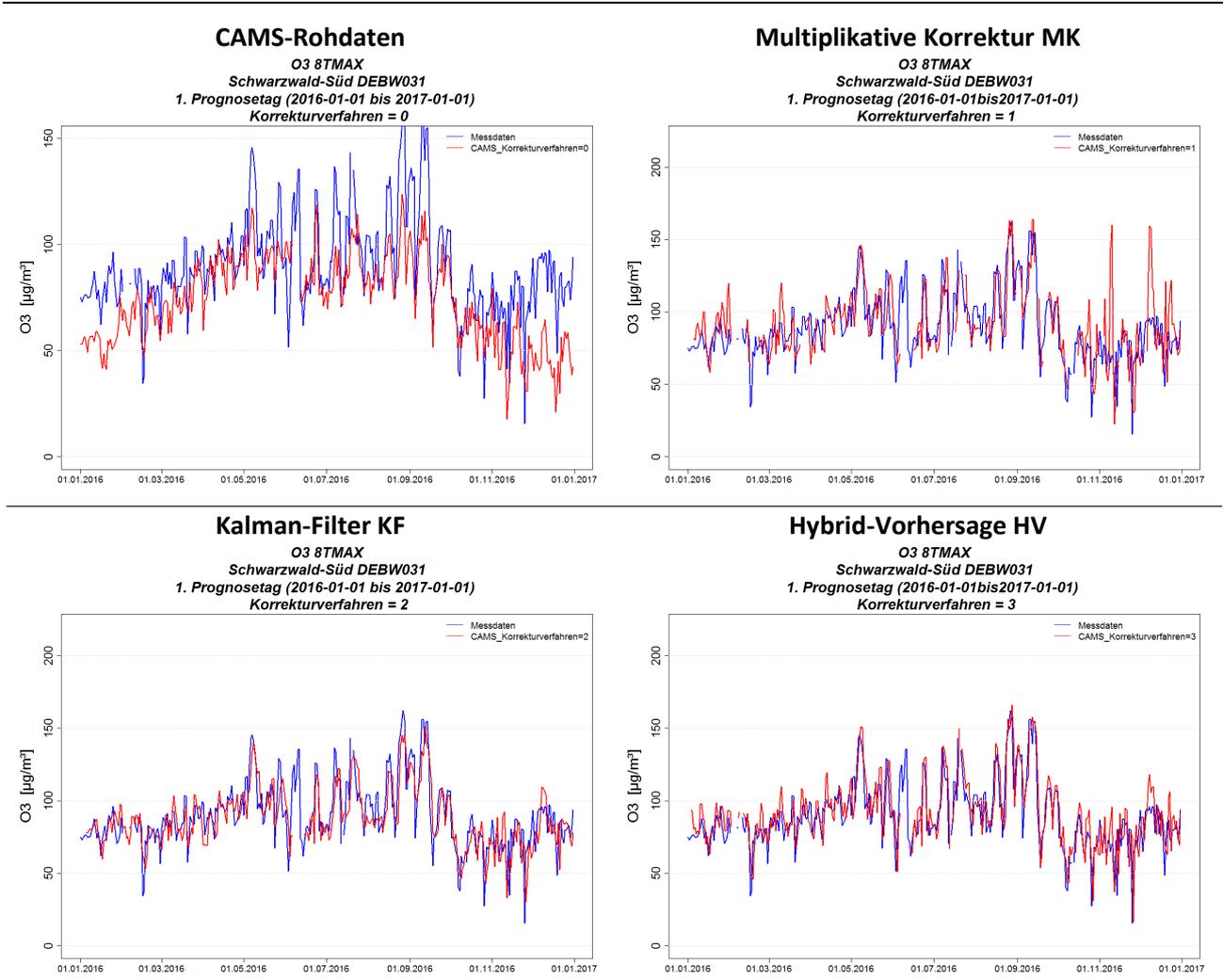
In Abbildung 5-4 und Abbildung 5-5 sind beispielhaft Zeitreihen der maximalen gleitenden 8-Stundenmittelwerte des Tages (8TMAX) über das Bezugsjahr 2016 für die CAMS-Rohdaten und die korrigierten Modellprognosen dargestellt. Abbildung 5-4 zeigt Zeitreihen der Station Schwarzwald-Süd (DEBW031), für die in den CAMS-Rohdaten die niedrigste Korrelation zu den Messdaten bestand (Abbildung 4-1), und Abbildung 5-5 zeigt Zeitreihen der Station Wesel-Feldmark (DENW030), für die die CAMS-Rohdaten die höchste Korrelation zu den Messdaten aufwiesen (Abbildung 4-1).

Für die Station DEBW031 zeigen die CAMS-Rohdaten eine deutliche Unterschätzung der 8TMAX der Messwerte, was möglicherweise auf die Höhenlage der Messstation (904 m.ü.NN) zurückzuführen ist. Die Unterschätzung wird durch alle drei Korrekturverfahren korrigiert, was aber insbesondere bei der Multiplikativen Korrektur vor allem zu Beginn und gegen Ende des Jahres 2016 zu ebenso deutlichen Überschätzungen führt. Kalman-Filter und Hybrid-Vorhersage liefern hier bessere Ergebnisse, wobei die Hybrid-Vorhersage eher zu leichten Überschätzungen neigt als der Kalman-Filter, dafür aber auch Episoden wie die um den 01.09.2016 besser abbildet.

Die 8TMAX der Messwerte der Station DENW030 werden durch die CAMS-Rohdaten bereits recht gut getroffen. Anwendung der Korrekturverfahren führt zu einer besseren Erfassung der Episoden z. B. im Mai und im Juli 2016, wobei wie zuvor der Kalman-Filter die meisten Maxima noch nicht ganz erreicht,

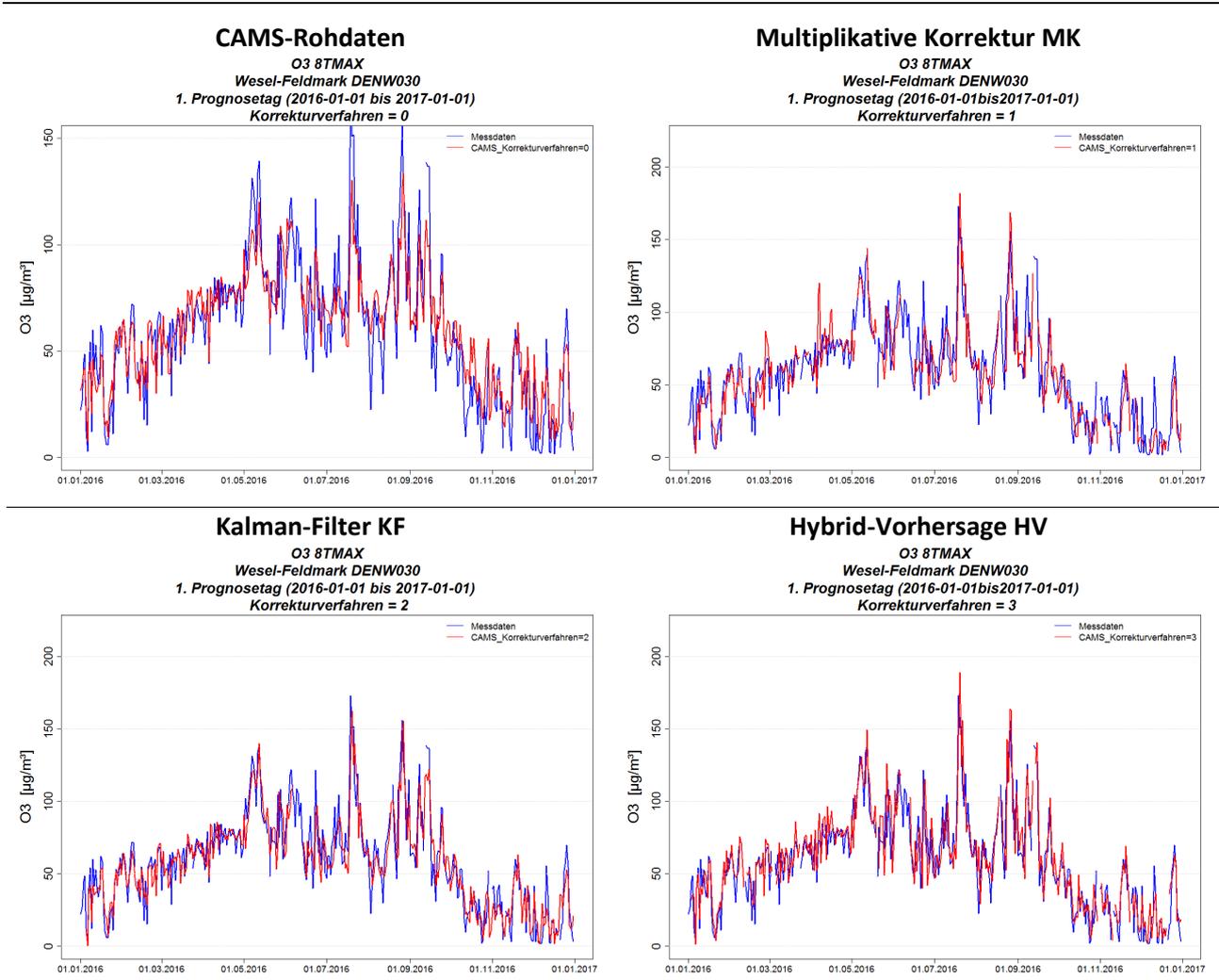
während die Hybrid-Vorhersage tendenziell eher zu leichter Überschätzung neigt. Die Multiplikative Korrektur führt z. B. Anfang März und im April 2016 zu deutlicheren Überschätzungen, die jedoch nicht so extrem sind wie für die Station DEBW031.

Abbildung 5-4: Zeitreihen der 8TMAX für O₃, Station DEBW031, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-5: Zeitreihen der 8TMAX für O₃, Station DENW030, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



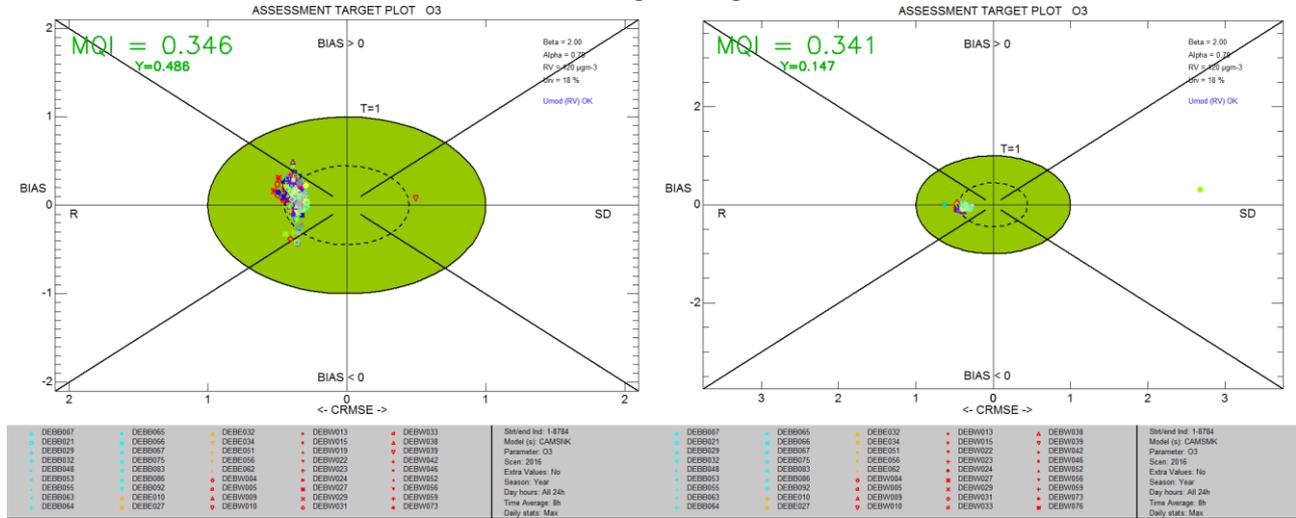
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

In Abbildung 5-6 und Abbildung 5-7 werden die Assessment Target Plots der CAMS-Rohdaten mit denen der korrigierten Verfahren für alle drei betrachteten Prognosetage miteinander verglichen. Die Auswertung bezieht sich dabei wie in Abschnitt 4.1 auf die 8TMAX. Wie in Abschnitt 4.1.2 ausgeführt, erfüllen für Ozon bereits die CAMS-Rohdaten für alle Stationen das vom DELTA-Tool definierte Modellqualitätsziel. Das 90 %-Perzentil der MQI aller Stationen für die CAMS-Rohdaten beträgt für den 1. Prognosetag 0.346. Durch Anwendung der Korrekturverfahren wird das 90 %-Perzentil der MQI aller Stationen nur noch geringfügig verringert, der niedrigste Wert von 0.301 am 1. Prognosetag ergibt sich bei Anwendung des Kalman-Filters. Alle Korrekturverfahren verringern, wie schon anhand der Metriken gezeigt, den Bias zwischen Messung und Modell erheblich, und der Anteil Stationen mit einem MQI < 0.5 ist für die korrigierten Modellprognosen höher als für die CAMS-Rohdaten. Auffällig ist, dass der Target Plot der Multiplikativen Korrektur am 1. Prognosetage eine andere Skalierung aufweist als die anderen Target Plots. Grund hierfür ist wiederum die Station DEHE052, im Target Plot als grüner Punkt am rechten Bildrand erkennbar, für die durch die Multiplikative Korrektur aufgrund ungünstiger Verhältnisse von Mess- zu Modellwerten der extrem hohe Wert von über 2500 µg/m³ berechnet wurde.

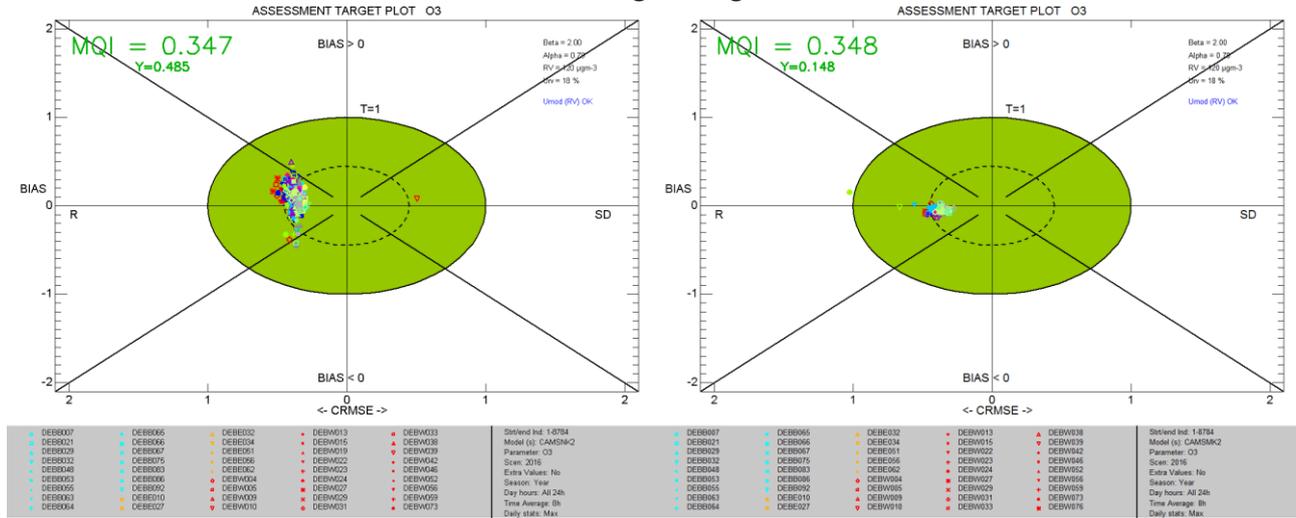
Werden die Assessment Target Plots der verschiedenen Prognosetage miteinander verglichen, so steigt der MQI mit der Vorhersagelänge nur geringfügig an, für die CAMS-Rohdaten von 0.346 auf 0.359 und für die mit dem Kalman-Filter korrigierten Prognosedaten von 0.301 auf 0.318. Der Einfluss der Vorhersagelänge auf den MQI ist damit auch für die korrigierten Modellprognosen vernachlässigbar.

Abbildung 5-6: Assessment Target Plots O₃ für CAMS-Rohdaten (links) und Multiplikative Korrektur (rechts), Bezugsjahr 2016, 1. bis 3. Prognosestag

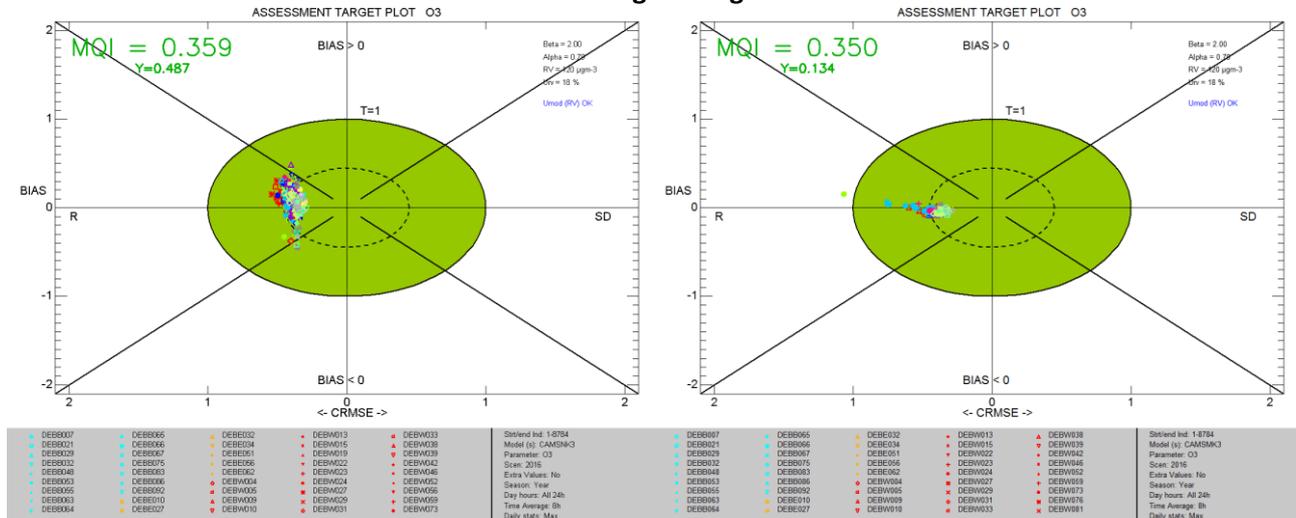
1. Prognosestag



2. Prognosestag



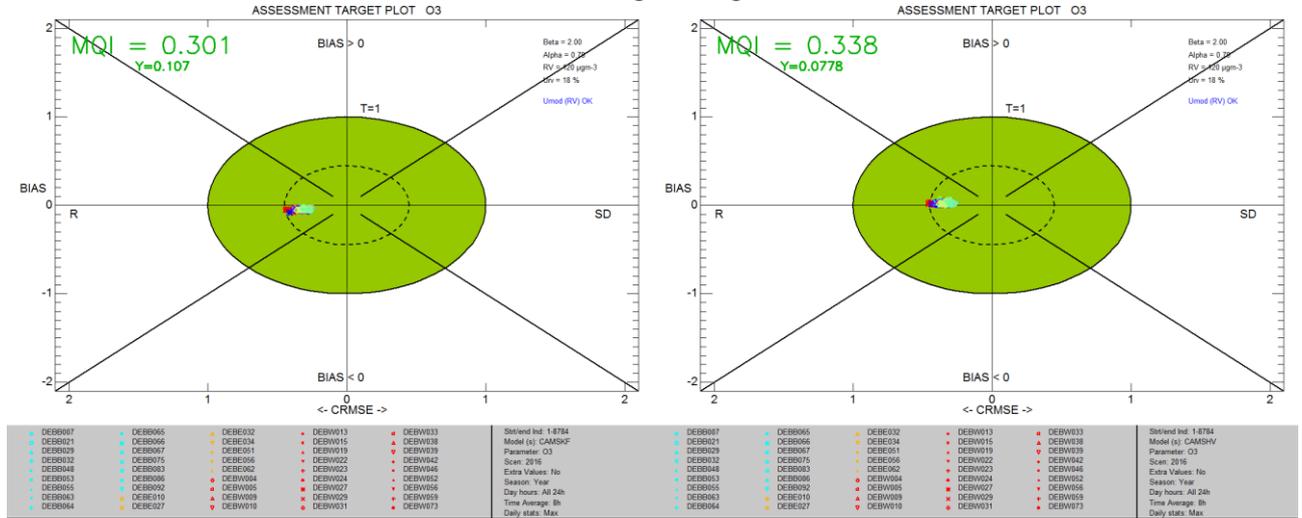
3. Prognosestag



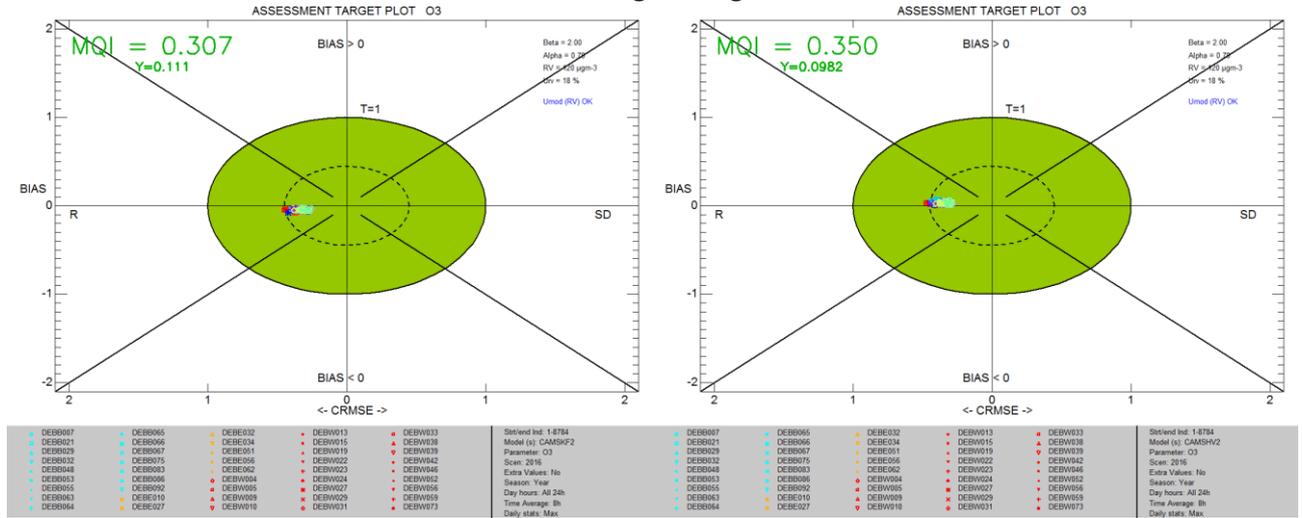
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-7: Assessment Target Plots O₃ für Kalman-Filter (links) und Hybrid-Vorhersage (rechts), Bezugsjahr 2016, 1. bis 3. Prognostetag

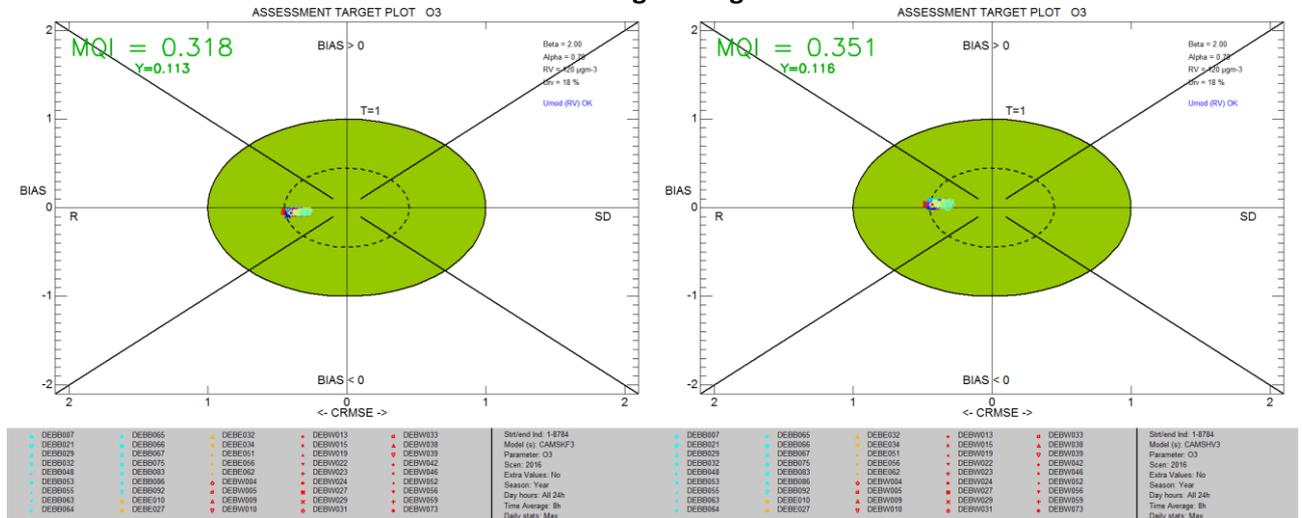
1. Prognostetag



2. Prognostetag



3. Prognostetag



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

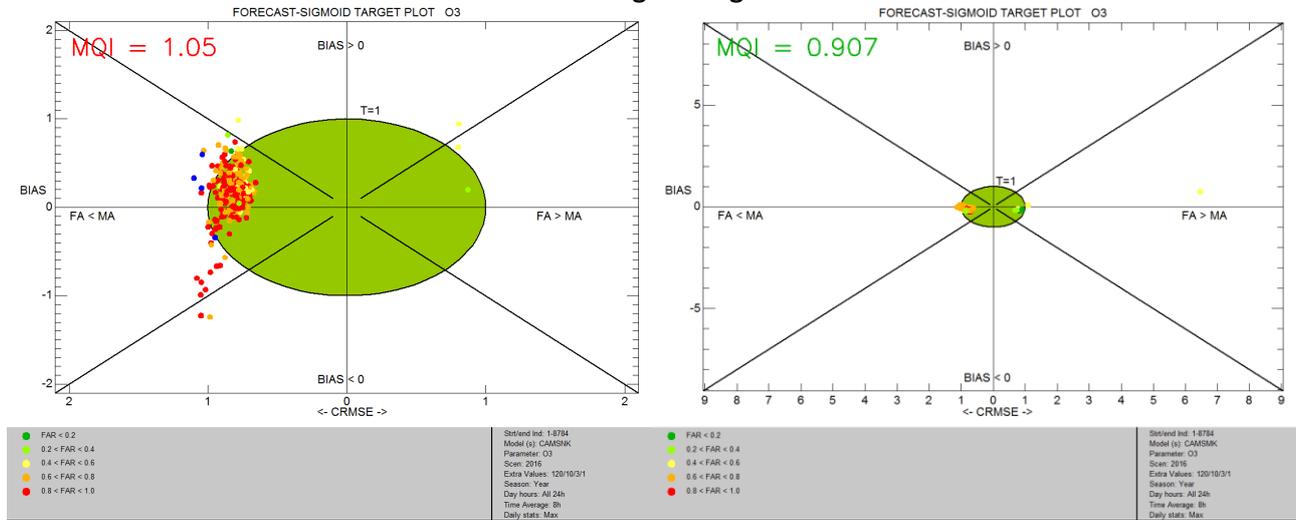
Im Forecast-Modus des DELTA-Tools zeigen die Target Plots die Güte der Modellprognosen im Vergleich zur Prognosegüte eines persistenten Modells. Die Auswertung bezieht sich wie im Benchmark-Modus auf die 8TMAX. Analog zu Abschnitt 4.1.2 ist der MQI sowohl der CAMS-Rohdaten als auch der korrigierten Prognosedaten deutlich schlechter als im Benchmark-Modus (Abbildung 5-8, Abbildung 5-9). Bei den CAMS-Rohdaten weist am 1. Prognosetag eine ganze Reihe von Stationen einen MQI > 1 auf und wird damit schlechter modelliert als mit dem persistenten Modell. Das 90%-Perzentil der MQI aller Stationen liegt für den 1. Prognosetag bei 1.05 und ist damit ebenfalls > 1.

Durch Anwendung der Korrekturverfahren werden MQI und Bias deutlich verringert, für Kalman-Filter und Hybrid-Vorhersage liegen am 1. Prognosetag die MQI aller Stationen unter 1. Für den Kalman-Filter gilt wie für die CAMS-Rohdaten, dass die Anzahl der FA bis auf wenige Ausnahmen kleiner ist als die Anzahl der MA, während für die Hybrid-Vorhersage deutlich mehr Stationen eine höhere Anzahl FA als MA aufweisen. Insgesamt weist der Kalman-Filter mit einem 90 %-Perzentil der MQI aller Stationen von 0.804 auch im Forecast-Mode den niedrigsten Wert aller Korrekturverfahren auf. Der Target Plot der Multiplikativen Korrektur zeigt als Ausreißer wieder die Station DEHE052 als gelben Punkt am rechten äußeren Rand.

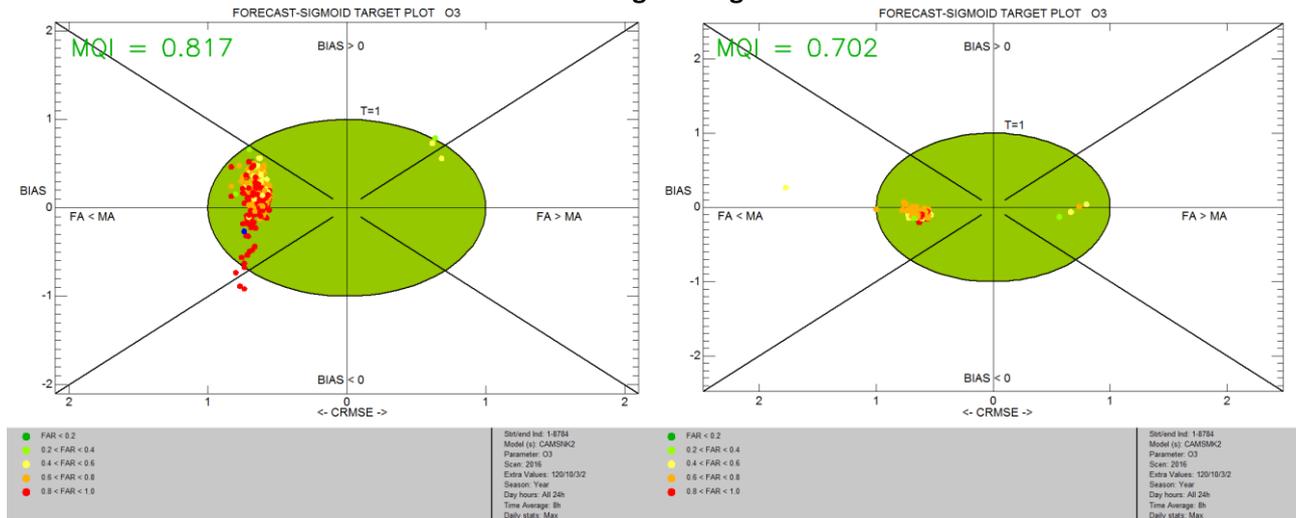
Werden die Forecast-Sigmoid Target Plots der verschiedenen Prognosetage miteinander verglichen, so zeigen sowohl die CAMS-Rohdaten als auch die korrigierten Prognosedaten eine umso bessere Performance gegenüber dem persistenten Modell, je länger die Vorhersage ist. Das 90 %-Perzentil der MQI aller Stationen für die CAMS-Rohdaten sinkt, wie bereits in Abschnitt 4.1.2 beschrieben, von 1.05 am 1. Prognosetag auf 0.746 am 3. Prognosetag. Für die mit dem Kalman-Filter korrigierten Prognosedaten sinkt das 90 %-Perzentil der MQI von 0.804 am 1. Prognosetag auf 0.572 am 3. Prognosetag. Die Korrektur mit dem Kalman-Filter erzeugt somit für alle drei Prognosetage das jeweils niedrigste 90 %-Perzentil der MQI aller Korrekturverfahren.

