



universität  
wien

# DIPLOMARBEIT

Titel der Diplomarbeit

## Qualitätskontrolle synoptischer Daten

angestrebter akademischer Grad

MAGISTRA DER NATURWISSENSCHAFTEN  
(MAG.RER.NAT.)

Verfasserin: Andrea Steiner  
Matrikel-Nummer: 0501629  
Studienrichtung: A 415 (Meteorologie)  
Betreuer: Univ.Prof.Dr. Reinhold Steinacker

Wien, November 2011



---

# Zusammenfassung

Die vorliegende Diplomarbeit behandelt die am Institut für Meteorologie und Geophysik der Universität Wien entwickelte Qualitätskontrolle (QC), die dem operationell ausgeführten Analyseverfahren VERA (Vienna Enhanced Resolution Analysis) vorgeschaltet ist und auch bei Feldexperimenten oder im Zuge der Modellverifikation zum Einsatz kommt.

Nach einem Überblick über die derzeit gängigen Qualitätskontrollmethoden wird der mathematische Hintergrund der sogenannten VERA-QC präsentiert. Für Parameter die eine hohe Redundanz in den Messwerten aufweisen wird die grundlegende Bedingung eines glatten Analysefeldes mit Hilfe einer Kostenfunktion formuliert und als Variationsproblem gelöst.

Weiters wird auf Besonderheiten eingegangen, welche die VERA-QC von herkömmlichen Qualitätskontrollen abheben und die Kontrolle von Daten in komplexem Gelände ermöglichen. Für die räumliche Konsistenzprüfung der Messwerte werden benachbarte Stationen mit der Methode der natürlichen Nachbarfindung definiert, welche sich automatisch an eine variierende Stationsdichte anpasst. Eine spezielle Cluster-Behandlung löst das Problem von zu eng benachbarten Stationen und eine Gross Error Erkennung, die ebenso auf der Minimierung der Krümmung des Analysefeldes basiert, verwirft Meldungen mit groben Fehlern.

Im Rahmen dieser Diplomarbeit wurde die VERA-QC um eine Bias-Korrektur erweitert, die systematische Fehler korrigiert und sich dynamisch auf neue Gegebenheiten bezüglich des Beobachtungsfeldes einstellt. Um zu überprüfen, ob eine Beobachtung im Vergleich zum lokalen Klima realistische Werte liefert, wird eine Methode vorgestellt wie aus Stationszeitreihen lokale klimatologisch mögliche Grenzwerte für einen Limit-Check ermittelt werden können. Da die räumliche Qualitätskontrolle von Temperaturwerten, vor allem in komplexem Gelände, aufgrund der reduzierten Talvolumina und der damit verbundenen Überhitzung problematisch ist, wurde in der VERA-QC die Möglichkeit implementiert, an Stelle der potentiellen Temperatur die Abweichung der potentiellen Temperatur zur entsprechenden Temperatur der Standardatmosphäre zu kontrollieren. Auch die Gross Error Erkennung wurde um variable stationsspezifische Grenzwerte, die aus den gesammelten Korrekturvorschlägen berechnet werden, erweitert. Die VERA-QC wird mit den beschriebenen Erweiterungen operationell zu jeder vollen Stunde ausgeführt und liefert zufriedenstellende Ergebnisse. Um einen vollständigen und einheitlichen Datensatz zu erhalten, wurden die dem Institut zur Verfügung stehenden Stationsmeldungen des Jahres 2010 nachträglich von der VERA-QC kontrolliert. Die gesammelten Beobachtungen und die ausgewerteten ermittelten Korrekturvorschläge dieses Zeitraums lassen weiteres Potential für interessante Untersuchungen erkennen.

Die mathematischen Grundlagen sowie ein Vergleich der VERA-QC mit anderen Qualitätskontrollmethoden wurden bereits im Artikel „Data Quality Control Based on Self-Consistency“ in der meteorologischen Fachzeitschrift *Monthly Weather Review* veröffentlicht. Eine zweite anwendungsorientierte Publikation ist bereits in Arbeit. Im September 2011 wurden die überzeugenden Ergebnisse der VERA-QC auf der meteorologischen Tagung ECAM (European Conference on Application of Meteorology) in Berlin und im November 2011 im Rahmen des 4. Österreichischen MeteorologInnentags in Klagenfurt vorgestellt.

---

# Abstract

This diploma thesis discusses the quality control (QC) method VERA-QC, which was developed at the Department of Meteorology and Geophysics at the University of Vienna. It is used as a preprocessing tool for VERA, the department's Vienna Enhanced Resolution Analysis of basic meteorological parameters.