Abbildung 5-8: Forecast-Sigmoid Target Plots O₃ für CAMS-Rohdaten (links) und Multiplikative Korrektur (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag

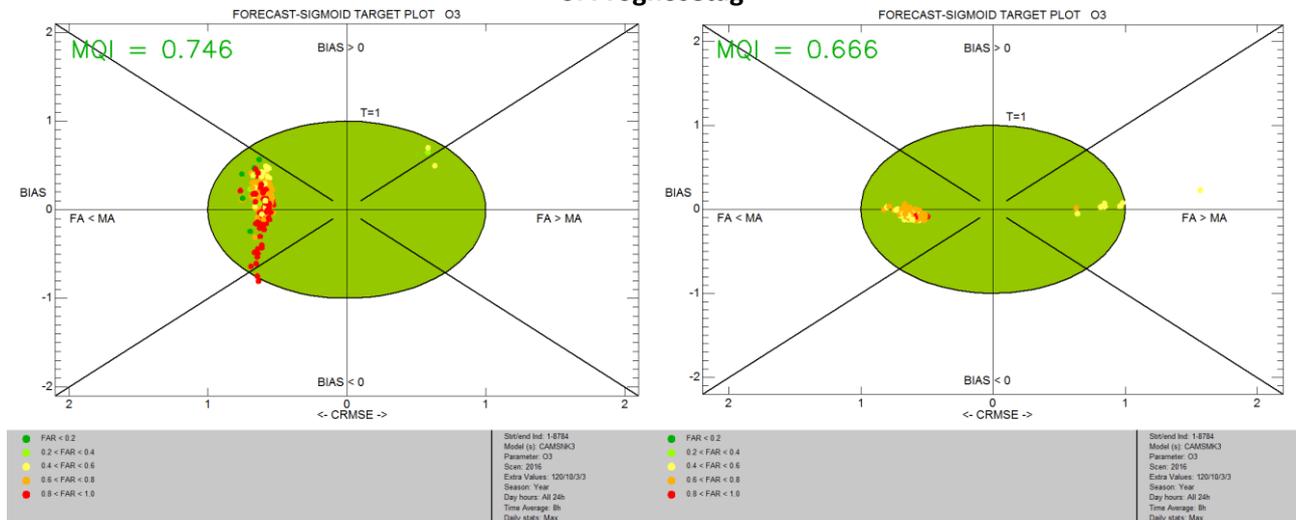
1. Prognosetag



2. Prognosetag



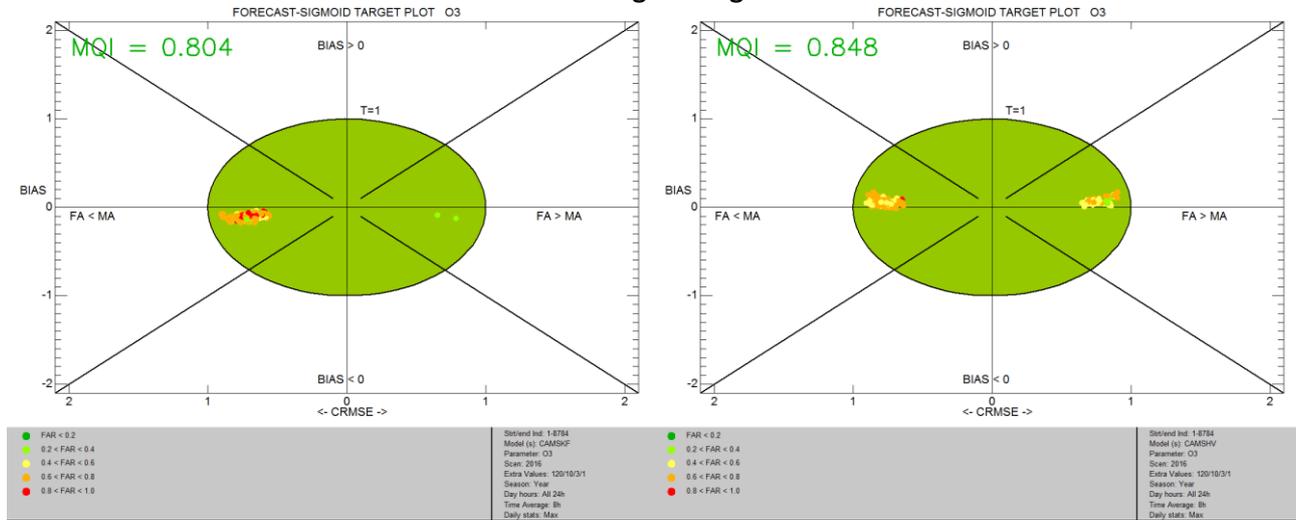
3. Prognosetag



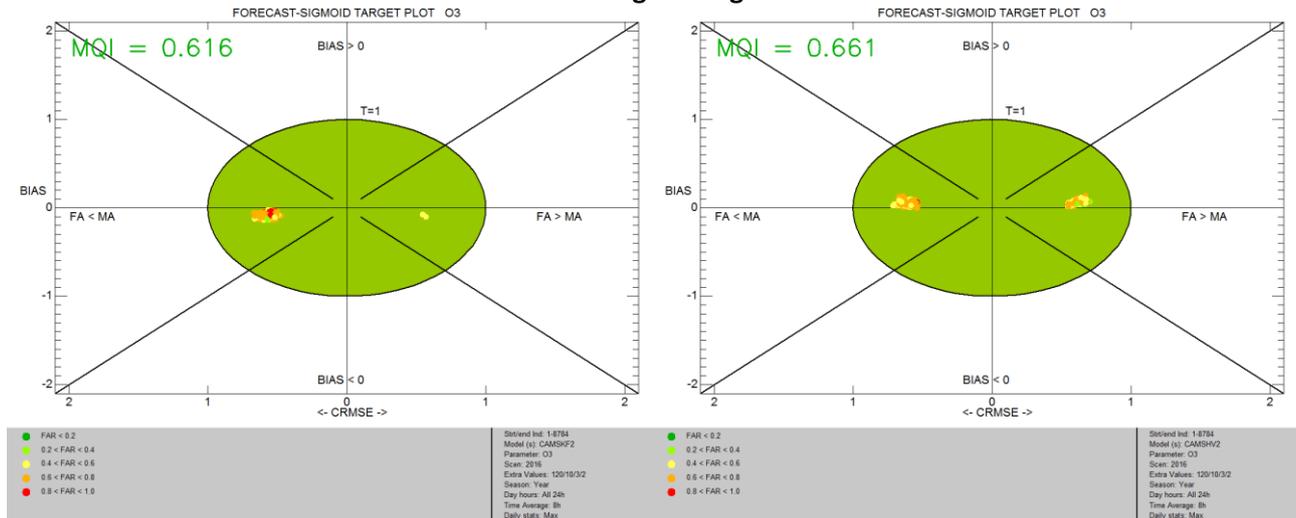
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-9: Forecast-Sigmoid Target Plots O₃ für Kalman-Filter (links) und Hybrid-Vorhersage (rechts), Bezugsjahr 2016, 1. bis 3. Prognostag

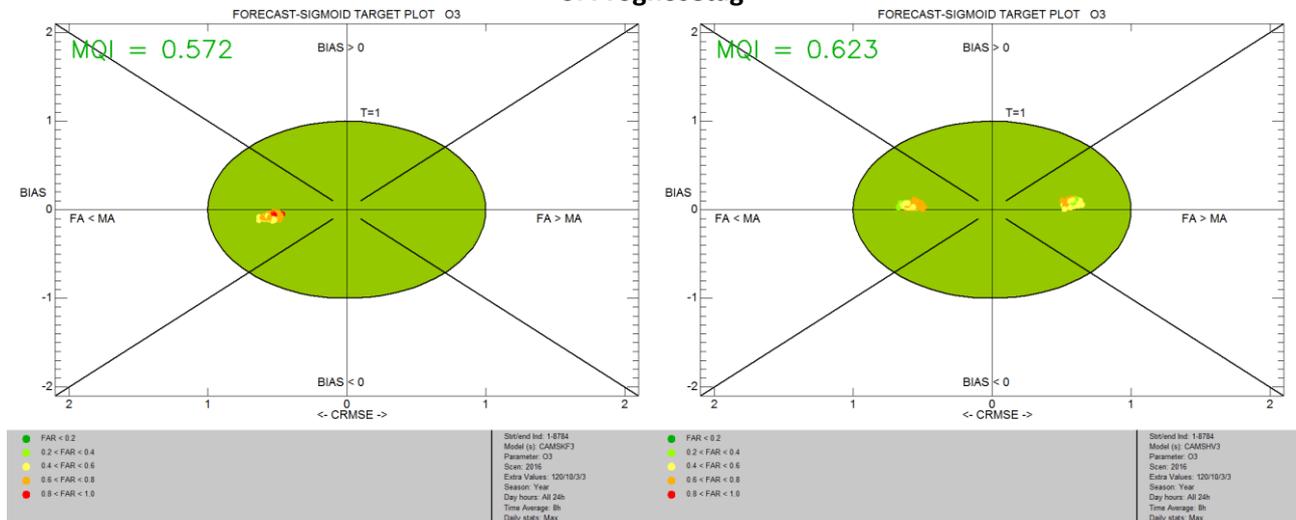
1. Prognostag



2. Prognostag



3. Prognostag



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

5.1.1.2 Saisonale Auswertung 2016

Um die Performanz der Korrekturverfahren auch saisonal erfassen zu können, wurden auch die Jahreszeitreihen der korrigierten CAMS-Prognosedaten 2016 analog zu Abschnitt 4.1.3 nach Jahreszeiten unterteilt, und anschließend wurden für jede Jahreszeit getrennt die Metriken nach Tabelle 3-5 bestimmt.

In Tabelle 5-4 und Tabelle 5-5 sind die Mediane der für alle Stationen berechneten Metriken für die vier Jahreszeiten für die CAMS-Rohdaten und die korrigierten CAMS-Daten aufgeführt, jeweils bezogen auf die Einstundenmittelwerte (1SMW). Fett hervorgehobene Zahlen markieren jeweils den Wert, der die größte Übereinstimmung zwischen Messung und Modell für diese Größe anzeigt.

Demnach ergibt sich im Regelfall auch bei saisonaler Betrachtung durch jedes der eingesetzten Korrekturverfahren eine Verbesserung der CAMS-Prognose. Wie bei der Jahresgesamtbetrachtung in Abschnitt 5.1.1.1 erzeugt der Kalman-Filter häufig die besten Ergebnisse. Ausnahmen hiervon bilden die Erfassung der Mittel- und Maximalwerte im Frühling, die durch die Multiplikative Korrektur besser wiedergegeben werden, und die Erfassung der Maximalwerte und der Standardabweichung im Sommer und im Herbst, die durch die Hybrid-Vorhersage besser erfasst werden. Nur im Winter werden die Maximalwerte am besten durch den Kalman-Filter erreicht. Korrelation und FGE zeigen zu allen Jahreszeiten für den Kalman-Filter bessere Werte als für die Hybrid-Vorhersage, umgekehrt ist der MFB für die Hybrid-Vorhersage zu allen Jahreszeiten bis auf den Herbst deutlich niedriger als für den Kalman-Filter. Im Winter werden die höchste Korrelation und der niedrigste FGE durch die CAMS-Rohdaten erreicht.

Für die maximalen gleitenden 8-Stundenmittelwerte eines Tages (8TMAX) in Tabelle 5-6 und Tabelle 5-7 ist das Verhältnis der Metriken der einzelnen Korrekturverfahren über die Jahreszeiten etwas konstanter als bei den 1SMW, entsprechend unterscheiden sich die Ergebnisse nicht wesentlich von der Auswertung der Jahresgesamtbetrachtung in Abschnitt 5.1.1.1. Hybrid-Vorhersage und Kalman-Filter erzeugen die besten Ergebnisse, wobei die Hybrid-Vorhersage im Regelfall die Maximalwerte und die Standardabweichung und anders als bei den 1SMW auch den Mittelwert am besten trifft, während der Kalman-Filter etwas bessere Werte bei Korrelation, RMSE, MFB und FGE aufweist. Aufgrund der besseren Erfassung der Maximalwerte erreicht die Hybrid-Vorhersage auch saisonal betrachtet mit 60 % bis 67 % die höchste Hit Rate. Gleichzeitig ist damit auch die FAR für die Hybrid-Vorhersage am höchsten, allerdings liegt der Wert mit 1.1 % (Frühling) bis 2.6 % (Sommer) auf niedrigem Niveau.

Für die maximalen Einstundenmittelwerte eines Tages (1TMAX) in Tabelle 5-8 und Tabelle 5-9 sind die berechneten Metriken für die vier Jahreszeiten deutlich heterogener als für die 8TMAX und auch als die 1SMW. Im Frühling erzielen die CAMS-Rohdaten die besten Metriken. Ausnahme sind die Maximalwerte und die Standardabweichung, für die mit den CAMS-Rohdaten die schlechtesten Werte berechnet werden. Multiplikative Korrektur und Kalman-Filter zeigen gute und mittlere Werte, das Hybrid-Verfahren schneidet bis auf die Standardabweichung, die korrekt erfasst wird, am schlechtesten ab. Im Sommer ergeben sich für die CAMS-Rohdaten die besten Werte für Korrelation und MFB, für den Kalman-Filter die besten Werte für RMSE und FGE, und für das Hybrid-Verfahren die besten Werte für Mittel- und Maximalwerte sowie Standardabweichung, allerdings auch die schlechtesten Werte für Korrelation, RMSE und FGE. Im Herbst ergeben sich für die CAMS-Rohdaten die schlechtesten Metriken im Vergleich mit den Korrekturverfahren, bis auf die Korrelation, die für die Rohdaten am besten ist. Die besten Metriken erzielt insgesamt der Kalman-Filter, bis auf die Maximalwerte und die Standardabweichung, die das Hybrid-Verfahren am besten erfasst. Im Winter werden die Mittelwerte am besten durch die CAMS-Rohdaten erfasst, die Maximalwerte durch den Kalman-Filter und die Standardabweichung durch das Hybrid-Verfahren. Für alle anderen Metriken erzielt der Kalman-Filter die besten Werte, bis auf den Bias, der bei den CAMS-Rohdaten noch niedriger liegt. Die Minimalwerte werden für alle Jahreszeiten am besten durch die Hybrid-Vorhersage wiedergegeben.

Zusammenfassend lässt sich sagen, dass in den überschreitungsrelevanten Jahreszeiten Frühjahr, Sommer und Herbst die überschreitungsrelevanten Größen, d. h. Maximalwerte und Standardabweichung, im Regelfall am besten durch die Hybrid-Vorhersage erfasst werden. Ausnahme davon ist der Frühling, in dem die Maximalwerte besser durch die Multiplikative Korrektur erfasst werden. Die Hybrid-Vorhersage führt hier im Median eher zu einer Überschätzung der Maximalwerte. Für Metriken wie Korrelation, RMSE, MFB und FGE erzielt häufig, aber nicht in allen Fällen der Kalman-Filter die besten Ergebnisse. Überschreitungen werden am besten durch die Hybrid-Vorhersage erfasst.

Tabelle 5-4: Metriken O₃ für 1SMW, 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Frühling und Sommer 2016 (jeweils Median über alle Stationen)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	1.25	1.31	1.31	1.25
Min Mod [$\mu\text{g}/\text{m}^3$]	11.65	4.78	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	57.60	58.95	58.15	58.11
Mean Mod [$\mu\text{g}/\text{m}^3$]	68.62	59.04	58.71	58.42
Max Obs [$\mu\text{g}/\text{m}^3$]	135.97	135.97	135.97	135.97
Max Mod [$\mu\text{g}/\text{m}^3$]	121.34	137.56	133.75	140.28
SD Obs [$\mu\text{g}/\text{m}^3$]	26.67	26.82	26.76	26.67
SD Mod [$\mu\text{g}/\text{m}^3$]	18.39	23.67	23.53	26.61
Korrelation	0.76	0.75	0.78	0.72
BIAS [$\mu\text{g}/\text{m}^3$]	8.73	0.22	0.55	0.34
RMSE [$\mu\text{g}/\text{m}^3$]	19.19	17.59	16.40	19.13
MFB [%]	23.11	5.54	5.25	-0.24
FGE [%]	32.43	30.96	29.86	36.59
Sommer				
Min Obs [$\mu\text{g}/\text{m}^3$]	1.19	1.32	1.20	1.20
Min Mod [$\mu\text{g}/\text{m}^3$]	31.79	4.87	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	55.96	56.70	55.91	56.24
Mean Mod [$\mu\text{g}/\text{m}^3$]	70.48	56.17	55.74	56.05
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	173.51	173.66	173.51
Max Mod [$\mu\text{g}/\text{m}^3$]	139.11	162.03	152.84	169.59
SD Obs [$\mu\text{g}/\text{m}^3$]	29.45	29.86	29.67	29.66
SD Mod [$\mu\text{g}/\text{m}^3$]	16.72	26.75	26.36	29.51
Korrelation	0.81	0.81	0.83	0.78
BIAS [$\mu\text{g}/\text{m}^3$]	13.36	-0.63	-0.33	-0.16
RMSE [$\mu\text{g}/\text{m}^3$]	23.05	17.70	16.77	19.80
MFB [%]	34.41	4.05	3.39	-0.91
FGE [%]	39.76	30.17	30.13	36.14

Tabelle 5-5: Metriken O₃ für 1SMW, 1. Prognostetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Herbst und Winter (jeweils Median über alle Stationen)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	0.60	0.60	0.60	0.60
Min Mod [$\mu\text{g}/\text{m}^3$]	1.88	0.41	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	33.40	33.01	32.20	33.26
Mean Mod [$\mu\text{g}/\text{m}^3$]	48.82	32.81	32.64	33.81
Max Obs [$\mu\text{g}/\text{m}^3$]	155.00	154.86	155.00	155.00
Max Mod [$\mu\text{g}/\text{m}^3$]	124.94	149.28	143.35	154.15
SD Obs [$\mu\text{g}/\text{m}^3$]	28.73	28.13	27.99	28.42
SD Mod [$\mu\text{g}/\text{m}^3$]	21.02	25.65	25.57	28.01
Korrelation	0.83	0.82	0.85	0.82
BIAS [$\mu\text{g}/\text{m}^3$]	13.50	0.03	0.58	0.71
RMSE [$\mu\text{g}/\text{m}^3$]	20.63	15.98	14.84	16.88
MFB [%]	57.70	13.09	3.69	-4.27
FGE [%]	64.89	55.38	60.31	67.37
Winter				
Min Obs [$\mu\text{g}/\text{m}^3$]	0.60	0.60	0.60	0.60
Min Mod [$\mu\text{g}/\text{m}^3$]	1.11	0.39	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	31.75	32.06	32.11	31.90
Mean Mod [$\mu\text{g}/\text{m}^3$]	38.92	31.36	32.17	32.75
Max Obs [$\mu\text{g}/\text{m}^3$]	83.28	83.17	83.28	83.21
Max Mod [$\mu\text{g}/\text{m}^3$]	74.75	90.33	80.39	91.88
SD Obs [$\mu\text{g}/\text{m}^3$]	21.96	22.20	22.23	22.04
SD Mod [$\mu\text{g}/\text{m}^3$]	15.15	18.32	17.80	21.01
Korrelation	0.76	0.68	0.75	0.70
BIAS [$\mu\text{g}/\text{m}^3$]	5.20	-0.78	0.04	0.66
RMSE [$\mu\text{g}/\text{m}^3$]	15.77	16.34	14.51	16.39
MFB [%]	36.78	11.44	6.90	-2.46
FGE [%]	55.79	56.38	57.81	65.12

Tabelle 5-6: Metriken O₃ für 8TMAX, 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Frühling und Sommer 2016 (jeweils Median über alle Stationen)

Metrik	CAMS	MK	KF	HV
	Frühling			
Min Obs [$\mu\text{g}/\text{m}^3$]	37.03	37.26	37.03	37.03
Min Mod [$\mu\text{g}/\text{m}^3$]	44.40	40.56	40.63	37.57
Mean Obs [$\mu\text{g}/\text{m}^3$]	79.18	79.99	80.01	79.43
Mean Mod [$\mu\text{g}/\text{m}^3$]	82.28	79.05	79.01	80.37
Max Obs [$\mu\text{g}/\text{m}^3$]	130.17	129.82	130.17	130.17
Max Mod [$\mu\text{g}/\text{m}^3$]	115.83	128.37	126.50	131.59
SD Obs [$\mu\text{g}/\text{m}^3$]	19.57	19.40	19.40	19.56
SD Mod [$\mu\text{g}/\text{m}^3$]	15.20	17.84	17.27	19.04
Korrelation	0.81	0.80	0.84	0.80
BIAS [$\mu\text{g}/\text{m}^3$]	2.13	-0.90	-0.94	0.91
RMSE [$\mu\text{g}/\text{m}^3$]	12.22	12.02	10.71	12.37
MFB [%]	4.04	-0.75	-0.62	1.27
FGE [%]	12.15	12.42	11.29	12.82
gemessene Überschreitungen	4 ¹	3 ¹	4 ¹	4 ¹
Hit Rate [%]	0.00 ¹	50.00 ¹	40.00 ¹	60.00¹
FAR [%]	0.00¹	0.00¹	0.00¹	1.14 ¹
Sommer				
Min Obs [$\mu\text{g}/\text{m}^3$]	39.02	39.90	39.14	39.14
Min Mod [$\mu\text{g}/\text{m}^3$]	65.36	46.63	47.67	39.69
Mean Obs [$\mu\text{g}/\text{m}^3$]	83.36	83.77	83.27	83.64
Mean Mod [$\mu\text{g}/\text{m}^3$]	85.76	81.22	81.12	83.09
Max Obs [$\mu\text{g}/\text{m}^3$]	154.62	152.66	154.62	153.99
Max Mod [$\mu\text{g}/\text{m}^3$]	130.41	147.63	141.63	153.26
SD Obs [$\mu\text{g}/\text{m}^3$]	22.29	22.59	22.41	22.41
SD Mod [$\mu\text{g}/\text{m}^3$]	13.07	20.16	18.89	22.55
Korrelation	0.81	0.79	0.82	0.80
BIAS [$\mu\text{g}/\text{m}^3$]	2.27	-2.59	-2.36	-0.51
RMSE [$\mu\text{g}/\text{m}^3$]	15.68	14.75	13.64	14.73
MFB [%]	5.23	-2.33	-1.70	-0.58
FGE [%]	14.90	14.05	13.03	14.30
gemessene Überschreitungen	7 ¹	6 ¹	6 ¹	7 ¹
Hit Rate [%]	28.57 ¹	50.00 ¹	50.00 ¹	63.64¹
FAR [%]	0.00¹	1.41 ¹	1.32 ¹	2.63 ¹

¹: Schwellenwert 120 $\mu\text{g}/\text{m}^3$

Tabelle 5-7: Metriken O₃ für 8TMAX, 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Herbst und Winter 2016 (jeweils Median über alle Stationen)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	4.42	4.90	4.44	4.47
Min Mod [$\mu\text{g}/\text{m}^3$]	22.35	10.57	8.04	5.10
Mean Obs [$\mu\text{g}/\text{m}^3$]	52.94	51.37	51.02	52.26
Mean Mod [$\mu\text{g}/\text{m}^3$]	61.03	49.27	50.12	53.57
Max Obs [$\mu\text{g}/\text{m}^3$]	140.54	140.12	140.54	140.51
Max Mod [$\mu\text{g}/\text{m}^3$]	115.72	133.18	129.14	139.04
SD Obs [$\mu\text{g}/\text{m}^3$]	31.54	30.59	30.87	31.17
SD Mod [$\mu\text{g}/\text{m}^3$]	21.84	29.18	28.46	29.97
Korrelation	0.91	0.89	0.91	0.91
BIAS [$\mu\text{g}/\text{m}^3$]	7.22	-1.60	-0.76	1.27
RMSE [$\mu\text{g}/\text{m}^3$]	16.52	13.66	12.17	12.81
MFB [%]	24.71	-0.82	1.45	4.40
FGE [%]	31.80	28.81	25.77	26.34
gemessene Überschreitungen	4 ¹	4 ¹	4 ¹	4 ¹
Hit Rate [%]	0.00 ¹	60.00 ¹	50.00 ¹	66.67¹
FAR [%]	0.00¹	0.00¹	0.00¹	1.16 ¹
Winter				
Min Obs [$\mu\text{g}/\text{m}^3$]	3.83	3.94	3.83	3.94
Min Mod [$\mu\text{g}/\text{m}^3$]	18.00	7.30	5.83	4.58
Mean Obs [$\mu\text{g}/\text{m}^3$]	43.45	43.77	43.97	43.68
Mean Mod [$\mu\text{g}/\text{m}^3$]	47.31	41.22	41.91	45.34
Max Obs [$\mu\text{g}/\text{m}^3$]	78.44	78.25	78.44	78.36
Max Mod [$\mu\text{g}/\text{m}^3$]	71.12	78.52	73.72	81.45
SD Obs [$\mu\text{g}/\text{m}^3$]	20.13	20.15	20.37	20.18
SD Mod [$\mu\text{g}/\text{m}^3$]	12.69	17.68	15.95	18.58
Korrelation	0.79	0.72	0.81	0.80
BIAS [$\mu\text{g}/\text{m}^3$]	2.55	-2.37	-1.70	1.90
RMSE [$\mu\text{g}/\text{m}^3$]	13.57	14.43	11.96	12.42
MFB [%]	16.43	-0.18	2.32	7.48
FGE [%]	32.22	32.55	29.01	28.42
gemessene Überschreitungen	0 ¹	0 ¹	0 ¹	0 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹

¹: Schwellenwert 120 $\mu\text{g}/\text{m}^3$

Tabelle 5-8: Metriken O₃ für 1TMAX, 1. Prognostag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Frühling und Sommer 2016 (jeweils Median über alle Stationen)

Metrik	CAMS	MK	KF	HV
	Frühling			
Min Obs [$\mu\text{g}/\text{m}^3$]	41.59	41.89	41.94	41.69
Min Mod [$\mu\text{g}/\text{m}^3$]	48.39	45.18	47.00	43.91
Mean Obs [$\mu\text{g}/\text{m}^3$]	85.89	86.37	86.62	86.16
Mean Mod [$\mu\text{g}/\text{m}^3$]	86.31	85.70	85.24	88.11
Max Obs [$\mu\text{g}/\text{m}^3$]	135.94	135.94	135.94	135.94
Max Mod [$\mu\text{g}/\text{m}^3$]	121.41	137.50	133.58	140.15
SD Obs [$\mu\text{g}/\text{m}^3$]	19.63	19.60	19.56	19.58
SD Mod [$\mu\text{g}/\text{m}^3$]	15.62	18.13	17.34	19.27
Korrelation	0.80	0.76	0.81	0.76
BIAS [$\mu\text{g}/\text{m}^3$]	-0.19	-0.73	-1.24	2.23
RMSE [$\mu\text{g}/\text{m}^3$]	12.18	13.16	11.60	13.62
MFB [%]	0.59	-0.54	-1.00	2.75
FGE [%]	11.16	12.30	10.89	12.52
gemessene Überschreitungen	0	0	0	0
Hit Rate [%]	0.00	0.00	0.00	0.00
FAR [%]	0.00	0.00	0.00	0.00
Sommer				
Min Obs [$\mu\text{g}/\text{m}^3$]	40.60	40.90	40.60	40.60
Min Mod [$\mu\text{g}/\text{m}^3$]	66.91	50.63	52.39	44.07
Mean Obs [$\mu\text{g}/\text{m}^3$]	91.12	91.20	90.89	91.35
Mean Mod [$\mu\text{g}/\text{m}^3$]	88.85	88.25	87.72	91.81
Max Obs [$\mu\text{g}/\text{m}^3$]	173.66	172.73	173.66	173.66
Max Mod [$\mu\text{g}/\text{m}^3$]	139.11	161.81	152.74	169.39
SD Obs [$\mu\text{g}/\text{m}^3$]	25.05	25.61	25.32	25.29
SD Mod [$\mu\text{g}/\text{m}^3$]	14.29	22.29	20.41	24.37
Korrelation	0.83	0.80	0.81	0.77
BIAS [$\mu\text{g}/\text{m}^3$]	-1.72	-3.48	-3.62	0.47
RMSE [$\mu\text{g}/\text{m}^3$]	16.42	16.09	15.32	16.74
MFB [%]	0.71	-3.02	-2.62	0.77
FGE [%]	14.16	14.12	13.35	14.53
gemessene Überschreitungen	0	0	0	0
Hit Rate [%]	0.00	0.00	0.00	0.00
FAR [%]	0.00	0.00	0.00	0.00

¹: Schwellenwert 180 $\mu\text{g}/\text{m}^3$ (Informationsschwelle)

Tabelle 5-9: Metriken O₃ für 1TMAX, 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Zeiträume Herbst und Winter 2016 (jeweils Median über alle Stationen)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	6.79	6.98	6.91	6.84
Min Mod [$\mu\text{g}/\text{m}^3$]	25.02	13.44	12.09	9.24
Mean Obs [$\mu\text{g}/\text{m}^3$]	61.51	59.90	59.47	60.89
Mean Mod [$\mu\text{g}/\text{m}^3$]	64.73	57.46	57.75	62.81
Max Obs [$\mu\text{g}/\text{m}^3$]	155.00	154.66	155.00	155.00
Max Mod [$\mu\text{g}/\text{m}^3$]	124.87	149.28	143.35	154.15
SD Obs [$\mu\text{g}/\text{m}^3$]	33.97	33.03	33.17	33.63
SD Mod [$\mu\text{g}/\text{m}^3$]	22.72	31.59	30.67	32.28
Korrelation	0.91	0.88	0.90	0.89
BIAS [$\mu\text{g}/\text{m}^3$]	2.95	-2.35	-1.73	1.97
RMSE [$\mu\text{g}/\text{m}^3$]	16.52	15.67	14.00	14.92
MFB [%]	13.45	-2.18	-0.86	5.01
FGE [%]	25.75	27.05	24.44	25.26
gemessene Überschreitungen	0	0	0	0
Hit Rate [%]	0.00	0.00	0.00	0.00
FAR [%]	0.00	0.00	0.00	0.00
Winter				
Min Obs [$\mu\text{g}/\text{m}^3$]	5.82	5.91	5.82	5.91
Min Mod [$\mu\text{g}/\text{m}^3$]	21.10	9.58	8.61	7.28
Mean Obs [$\mu\text{g}/\text{m}^3$]	49.47	49.68	50.00	49.60
Mean Mod [$\mu\text{g}/\text{m}^3$]	50.41	47.47	47.33	52.16
Max Obs [$\mu\text{g}/\text{m}^3$]	83.24	83.17	83.24	83.24
Max Mod [$\mu\text{g}/\text{m}^3$]	74.75	90.33	80.46	91.88
SD Obs [$\mu\text{g}/\text{m}^3$]	20.26	20.40	20.43	20.29
SD Mod [$\mu\text{g}/\text{m}^3$]	12.44	18.78	16.13	19.19
Korrelation	0.75	0.69	0.77	0.75
BIAS [$\mu\text{g}/\text{m}^3$]	-0.19	-2.25	-2.14	2.87
RMSE [$\mu\text{g}/\text{m}^3$]	14.12	15.74	12.98	14.08
MFB [%]	8.01	-1.16	-0.04	7.87
FGE [%]	28.30	29.44	26.39	27.11
gemessene Überschreitungen	0	0	0	0
Hit Rate [%]	0.00	0.00	0.00	0.00
FAR [%]	0.00	0.00	0.00	0.00

¹: Schwellenwert 180 $\mu\text{g}/\text{m}^3$ (Informationsschwelle)

5.1.1.3 Episode 2015

Aufgrund der geringen Anzahl Schwellenwertüberschreitungen für Ozon im Jahr 2016 wurden die Korrekturverfahren auf eine Episode im Jahr 2015 (11. Juni bis 9. Juli 2015) angewendet, die zumindest eine Reihe von Überschreitungen des Schwellenwerts von $120 \mu\text{g}/\text{m}^3$ für die 8TMAX aufweist.

Laut DWD war der Juni 2015 eher wechselhaft. Sommerliche und hochsommerliche Temperaturen traten nur um den 5. Juni und dann wieder um den 12. Juni auf. Gleichzeitig war der Juni 2015 eher trocken, auch wenn es vereinzelt zu Starkregenereignissen vor allem in Süddeutschland kam (DWD, 2015a). Im Gegensatz dazu brachte der Juli 2015 in seinem ersten Drittel in weiten Teilen Deutschlands eine der heftigsten Hitzewellen seit Beginn der Wetteraufzeichnungen. Ausreichend Regen fiel nur in der Nordhälfte Deutschlands, insbesondere ein breiter Streifen in der Mitte Deutschlands war erheblich von Trockenheit betroffen. Der Juli 2015 verlief sehr sonnenscheinreich (DWD, 2015b).

Für die 1SMW in Tabelle 5-10 sind die Ergebnisse ähnlich wie für das Jahr 2016 in Tabelle 5-1, d. h. der Kalman-Filter erzeugt insgesamt gute Metriken, wird allerdings hinsichtlich Maximalwerten und Standardabweichung deutlich von der Hybrid-Vorhersage übertroffen, und im Gegensatz zu 2016 weist die Multiplikative Korrektur etwas bessere Werte hinsichtlich Korrelation und RMSE auf.

Für die 8TMAX in Tabelle 5-11 sind Kalman-Filter und Hybrid-Vorhersage ähnlich gut; die Hybrid-Vorhersage überschätzt und der Kalman-Filter unterschätzt etwas die gemessenen Maximalwerte. Die Hit Rate ist am 1. Prognosetag (Tabelle 5-11) für KF und HV identisch und mit im Median 78 % deutlich höher als für die CAMS-Rohdaten mit im Median 20%. Für den 2. und 3. Prognosetag (hier nicht gezeigt) ist die Hit Rate für den Kalman-Filter mit im Median 75% etwas niedriger als für die Hybrid-Vorhersage mit im Median 78 %. Die FAR ist an allen Prognosetagen für die Hybrid-Vorhersage etwas größer als für den Kalman-Filter und beträgt am 3. Prognosetag im Median rund 7 %.

Für die 1TMAX in Tabelle 5-12 zeigt die Hybrid-Vorhersage in allen relevanten Größen die besten bzw. für Bias und MFB die zweitbesten Werte. Ausnahme ist die Korrelation, die hier für die CAMS-Rohdaten am höchsten ist, allerdings für die Korrekturverfahren auch nicht wesentlich geringer ist. Von den im Median gemessenen Überschreitungen der Informationsschwelle von $180 \mu\text{g}/\text{m}^3$ wird im Median keine erkannt, weder durch die CAMS-Rohdaten noch durch die korrigierten Modellprognosen. Auswertungen der Box-Whisker-Plots der Hit Rate (Abbildung 5-10) zeigen aber, dass im oberen Quartil bis zu 67 % und im Maximum bis zu 100 % der Überschreitungen der Informationsschwelle durch die Korrekturverfahren erkannt werden. Den niedrigsten Wert im oberen Quartil weist hier mit 50 % der Kalman-Filter auf. Durch die unkorrigierten CAMS-Prognosen werden keine Überschreitungen der Informationsschwelle erkannt.

Tabelle 5-10: Metriken O₃, Stundenmittelwerte 1SMW Episode 2015 für den 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	3.25	4.17	3.50	3.62
Min Mod [$\mu\text{g}/\text{m}^3$]	36.38	14.31	5.45	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	71.62	73.14	71.14	71.55
Mean Mod [$\mu\text{g}/\text{m}^3$]	72.42	72.40	71.41	71.41
Max Obs [$\mu\text{g}/\text{m}^3$]	193.00	192.16	193.00	192.65
Max Mod [$\mu\text{g}/\text{m}^3$]	134.24	188.50	182.66	194.19
SD Obs [$\mu\text{g}/\text{m}^3$]	36.07	36.78	36.53	35.95
SD Mod [$\mu\text{g}/\text{m}^3$]	17.36	34.50	34.35	36.29
Korrelation	0.82	0.84	0.83	0.81
BIAS [$\mu\text{g}/\text{m}^3$]	-0.32	-0.85	0.32	-0.39
RMSE [$\mu\text{g}/\text{m}^3$]	24.48	20.76	21.10	22.43
MFB [%]	9.40	1.78	2.04	-1.57
FGE [%]	30.10	26.37	27.40	30.27

Tabelle 5-11: Metriken O₃, maximale gleitende 8-Stundenmittelwerte eines Tages 8TMAX Episode 2015 für den 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	54.54	55.25	55.07	55.07
Min Mod [$\mu\text{g}/\text{m}^3$]	62.64	55.20	53.82	52.89
Mean Obs [$\mu\text{g}/\text{m}^3$]	102.43	101.57	101.52	102.29
Mean Mod [$\mu\text{g}/\text{m}^3$]	87.60	99.76	100.42	102.17
Max Obs [$\mu\text{g}/\text{m}^3$]	174.25	171.73	174.25	173.68
Max Mod [$\mu\text{g}/\text{m}^3$]	126.86	173.63	170.52	177.55
SD Obs [$\mu\text{g}/\text{m}^3$]	32.84	33.90	33.75	32.58
SD Mod [$\mu\text{g}/\text{m}^3$]	15.58	33.00	31.40	32.75
Korrelation	0.88	0.87	0.89	0.88
BIAS [$\mu\text{g}/\text{m}^3$]	-14.28	-2.09	-1.07	-0.28
RMSE [$\mu\text{g}/\text{m}^3$]	25.10	16.58	15.50	15.42
MFB [%]	-11.41	-1.61	-0.37	-0.09
FGE [%]	20.44	13.82	12.99	12.63
gemessene Überschreitungen	8 ¹	5 ¹	7 ¹	7 ¹
Hit Rate [%]	20.00 ¹	75.00 ¹	77.78¹	77.78¹
FAR [%]	0.00¹	0.00¹	4.55 ¹	5.00 ¹

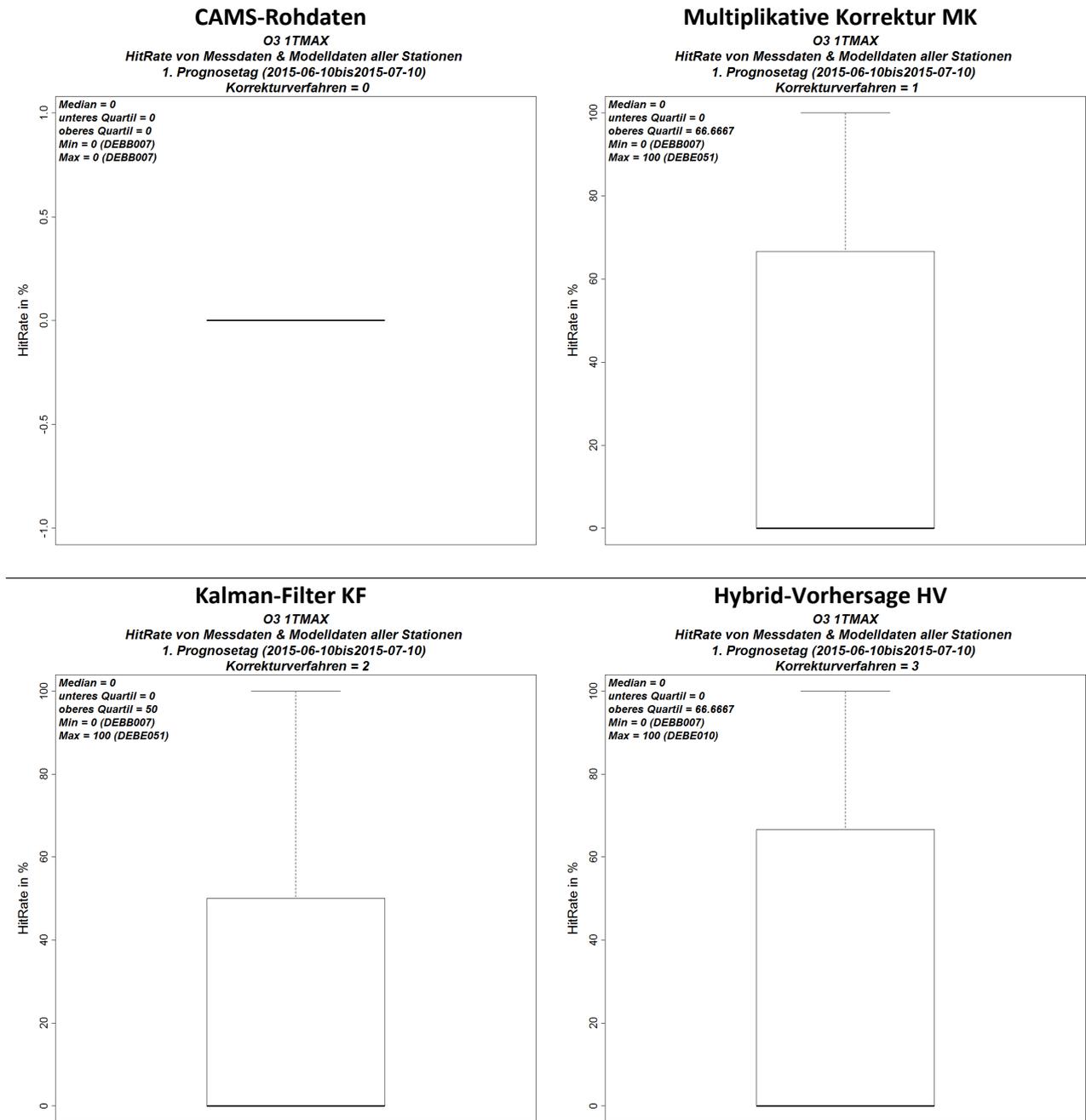
¹: Schwellenwert 120 $\mu\text{g}/\text{m}^3$

Tabelle 5-12: Metriken O₃, maximale Stundenmittelwerte eines Tages 1TMAX Episode 2015 für den 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	59.83	60.49	60.27	60.27
Min Mod [$\mu\text{g}/\text{m}^3$]	61.40	58.61	59.01	58.07
Mean Obs [$\mu\text{g}/\text{m}^3$]	110.10	109.76	108.68	108.74
Mean Mod [$\mu\text{g}/\text{m}^3$]	90.60	107.94	108.34	109.96
Max Obs [$\mu\text{g}/\text{m}^3$]	192.65	191.02	192.65	192.02
Max Mod [$\mu\text{g}/\text{m}^3$]	133.38	187.21	181.98	193.13
SD Obs [$\mu\text{g}/\text{m}^3$]	35.73	36.50	36.80	35.68
SD Mod [$\mu\text{g}/\text{m}^3$]	17.12	36.06	33.38	35.27
Korrelation	0.89	0.87	0.86	0.85
BIAS [$\mu\text{g}/\text{m}^3$]	-18.36	-2.86	-0.67	1.03
RMSE [$\mu\text{g}/\text{m}^3$]	28.84	18.93	18.95	18.84
MFB [%]	-14.21	-1.96	0.51	1.21
FGE [%]	21.81	14.41	14.25	14.02
gemessene Überschreitungen	1 ¹	1 ¹	1 ¹	1 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹

¹: Schwellenwert 180 $\mu\text{g}/\text{m}^3$ (Informationsschwelle)

Abbildung 5-10: Box-Whisker-Plots der Hit Rate bezüglich der Informationsschwelle für O₃ an den Stationsorten, Episode 2015, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)

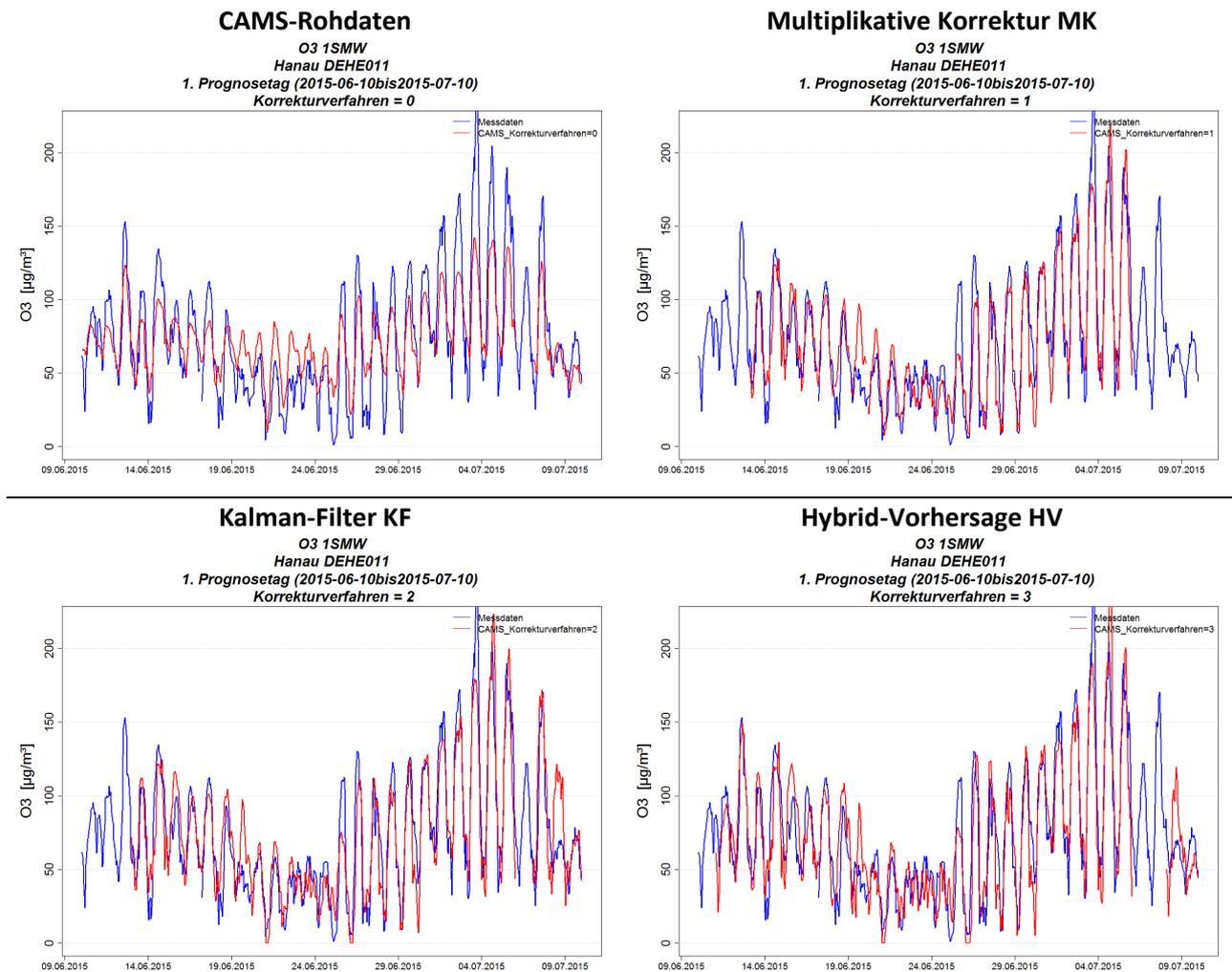


Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-11 zeigt beispielhaft Zeitreihen der Stundenmittelwerte für die Episode 2015 für die Station DEHE011, die mit 0.9 den höchsten Korrelationskoeffizienten aller Stationen für diese Episode aufweist. Die CAMS-Rohdaten erfassen den Tagesgang recht gut, unterschätzen aber deutlich die Maxima und überschätzen die Minima. Insbesondere die hohen Konzentrationswerte um den 04.07.2015 werden durch das Modell kaum erkannt. Alle drei Korrekturverfahren erfassen die Amplitude des Ta-

gangs und den Anstieg der Konzentrationen um den 04.07.2015 deutlich besser, wobei die korrigierten Modellprognosen ihre höchsten Werte erst nach dem absoluten Maximum der Messwerte erreichen.

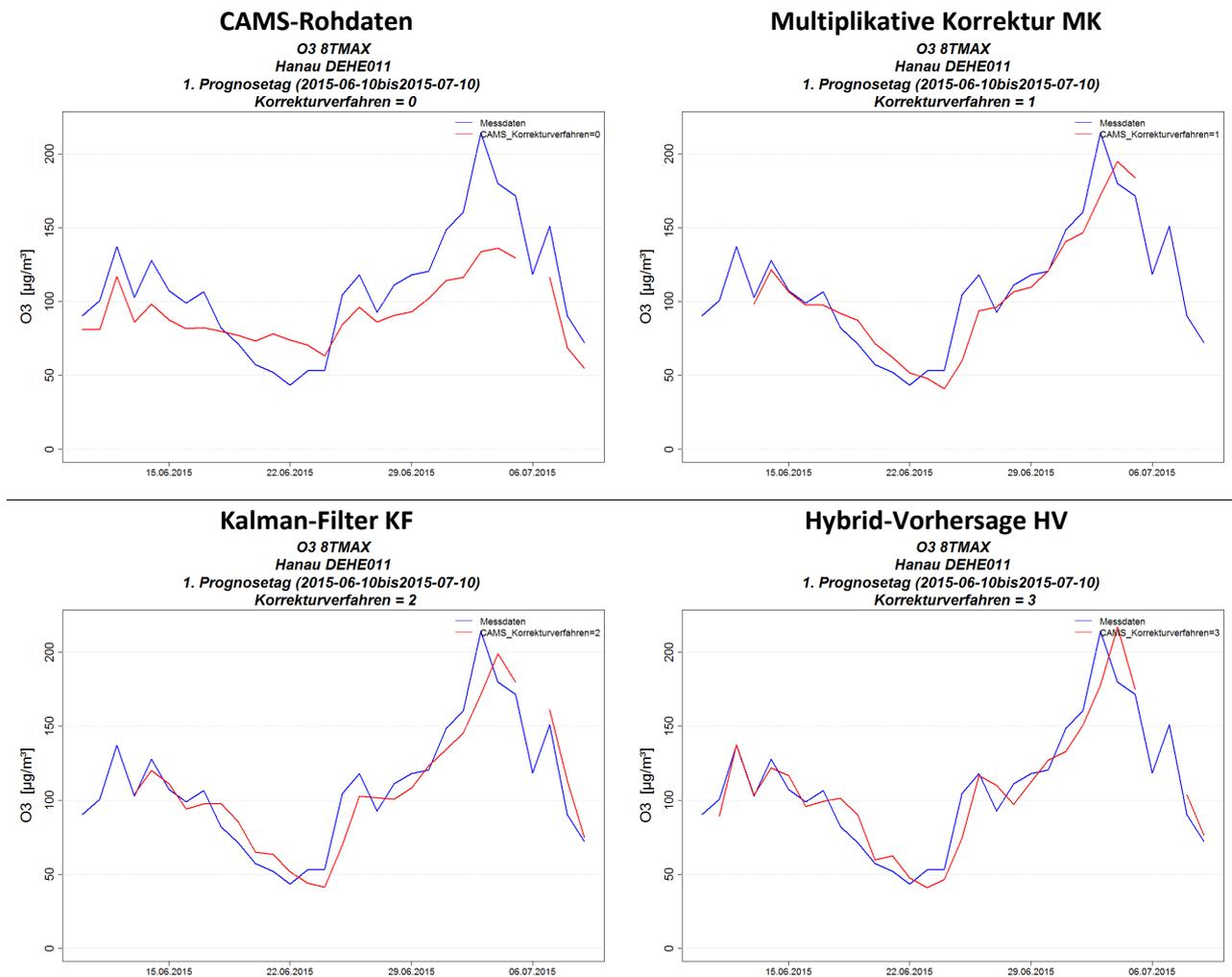
Abbildung 5-11: Zeitreihen der 1SMW für O₃, Station DEHE011, Episode 2015, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

In Abbildung 5-12 sind die zugehörigen Zeitreihen der 8TMAX für diese Episode dargestellt. Die CAMS-Rohdaten zeigen zwar gegen Anfang Juli einen leichten Anstieg, verfehlen das eigentliche Maximum jedoch deutlich, während die Minimalwerte um den 22.06.2015 durch die CAMS-Rohdaten überschätzt werden. Alle drei Korrekturverfahren führen hier zu einer deutlichen Verbesserung im Konzentrationsverlauf der 8TMAX-Werte, wobei der Verlauf in allen drei Fällen zeitlich versetzt ist zum Verlauf der 8TMAX-Werte der Messdaten. Grund dafür ist, dass die Korrektur in der Prognose durchgeführt wird auf Basis von Mess- und Modellwerten von bis zu drei vorangegangenen Tagen. Die Maximalwerte werden am besten von der Hybrid-Vorhersage getroffen.

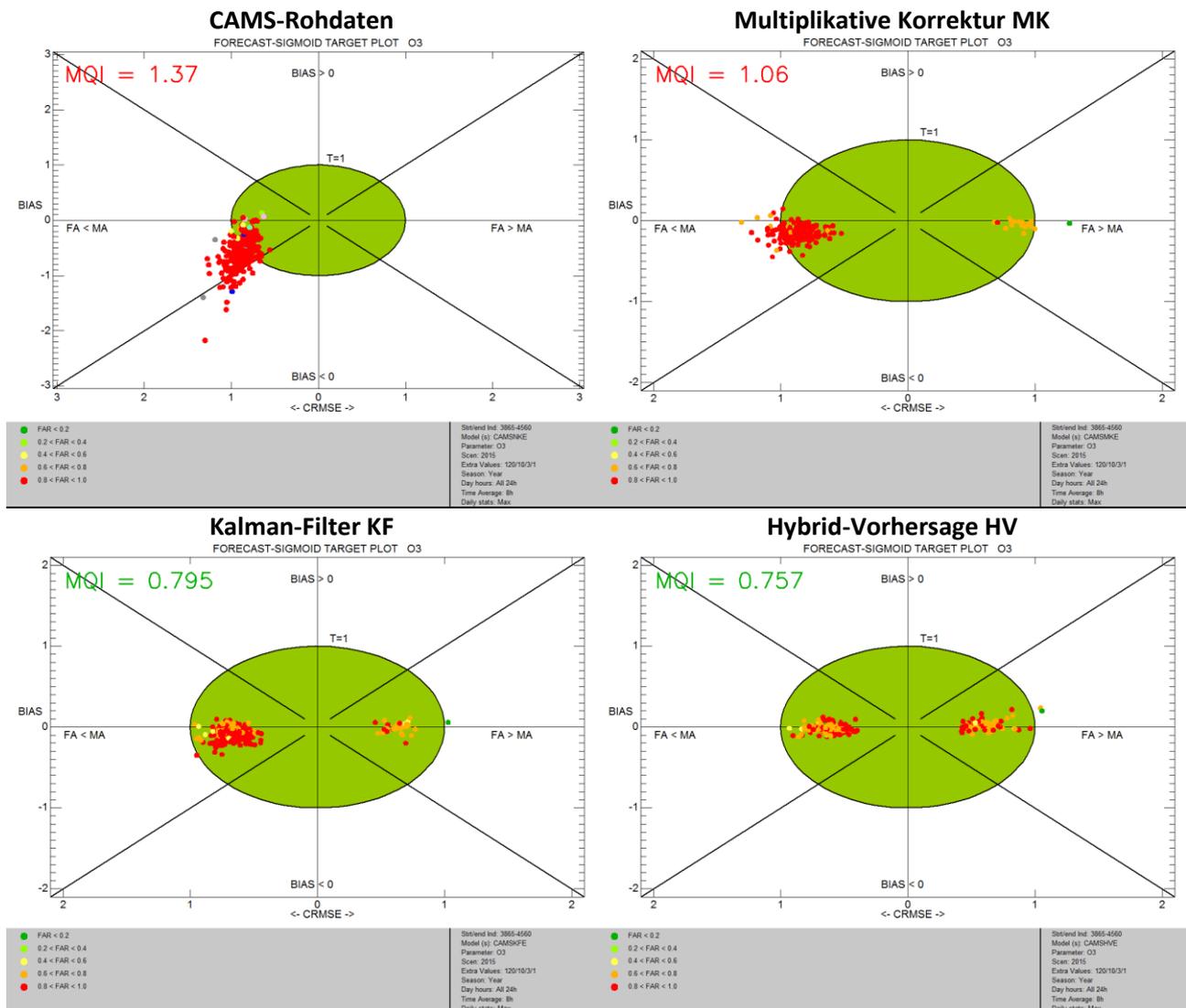
Abbildung 5-12: Zeitreihen der 8TMAX für O₃, Station DEHE011, Episode 2015, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-13 zeigt die im Forecast-Modus des DELTA-Tools erstellten Target Plots für die Episode 2015 für den 1. Prognosetag. Wie schon 2016 ist der MQI der CAMS-Rohdaten für eine ganze Reihe von Stationen größer 1, d. h. die CAMS-Rohdaten sind dort schlechter als eine Modellierung mit dem persistenten Modell. Das 90%-Perzentil der MQI aller Stationen liegt entsprechend mit 1.37 deutlich über 1. Die Multiplikative Korrektur zeigt in der Episode keine erkennbaren Ausreißer, ihr 90%-Perzentil der MQI aller Stationen liegt aber mit 1.06 immer noch höher als 1 und ist damit schlechter als das persistente Modell. Das niedrigste 90%-Perzentil der MQI aller Stationen weist in der Episode das Hybrid-Verfahren mit 0.757 auf.

Abbildung 5-13: Forecast-Sigmoid Target Plots für O₃, Episode 2015, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

5.1.2 PM10

5.1.2.1 Jahreszeitreihe 2016

Die Jahreszeitreihen der CAMS-PM10-Prognosen für 2016 an den Messorten wurden analog zu Ozon für die drei betrachteten Prognosetage einer Korrektur durch die in Abschnitt 3.3 beschriebenen Korrekturverfahren unterzogen. Für die korrigierten CAMS-PM10-Prognosen wurde dann ebenfalls eine statistische Auswertung mit Hilfe der in Abschnitt 3.4.1 aufgeführten Metriken und auf Basis der in Abschnitt 3.4.2 erläuterten Target-Plots des DELTA-Tools durchgeführt.

In Tabelle 5-13 sind die für den 1. Prognosetag für die 1SMW berechneten Metriken aufgeführt. Es zeigt sich, dass die CAMS-Rohdaten für PM10 die Messwerte im Mittel und in den Maximalwerten stark unterschätzen. Dies steht im Kontrast zu Ozon (Abschnitt 5.1.1), wo die CAMS-Rohdaten die Messwerte im Mittel über- und nur in den Maximalwerten unterschätzen. Die Ergebnisse bezüglich der Metriken sind ähnlich wie für Ozon im Jahr 2016 in Tabelle 5-1. Grundsätzlich führt jedes Korrektur-

verfahren zu einer Verbesserung der CAMS-Prognosen. Der Kalman-Filter erzeugt insgesamt gute Metriken, wird allerdings hinsichtlich der Maximalwerte von der Hybrid-Vorhersage übertroffen, und die besten Metriken für die Standardabweichung zeigt hier die Multiplikative Korrektur.

Tabelle 5-13: Metriken PM10, Stundenmittelwerte 1SMW für den 1. Prognosetag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	0.50	0.50	0.50	0.50
Min Mod [$\mu\text{g}/\text{m}^3$]	0.91	0.65	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.24	16.16	16.13	16.21
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.78	16.37	16.17	16.31
Max Obs [$\mu\text{g}/\text{m}^3$]	141.93	127.50	128.45	128.45
Max Mod [$\mu\text{g}/\text{m}^3$]	50.67	116.15	90.12	137.48
SD Obs [$\mu\text{g}/\text{m}^3$]	11.57	11.29	11.30	11.36
SD Mod [$\mu\text{g}/\text{m}^3$]	6.29	11.41	10.10	11.89
Korrelation	0.50	0.56	0.58	0.53
BIAS [$\mu\text{g}/\text{m}^3$]	-3.79	0.28	0.10	0.11
RMSE [$\mu\text{g}/\text{m}^3$]	10.87	10.55	9.68	11.09
MFB [%]	-16.34	3.78	2.23	-4.43
FGE [%]	48.65	44.02	45.95	52.05

Tabelle 5-14 zeigt für den 1. Prognosetag die für die Tagesmittelwerte 1TMW berechneten Metriken. Für die meisten Größen zeigt der Kalman-Filter die besten Werte, die beste Erfassung der Maximalwerte und der Standardabweichung zeigt diesmal aber nicht die Hybrid-Vorhersage, sondern die Multiplikative Korrektur. Die im Median gemessenen Überschreitungen des Schwellenwerts von $50 \mu\text{g}/\text{m}^3$ werden von keinem der Verfahren angezeigt. Auswertungen der Box-Whisker-Plots der Hit Rate (hier nicht gezeigt) zeigen aber, dass im oberen Quartil bis zu 50 % und im Maximum bis zu 100 % der Überschreitungen des Schwellenwerts durch die Korrekturverfahren erkannt werden. Die unkorrigierten CAMS-Rohdaten erkennen im Maximum 25% der Überschreitungen.

Tabelle 5-14: Metriken PM10, Tagesmittelwerte 1TMW für den 1. Prognosetag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	2.02	2.05	2.02	2.03
Min Mod [$\mu\text{g}/\text{m}^3$]	3.10	2.79	1.97	1.41
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.27	16.18	16.15	16.22
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.77	16.38	16.16	16.34
Max Obs [$\mu\text{g}/\text{m}^3$]	72.84	66.99	67.31	67.31
Max Mod [$\mu\text{g}/\text{m}^3$]	35.88	65.94	58.33	71.18
SD Obs [$\mu\text{g}/\text{m}^3$]	9.20	9.05	8.93	9.06
SD Mod [$\mu\text{g}/\text{m}^3$]	5.48	9.19	8.23	9.24
Korrelation	0.58	0.69	0.72	0.72
BIAS [$\mu\text{g}/\text{m}^3$]	-3.80	0.29	0.09	0.11
RMSE [$\mu\text{g}/\text{m}^3$]	8.29	7.25	6.27	6.77
MFB [%]	-21.50	2.49	2.47	0.33
FGE [%]	39.01	29.64	27.62	28.81
gemessene Überschreitungen	2 ¹	2 ¹	2 ¹	2 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00¹	0.59 ¹	0.28 ¹	0.28 ¹

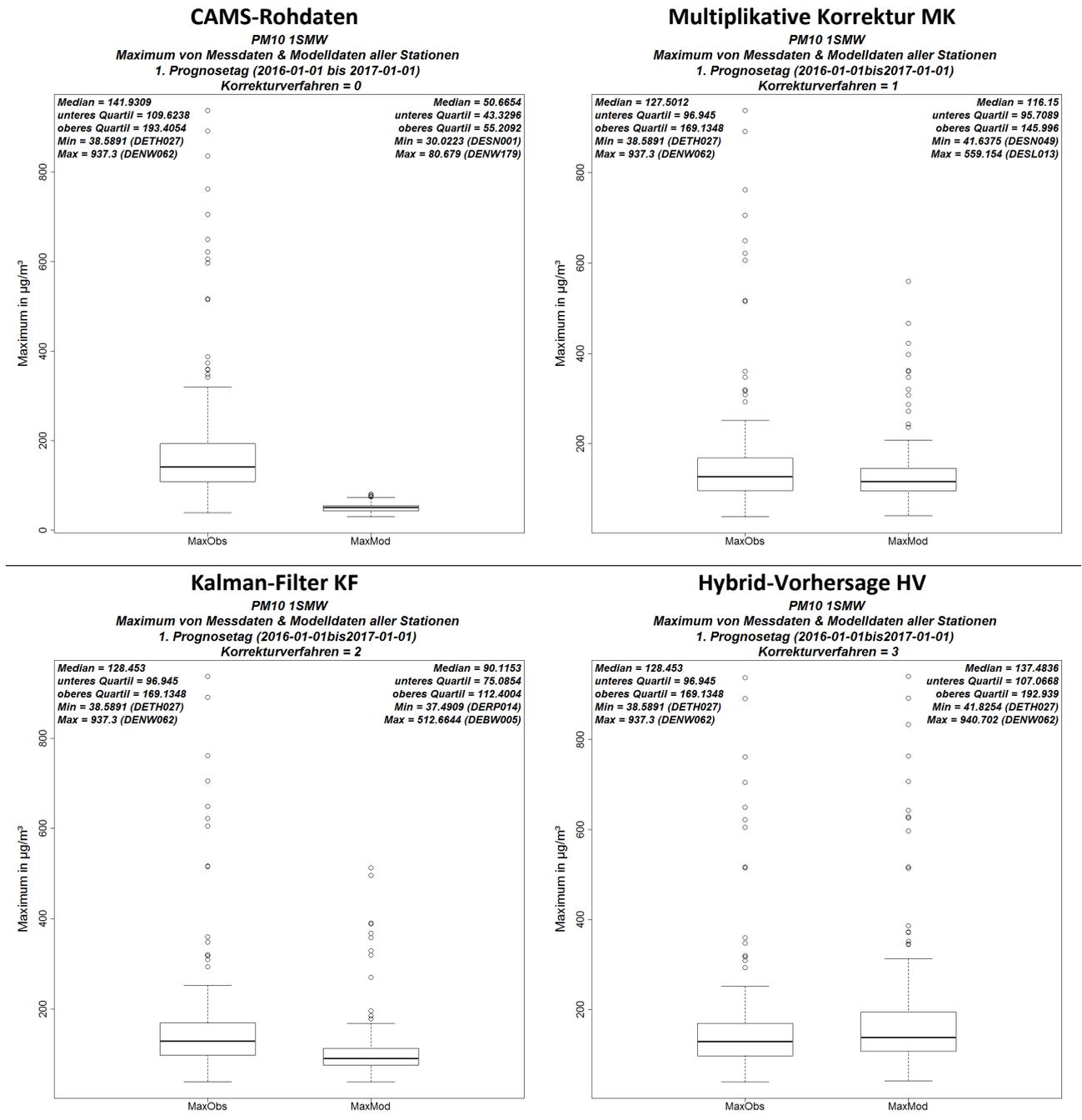
¹: Schwellenwert 50 $\mu\text{g}/\text{m}^3$

Abbildung 5-14 zeigt für PM10 beispielhaft für den 1. Prognosetag Box-Whisker-Plots der Maxima der 1SMW an den Stationsorten, für die CAMS-Rohdaten und für die drei Korrekturverfahren, jeweils im Vergleich mit den Messwerten. Der Box-Whisker-Plot für die CAMS-Rohdaten verdeutlicht die starke Unterschätzung der Maximalwerte durch die unkorrigierten Modellprognosen. Alle drei Korrekturverfahren führen zu einer verbesserten Erfassung der Maximalwerte. Die Hybrid-Vorhersage trifft das Minimum und das Maximum der Maximalwerte stationsgenau, führt aber im Median und in den Quartilen zu einer leichten Überschätzung der Maximalwerte.

Abbildung 5-15 zeigt die zugehörigen Box-Whisker-Plots der Standardabweichungen. Die Standardabweichungen werden wie schon die Maximalwerte in Abbildung 5-14 durch die CAMS-Rohdaten stark unterschätzt. Die beste Korrektur erfolgt hier durch das Verfahren der Multiplikativen Korrektur. Kalman-Filter und Hybrid-Vorhersage verbessern die Standardabweichung ebenfalls deutlich, wobei die Anwendung des Kalman-Filters die Standardabweichung der Messdaten leicht unterschätzt und die Anwendung der Hybrid-Vorhersage die Standardabweichung der Messdaten leicht überschätzt.

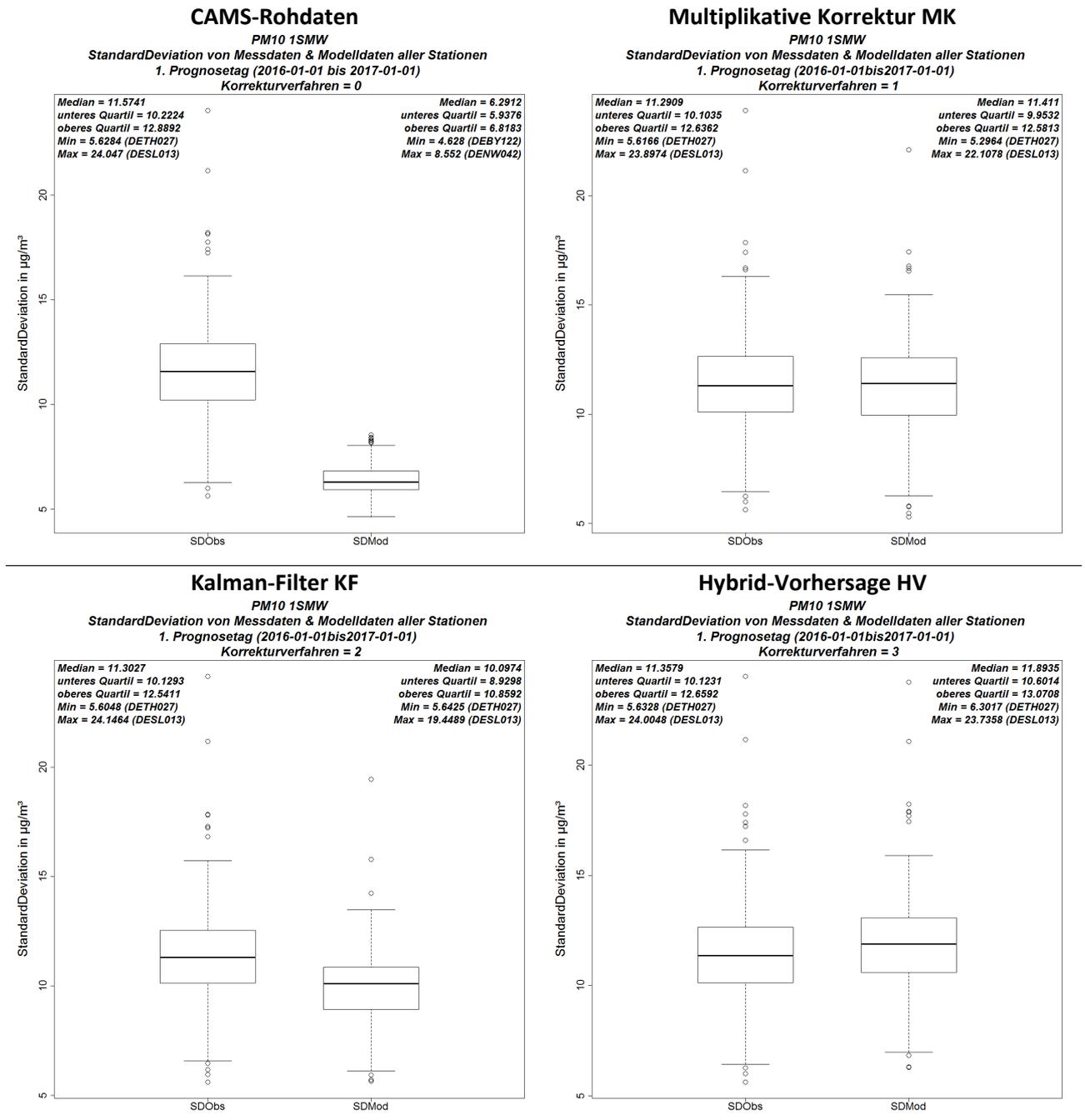
In Abbildung 5-16 sind für den 1. Prognosetag die Soccer-Plots der 1SMW für die CAMS-Rohdaten und die drei Korrekturverfahren dargestellt. Demnach weisen die CAMS-Rohdaten MFB-Werte bis $\pm 50\%$ und FGE-Werte bis zu knapp 90 % auf; eine lineare Gruppierung der Werte liegt auch hier vor, ist jedoch weniger stark ausgeprägt als bei Ozon (Abbildung 5-3). Die Anwendung der Korrekturverfahren verringert in jedem Fall den MFB und in geringerem Maße auch den FGE.

Abbildung 5-14: Box-Whisker-Plots der Maxima an den Stationsorten für PM10, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts), jeweils im Vergleich mit den Messwerten



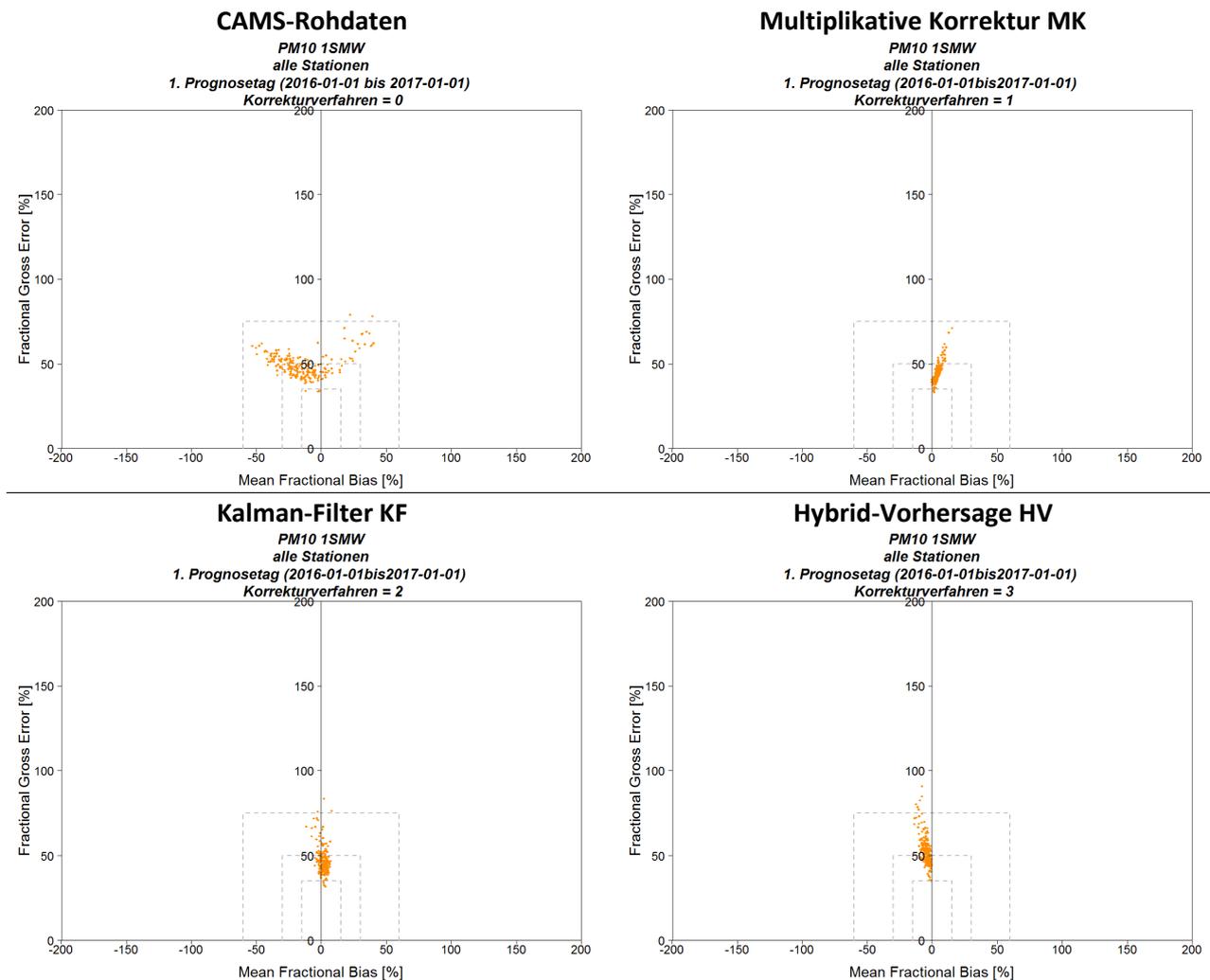
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-15: Box-Whisker-Plots der Standardabweichung an den Stationsorten für PM10, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts), jeweils im Vergleich mit den Messwerten



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-16: Soccer-Plots für PM10, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

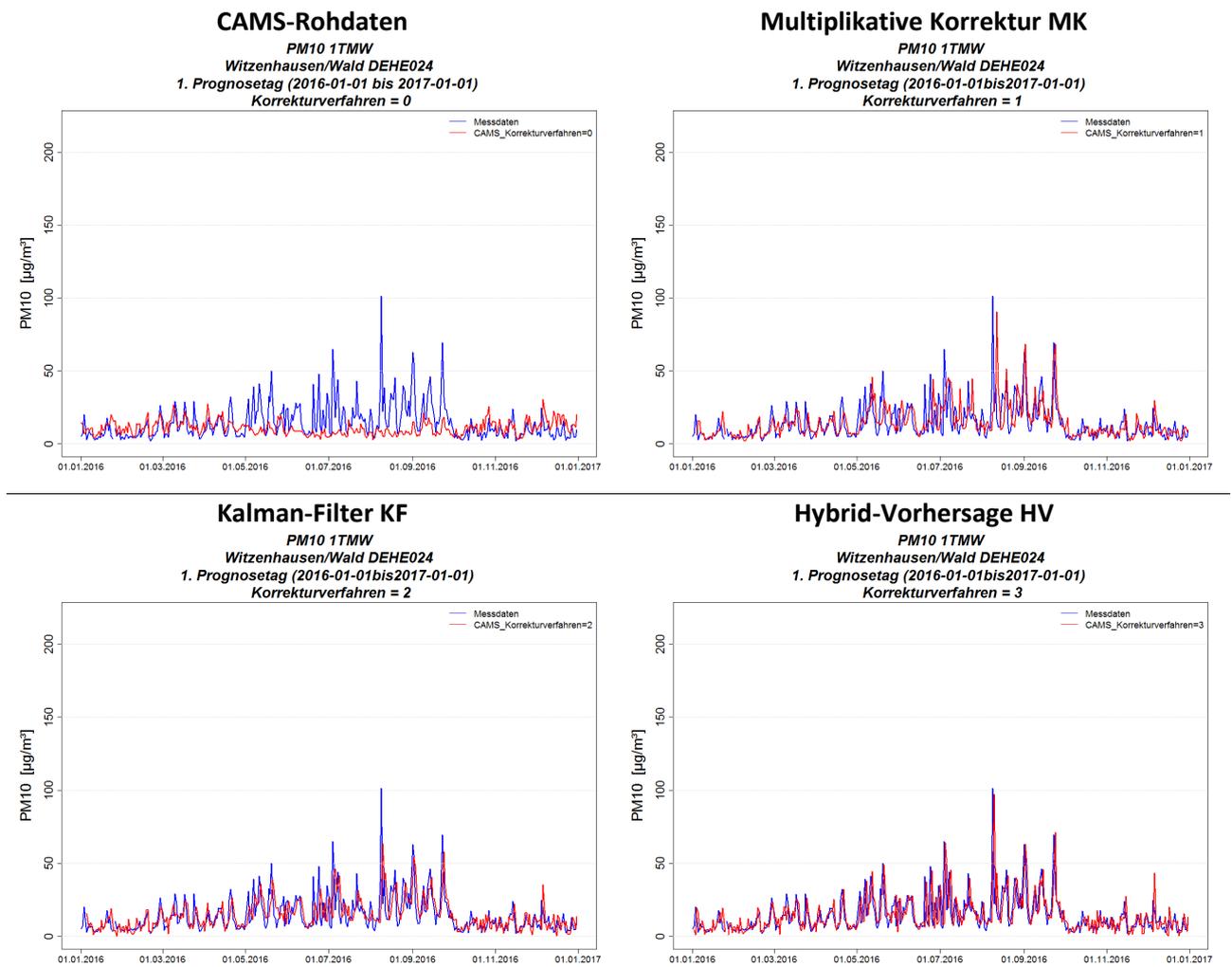
In Abbildung 5-17 und Abbildung 5-18 sind beispielhaft Zeitreihen der Tagesmittelwerte (1TMW) über das Bezugsjahr 2016 für die CAMS-Rohdaten und die korrigierten Modellprognosen dargestellt. Abbildung 5-17 zeigt Zeitreihen der Station Witzenhausen/Wald (DEHE024), für die in den CAMS-Rohdaten die niedrigste Korrelation zu den Messdaten bestand (Abbildung 4-8), und Abbildung 5-18 zeigt Zeitreihen der Station Kehl (DEBW022), für die die CAMS-Rohdaten die höchste Korrelation zu den Messdaten aufwiesen (Abbildung 4-8).

Die Station DEHE024 (Abbildung 5-17) zeigt von Anfang Mai bis Ende September 2016 hohe Tagesmittelwerte, die von den CAMS-Rohdaten in keiner Weise erfasst werden. Alle drei Korrekturverfahren passen die CAMS-Rohdaten an die hohen Tagesmittelwerte der Messdaten an. Den geringsten zeitlichen Versatz ergibt sich dabei durch die Hybrid-Vorhersage, mit diesem Verfahren werden auch die Maxima am besten getroffen. Die Multiplikative Korrektur mit 3 Tagen Vorlaufzeit zeigt den deutlichsten zeitlichen Versatz, die Maxima werden zum Teil unterschätzt. Der Kalman-Filter unterschätzt die Maxima noch stärker.

Die Tagesmittelwerte der Messdaten an der Station DEBW022 werden durch die CAMS-Rohdaten ebenfalls unterschätzt, wenn auch nicht so stark wie an der Station DEHE024. Auch hier führen alle drei Korrekturverfahren zu einer verbesserten Erfassung der Tagesmittelwerte. Die Multiplikative

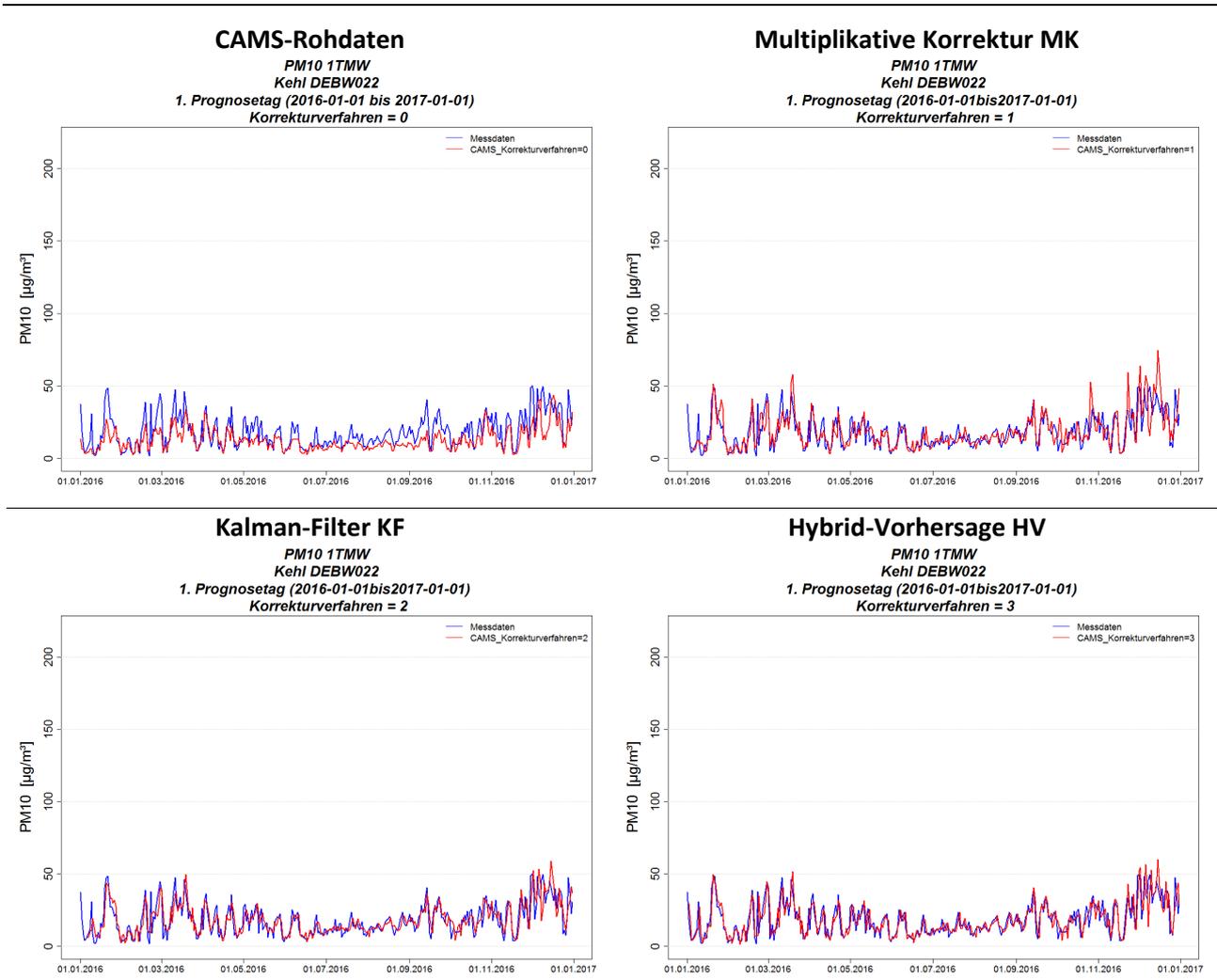
Korrektur erzeugt zu Beginn und gegen Ende des Jahres einige Überschätzungen, die deutlich höher sind als die entsprechenden Werte bei Anwendung der anderen beiden Korrekturverfahren. Die beste Korrektur erfolgt hier wie zuvor bei Station DEHE024 durch die Hybrid-Vorhersage.

Abbildung 5-17: Zeitreihen der 1TMW für PM10, Station DEHE024, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-18: Zeitreihen der der 1TMW für PM10, Station DEBW022, Bezugsjahr 2016, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



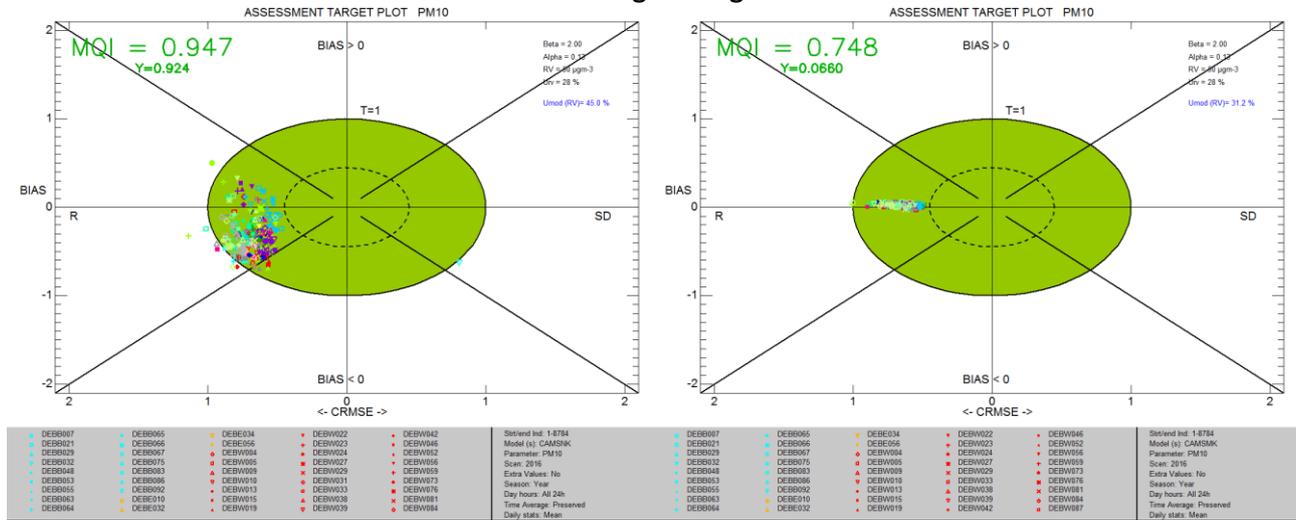
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-19 und Abbildung 5-20 zeigen die Assessment Target Plots für PM10 für die CAMS-Rohdaten und die korrigierten Prognosedaten für alle drei betrachteten Prognosetage. Die Auswertung bezieht sich dabei ausschließlich auf die 1TMW. Wie in Abschnitt 4.2.1 ausgeführt, liegt das 90%-Perzentil der MQI aller Stationen für die CAMS-Rohdaten am 1. Prognosetage mit 0.947 knapp unter 1. Dabei ist der MQI einzelner Stationen größer 1, die CAMS-Rohdaten erfüllen dort also die Qualitätskriterien des DELTA-Tools nicht. Anwendung der Korrekturverfahren führt in jedem Fall zu einer Verbesserung der Kennwerte für PM10. Alle Korrekturverfahren verringern, wie schon anhand der Metriken gezeigt, den Bias zwischen Messung und Modell erheblich. Der MQI aller Stationen liegt nach Anwendung der Korrekturverfahren bis auf wenige Ausnahmen (Multiplikative Korrektur) unter 1, und das 90%-Perzentil der MQI aller Stationen sinkt am 1. Prognosetage bis auf 0.703 (Kalman-Filter).

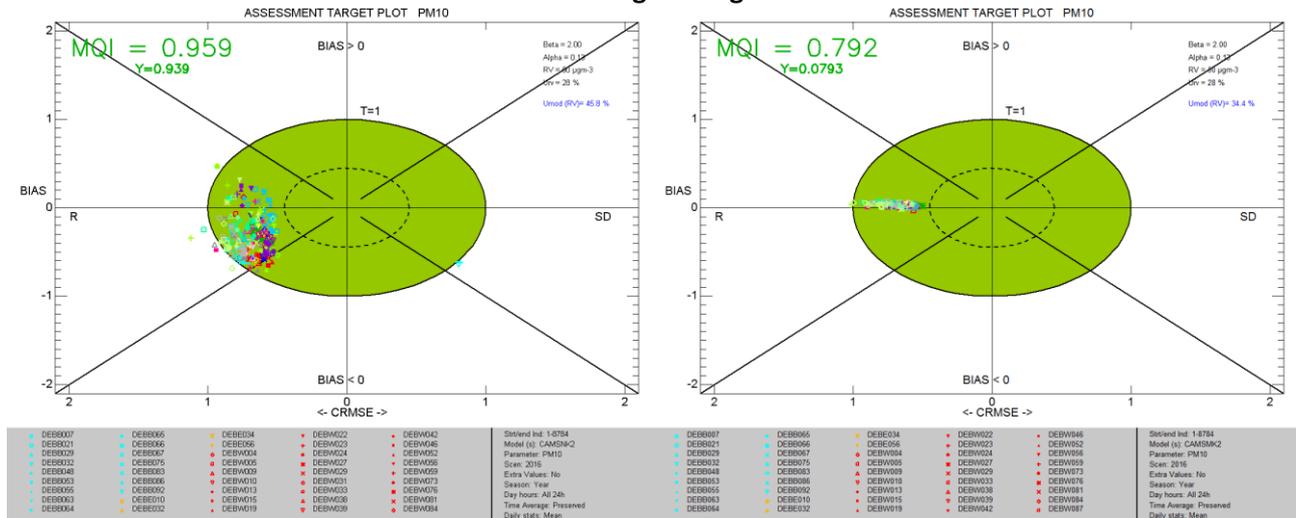
Werden die Assessment Target Plots der verschiedenen Prognosetage miteinander verglichen, so steigt wie bei O₃ der MQI mit der Vorhersagelänge nur geringfügig an, für die CAMS-Rohdaten von 0.947 auf 0.972 und für die mit dem Kalman-Filter korrigierten Prognosedaten von 0.703 auf 0.728. Der Einfluss der Vorhersagelänge auf den MQI ist damit auch bei PM10 für die korrigierten Modellprognosen vernachlässigbar.

Abbildung 5-19: Assessment Target Plots PM10 für CAMS-Rohdaten (links) und Multiplikative Korrektur (rechts), Bezugsjahr 2016, 1. bis 3. Prognosezeit

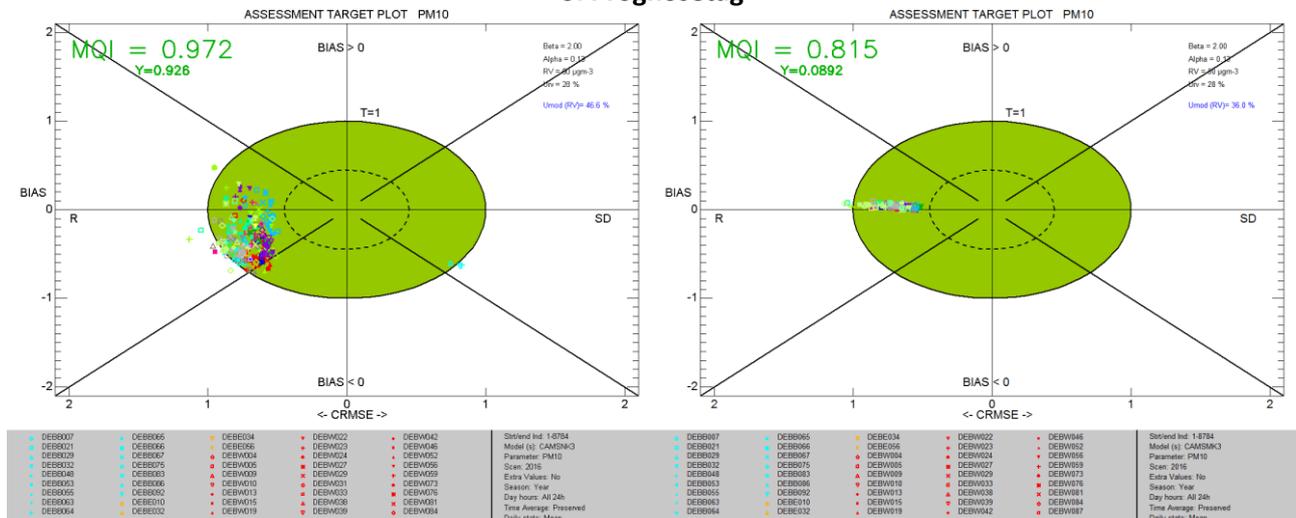
1. Prognosezeit



2. Prognosezeit



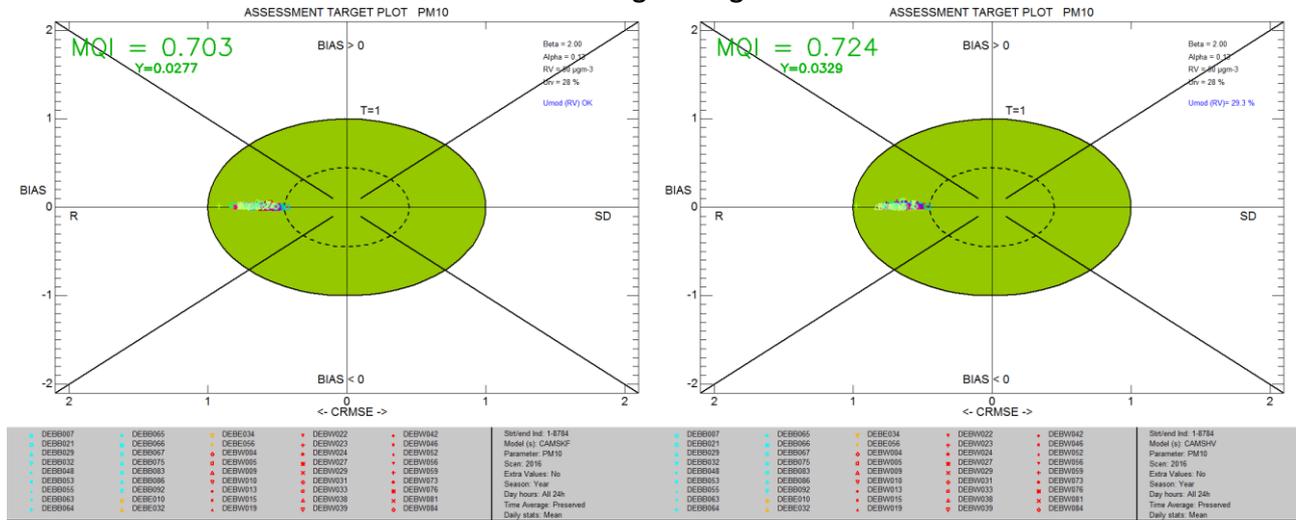
3. Prognosezeit



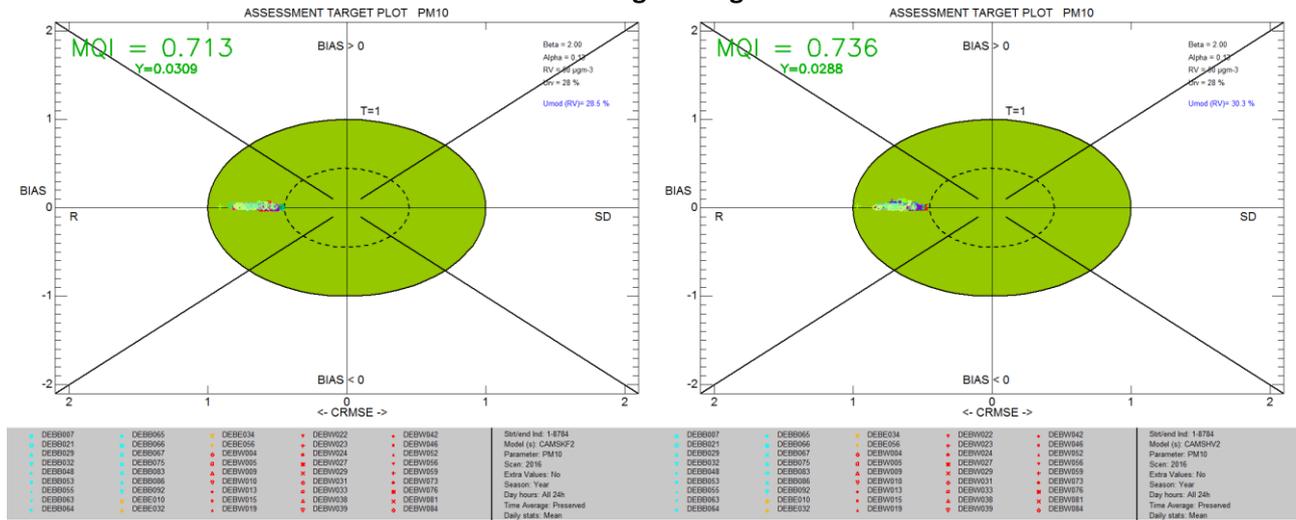
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-20: Assessment Target Plots PM10 für Kalman-Filter (links) und Hybrid-Vorhersage (rechts), Bezugsjahr 2016, 1. bis 3. Prognosetag

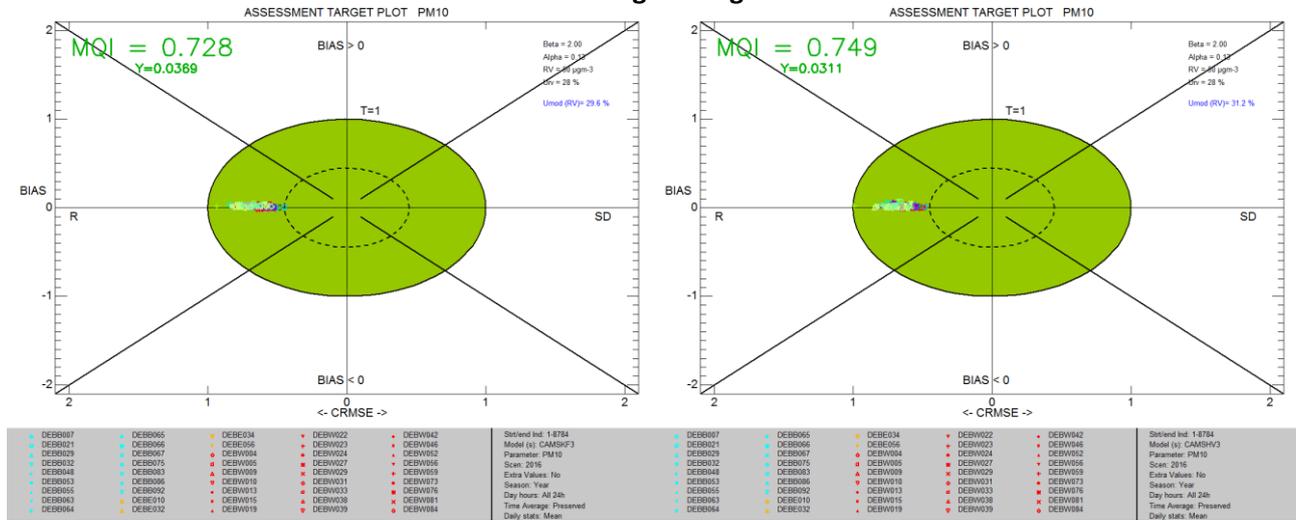
1. Prognosetag



2. Prognosetag



3. Prognosetag



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

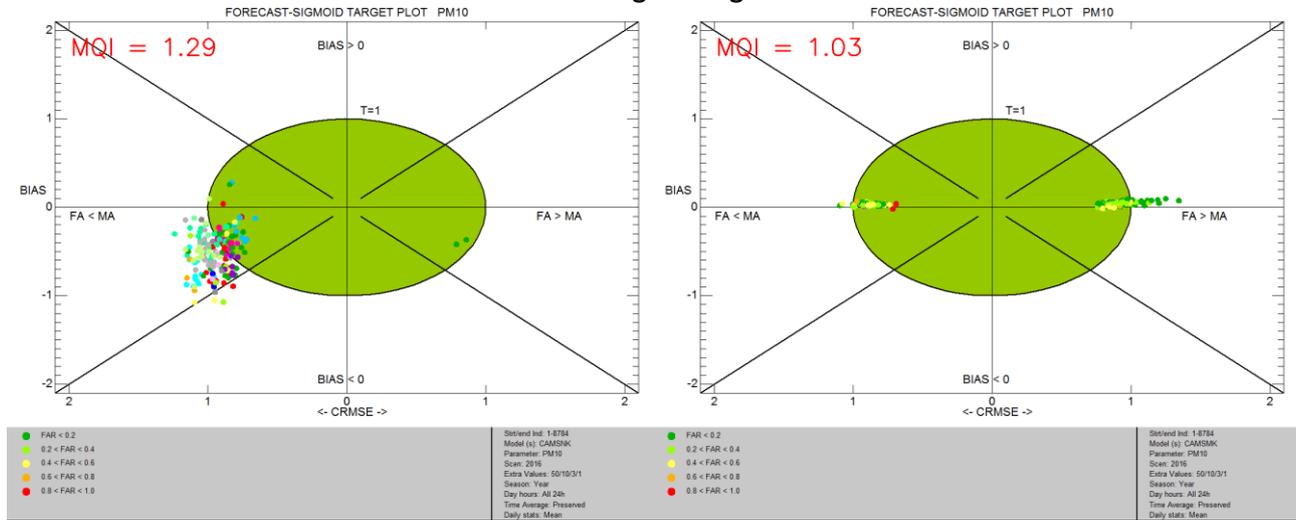
Im Forecast-Modus ist der MQI sowohl der CAMS-Rohdaten als auch der korrigierten Prognosedaten wie schon bei O₃ deutlich schlechter als im Benchmark-Modus (Abbildung 5-21, Abbildung 5-22). Bei den CAMS-Rohdaten weisen viele Stationen einen MQI > 1 auf und werden damit schlechter modelliert als mit dem persistenten Modell. Wie in Abschnitt 4.2.1 beschrieben, liegt das 90%-Perzentil der MQI aller Stationen für den 1. Prognosetag bei 1.29 und damit deutlich über 1.

Durch Anwendung der Korrekturverfahren werden MQI und Bias deutlich verringert, für Kalman-Filter und Hybrid-Vorhersage liegen am 1. Prognosetag die MQI der meisten, aber nicht aller Stationen unter 1. Für die Multiplikative Korrektur ergibt sich am 1. Prognosetag ein 90 %-Perzentil der MQI von 1.03. Die Korrektur durch dieses Verfahren erzeugt also bei PM10 Prognosedaten, die immer noch schlechter sind als die Prognose eines persistenten Modells. Insgesamt weist der Kalman-Filter mit einem 90 %-Perzentil der MQI aller Stationen von 0.900 auch bei PM10 im Forecast-Mode den niedrigsten Wert aller Korrekturverfahren auf. Im Gegensatz zu O₃ führt die Korrektur bei PM10 bei allen Korrekturverfahren dazu, dass an deutlich mehr Stationen als bei den CAMS-Rohdaten die Anzahl der FA die Anzahl der MA übersteigt.

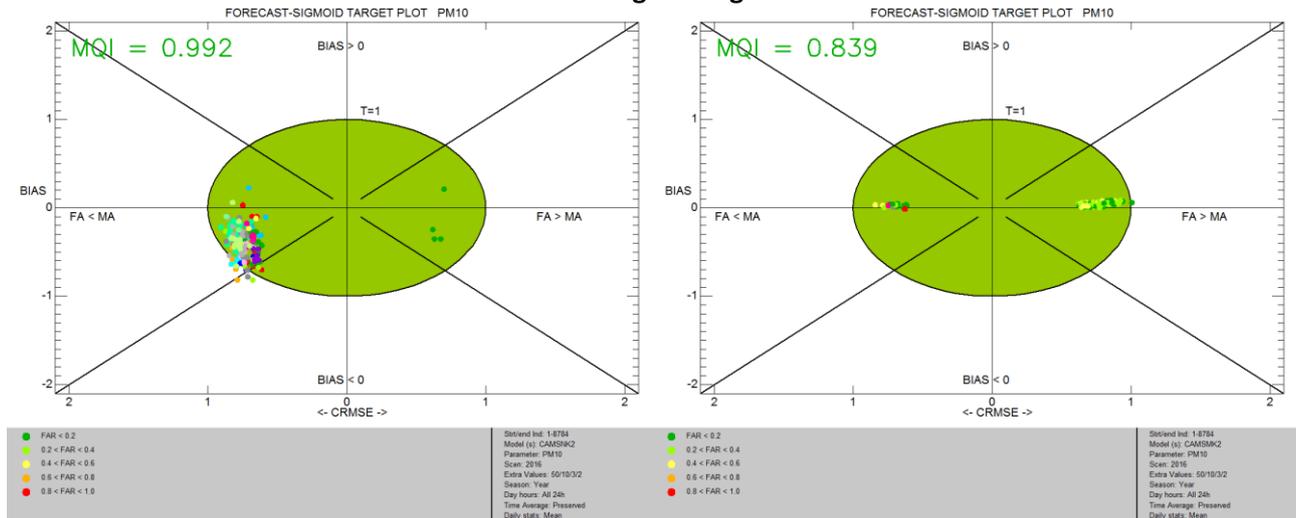
Werden die Forecast-Sigmoid Target Plots der verschiedenen Prognosetage miteinander verglichen, so zeigen sowohl die CAMS-Rohdaten als auch die korrigierten Prognosedaten eine umso bessere Performance gegenüber dem persistenten Modell, je länger die Vorhersage ist. Das 90 %-Perzentil der MQI aller Stationen für die CAMS-Rohdaten sinkt, wie bereits in Abschnitt 4.1.2 beschrieben, von 1.29 am 1. Prognosetag auf 0.907 am 3. Prognosetag. Für die mit dem Kalman-Filter korrigierten Prognosedaten sinkt das 90 %-Perzentil der MQI von 0.900 am 1. Prognosetag auf 0.685 am 3. Prognosetag. Die Korrektur mit dem Kalman-Filter erzeugt somit für alle drei Prognosetage das jeweils niedrigste 90 %-Perzentil der MQI aller Korrekturverfahren.

Abbildung 5-21: Forecast-Sigmoid Target Plots PM10 für CAMS-Rohdaten (links) und Multiplikative Korrektur (rechts), Bezugsjahr 2016, 1. bis 3. Prognostag

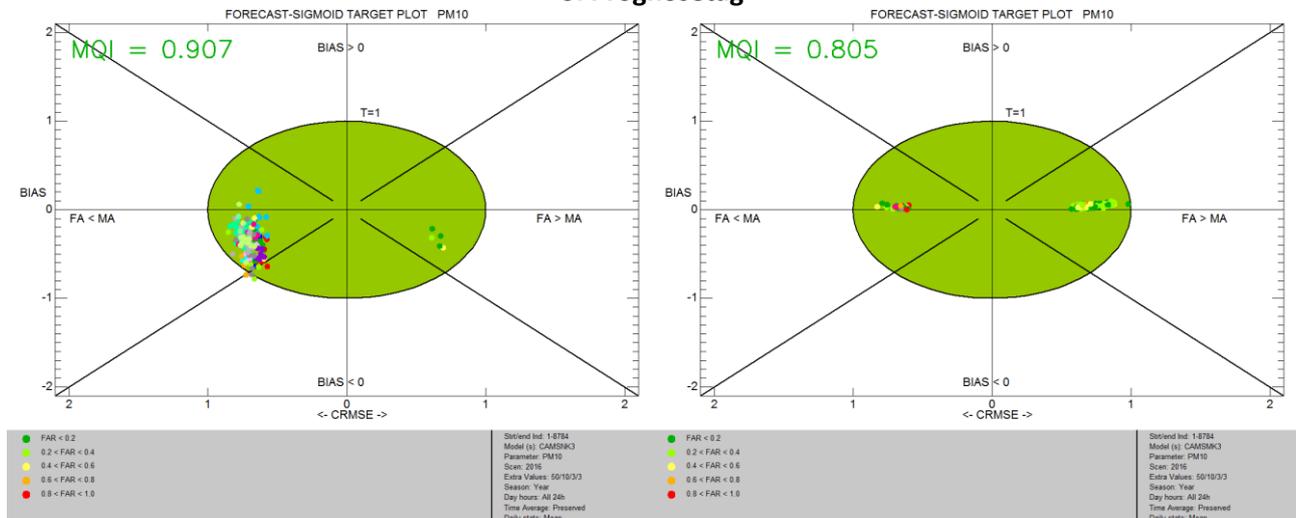
1. Prognostag



2. Prognostag



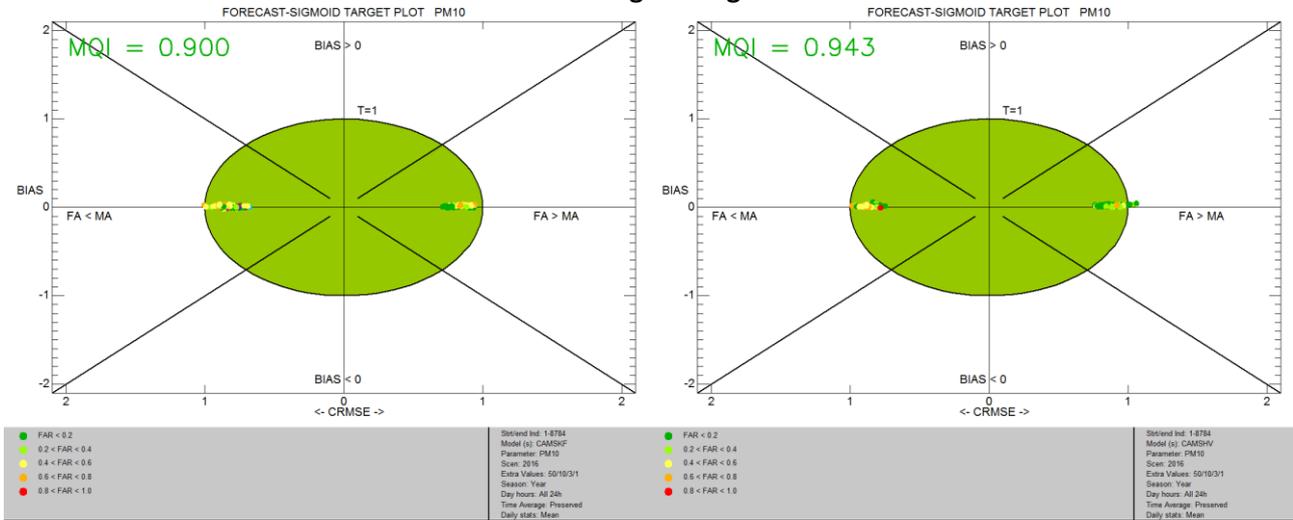
3. Prognostag



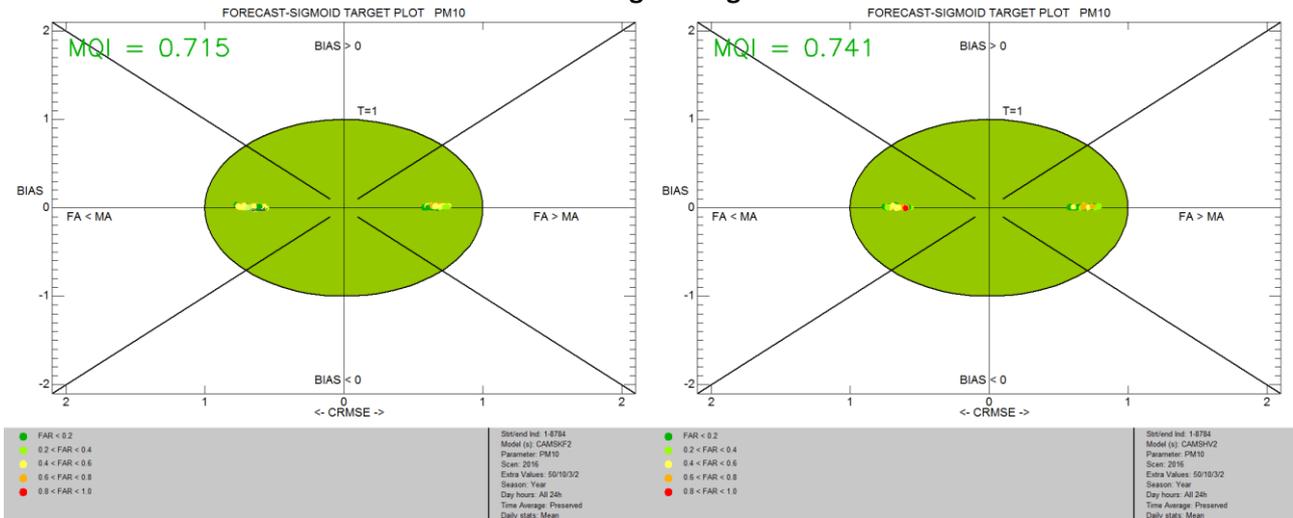
Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-22: Forecast-Sigmoid Target Plots PM10 für Kalman-Filter (links) und Hybrid-Vorhersage (rechts), Bezugsjahr 2016, 1. bis 3. Prognostag

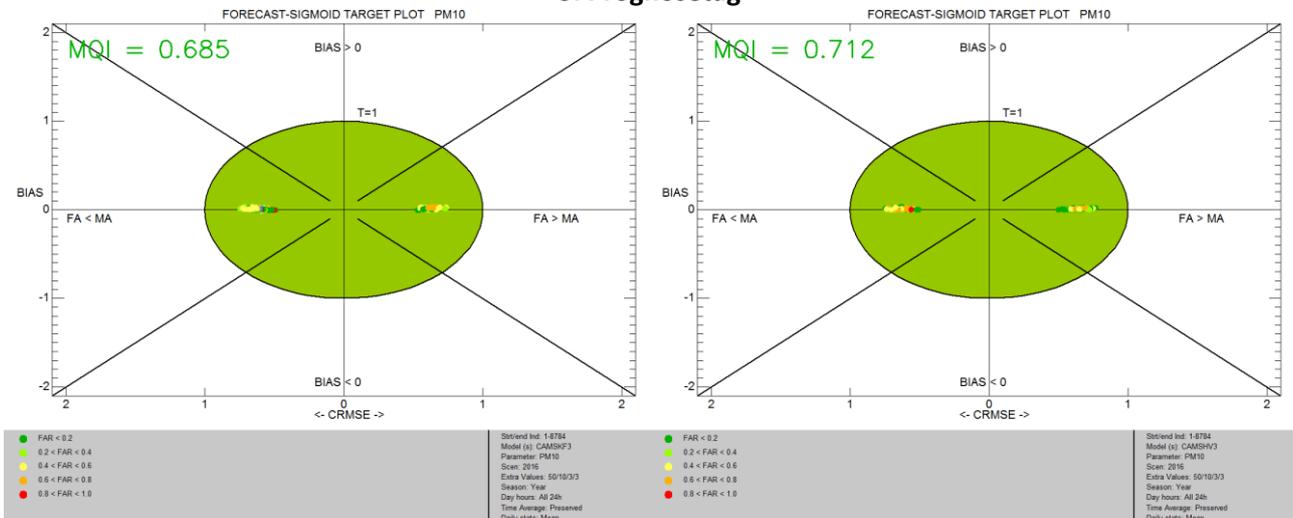
1. Prognostag



2. Prognostag



3. Prognostag



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

5.1.2.2 Saisonale Auswertung 2016

Analog zu O_3 wurden auch für PM_{10} die Jahreszeitreihen 2016 der korrigierten CAMS-Prognosen für die drei betrachteten Korrekturverfahren nach Jahreszeiten getrennt untersucht. In Tabelle 5-15 und Tabelle 5-16 sind die Mediane der für alle Stationen berechneten Metriken für die vier Jahreszeiten für die CAMS-Rohdaten und die korrigierten CAMS-Daten aufgeführt, jeweils bezogen auf die Einstundenmittelwerte (1SMW).

Grundsätzlich gilt das gleiche wie für O_3 , d. h. das auch bei saisonaler Betrachtung die Anwendung eines Korrekturverfahrens im Regelfall zu einer Verbesserung der Prognose führt. Im Frühling zeigen die Metriken der CAMS-Rohdaten gute Werte für Korrelation, RMSE und MFB, in den anderen Jahreszeiten und für andere Größen erzielen die korrigierten Daten bessere Metriken als die Rohdaten. Dies gilt insbesondere für das Hybrid-Verfahren, das bis auf wenige Ausnahmen am besten oder gut die Mittel- und Maximalwerte sowie die Standardabweichung erfasst, und den Kalman-Filter, der außer im Frühling die besten Werte für Korrelation, RMSE und MFB erzielt. Die Multiplikative Korrektur erfasst die Minimalwerte am besten und erzielt im Regelfall den niedrigsten FGE, in Sommer und Winter erfasst sie auch die Standardabweichung und im Winter die Maximalwerte am besten.

Für die Tagesmittelwerte 1TMW in Tabelle 5-17 und Tabelle 5-18 ist das Verhältnis der Metriken der einzelnen Korrekturverfahren über die Jahreszeiten etwas konstanter als bei den 1SMW. Mit den CAMS-Rohdaten werden zu allen Jahreszeiten und für alle betrachteten Metriken im Median die schlechtesten Werte erzielt. Der Kalman-Filter erzeugt bis auf eine Ausnahme den höchsten Korrelationskoeffizienten und die niedrigsten Werte für RMSE und FGE. Mittel- und Maximalwerte sowie die Standardabweichung werden im Regelfall am besten durch die Hybrid-Vorhersage erfasst, die auch den niedrigsten MFB erzielt. Ausnahme davon ist wie bei den 1SMW der Winter, in dem Maximalwerte und Standardabweichung am besten durch die Multiplikative Korrektur erfasst werden, sowie der Sommer, in dem ebenfalls die Multiplikative Korrektur die Standardabweichung am besten trifft.

Zusammenfassend lässt sich sagen, dass in der überschreitungsrelevanten Jahreszeit Winter die überschreitungsrelevanten Größen, d. h. Maximalwerte und Standardabweichung, am besten durch die Multiplikative Korrektur und am zweitbesten durch das Hybrid-Verfahren erfasst werden. Das Hybrid-Verfahren führt im Winter bei den Maximalwerten im Median zu Überschätzungen. Damit verhalten sich die Metriken von Multiplikativer Korrektur und Hybrid-Vorhersage umgekehrt zu den für die Episode 2017 bestimmten Metriken (Abschnitt 5.1.2.3), bei denen das Hybrid-Verfahren Maximalwerte und Standardabweichung am besten erfasst und die Multiplikative Korrektur am zweitbesten. Im Herbst 2016, in dem ebenfalls an einigen Stationen Überschreitungen des Schwellenwerts von $50 \mu\text{g}/\text{m}^3$ gemessen wurden, werden Maximalwerte und Standardabweichung durch das Hybrid-Verfahren am besten erfasst.

Tabelle 5-15: Metriken PM10 für 1SMW, 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Frühling und Sommer 2016 (jeweils Median über alle Stationen)

Metrik	CAMS	MK	KF	HV
	Frühling			
Min Obs [$\mu\text{g}/\text{m}^3$]	1.00	1.00	1.00	1.00
Min Mod [$\mu\text{g}/\text{m}^3$]	2.00	1.43	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	15.95	15.95	15.89	15.85
Mean Mod [$\mu\text{g}/\text{m}^3$]	13.80	15.89	15.84	15.98
Max Obs [$\mu\text{g}/\text{m}^3$]	79.01	78.56	78.91	79.01
Max Mod [$\mu\text{g}/\text{m}^3$]	44.86	69.37	58.40	77.42
SD Obs [$\mu\text{g}/\text{m}^3$]	10.55	10.49	10.53	10.49
SD Mod [$\mu\text{g}/\text{m}^3$]	6.76	9.36	9.11	10.60
Korrelation	0.56	0.54	0.56	0.49
BIAS [$\mu\text{g}/\text{m}^3$]	-1.51	-0.02	-0.04	0.20
RMSE [$\mu\text{g}/\text{m}^3$]	8.86	9.17	9.00	10.31
MFB [%]	0.22	3.82	1.09	-3.99
FGE [%]	43.35	42.55	45.31	52.04
Sommer				
Min Obs [$\mu\text{g}/\text{m}^3$]	1.04	1.09	1.10	1.07
Min Mod [$\mu\text{g}/\text{m}^3$]	1.68	1.58	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	13.44	13.44	13.41	13.43
Mean Mod [$\mu\text{g}/\text{m}^3$]	8.10	13.59	13.34	13.49
Max Obs [$\mu\text{g}/\text{m}^3$]	71.89	71.20	71.20	71.25
Max Mod [$\mu\text{g}/\text{m}^3$]	24.55	64.92	49.80	71.78
SD Obs [$\mu\text{g}/\text{m}^3$]	7.96	7.85	7.78	7.92
SD Mod [$\mu\text{g}/\text{m}^3$]	3.32	7.76	6.49	8.18
Korrelation	0.43	0.46	0.49	0.44
BIAS [$\mu\text{g}/\text{m}^3$]	-4.78	0.19	-0.02	0.01
RMSE [$\mu\text{g}/\text{m}^3$]	8.69	8.15	7.36	8.39
MFB [%]	-34.70	2.18	2.16	-3.07
FGE [%]	51.44	40.73	39.72	45.64

Tabelle 5-16: Metriken PM10 für 1SMW, 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Herbst und Winter (jeweils Median über alle Stationen)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	1.00	1.00	1.00	1.00
Min Mod [$\mu\text{g}/\text{m}^3$]	1.40	1.22	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.78	16.71	16.76	16.76
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.80	17.22	16.72	16.76
Max Obs [$\mu\text{g}/\text{m}^3$]	74.21	72.05	73.55	73.15
Max Mod [$\mu\text{g}/\text{m}^3$]	39.58	80.41	58.62	71.64
SD Obs [$\mu\text{g}/\text{m}^3$]	10.29	10.31	10.37	10.30
SD Mod [$\mu\text{g}/\text{m}^3$]	6.16	11.16	9.34	10.79
Korrelation	0.58	0.59	0.60	0.55
BIAS [$\mu\text{g}/\text{m}^3$]	-4.63	0.38	0.00	0.09
RMSE [$\mu\text{g}/\text{m}^3$]	9.77	9.82	8.86	10.02
MFB [%]	-22.63	2.74	1.65	-4.76
FGE [%]	48.31	41.97	44.48	50.97
Winter				
Min Obs [$\mu\text{g}/\text{m}^3$]	0.50	0.50	0.50	0.50
Min Mod [$\mu\text{g}/\text{m}^3$]	0.98	0.73	0.00	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	18.52	18.09	17.98	18.41
Mean Mod [$\mu\text{g}/\text{m}^3$]	13.56	18.54	18.19	18.51
Max Obs [$\mu\text{g}/\text{m}^3$]	122.70	102.74	103.04	102.85
Max Mod [$\mu\text{g}/\text{m}^3$]	39.88	106.97	79.64	117.93
SD Obs [$\mu\text{g}/\text{m}^3$]	14.66	14.16	14.11	14.15
SD Mod [$\mu\text{g}/\text{m}^3$]	6.55	14.54	12.54	15.13
Korrelation	0.45	0.57	0.59	0.56
BIAS [$\mu\text{g}/\text{m}^3$]	-4.50	0.41	0.33	0.26
RMSE [$\mu\text{g}/\text{m}^3$]	13.74	13.25	11.89	13.74
MFB [%]	-9.35	5.57	3.21	-5.22
FGE [%]	49.43	49.03	51.92	58.22

Tabelle 5-17: Metriken PM10 für 1TMW, 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Frühling und Sommer 2016 (jeweils Median über alle Stationen)

Metrik	CAMS	MK	KF	HV
	Frühling			
Min Obs [$\mu\text{g}/\text{m}^3$]	3.81	3.81	3.81	3.81
Min Mod [$\mu\text{g}/\text{m}^3$]	4.52	4.78	3.77	3.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	15.95	15.92	15.89	15.84
Mean Mod [$\mu\text{g}/\text{m}^3$]	13.80	15.87	15.84	15.90
Max Obs [$\mu\text{g}/\text{m}^3$]	43.47	43.33	43.38	43.38
Max Mod [$\mu\text{g}/\text{m}^3$]	32.81	39.28	37.21	41.45
SD Obs [$\mu\text{g}/\text{m}^3$]	7.73	7.70	7.72	7.66
SD Mod [$\mu\text{g}/\text{m}^3$]	5.60	7.06	6.87	7.61
Korrelation	0.66	0.67	0.71	0.69
BIAS [$\mu\text{g}/\text{m}^3$]	-1.47	-0.01	-0.04	0.19
RMSE [$\mu\text{g}/\text{m}^3$]	6.14	5.81	5.42	5.83
MFB [%]	-5.22	1.72	1.53	1.05
FGE [%]	31.74	28.85	27.40	28.50
gemessene Überschreitungen	0 ¹	0 ¹	0 ¹	0 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
Sommer				
Min Obs [$\mu\text{g}/\text{m}^3$]	4.95	5.10	4.98	5.05
Min Mod [$\mu\text{g}/\text{m}^3$]	3.53	4.87	5.25	4.38
Mean Obs [$\mu\text{g}/\text{m}^3$]	13.43	13.48	13.44	13.44
Mean Mod [$\mu\text{g}/\text{m}^3$]	8.09	13.62	13.34	13.49
Max Obs [$\mu\text{g}/\text{m}^3$]	29.71	29.31	28.88	29.64
Max Mod [$\mu\text{g}/\text{m}^3$]	15.60	31.57	26.31	29.16
SD Obs [$\mu\text{g}/\text{m}^3$]	5.06	5.05	5.02	5.07
SD Mod [$\mu\text{g}/\text{m}^3$]	2.48	5.06	4.27	5.05
Korrelation	0.54	0.57	0.65	0.65
BIAS [$\mu\text{g}/\text{m}^3$]	-4.78	0.20	-0.03	0.01
RMSE [$\mu\text{g}/\text{m}^3$]	6.32	4.69	3.99	4.29
MFB [%]	-40.16	1.34	1.05	-0.19
FGE [%]	44.44	26.36	23.23	24.50
gemessene Überschreitungen	0 ¹	0 ¹	0 ¹	0 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹

¹: Schwellenwert 50 $\mu\text{g}/\text{m}^3$

Tabelle 5-18: Metriken PM10 für 1TMW, 1. Prognosetag, CAMS-Rohdaten und Korrekturverfahren MK, KF und HV, Herbst und Winter 2016 (jeweils Median über alle Stationen)

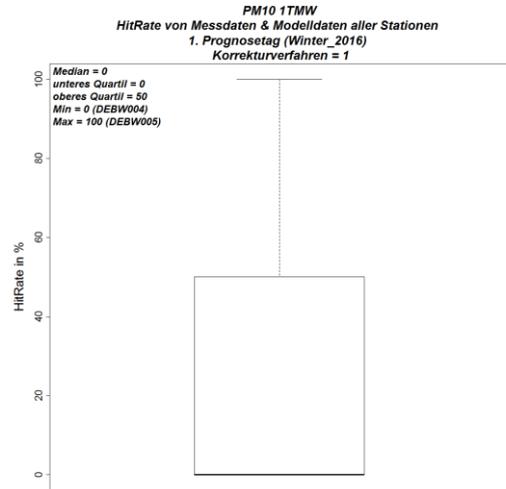
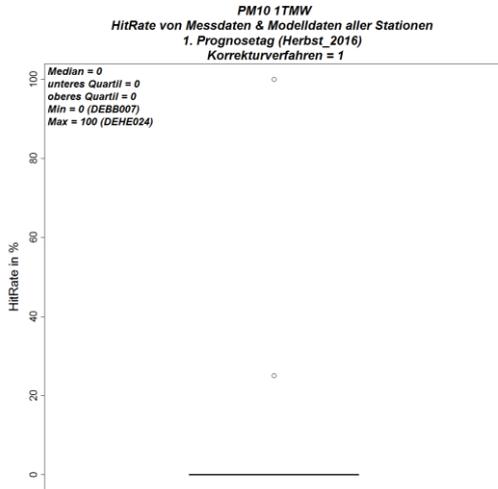
Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	3.41	3.42	3.41	3.42
Min Mod [$\mu\text{g}/\text{m}^3$]	3.20	3.24	2.99	2.06
Mean Obs [$\mu\text{g}/\text{m}^3$]	16.80	16.74	16.77	16.77
Mean Mod [$\mu\text{g}/\text{m}^3$]	11.80	17.23	16.71	16.77
Max Obs [$\mu\text{g}/\text{m}^3$]	38.78	38.59	38.78	38.66
Max Mod [$\mu\text{g}/\text{m}^3$]	28.23	46.14	36.94	39.50
SD Obs [$\mu\text{g}/\text{m}^3$]	7.92	7.91	7.99	7.88
SD Mod [$\mu\text{g}/\text{m}^3$]	5.31	8.97	7.50	8.12
Korrelation	0.69	0.73	0.76	0.74
BIAS [$\mu\text{g}/\text{m}^3$]	-4.62	0.37	-0.01	0.08
RMSE [$\mu\text{g}/\text{m}^3$]	7.51	6.37	5.54	5.82
MFB [%]	-28.21	1.37	1.49	-0.60
FGE [%]	38.27	28.84	27.63	28.77
gemessene Überschreitungen	0 ¹	0 ¹	0 ¹	0 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
Winter				
Min Obs [$\mu\text{g}/\text{m}^3$]	2.28	2.32	2.28	2.32
Min Mod [$\mu\text{g}/\text{m}^3$]	4.02	3.62	2.83	2.15
Mean Obs [$\mu\text{g}/\text{m}^3$]	18.53	18.11	17.99	18.30
Mean Mod [$\mu\text{g}/\text{m}^3$]	13.56	18.54	18.22	18.46
Max Obs [$\mu\text{g}/\text{m}^3$]	71.40	65.77	66.78	66.78
Max Mod [$\mu\text{g}/\text{m}^3$]	30.50	65.11	57.81	69.94
SD Obs [$\mu\text{g}/\text{m}^3$]	12.57	12.28	12.01	12.19
SD Mod [$\mu\text{g}/\text{m}^3$]	5.58	12.34	11.01	12.45
Korrelation	0.53	0.68	0.70	0.72
BIAS [$\mu\text{g}/\text{m}^3$]	-4.60	0.41	0.31	0.27
RMSE [$\mu\text{g}/\text{m}^3$]	11.73	9.68	8.59	9.43
MFB [%]	-13.03	4.35	4.65	1.30
FGE [%]	38.64	34.60	33.60	34.53
gemessene Überschreitungen	2 ¹	1 ¹	1 ¹	1 ¹
Hit Rate [%]	0.00 ¹	0.00 ¹	0.00 ¹	0.00 ¹
FAR [%]	0.00¹	2.35 ¹	1.18 ¹	1.16 ¹

¹: Schwellenwert 50 $\mu\text{g}/\text{m}^3$

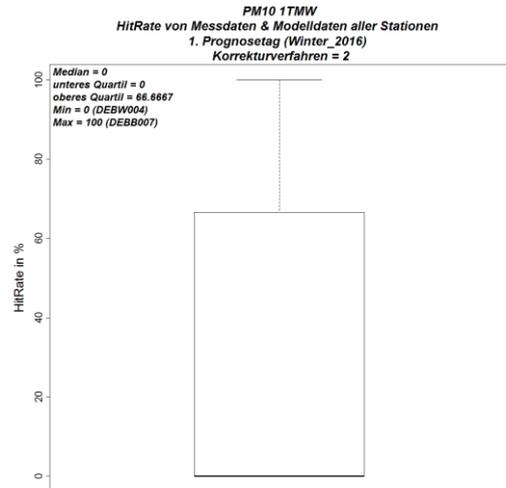
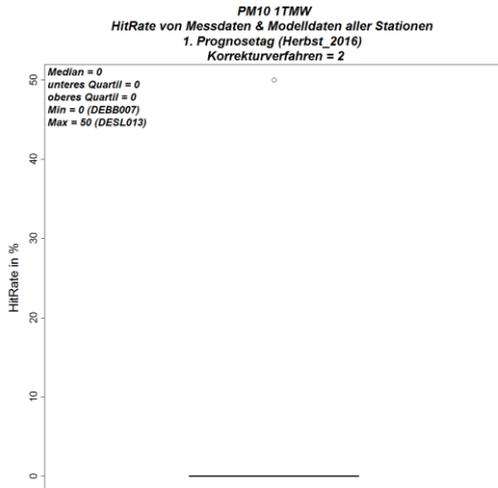
Die gemessenen Überschreitungen des Schwellenwerts von $50 \mu\text{g}/\text{m}^3$ werden im Median von keinem der Verfahren angezeigt. Auswertungen der Box-Whisker-Plots der Hit Rate (Abbildung 5-23) zeigen aber, dass nach Anwendung der Korrekturverfahren im Herbst an einzelnen Stationen zwischen 50 % (Kalman-Filter) und 100 % (Multiplikative Korrektur) der Überschreitungen korrekt erkannt werden. Im Winter werden im oberen Quartil bis zu 67 % (Kalman-Filter und Hybrid-Vorhersage) der Überschreitungen korrekt erkannt. Die unkorrigierten CAMS-Rohdaten erkennen im Herbst keine und im Winter an einzelnen Stationen im Maximum 50% der Überschreitungen (ohne Abbildung).

Abbildung 5-23: Box-Whisker-Plots der Hit Rate an den Stationsorten für PM10, Herbst (links) und Winter (rechts) 2016, Korrekturverfahren MK (oben), KF (Mitte) und HV (unten)

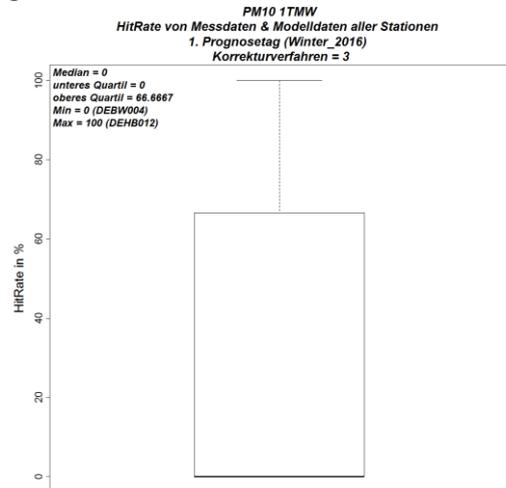
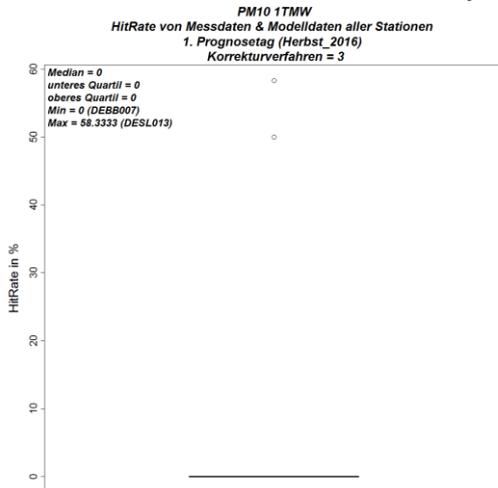
Multiplikative Korrektur MK



Kalman-Filter KF



Hybrid-Vorhersage HV



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

5.1.2.3 Episode 2017

Aufgrund der geringen Anzahl Schwellenwertüberschreitungen für PM10 im Jahr 2016 wurden die Korrekturverfahren auf eine Episode im Jahr 2017 (16. Januar bis 20. Februar 2017) angewendet, die zumindest eine Reihe von Überschreitungen des Schwellenwerts von $50 \mu\text{g}/\text{m}^3$ für die 1TMW aufweist.

Laut DWD befand sich Deutschland im Januar 2017 im Bereich von Tiefdruckgebieten mit teilweise kräftigen Niederschlägen, die überwiegend als Schnee fielen. Diese zogen aber meist recht schnell wieder ab und machten Platz für Hochdruckgebiete, in denen die zuvor eingeflossene Luft stark auskühlen konnte. Das ergab insgesamt einen kalten und trockenen Januar mit viel Sonnenschein. Zu Beginn der zweiten Januarhälfte geriet Deutschland unter den Einfluss der Hochdruckgebiete „Brigitta“, „Christa“ und „Doris“. Dadurch herrschte über den verschneiten Gebieten in der Mitte und besonders im Süden strenge Winterkälte. Bedingt durch viele Tage unter Hochdruckeinfluss gehört der Januar 2017 zu den vier sonnenscheinreichsten Januarmonaten seit Beginn solcher Messungen im Jahr 1951 (DWD, 2017a). Der Februar 2017 begann zunächst im Nordosten unter Hochdruckeinfluss noch kalt. Den Süden und Westen streiften aber bereits Tiefausläufer mit milderer Luft. Ab Mitte des Monats befanden sich dann sämtliche Landesteile in einer niederschlagsreichen, kräftigen westlichen Strömung. Diese erreichte am 23. Februar mit Sturmböen vor allem im Westen und frühlingshaften Temperaturen im Süden ihren Höhepunkt. Insgesamt verlief der Februar deutlich zu warm mit nahezu ausgeglichenem Niederschlag und Sonnenschein (DWD, 2017b).

Die Ergebnisse für die 1SMW in Tabelle 5-19 und für die 1TMW in Tabelle 5-20 unterscheiden sich insofern von den bisherigen Ergebnissen, als die besten Metriken bis auf wenige Ausnahmen von der Hybrid-Vorhersage erzielt werden. An zweiter Stelle stehen die Ergebnisse des Kalman-Filters, dann folgen die Ergebnisse der Multiplikativen Korrektur, die auch die höchste FAR aufweisen. Die Aussagen lassen sich uneingeschränkt auf den 2. und den 3. Prognosetag übertragen (hier nicht dargestellt). Alle drei Verfahren stellen aber weiterhin eine Verbesserung gegenüber den CAMS-Rohdaten dar.

Tabelle 5-19: Metriken PM10, Stundenmittelwerte 1SMW Episode 2017 für den 1. Prognosetag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	1.64	1.74	1.74	1.74
Min Mod [$\mu\text{g}/\text{m}^3$]	2.29	3.87	5.17	0.00
Mean Obs [$\mu\text{g}/\text{m}^3$]	33.74	35.57	35.85	34.45
Mean Mod [$\mu\text{g}/\text{m}^3$]	17.04	36.43	36.02	34.48
Max Obs [$\mu\text{g}/\text{m}^3$]	102.00	102.00	102.00	102.00
Max Mod [$\mu\text{g}/\text{m}^3$]	41.26	115.24	87.37	103.52
SD Obs [$\mu\text{g}/\text{m}^3$]	19.27	19.19	19.32	19.07
SD Mod [$\mu\text{g}/\text{m}^3$]	6.99	19.56	15.63	19.17
Korrelation	0.57	0.57	0.58	0.61
BIAS [$\mu\text{g}/\text{m}^3$]	-15.93	0.97	0.20	0.05
RMSE [$\mu\text{g}/\text{m}^3$]	22.94	18.19	16.48	17.31
MFB [%]	-51.64	4.98	6.74	-0.21
FGE [%]	61.50	42.34	40.85	43.78

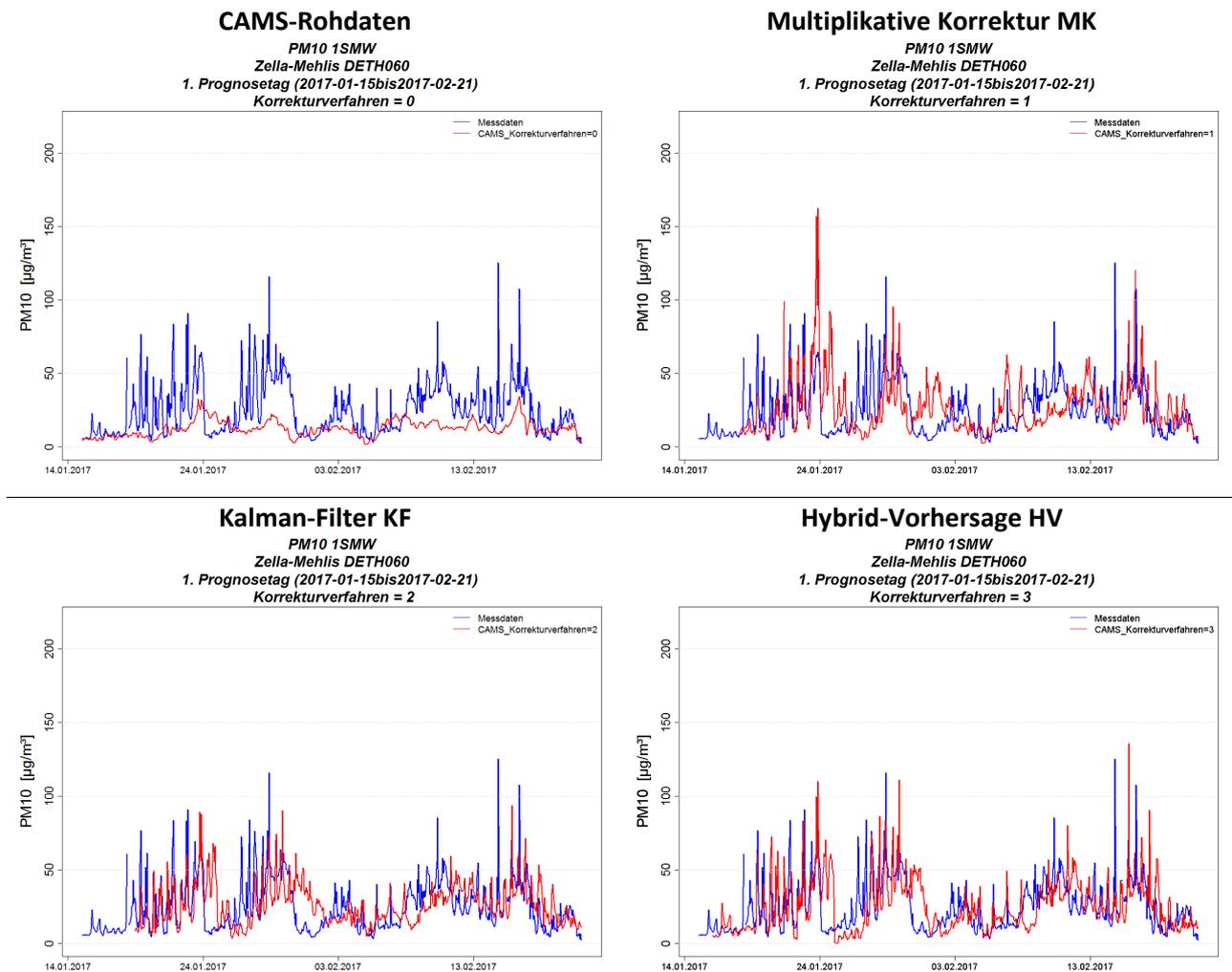
Tabelle 5-20: Metriken PM10, Tagesmittelwerte 1TMW für den 1. Prognosetag, CAMS (Rohdaten) und Korrekturverfahren MK, KF und HV (fett hervorgehoben: Werte mit der größten Übereinstimmung zwischen Messung und Modell)

Metrik	CAMS	MK	KF	HV
Min Obs [$\mu\text{g}/\text{m}^3$]	6.78	8.90	8.88	8.27
Min Mod [$\mu\text{g}/\text{m}^3$]	5.15	9.96	13.31	8.50
Mean Obs [$\mu\text{g}/\text{m}^3$]	33.71	35.52	35.74	34.50
Mean Mod [$\mu\text{g}/\text{m}^3$]	17.02	36.43	36.01	34.51
Max Obs [$\mu\text{g}/\text{m}^3$]	74.02	74.02	74.02	74.02
Max Mod [$\mu\text{g}/\text{m}^3$]	31.49	80.13	67.58	73.53
SD Obs [$\mu\text{g}/\text{m}^3$]	17.07	16.87	16.71	16.72
SD Mod [$\mu\text{g}/\text{m}^3$]	6.31	17.08	14.09	16.79
Korrelation	0.65	0.67	0.70	0.76
BIAS [$\mu\text{g}/\text{m}^3$]	-15.90	0.94	0.18	0.04
RMSE [$\mu\text{g}/\text{m}^3$]	21.13	13.42	12.23	11.62
MFB [%]	-55.22	3.83	4.89	0.57
FGE [%]	57.55	32.08	29.76	29.74
gemessene Überschreitungen	6 ¹	6 ¹	6 ¹	6 ¹
Hit Rate [%]	0.00 ¹	53.85 ¹	50.00 ¹	60.00¹
FAR [%]	0.00¹	9.38 ¹	7.41 ¹	7.41 ¹

¹: Schwellenwert 50 $\mu\text{g}/\text{m}^3$

Abbildung 5-24 zeigt beispielhaft Zeitreihen der Stundenmittelwerte für die Episode 2017 für die Station DETH060, die mit 0.38 den niedrigsten Korrelationskoeffizienten aller Stationen für diese Episode aufweist. Die CAMS-Rohdaten erkennen die Perioden mit hohen PM10-Konzentrationen kaum und unterschätzen die Maxima deutlich. Nach Anwendung der Korrekturverfahren sind die Perioden mit hohen PM10-Konzentrationen gut in den korrigierten Modellprognosen zu erkennen, allerdings sind die Maxima zeitlich deutlich versetzt. Am stärksten erscheint der zeitliche Versatz bei der Multiplikativen Korrektur, dem Verfahren mit der längsten Vorlaufzeit (3 Tage). Die Maximalwerte selbst werden am besten von der Hybrid-Vorhersage getroffen, der Kalman-Filter unterschätzt die Maximalwerte im Regelfall.

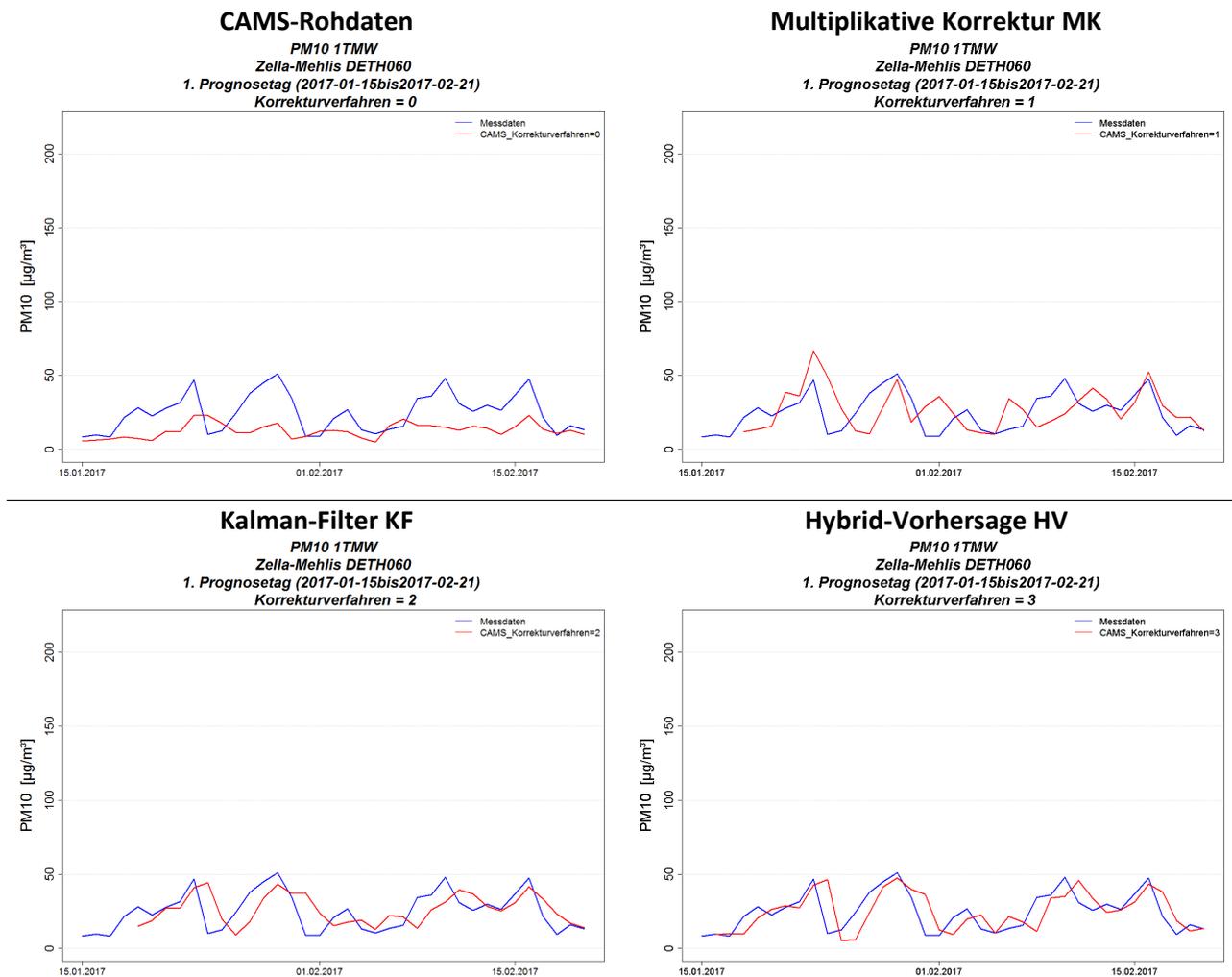
Abbildung 5-24: Zeitreihen der 1SMW für PM10, Station DETH060, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

In Abbildung 5-25 sind die zugehörigen Zeitreihen der 1TMW für diese Episode dargestellt. Die CAMS-Rohdaten geben auch hier den Verlauf der Konzentrationswerte nicht wieder. Alle drei Korrekturverfahren führen zu einer Verbesserung im Konzentrationsverlauf der 1TMW-Werte, wobei der Verlauf in allen drei Fällen zeitlich versetzt ist zum Verlauf der 1TMW-Werte der Messdaten. Diese Phasenverschiebung erscheint bei der Multiplikativen Korrektur mit 3 Vorlauf-Tagen am größten und bei der Hybrid-Vorhersage mit nur einem Vorlauf-Tag am geringsten zu sein. Die Maximalwerte werden ebenfalls am besten von durch die Hybrid-Vorhersage getroffen.

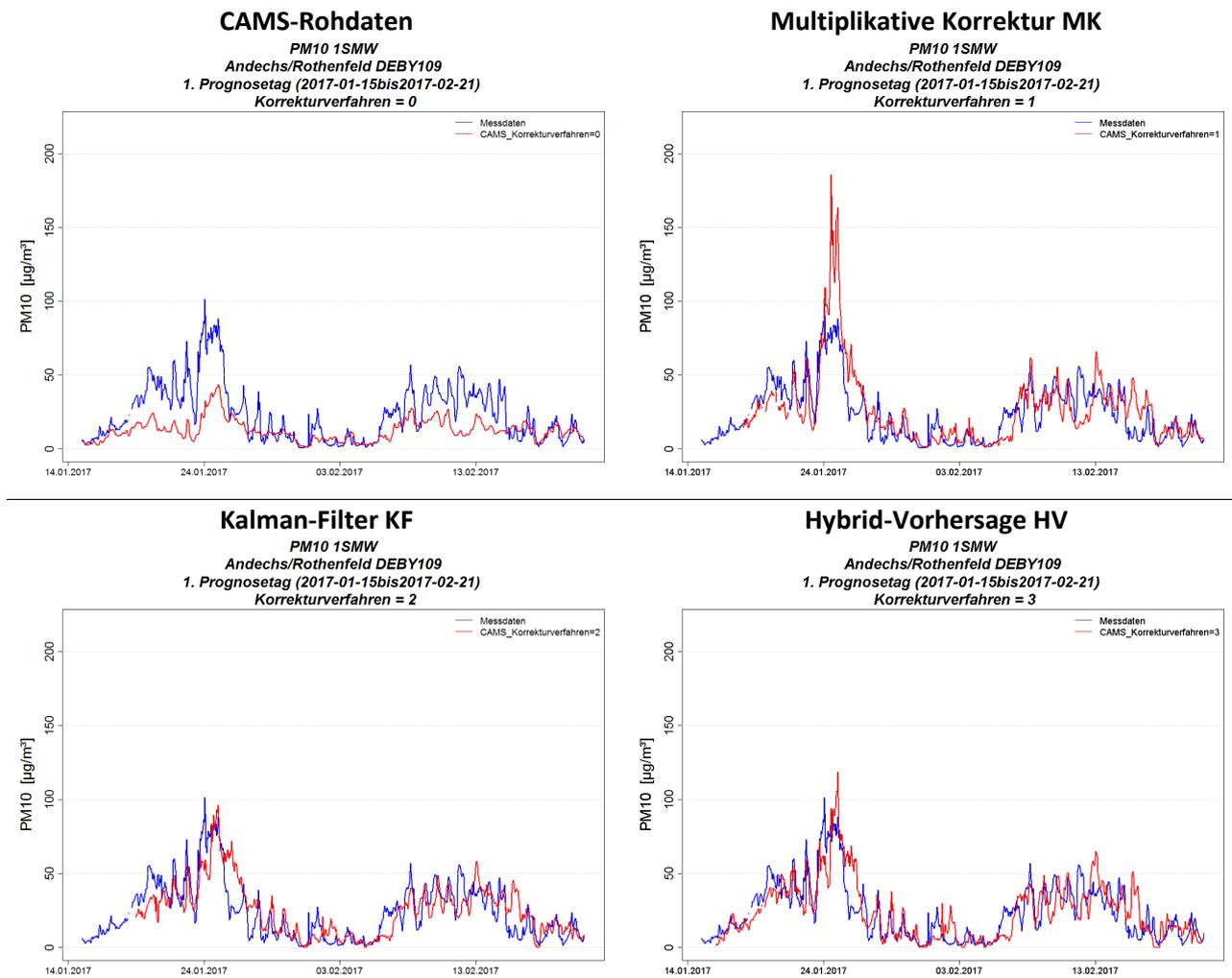
Abbildung 5-25: Zeitreihen der 1TMW für PM10, Station DETH060, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Im Vergleich zu Abbildung 5-24 zeigt Abbildung 5-26 die Zeitreihen der Stundenmittelwerte für die Episode 2017 für die Station DEBY109, die mit 0,77 einen der höchsten Korrelationskoeffizienten aller Stationen für diese Episode aufweist. Die CAMS-Rohdaten erkennen die Perioden mit hohen PM10-Konzentrationen etwas besser als an der Station DETH060, insbesondere zeigen die Modelldaten einen deutlichen Konzentrationsanstieg um den 24./25. Januar und ein Absinken der Konzentrationen etwa zur gleichen Zeit wie die Messdaten. Die Maxima werden jedoch auch hier z. T. deutlich unterschätzt. Nach Anwendung der Korrekturverfahren sind die Perioden mit hohen PM10-Konzentrationen gut in den korrigierten Modellprognosen zu erkennen, allerdings sind die Maxima analog zu DETH060 zeitlich versetzt zu den Messdaten. Am stärksten erscheint der zeitliche Versatz bei der Multiplikativen Korrektur, dem Verfahren mit der längsten Vorlaufzeit (3 Tage). Zudem zeigt die Multiplikative Korrektur um den 25. Januar herum eine sehr deutliche Überschätzung der Konzentrationswerte, die wie in Abschnitt 5.1.1.1 erläutert auf sehr ungünstige Verhältnisse der Mess- zu den Modellwerten an den Vorlauf Tagen und darauf basierend hohe Korrekturfaktoren zurückzuführen ist. Kalman-Filter und Hybrid-Vorhersage treffen den grundsätzlichen Konzentrationsverlauf und die Maxima recht gut, wobei der Kalman-Filter eher etwas zur Unterschätzung und die Hybrid-Vorhersage eher etwas zur Überschätzung neigt.

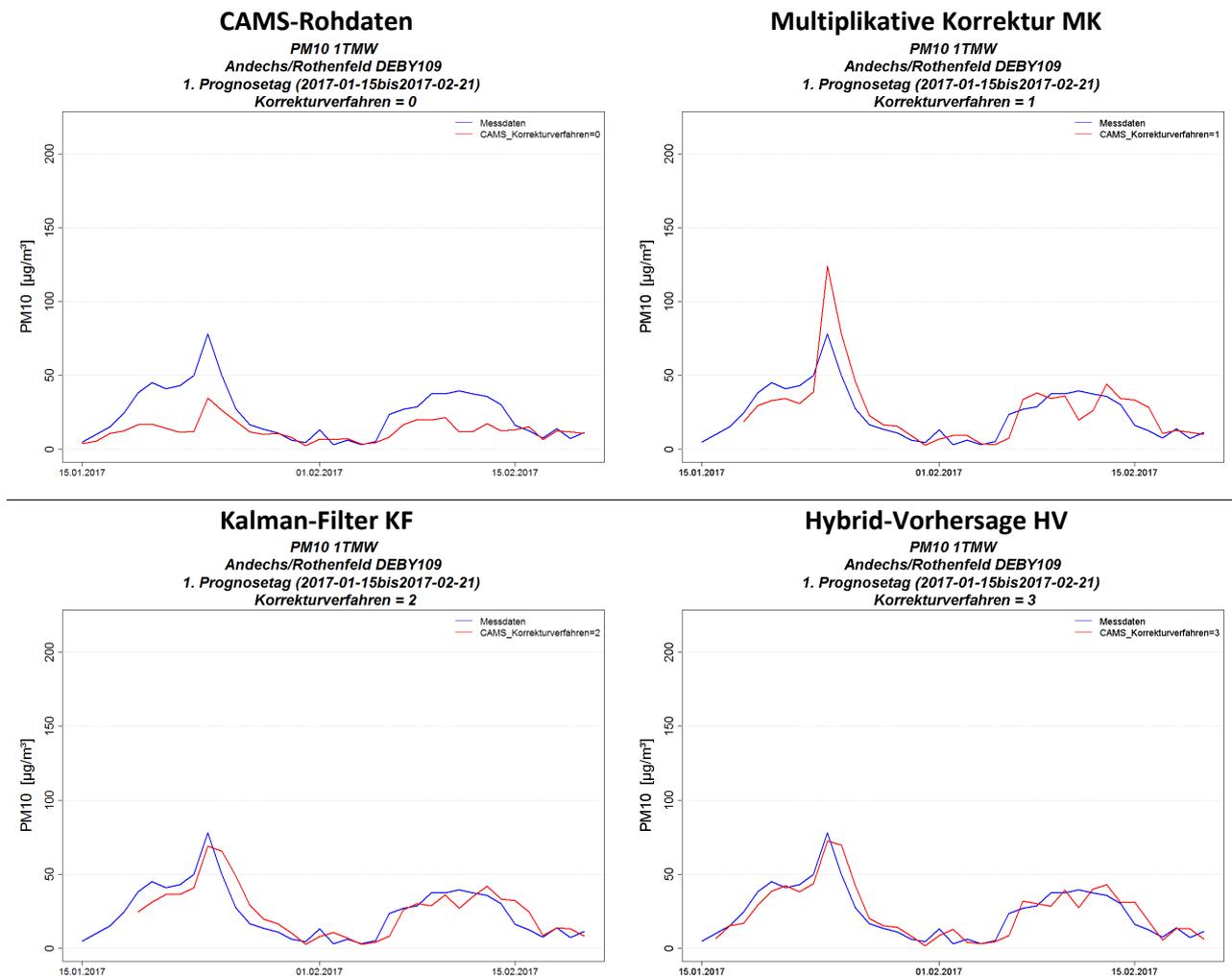
Abbildung 5-26: Zeitreihen der 1SMW für PM10, Station DEBY109, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

In Abbildung 5-27 sind die zugehörigen Zeitreihen der 1TMW für die Station DEBY109 dargestellt. Die CAMS-Rohdaten geben hier den Verlauf der Konzentrationswerte ansatzweise und damit etwas besser als an der Station DETH060 in Abbildung 5-25 wieder, jedoch werden auch hier die Maximalwerte nicht erreicht. Alle drei Korrekturverfahren führen zu einer verbesserten Erfassung der 1TMW-Werte. Der zeitliche Versatz der korrigierten Daten zum Verlauf der 1TMW-Werte der Messdaten erscheint auf Grund der glatteren Konzentrationskurven nicht ganz so stark ausgeprägt wie an der Station DETH060, ist jedoch auch hier vorhanden. Er ist bei der Multiplikativen Korrektur mit 3 Vorlauf-Tagen am größten und bei der Hybrid-Vorhersage mit nur einem Vorlauf-Tag am geringsten. Die Maximalwerte werden wie bei den Stundenmittelwerten am besten vom Kalman-Filter und von der Hybrid-Vorhersage getroffen, während sich bei der Multiplikativen Korrektur wie schon bei den Stundenmittelwerten eine deutliche Überschätzung der Konzentrationswerte um den 25. Januar herum zeigt.

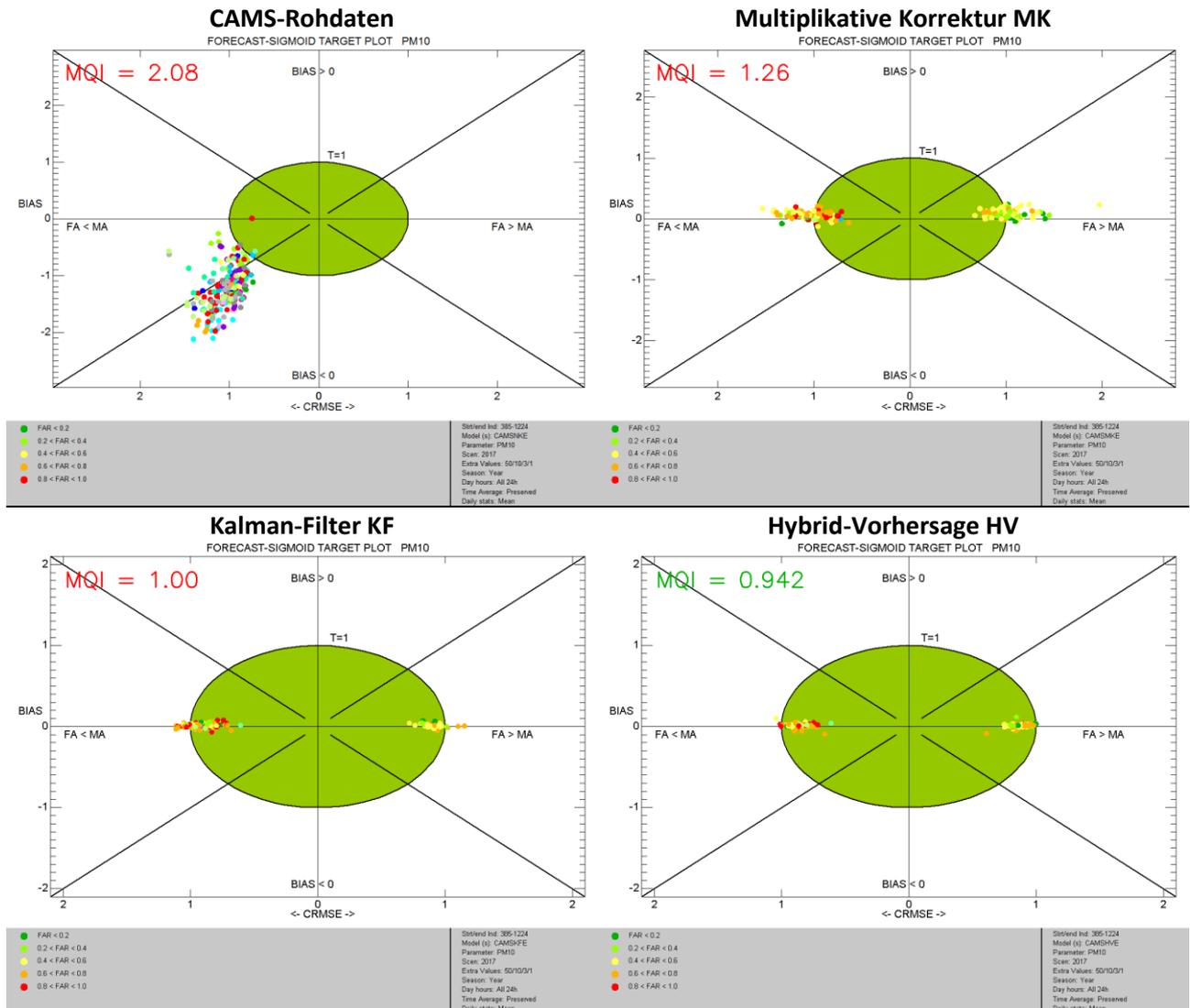
Abbildung 5-27: Zeitreihen der 1TMW für PM10, Station DEBY109, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Abbildung 5-28 zeigt die im Forecast-Modus des DELTA-Tools erstellten Target Plots für die Episode 2017. Für die CAMS-Rohdaten ist der MQI für die meisten Stationen größer 1, d. h. die CAMS-Rohdaten sind dort schlechter als eine Modellierung mit dem persistenten Modell. Das 90%-Perzentil der MQI aller Stationen liegt mit 2.08 deutlich über 1. Die Multiplikative Korrektur erscheint als Korrekturverfahren für diese Episode nicht tauglich zu sein, denn auch hier liegt das 90%-Perzentil der MQI aller Stationen mit 1.26 deutlich über 1. Sogar der Kalman-Filter erreicht für diese Episode nur ein 90%-Perzentil der MQI aller Stationen von 1.00 und ist damit gerade ebenso gut wie das persistente Modell. Das niedrigste 90%-Perzentil der MQI aller Stationen weist in der Episode das Hybrid-Verfahren mit 0.942 auf.

Abbildung 5-28: Forecast-Sigmoid Target Plots für PM10, Episode 2017, CAMS-Rohdaten (oben links) und Korrekturverfahren MK (oben rechts), KF (unten links) und HV (unten rechts)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

5.2 Fazit

In Abschnitt 4 wurde anhand von Mess- und CAMS-Prognosedaten für das Jahr 2016 gezeigt, dass die CAMS-Prognosedaten für O₃ die Messdaten im Mittel über- und die PM10-Daten im Mittel deutlich unterschätzen. Maximalwerte und Standardabweichung der Messdaten werden von den CAMS-Prognosedaten sowohl für O₃ als auch für PM10 unterschätzt.

Es wurden drei verschiedene Korrekturverfahren untersucht, um die CAMS-Prognosedaten für O₃ und PM10 im Vergleich mit den Messdaten zu verbessern. Dazu wurden in Abschnitt 5 die in Abschnitt 3.4.1 beschriebenen Metriken für alle Korrekturverfahren bestimmt. Zusätzlich wurde das DELTA-Tool (Abschnitt 3.4.2) auf die CAMS-Rohdaten und die korrigierten Prognosedaten angewendet und die statistische Auswertung anhand der mit dem DELTA-Tool erzeugten Target Plots erweitert. Die Untersuchung wurde sowohl für die gesamten Jahreszeitreihen 2016 als auch getrennt nach Jahreszeiten durchgeführt. Zusätzlich wurde für O₃ eine Episode im Jahr 2015 und für PM10 eine Episode im Jahr 2017 ausgewertet.

Von den untersuchten Korrekturverfahren haben der Kalman-Filter (Abschnitt 3.3.3) und die Hybrid-Vorhersage (Abschnitt 3.3.2) die besten Prognosedaten im Vergleich mit den Messdaten erzeugt. Dabei haben beide Verfahren ihre Vor- und Nachteile. Insgesamt erzielt der Kalman-Filter häufiger bessere Ergebnisse im Sinne der betrachteten Metriken, d. h. für die damit korrigierten Prognosedaten ergeben sich im Regelfall höhere Korrelationskoeffizienten und niedrigere Werte für RMSE, MFB und/oder FGE. Auch das 90 %-Perzentil der MQI aller Stationen fällt sowohl im Benchmark-Modus als auch im Forecast-Modus für den Kalman-Filter am niedrigsten aus, d. h. nach der Korrektur mit dem Kalman-Filter erfüllen die Prognosedaten am besten die Qualitätskriterien des DELTA-Tools. Allerdings zeigt die Hybrid-Vorhersage häufig ein besseres Verhalten, wenn es um die Erfassung von Maximalwerten und die Überschreitung von Schwellenwerten geht. Dies entspricht auch den Ergebnissen von Kang et al. (2008). Insbesondere bei den zusätzlich betrachteten Schadstoffepisoden in den Jahren 2015 und 2017 erzielt die Korrektur mit der Hybrid-Vorhersage bessere Ergebnisse als mit dem Kalman-Filter.

Die Multiplikative Korrektur (Abschnitt 3.3.1) liegt bei den Ergebnissen der statistischen Auswertung im Regelfall im Mittelfeld, in einigen Fällen erzielt sie auch bessere Ergebnisse als der Kalman-Filter und die Hybrid-Vorhersage. Allerdings ist der Einfluss der nach oben unbegrenzten Korrekturfaktoren (Abschnitt 3.3.1) ein Unsicherheitsfaktor, der in Einzelfällen zu extrem hohen Konzentrationen führen kann, wie am Beispiel der Station DEHE052 verdeutlicht wurde (Abschnitt 5.1.1.1). In anderen Fällen wie der betrachteten PM10-Episode 2017 (Abschnitt 5.1.2.3) schneidet das Verfahren auch dann schlecht ab, wenn keine erkennbar falschen Konzentrationswerte berechnet wurden. Insofern sollte das Verfahren nicht ohne weitere Untersuchungen zur sinnvollen Beschränkung des Korrekturfaktors, z. B. in Abhängigkeit von Schadstoff, Standort und/oder Konzentrationsniveau, angewendet werden.

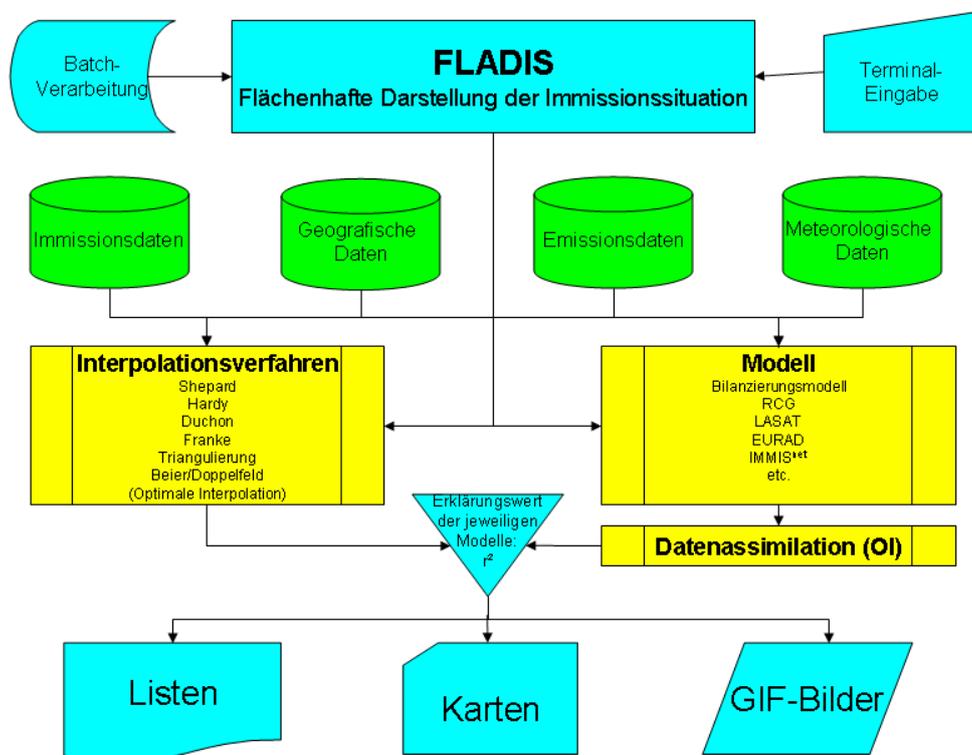
6 Korrektur in der Fläche

6.1 Systembeschreibung FLADIS

FLADIS ist ein Programmsystem zur flächenhaften Darstellung der Immissionsituation in einem vorgegebenen Gebiet. Die Grundfunktionalität besteht darin, Punktmessdaten durch reine Interpolation oder durch Kopplung der Interpolationsergebnisse mit einem Modellhintergrund in die Fläche zu übertragen. Punktmessdaten sind dabei in erster Linie Luftschadstoffe, es können aber auch meteorologische Daten oder andere Größen visualisiert werden. Daneben stellt FLADIS eine Reihe von Werkzeugen zur Verfügung, um die eingelesenen Daten zu analysieren und zu verarbeiten.

Abbildung 6-1 zeigt das Ablaufschema von FLADIS. Für jeden Zeitschritt, für den Daten vorliegen, wird wahlweise eine reine Interpolation von Messdaten oder eine gewichtete Kopplung der Interpolationsergebnisse mit Modellergebnissen durchgeführt. Es stehen verschiedene Interpolationsverfahren zur Verfügung, von denen die am häufigsten eingesetzten Verfahren in Abschnitt 6.1.1 kurz erläutert werden. Modellergebnisse können sowohl von internen als auch von externen Modellen bezogen werden. Als interne Modelle sind ein Bilanzierungsansatz und ein lineares statistisches Verfahren implementiert. Externe Modelle wie z. B. RCG, LASAT, EURAD oder IMMIS^{net} können über definierte Schnittstellen angebunden werden. Werden interpolierte Messdaten und Modelldaten miteinander kombiniert, so stellt FLADIS die Option zur Verfügung, vor der Kopplung eine Datenassimilation für die Modelldaten durchzuführen, d. h. die Modelldaten gegen die Messdaten zu ziehen. Als Assimilationsverfahren wird in FLADIS die Optimale Interpolation eingesetzt. Um das für die gegebene Datengrundlage am besten geeignete Interpolationsverfahren ermitteln zu können, ist in FLADIS ein Kreuzvalidierungsverfahren entsprechend der VDI-Richtlinie 4280 Blatt 5 (KRdL, 2009b) implementiert (Abschnitt 6.1.2).

Abbildung 6-1: Flussdiagramm des Programmsystems FLADIS



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

Im vorliegenden Projekt wurde FLADIS eingesetzt, um ein geeignetes Verfahren zur Interpolation der Korrekturfaktoren an den Messorten auf das CAMS-Modellgitter zu bestimmen. Dazu wurden die Korrekturfaktoren, die von der Testumgebung aus Abschnitt 3 als stationsbezogene Zeitreihen im FLADIS-Format abgelegt wurden, zusammen mit den zugehörigen Metadaten (Abschnitt 3.2.1) als Stationsdaten in FLADIS eingelesen.

Anschließend wurde die Interpolation der Korrekturfaktoren mit den verschiedenen in FLADIS implementierten Interpolationsverfahren vorgenommen. Abweichend vom üblichen Vorgehen in FLADIS war das Zielgitter der Interpolation dabei nicht das FLADIS-interne Gitter, sondern das ursprüngliche, durch die CAMS-Prognosedaten vorgegebene Gitter.

Die Eignung der verschiedenen Interpolationsverfahren für die Interpolation der Korrekturfaktoren wurde dann mit der Kreuzvalidierung entsprechend VDI 4280 Blatt 5 (KRdL, 2009b) bewertet und das geeignete Interpolationsverfahren ausgewählt.

6.1.1 Interpolationsverfahren

Als Interpolationsverfahren für die räumliche Verteilung der Messwerte bietet FLADIS verschiedene Optionen an, u. a.

- ▶ Shepard-Verfahren,
- ▶ Hardy-Verfahren,
- ▶ Thin Plate Splines nach Duchon,
- ▶ Thin Plate Splines nach Franke
- ▶ lineare Interpolation nach Triangulierung und
- ▶ Optimale Interpolation.

Bei den Shepard-Verfahren (Shepard, 1968), auch als IDW (Inverse Distance Weighting)-Verfahren bezeichnet, ergibt sich der Wert an einem beliebigen Ort im Interpolationsgebiet als Summe der Messwerte im Gebiet, die jeweils mit einer Wichtungsfunktion multipliziert werden. Die Wichtungsfunktionen sind so konstruiert, dass der Einfluss eines Messwertes auf einen Interpolationspunkt mit zunehmender Entfernung abnimmt. Es werden globale und lokale Ansätze unterschieden. Im globalen Fall wirkt eine Wichtungsfunktion im gesamten Modellgebiet, im lokalen Fall nur im vorgegebenen Radius.

Die Hardy'sche Multiquadrikenmethode (Hardy, 1990) und die Thin Plate Splines (TPS) nach Duchon (Duchon, 1977) gehören zu den radialen Basisfunktionsmethoden, die auf sogenannten radialen Funktionen beruhen, die vom Abstand zwischen Interpolationspunkt und Messpunkt abhängig sind. Die für die Interpolation erforderlichen Koeffizienten werden durch Lösen eines linearen Gleichungssystems ermittelt, dessen Größe von der Anzahl der betrachteten Datenpunkte (Messpunkte) im Modellgebiet abhängt.

Die Thin Plate Splines nach Franke (Franke & Nielson, 1980) gehören ebenfalls zu den radialen Basisfunktionsmethoden. Im Gegensatz zur Hardy'schen Multiquadrikenmethoden und zu den TPS nach Duchon wird das Modellgebiet bei Franke in Abhängigkeit von der räumlichen Verteilung der Datenpunkte in Teilgebiete zerlegt, so dass das Verfahren einen eher lokalen Charakter aufweist und auch für größere Datenmengen geeignet ist.

Bei der Interpolation nach Triangulierung wird zunächst für das Interpolationsgebiet eine Delaunay-Triangulierung der Messpunkte durchgeführt, die auf Voronoi-Diagrammen basiert (Abramowski & Müller, 1991), und anschließend werden die Messwerte auf den so entstandenen Dreiecksflächen linear interpoliert.

Die Optimale Interpolation (OI) schließlich ist kein Interpolationsverfahren im eigentlichen Sinne, d. h., sie zielt nicht auf die Wiedergabe stützstellentreuer Werte, sondern auf die Beschreibung der

mittleren Feldstrukturen der darzustellenden Größe in der Skala des Interpolationsrasters. Die Wiedergabe stützstellentreuer Werte ist gegebenenfalls möglich, jedoch im Allgemeinen weder zu erwarten noch vom Verfahren zu gewährleisten. Die OI kommt aus dem Bereich der Geostatistik und wurde von Gandin (1965) in die Meteorologie eingeführt, um Beobachtungswerte in Modellrechnungen einfließen zu lassen, sie liefert daher nur in Kombination mit einem Modellhintergrund sinnvolle Ergebnisse. Die OI zieht gewissermaßen die Modellergebnisse gegen die Messwerte und wird daher häufig auch zur Datenassimilation verwendet.

Die Optimale Interpolation beruht auf der Theorie der Prognose stochastischer Prozesse und erlaubt es, die räumlichen Beziehungen der Messwerte untereinander zu quantifizieren. Gleichzeitig kann ein möglicher Beobachtungsfehler an den Stationen berücksichtigt werden. Während bei einer Kombination von Interpolations- und Modellergebnissen z. B. über eine lineare Wichtungsfunktion die Wichtung einheitlich für das gesamte Feld durchgeführt wird, korrigiert die OI die Modellergebnisse räumlich differenziert in Abhängigkeit von Struktur und Einflussbereich der Messdaten.

Die in FLADIS implementierte OI beruht auf der Arbeit von Flemming (Flemming, 2003). Die räumliche Interpolation findet auf der Grundlage eines empirisch geschätzten Kovarianzmodells statt. Für die Bestimmung der empirischen Kovarianzwerte stehen die klassische Momentenmethode und eine gegenüber Ausreißern und Abweichungen von der Normalverteilung robustere Schätzung nach Cressie (1993) zur Verfügung. Die die Schätzwerte approximierende Kovarianzfunktion kann exponentiell, Gauß'schen oder sphärischen Typs sein.

Im vorliegenden Projekt wurden der globale und der lokale Ansatz des Shepard-Verfahrens, das Hardy-Verfahren, die TPS nach Duchon und die lineare Interpolation nach Triangulierung als Interpolationsverfahren auf die Korrekturfaktoren angewendet und mit Hilfe der Kreuzvalidierung (Abschnitt 6.1.2) bewertet. Für das Shepard-Verfahren wurde der Exponent der Wichtungsfunktion sowohl für den globalen als auch für den lokalen Ansatz zwischen 2 und 4 variiert. Für den lokalen Ansatz wurde ein Radius von 130 km für O_3 und von 160 km für PM10 für die Wichtungsfunktion vorgegeben, um die Fläche der Bundesrepublik Deutschland lückenlos abdecken zu können.

Die TPS nach Franke und die OI wurden im vorliegenden Projekt nicht verwendet. Die TPS nach Franke haben in der Vergangenheit im Rahmen der Kreuzvalidierung keine guten Ergebnisse erzielt, im Regelfall weist das Verfahren deutlich höhere Werte für RMSE und MRA (siehe Abschnitt 6.1.2) auf als die anderen eingesetzten Verfahren. Die OI hingegen ist ein Verfahren, dessen Ziel es ist, Beobachtungswerte in Modellrechnungen einfließen zu lassen. Da im vorliegenden Fall Korrekturfaktoren an Stationsorten betrachtet werden, für die keine Modellergebnisse existieren, ist ein Einsatz der OI inhaltlich nicht begründbar.

Die im vorliegenden Projekt verwendeten Interpolationsverfahren sind stützstellentreu auf den Interpolationspunkt. Liegt eine Messstation, für die ein Korrekturfaktor existiert, in einer Gitterzelle, so entspricht der interpolierte Korrekturfaktor für diese Gitterzelle genau dem Korrekturfaktor der Messstation, wenn die Messstation sich nahe am Mittelpunkt der Gitterzelle befindet. Weiterhin wurde bei der Bestimmung der Korrekturfaktoren nach Abschnitt 3.3 in Abstimmung mit dem Auftraggeber davon ausgegangen, dass als Modellwert am Messort direkt und ohne Interpolation auf den Messort der Gitterwert derjenigen Gitterzelle verwendet wird, in der die Messstation liegt (Abschnitt 3.1). Für die Korrektur des Modellwerts am Messort wäre es demnach sinnvoll, wenn der Korrekturfaktor für die Gitterzelle dem Korrekturfaktor am Messort entspricht.

Dies kann theoretisch erreicht werden, wenn vor der Interpolation der Korrekturfaktoren die Messorte auf den Mittelpunkt der jeweils zugehörigen Gitterzelle verschoben werden. In der Praxis ergeben sich jedoch Probleme, wenn zwei oder mehr Stationen in einer Gitterzelle liegen, für die zudem i. d. R. unterschiedliche Korrekturfaktoren berechnet wurden. Das Shepard-Verfahren erlaubt diese Vorge-

hensweise, da die Einflüsse der Messorte auf eine Gitterzelle der Reihe nach abgearbeitet werden. Liegen zwei oder mehr Stationen in einer Gitterzelle, so wird für diese Gitterzelle implizit der Mittelwert der Korrekturfaktoren der betroffenen Stationen gebildet. Für die radialen Basisfunktionsmethoden (Hardy, Duchon) und die lineare Interpolation nach Triangulierung hingegen ist das Verschieben zwei oder mehrerer Stationen auf ein und denselben Ort (Mittelpunkt der zugehörigen Gitterzelle) nicht zulässig. Im Fall der radialen Basisfunktionsmethoden ergibt sich dann ein Gleichungssystem, das zwei oder mehr Gleichungen aufweist, deren linke Seite identisch ist und deren rechte Seite sich i. d. R. unterscheidet (unterschiedliche Korrekturfaktoren). Dieses Gleichungssystem hat keine Lösung. Im Fall der linearen Interpolation nach Triangulierung ist die zugrunde liegende Triangulierung nicht möglich, wenn zwei oder mehr Messorte die gleiche Position haben. Im Rahmen des Projekts wurden daher für das Shepard-Verfahren die Stationsorte vor der Interpolation auf den Mittelpunkt der jeweils zugehörigen Gitterzellen verschoben. Für die anderen Interpolationsverfahren wurden die Stationsorte an ihrer ursprünglichen Position belassen und die Abweichung des Korrekturfaktors der Gitterzelle vom Korrekturfaktor der darin liegenden Station in Kauf genommen.

6.1.2 Kreuzvalidierung

Mit Hilfe des statistischen Verfahrens der Kreuzvalidierung ermöglicht FLADIS zum einen Aussagen über die Güte der berechneten flächenhaften Darstellung in Abhängigkeit vom angewendeten Berechnungsverfahren. Zum andern liefert es Angaben zum Einfluss einzelner Stationen auf die flächenhafte Darstellung und unterstützt so die Optimierung des in die Berechnung eingeflossenen Messnetzes.

In FLADIS ist das „leave-one-out“-Verfahren zur Kreuzvalidierung implementiert. Die angewendete Methodik entspricht der VDI Richtlinie 4280 Blatt 5 „Ermittlung der Unsicherheit räumlicher Beurteilungen der Luftqualität“ (KRdL, 2009b).

Das „leave-one-out“-Verfahren beruht darauf, dass für jeden betrachteten Zeitschritt reihum jede Station aus dem Datensatz ausgeschlossen und der Wert an ihrer Stelle durch das jeweilige Interpolationsverfahren bestimmt wird. Aus der Differenz zwischen dem Wert, den das Modell an der Stelle der ausgelassenen Station im Mittel über alle betrachteten Zeitschritte errechnet, und dem entsprechend gemittelten Messwert der ausgelassenen Station wird die relative Abweichung an der Station berechnet. Je größer die relative Abweichung an einem Stationsort, desto größer ist der Einfluss der Station auf die flächenhafte Darstellung. Abschließend wird über alle Stationen die Wurzel des mittleren quadratischen Fehlers (RMSE) sowie die mittlere relative Abweichung (MRA) ermittelt. Je kleiner diese beiden Werte sind, desto höher ist (im Vergleich mit den Kreuzvalidierungsergebnissen für andere Interpolationsverfahren) die Güte der berechneten flächenhaften Darstellung einzuordnen.

Im vorliegenden Projekt ist die Zielgröße der Kreuzvalidierung kein Messwert, sondern ein Korrekturfaktor. Dieser unterscheidet sich in zwei Punkten von einem Messwert:

- d) er kann sinnvoll sehr kleine Werte nahe Null annehmen
- e) er kann sinnvoll negative Werte annehmen.

Da die relative Abweichung an einer Station auf den Korrekturfaktor dieser Station bezogen wird, kann sie auf Grund von a) sehr hohe Werte annehmen, wenn der Korrekturfaktor sehr klein ist, selbst wenn die absolute Differenz zwischen modelliertem und tatsächlichem Korrekturfaktor als gering einzustufen ist. Diese sehr hohen relativen Abweichungen an einzelnen Stationen beeinflussen entsprechend die mittlere relative Abweichung über alle Stationen. Für die Bewertung der einzelnen Interpolationsverfahren hinsichtlich ihrer Eignung für die Interpolation der Korrekturfaktoren in die Fläche wurde daher im vorliegenden Projekt vor allem der RMSE herangezogen.

6.2 Ergebnisse der Interpolation

Im Folgenden wird der Einfluss des Interpolationsverfahrens auf die Interpolation der an den Messorten berechneten Korrekturfaktoren auf das Modellgitter diskutiert und bewertet, um die Wahl des in das Online-Tool zur Korrektur der Prognosedaten aus dem CAMS-Ensemble zu implementierenden Interpolationsverfahrens zu begründen.

Basis der Bewertung der Interpolationsverfahren ist, wie in Abschnitt 6.1.2 ausgeführt, vor allem der RMSE. In Tabelle 6-1 und Tabelle 6-2 ist der RMSE zusammen mit der MRA für die betrachteten Schadstoffe O₃ und PM10 aufgeführt. Die der Interpolation zu Grunde liegenden Korrekturfaktoren an den Stationsorten wurden entsprechend Abschnitt 5.2 mit der Hybrid-Vorhersage bestimmt.

Eine Auswertung der in Tabelle 6-1 und Tabelle 6-2 angegebenen Kenngrößen der Kreuzvalidierung mit Schwerpunkt auf dem RMSE zeigt für O₃, dass das globale Shepard-Verfahren mit Exponent = 2 die niedrigsten Kennwerte aufweist, gefolgt vom lokalen Shepard-Verfahren mit Exponent = 2 und dem Hardy-Verfahren. Die Triangulierung zeigt ebenfalls eine gute MRA, liegt aber beim RMSE nur auf Platz vier. Für PM10 zeigt sich beim RMSE das gleiche Bild wie für O₃, die MRA ist jedoch für die beiden Shepard-Varianten mit Exponent = 4 und für die Triangulierung besser. Da entsprechend den Überlegungen in Abschnitt 6.1.2 der RMSE die relevante Beurteilungsgröße ist, wurde das globale Shepard-Verfahren mit Exponent = 2 zur Implementierung in das Online-Tool ausgewählt.

Tabelle 6-1: O₃, Kenngrößen der Kreuzvalidierung für die Korrekturfaktoren der Hybrid-Vorhersage

Interpolationsverfahren	RMSE [$\mu\text{g}/\text{m}^3$]	MRA [%]
Duchon Thin Plate Splines	9.6	1.8
Hardy-Verfahren	8.2	1.6
Shepard global, Exponent = 2	7.0	1.5
Shepard global, Exponent = 4	8.4	1.6
Shepard lokal, Exponent = 2, Radius = 130 km	7.9	1.6
Shepard lokal, Exponent = 4, Radius = 130 km	8.8	1.7
Triangulierung	8.3	1.5

Tabelle 6-2: PM10, Kenngrößen der Kreuzvalidierung für die Korrekturfaktoren der Hybrid-Vorhersage

Interpolationsverfahren	RMSE [$\mu\text{g}/\text{m}^3$]	MRA [%]
Duchon Thin Plate Splines	3.1	3.3
Hardy-Verfahren	2.7	3.4
Shepard global, Exponent = 2	2.4	3.2
Shepard global, Exponent = 4	2.7	2.8
Shepard lokal, Exponent = 2, Radius = 160 km	2.5	3.0
Shepard lokal, Exponent = 4, Radius = 160 km	2.8	2.5
Triangulierung	2.7	2.9

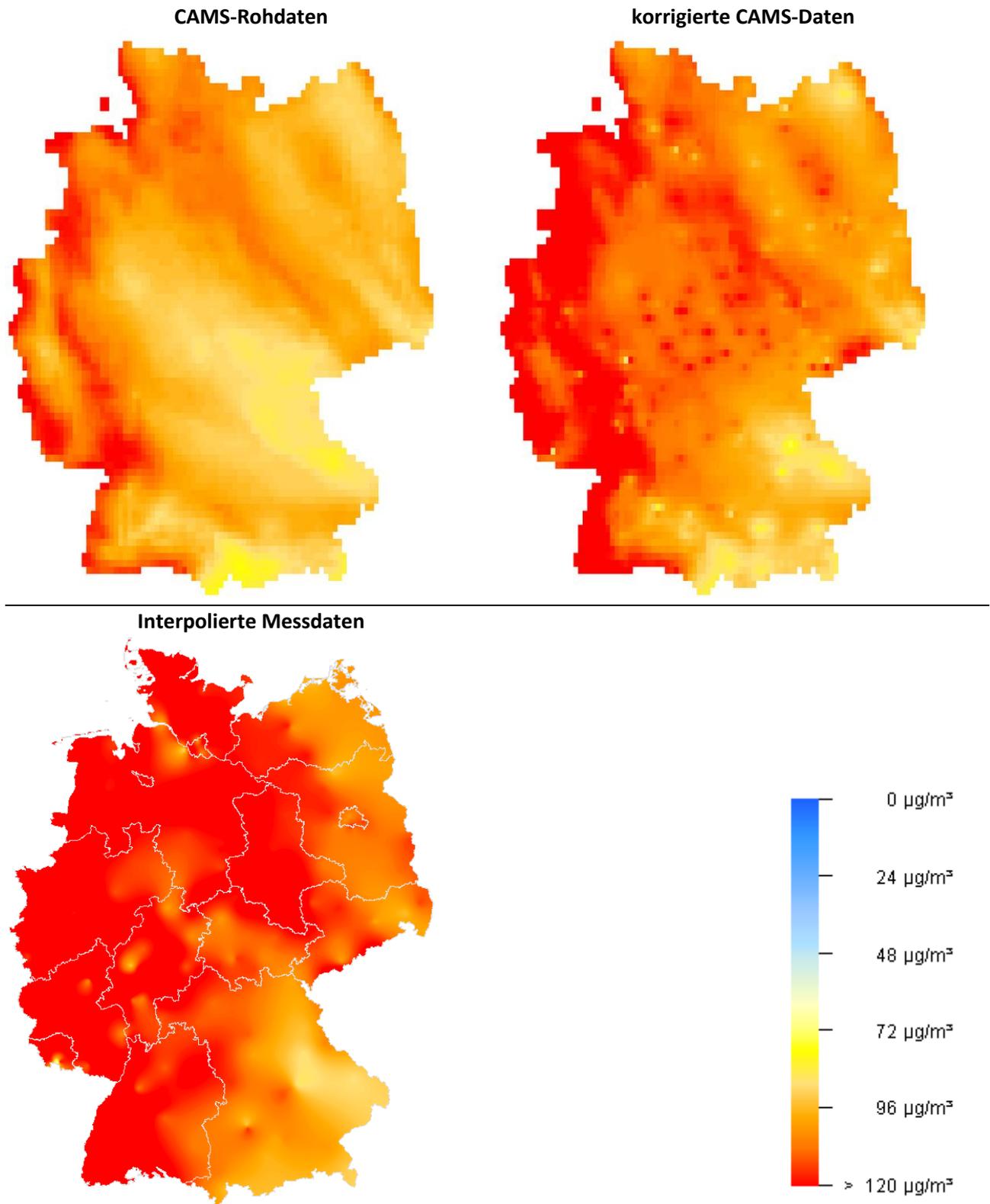
In Abbildung 6-2 ist das Ergebnis der flächenhaften Korrektur einer CAMS-Ozonprognose am Beispiel der 8TMAX-Werte für den 25.08.2016 dargestellt. Abbildung 6-2, oben links, zeigt das Raster der unkorrigierten CAMS-Rohdaten für den 1. Prognosetag. Es wurden großräumig Ozonkonzentrationen

zwischen $70 \mu\text{g}/\text{m}^3$ und $90 \mu\text{g}/\text{m}^2$ berechnet. Im westlichen Deutschland, insbesondere an der Grenze zu den Niederlanden, Belgien und Frankreich, wurden Konzentrationen nahe und oberhalb des Schwellenwertes von $120 \mu\text{g}/\text{m}^3$ prognostiziert, während entlang des Rheins Konzentrationen um die $90 \mu\text{g}/\text{m}^3$ berechnet wurden.

Ein Vergleich mit den interpolierten 8TMAX der Messdaten für den 25.08.2016 (Abbildung 6-2, unten links) zeigt, dass der Schwellenwert von $120 \mu\text{g}/\text{m}^3$ für die 8TMAX an diesem Tag in Deutschland großräumig überschritten wurde. Niedrigere Konzentrationen zwischen $80 \mu\text{g}/\text{m}^3$ und $100 \mu\text{g}/\text{m}^3$ wurden nur im Osten Deutschlands, in Teilen von Mecklenburg-Vorpommern, Brandenburg, Sachsen, Thüringen und Bayern, sowie im Bereich einiger Städte wie z. B. Hamburg oder Kassel beobachtet. Die 8TMAX für den 25.08.2016 wurden demnach von der unkorrigierten CAMS-Prognose deutlich unterschätzt, und insbesondere wurden die Bereiche mit Schwellenwertüberschreitungen nicht hinreichend erfasst.

Die Korrektur der CAMS-Prognose durch Anwendung der Hybrid-Vorhersage an den Stationsorten, Interpolation der an den Stationsorten ermittelten Korrekturfaktoren auf das Raster der Modelldaten und Anwendung der flächenhaften Korrektur auf die CAMS-Rohdaten (Abbildung 6-2, oben rechts) führt zu einem deutlichen Anstieg des Konzentrationsniveaus gegenüber den CAMS-Rohdaten. Insbesondere im Westen Deutschlands, in Teilen Niedersachsens und in Sachsen-Anhalt werden nun die Überschreitungen des Schwellenwertes von $120 \mu\text{g}/\text{m}^3$ großräumig erfasst, und auch im mittleren Teil Deutschlands, in Teilen Niedersachsens und in Hessen, werden die prognostizierten Konzentrationen durch die Korrektur deutlich angehoben, wenn auch die beobachteten Schwellenwertüberschreitungen nicht erfasst werden. Die Konzentrationen in Thüringen und Bayern im Konzentrationsbereich $80 \mu\text{g}/\text{m}^3$ bis $100 \mu\text{g}/\text{m}^3$ werden nach der Korrektur ebenfalls besser wiedergegeben.

Abbildung 6-2: Beispiel zur flächenhaften Korrektur: Ozon-Prognose 8TMAX für den 25.08.2016. CAMS-Rohdaten (oben links), nach Korrektur mit der Hybrid-Vorhersage, Interpolation der Korrekturfaktoren sowie Anwendung der flächenhaften Korrektur auf die CAMS-Rohdaten (oben rechts) sowie interpolierte Messdaten (unten links, vom Auftraggeber zur Verfügung gestellt)



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

7 Online-Tool

7.1 Allgemeines

Das Online-Tool Ozonprognose für die Korrektur der CAMS-Prognosedaten im operationellen Betrieb wurde in der Skriptsprache „R“ entwickelt. Das Online-Tool kann über eine Batchdatei oder über die Kommandozeile gestartet werden und korrigiert dann die CAMS-Prognosedaten für den heutigen oder einen vorgegebenen Tag auf Basis der Mess- und CAMS-Prognosedaten der vorangegangenen Tage. Erforderliche Einstellungen für das Online-Tool können über eine Konfigurationsdatei vorgenommen werden.

Im Rahmen des vorliegenden Schlussberichtes wird auf den Aufbau des Online-Tools und die Wahl der darin implementierten Verfahren eingegangen. Für eine Beschreibung der Eingangs- und Ausgabedaten, der erforderlichen Dateiformate und der Konfigurationsoptionen wird auf die Benutzeranleitung (IVU Umwelt, 2018) verwiesen.

7.2 Aufbau

Abbildung 7-1 zeigt das Ablaufschema des Online-Tools. Zunächst werden die zu untersuchenden Daten eingelesen. Dies sind zum einen die Messdaten der zur Korrektur herangezogenen vorangegangenen Tage für den zu untersuchenden Schadstoff zusammen mit den Stationsmetadaten und zum anderen die erforderlichen CAMS-Prognosedaten. Optional (s. u.) kann zudem eine Datei mit stationsbezogenen Nachweisgrenzen eingelesen werden.

Die Messdaten werden eingelesen und von MEZ nach UTC konvertiert, um den Bezug zu den CAMS-Prognosedaten herzustellen. Fehlwerte (Messwert -999) werden als solche abgelegt. Dadurch ist sichergestellt, dass alle weiteren Berechnungen, die Fehlwerte enthalten, im Ergebnis auch Fehlwerte liefern. Fehlt demnach einer der benötigten Werte (Messung oder Modell), so wird für die jeweilige Station und Stunde kein Korrekturfaktor gebildet und die Station von der folgenden räumlichen Interpolation der Korrekturfaktoren ausgeschlossen. Eventuell auftretende negative Messwerte werden im Online-Tool standardmäßig auf Null gesetzt, da einige Korrekturverfahren bei negativen Messwerten negative korrigierte Prognosedaten erzeugen. Alternativ kann über die Konfigurationsdatei eine Datei mit stationsbezogenen Nachweisgrenzen eingelesen werden. Dann werden Messwerte, die unterhalb der Nachweisgrenze liegen, auf einen ebenfalls konfigurierbaren Wert (z. B. 0 oder Fehlwert) gesetzt.

Die CAMS-Prognosedaten liegen stundenweise als NetCDF-Dateien vor und werden als georeferenzierter Rasterdatensatz eingelesen.

Wie in Abschnitt 3.2 erläutert, liegen die Messdaten als Stundenmittelwerte und die CAMS-Prognosedaten als Momentanwerte zu jeder Stunde vor. In Abschnitt 4.1.1 wurde untersucht, welchen Einfluss die Verwendung stundengemittelter CAMS-Prognosedaten auf die statistische Auswertung hat. Es wurde gezeigt, dass die betrachteten statistischen Metriken für die stundengemittelten und die Momentanwerte sich nur marginal unterscheiden. Da die Verwendung stundengemittelter CAMS-Daten den Stundenmittelwerten der Messdaten inhaltlich näher kommt, wird diese Vorgehensweise für die Anwendung des Online-Tools empfohlen. Dazu wird aus dem Momentanwerten jeweils zu Anfang und zu Ende einer Stunde der Mittelwert bestimmt.

Für jede Stunde an jedem Messort wird mit dem ausgewählten Korrekturverfahren ein Korrekturfaktor bestimmt. Der Korrektur liegen die stundengemittelten CAMS-Prognosewerte der vorangegangenen Tage an den Messorten sowie die zugehörigen Messwerte zugrunde. Dabei wird pro Station der Konzentrationswert derjenigen Gitterzelle des CAMS-Modells verwendet, in der die Messstation verortet ist; eine Interpolation auf den Messort innerhalb der Gitterzelle wird nicht vorgenommen.

In das Online-Tool Ozonprognose wurden 2 Korrekturverfahren implementiert:

- ▶ Multiplikative Korrektur (Abschnitt 3.3.1)
- ▶ Hybrid-Vorhersage (Abschnitt 3.3.2)

Das zu verwendende Korrekturverfahren kann im Online-Tool über die Konfigurationsdatei ausgewählt werden. Die Anzahl der zur Korrektur herangezogenen vorangegangenen Tage richtet sich wie in der Testumgebung nach dem Korrekturverfahren (Abschnitt 3.1), sie kann ebenfalls konfiguriert werden.

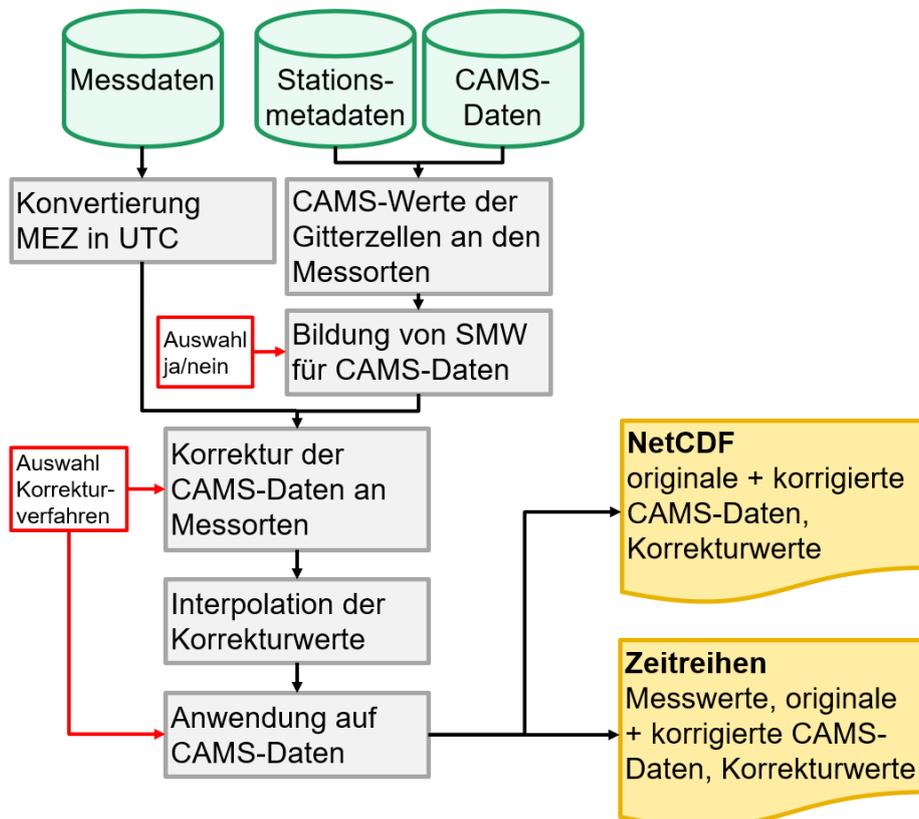
Anschließend werden für jede Stunde die an den Messorten bestimmten Korrekturfaktoren mit einem globalen Shepard-Verfahren (Abschnitt 6.2) auf ein Raster mit den gleichen Dimensionen wie das Ergebnis raster interpoliert. Vor der Interpolation werden die Messorte auf den Mittelpunkt derjenigen Gitterzelle verschoben, in der sie jeweils verortet sind (Abschnitt 7.3.2). Das Korrekturraster wird dann nach den Vorgaben des verwendeten Korrekturverfahrens auf die CAMS-Prognosedaten (je nach Auswahl auf die stundengemittelten Werte oder die Momentanwerte) angewendet, um den korrigierten Prognosedatensatz zu erhalten.

Als letzter Schritt werden die Ausgabedateien geschrieben. Für jede Prognosestunde werden die folgenden Datensätze als NetCDF-Rasterdatei geschrieben:

- ▶ CAMS-Prognosedaten in UTC (stundengemittelte Werte oder Momentanwerte)
- ▶ Korrekturraster
- ▶ Korrigierter Prognosedatensatz in UTC (stundengemittelte Werte oder Momentanwerte)

Optional können auch Zeitreihen für jede Messstation geschrieben werden, die pro Stunde die folgenden Werte enthalten: Prognosestunde, Stationskennung, CAMS-Prognosewert (SMW oder Momentanwert), korrigierter Prognosewert, Korrekturwert, Korrekturverfahren und die zugrundeliegenden CAMS-Prognosewerte (SMW oder Momentanwerte) und Messwerte der Vortage.

Abbildung 7-1: Ablaufschema Online-Tool



Quelle: Eigene Darstellung (IVU Umwelt GmbH)

7.3 Implementierte Verfahren

7.3.1 Korrekturverfahren

Im Rahmen des Projekts wurden die in Abschnitt 3.3 beschriebenen Verfahren in die Testumgebung implementiert und hinsichtlich ihrer Eignung zur Korrektur der CAMS-Prognosedaten an den Messorten untersucht. Davon wurden in Abstimmung mit dem Auftraggeber die

- ▶ Hybrid-Vorhersage (Kang et al., 2008) und die
- ▶ Multiplikative Korrektur (Borrego et al., 2011)

in das Online-Tool zur Ozonprognose implementiert.

Mit der Hybrid-Vorhersage wurde dasjenige Korrekturverfahren implementiert, das häufig ein besseres Verhalten zeigt, wenn es um die Erfassung von Maximalwerten und die Überschreitung von Schwellenwerten geht (Abschnitt 5.2). Dieses Verfahren wird für die standardmäßige Anwendung des Online-Tools empfohlen.

Die Multiplikative Korrektur wurde aufgrund guter Ergebnisse in der Literatur (Abschnitt 2) und als in der Leistungsbeschreibung vorgegebenes Verfahren zu einem frühen Zeitpunkt im Projektablauf in das Online-Tool implementiert. Im Laufe der systematischen Evaluierung der Korrekturverfahren in Abschnitt 5 wurde deutlich, dass die Multiplikative Korrektur in einigen Fällen durchaus gute Ergebnisse erzielt, dass der Einfluss der nach oben unbegrenzten Korrekturfaktoren (Abschnitt 3.3.1) jedoch ein Unsicherheitsfaktor ist, der in Einzelfällen zu extrem hohen Konzentrationen führen kann, wie am Beispiel der Station DEHE052 verdeutlicht wurde (Abschnitt 5.1.1.1). In anderen Fällen wie der betrachteten PM10-Episode 2017 (Abschnitt 5.1.2.3) schneidet das Verfahren auch dann schlecht ab, wenn keine erkennbar falschen Konzentrationswerte berechnet wurden. Insofern sollte das Verfahren nicht ohne weitere Untersuchungen zur sinnvollen Beschränkung des Korrekturfaktors, z. B. in Abhängigkeit von Schadstoff, Standort und/oder Konzentrationsniveau, angewendet werden.

7.3.2 Interpolationsverfahren

Als Verfahren zur Interpolation der an den Messorten bestimmten Korrekturfaktoren in die Fläche bzw. auf das Raster der CAMS-Prognosedaten wurde entsprechend den Ergebnissen in Abschnitt 6.2 ein globales Shepard-Verfahren mit Exponent=2 implementiert.

Das Shepard-Verfahren ist stützstellentreu auf den Interpolationspunkt. Liegt eine Messstation, für die ein Korrekturfaktor existiert, in einer Gitterzelle, so entspricht der interpolierte Korrekturfaktor für diese Gitterzelle genau dem Korrekturfaktor der Messstation, wenn die Messstation sich nahe am Mittelpunkt der Gitterzelle befindet. Weiterhin wurde bei der Bestimmung der Korrekturfaktoren nach Abschnitt 3.3 in Abstimmung mit dem Auftraggeber davon ausgegangen, dass als Modellwert am Messort direkt und ohne Interpolation auf den Messort der Gitterwert derjenigen Gitterzelle verwendet wird, in der die Messstation liegt (Abschnitt 3.1). Für die Korrektur des Modellwerts am Messort wäre es demnach sinnvoll, wenn der Korrekturfaktor für die Gitterzelle dem Korrekturfaktor am Messort entspricht.

Dies kann theoretisch erreicht werden, wenn vor der Interpolation der Korrekturfaktoren die Messorte auf den Mittelpunkt der jeweils zugehörigen Gitterzelle verschoben werden. In der Praxis ergeben sich für einige Interpolationsverfahren Probleme, wenn zwei oder mehr Stationen in einer Gitterzelle liegen, für die zudem i. d. R. unterschiedliche Korrekturfaktoren berechnet wurden (Abschnitt 6.1.1). Das Shepard-Verfahren hingegen erlaubt diese Vorgehensweise und wurde entsprechend in das Online-Tool implementiert. Liegen zwei oder mehr Stationen in einer Gitterzelle, so wird für diese Gitterzelle implizit der Mittelwert der Korrekturfaktoren der betroffenen Stationen gebildet.

8 Schlussbemerkungen und Ausblick

Als Standard-Verfahren für die Anwendung des Online-Tools im operationellen Betrieb wird uneingeschränkt die Hybrid-Vorhersage empfohlen, die im Rahmen der Evaluierung häufig ein besseres Verhalten als die anderen beiden untersuchten Verfahren zeigt, wenn es um die Erfassung von Maximalwerten und die Überschreitung von Schwellenwerten geht (Abschnitt 5.2).

Historisch bedingt (Abschnitt 7.3.1) wurde als alternatives zweites Verfahren die Multiplikative Korrektur in das Online-Tool implementiert. Diese sollte in der aktuellen Version nur unter sorgfältiger Prüfung der Ergebnisse eingesetzt werden. Insbesondere ist der Einfluss der nach oben unbegrenzten Korrekturfaktoren (Abschnitt 3.3.1) ein Unsicherheitsfaktor, der in Einzelfällen zu extrem hohen Konzentrationen führen kann, wie am Beispiel der Station DEHE052 verdeutlicht wurde (Abschnitt 5.1.1.1). Für eine Verbesserung und Absicherung des Verfahrens sollte untersucht werden, inwieweit sich der Korrekturfaktor sinnvoll begrenzen lässt, z. B. in Abhängigkeit von Schadstoff, Standort und/oder Konzentrationsniveau, und wie sich diese Begrenzung auf die Korrektur der Modellprognosen auswirkt.

Alternativ könnte der im Rahmen des Projekts ebenfalls betrachtete Kalman-Filter in das Online-Tool implementiert werden, dessen Anwendung in der Evaluierung häufig zu den höchsten Korrelationskoeffizienten führte und mit dem am besten die Qualitätskriterien des DELTA-Tools erfüllt werden konnten. Dabei ist mit der Hybrid-Vorhersage eine Form des Kalman-Filters im weitesten Sinne bereits implementiert, denn wie in Abschnitt 3.3.3 gezeigt wurde, konvergiert das Ergebnis des Kalman-Filters gegen das Ergebnis der Hybrid-Vorhersage, wenn das Verhältnis der im Kalman-Filter verwendeten Fehlervarianzen hinreichend groß ist.

Ein anderer Punkt betrifft generell das zeitliche Verhalten der Korrekturverfahren. Bei allen drei betrachteten Verfahren wird von einem Δt von 24 h ausgegangen, d. h. ein modellierter Prognosewert zu einer bestimmten Stunde X wird korrigiert mit den Mess- und Modellwerten der gleichen Stunde X an *nVorlauf* Tagen. Damit wird unterstellt, dass die gemessenen Konzentrationswerte einem Tagesgang unterliegen, der durch die Anwendung des Korrekturverfahrens auf die Prognose übertragen wird. Dies mag insbesondere für Ozon bis zu einem gewissen Grad seine Berechtigung haben, da die Ozonkonzentrationen durch den Tagesgang der Strahlung beeinflusst werden, der vor allem an Tagen mit hohen Ozonkonzentrationen hinreichend regelmäßig ist. Für PM10 hingegen ist diese Regelmäßigkeit des Tagesgangs bereits weniger gegeben.

Auf jeden Fall hat diese Vorgehensweise zur Folge, dass gemessene Spitzenkonzentrationen mit einem zeitlichen Versatz in die Korrektur eingehen, so dass Maximalwerte und Schwellenwertüberschreitungen nicht korrekt erfasst werden, insbesondere zu Beginn und zum Ende von Schadstoffepisoden. Dieser zeitliche Versatz ist umso größer, je länger *nVorlauf* ist, wie z. B. in Abbildung 5-24 und Abbildung 5-25 gut zu beobachten ist.

Es wäre demnach zu überlegen, ob die hier betrachteten Verfahren hinsichtlich der Erfassung von Schwellenwertüberschreitungen durch Einbeziehung kürzer als 24 h zurückliegender Zeitpunkte verbessert werden könnten. Für Ozon liegen an potentiellen Überschreitungstagen die Maximalwerte in der Regel am Nachmittag, d. h. zwischen 14:00 und 16:00. Es wäre zu untersuchen, ob nach der ersten Anwendung des Korrekturverfahrens am Vormittag eine zweite Anwendung am frühen Nachmittag sinnvoll wäre, die aus dem Vergleich des am Vormittag für diesen Vormittag prognostizierten Anstiegs der Ozonwerte und dem tatsächlichen Anstieg der dann vorliegenden gemessenen Ozonwerte ein Korrekturfaktor für die Prognose der Maximalkonzentrationen am Nachmittag ableitet. Für PM10 wäre möglicherweise eine einfache Verkürzung des Δt von 24 h auf z. B. 12 h oder 6 h sinnvoll, das wäre anhand von Testrechnungen zu untersuchen. Sowohl für Ozon als auch für PM10 verringert sich mit den beschriebenen Vorgehensweisen allerdings auch die Vorwarnzeit bei Schwellenwertüberschreitungen.

Alternativ oder in Kombination mit dieser Vorgehensweise wäre zu untersuchen, inwieweit die Verwendung meteorologischer Größen aus der Kürzest- oder Kurzfristvorhersage zu einer Verbesserung der Schadstoffprognose führen kann. Dabei deckt die meteorologische Kürzestfristvorhersage den Prognosezeitraum bis 12 h und die Kurzfristvorhersage den Prognosezeitraum 12 h bis 72h ab. Solche Prognosen werden beispielsweise vom DWD alle drei Stunden, ausgehend von 00, 03, 06, 09, 12, 15, 18 und 21 UTC, erstellt. Relevante meteorologische Parameter könnten für die Ozonprognose z. B. Wolkenbedeckung oder Strahlungsflüsse sein, für die PM10-Prognose z. B. vertikale Temperaturprofile (Inversionswetterlagen), Windgeschwindigkeit oder Niederschlag. Der Einfluss der meteorologischen Parameter wäre basierend auf Veröffentlichungen und/oder Untersuchungen vorliegender Prognosedaten zu quantifizieren und in ein geeignetes Korrekturverfahren zu überführen.

9 Quellenverzeichnis

- Abramowski, S.; Müller, H. (1991): Geometrisches Modellieren, BI Wissenschaftsverlag, 1991.
- Borrego, C.; Monteiro, A.; Pay, M.T.; Ribeiro, I.; Miranda, A.I.; Basart, S.; Baldasano, J.M. (2011): How bias-correction can improve air quality forecasts over Portu-gal. In: Atmospheric Environment 45 (2011) Nr. 37, S. 6629-6641.
- Clappier, A.; Pisoni, E.; Thunis, P. (2015): A new approach to design source-receptor relationships for air quality modelling. Environmental Modelling & Software 74 S. 66-74. 2015.
- Cressie, N. (1993): Statistics for Spatial Data. Wiley-Interscience. 1993.
- Delle Monache, L.; Nipen, T.; Deng, X.; Zhou, Y.; Stull, R. (2006): Ozone ensemble forecasts: 2. A Kalman filter predictor bias correction. Journal of Geophysical Research, 111, D05308. 2006.
- Delle Monache, L.; Wilczak, J.; McKeen, S.; Grell, G.; Pagowski, M.; Peckham, S.; Stull, R.; McHenry, J.; McQueen, J. (2008): A Kalman-filter bias correction method applied to deterministic, ensemble averaged and probabilistic forecasts of surface ozone. Tellus B, Vol. 60, Issue 2, S. 238-249. <https://doi.org/10.1111/j.1600-0889.2007.00332.x>. 2008.
- Diegmann, V.; Wiegand, G. (2000): FLADIS - A system for extending air pollution point data to continuous spatial information (according to EU Council Directive 96/62/EC). WIT Press, 2000. - 1-85312-822-8.
- Duchon J. (1977): Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In: Schempp W., Zeller K. (Hrsg.): Constructive Theory of Functions of Several Variables. Lecture Notes in Mathematics, Vol. 571. Springer, Berlin, Heidelberg, 1977.
- DWD (2015a): Deutschlandwetter im Juni 2015. Deutscher Wetterdienst, Pressemitteilung 29.06.2015.
- DWD (2015b): Deutschlandwetter im Juli 2015. Deutscher Wetterdienst, Pressemitteilung 30.07.2015.
- DWD (2017a): Deutschlandwetter im Januar 2017. Deutscher Wetterdienst, Pressemitteilung 30.01.2017.
- DWD (2017b): Deutschlandwetter im Februar 2017. Deutscher Wetterdienst, Pressemitteilung 27.02.2017.
- EU (2008): Richtlinie 2008/50/EG des Europäischen Parlaments und des Rates vom 21. Mai 2008 über Luftqualität und saubere Luft für Europa. Amtsblatt der Europäischen Union Nr. L 152, S. 1. 2008.
- FAIRMODE (2016): <http://fairmode.jrc.ec.europa.eu/>
- FAIRMODE (2017): Guidance Document on Model Quality Objectives and Benchmarking. Version 2.1. Forum for Air Quality Modelling in Europe (FAIRMODE). 2017.
- Flemming, J. (2003): Immissionsfelder aus Beobachtung, Modellierung und deren Kombination; Dissertation, Freie Universität Berlin, Fachbereich Geowissenschaften, <http://www.diss.fu-berlin.de/2003/71>, 2003.
- Franke, R.; Nielson, G. (1980): Smooth interpolation of large sets of scattered data; Inter. J. for Num. Meth. in Eng.; Vol.8, No.4, 1980.
- Gandin, L. S. (1965): Objective Analysis of meteorological fields. Translated 1965 from Russian by Israeli Program for Scientific Translations. Jerusalem, 1965.
- Hardy, R. L. (1990): Theory and application of the multiquadratic-biharmonic method, Comp. Math. Appl., Vol.19, No.8/9, 1990.
- Inness, A.; Baier, F.; Benedetti, A.; Bouarar, I.; Chabrilat, S.; Clark, H.; Clerbaux, C.; Coheur, P.; Engelen, R. J.; Errera, Q.; Flemming, J.; George, M.; Granier, C.; Hadji-Lazaro, J.; Huijnen, V.; Hurtmans, D.; Jones, L.; Kaiser, J. W.; Kapsomenakis, J.; Lefever, K.; Leitão, J.; Razinger, M.; Richter, A.; Schultz, M. G.; Simmons, A. J.; Suttie, M.; Stein, O.; Thépaut, J.-N.; Thouret, V.; Vrekoussis, M.; Zerefos, C.; the MACC team (2013): The MACC reanalysis: an 8 yr data set of atmospheric composition. Atmos. Chem. Phys. Vol. 13, pp. 4073-4109, 2013. <https://doi.org/10.5194/acp-13-4073-2013>.
- IVU GmbH (1992): Flächenhafte Darstellung der Immissionssituation. Im Auftrag der Hessischen Landesanstalt für Umwelt. 1992.
- IVU GmbH (1996): Flächenhafte Darstellung der Immissionssituation. In: HLFU (Hrsg.): Umweltplanung, Arbeits- und Umweltschutz. Schriftenreihe der Hessischen Landesanstalt für Umwelt, Heft 201. 1996, - 3-89026-212-0,
- IVU Umwelt (2011): Tendenzen und Verursacher für die NO₂-Belastung in Sachsen. Im Auftrag des Sächsischen Landesamts für Umwelt, Landwirtschaft und Geologie. Schriftenreihe des LfULG, Heft 5/2011. 2011.

- IVU Umwelt (2014): Fehlwertersetzung in Messzeitreihen der LUBW. Im Auftrag der Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg. 2014.
- IVU Umwelt (2016): Benutzeranleitung FLADIS 4.2. 2016.
- IVU Umwelt (2018): Benutzeranleitung Online-Tool Ozonprognose. Mai 2018.
- Kang, D.; Mathur, R.; Rao, S.T.; Yu, S. (2008): Bias adjustment techniques for improving ozone air quality forecasts, *Journal of Geophysical Research*, 113, D23308. 2008.
- KRdL (2009a): VDI Richtlinie 4280 Blatt 4 " Planung von Immissionsmessungen - Ersetzung von Fehlwerten in Messreihen zur Ermittlung der Luftqualität." . Kommission Reinhaltung der Luft im VDI und DIN. 2009.
- KRdL (2009b): VDI Richtlinie 4280 Blatt 5 " Planung von Immissionsmessungen - Ermittlung der Unsicherheit räumlicher Beurteilungen der Luftqualität." . Kommission Reinhaltung der Luft im VDI und DIN. 2009.
- McKeen, S.; Wilczak, J.; Grell, G.; Djalalova, I.; Peckham, S.; Hsie, E.-Y.; Gong, W.; Bouchet, V.; Menard, S.; Moffet, R.; McHenry, J.; McQueen, J.; Tang, Y.; Carmichael, G. R.; Pagowski, M.; Chan, A.; Dye, T.; Frost, G.; Lee, P.; Mathur, R. (2005): Assessment of an ensemble of seven real-time ozone forecasts over eastern North America during the summer of 2004. *Journal of Geophysical Research*. 110, D21307. 2005.
- Neal, L. S.; Agnew, P.; Moseley, S.; Ordóñez, C.; Savage, N.H.; Tilbee, M. (2014): Application of a statistical post-processing technique to a gridded, operational, air quality forecast. In: *Atmospheric Environment* 98 (2014) S. 385-393, 2014.
- Pfäfflin, F.; Diegmann, V.; Stern, R.; Scholz, W. (2015): Flächendeckende Ermittlung der Immissions-Vorbelastung für Baden-Württemberg. Tagungsband zur Fachtagung Mettools IX, 17.-19.3.2015 in Offenbach. 2015.
- Pfäfflin, F. (2016): Luftqualitätsmodellierung und Exposition - Effekte der Berechnungsskalen. 2. Deutsches FAIRMODE Treffen. Veranstalter: UBA, 25.-26.10.2016 in Berlin. 2016.
- R: The R Project for Statistical Computing. <http://www.r-project.org>.
- RLuS (2012): Richtlinien zur Ermittlung der Luftqualität an Straßen ohne oder mit lockerer Randbebauung – RLuS 2012. Forschungsgesellschaft für Straßen- und Verkehrswesen, Arbeitsgruppe Straßenentwurf. Ausgabe 2012.
- Savage, N. H.; Agnew, P.; Davis, L. S.; Ordóñez, C.; Thorpe, R.; Johnson, C. E.; O'Connor, F. M.; Dalvi, M. (2013): Air quality modelling using the Met Office Unified Model (AQUA OS24-26): model description and initial evaluation. In: *Geosci. Model Dev.*, 6, S. 353–372, <https://doi.org/10.5194/gmd-6-353-2013>, 2013.
- Shepard, D. (1968): A two dimensional interpolation function for irregular spaced data; *ACM '68 Proceedings of the 1968 23rd ACM national conference*, pp. 517-524, 1968.
- Thunis, P.; Pederzoli, A.; Pernigotti, D., (2012): Performance criteria to evaluate air quality modeling applications. *Atmospheric Environment* 59 S. 476–482. 2012.
- Thunis, P.; Cuvelier, C. (2017): Fairmode DELTA Version 5.5. Concepts / User's Guide / Diagrams. Joint Research Centre, Ispra. 2017.
- Tusell, F. (2011): Kalman Filtering in R. *Journal of Statistical Software*, Vol. 39, Issue 2. 2011.