After summarizing the most commonly used quality control methods, the mathematical background of VERA-QC is explained. For measurements exhibiting a high redundancy, the analysis field of the considered parameter is required to be distributed smoothly. This condition is formulated in terms of a cost function that, after minimization, leads to a variational problem which can be solved as a matrix equation.

The VERA-QC offers some notable features and therefore is preferable compared to other common QC-methods. Using the concept of finding natural neighbors, VERA-QC can automatically adapt to different station densities as it is often the case in complex terrain or in coastal areas. There is also a special cluster treatment and an advanced gross error recognition method based on the minimization of the analysis field's curvature as well.

Furthermore, additional improvements, which enhance the performance of the operational VERA-QC, have been implemented. By collecting the proposed corrections (the so called deviations), a bias correction can be computed. It handles systematic errors and automatically adapts to changes in the observation field. As for the fact that the assumption of a dry adiabatic lapse rate (used when computing potential temperature) is, in general, not fulfilled in e.g. mountainous regions, the spatial comparison of potential temperature values is challenging. Therefore, instead of checking the potential temperatures for their spatial consistency, the departure from the corresponding standard atmosphere's potential temperature is controlled by the VERA-QC. Moreover, the criteria for recognizing a gross error have been expanded by variable station specific limits. A method to calculate climatologically reasonable values based on station time series is also presented.

Observations and quality controlled data covering the period of the year 2010 have been computed and evaluated. The results are promising and indicate that there is still potential for further interesting research. By now, the VERA-QC is running operationally on an hourly basis and performs satisfactorily.

Concluding, it should be remarked that the mathematical basics and a comparison of the VERA-QC to other QC methods is introduced in the article „Data Quality Control Based on Self-Consistency“ which is published in the meteorological journal *Monthly Weather Review*. Another article focusing on the operational outcome of VERA-QC is currently in preparation. Both were presented in September 2011 in Berlin at the meteorological conference ECAM (European Conference on Application of Meteorology) and in November 2011 in Klagenfurt at the conference 4. Österreichischer MeteorologInnentag.

# Inhaltsverzeichnis

Abkürzungsverzeichnis	vii
Abbildungsverzeichnis	viii
Tabellenverzeichnis	xi
<b>1 Einleitung</b>	<b>1</b>
1.1 Meteorologische Messungen und mögliche Fehler . . . . .	1
1.2 Begriffsdefinition: Qualitätssicherung und Qualitätskontrolle . . . . .	3
1.2.1 Automatische Qualitätskontrolle . . . . .	4
1.3 Gängige Qualitätskontrollmethoden . . . . .	4
<b>2 Die Qualitätskontrolle für VERA: VERA-QC</b>	<b>11</b>
2.1 Problemstellung und mathematische Formulierung . . . . .	11
2.1.1 Das Konzept der natürlichen Nachbarfindung . . . . .	12
2.1.2 Aufstellung der Kostenfunktion . . . . .	13
2.1.3 Minimierung der Kostenfunktion . . . . .	16
2.1.4 Diskretisierung . . . . .	16
2.1.5 Kostenfunktionsreduktion . . . . .	18
2.1.6 Beurteilung der Stationsmeldungen (akzeptieren, korrigieren oder verwerfen) . . . . .	20
2.1.7 Behandlung von Stationscluster . . . . .	20
2.1.8 Ablauf der VERA-QC mit Cluster-Behandlung und Gross Error Erkennung . . . . .	21
<b>3 Die Bias-Korrektur</b>	<b>23</b>
3.1 Begriffsdefinition . . . . .	23
3.2 Das Problem systematischer Fehler und deren Lösung: die Bias- Korrektur . . . . .	23
3.3 Auswertung der Bias-Korrekturen . . . . .	27
3.3.1 Reduzierter Bodendruck . . . . .	27
3.3.2 Potentielle Temperatur . . . . .	29
3.3.3 Äquivalentpotentielle Temperatur . . . . .	29

---

3.3.4	Windgeschwindigkeit . . . . .	31
<b>4</b>	<b>Der Limit-Check</b>	<b>35</b>
4.1	Einführung klimatologischer Limits . . . . .	35
4.1.1	Limit-Check . . . . .	35
4.1.2	Klimatologische Grenzwerte . . . . .	36
4.1.3	Auswertung der klimatologischen Grenzwerte . . . . .	41
4.1.4	Alternative Berechnung klimatologischer Limits . . . . .	45
<b>5</b>	<b>Qualitätskontrolle der potentiellen Temperatur</b>	<b>49</b>
5.1	Die potentielle Temperatur . . . . .	49
5.2	Problemstellung . . . . .	50
5.3	Lösungsvorschlag . . . . .	53
<b>6</b>	<b>Die variable Gross Error Schwelle</b>	<b>59</b>
6.1	Definition einer stationspezifischen variablen Gross Error Schwelle . . . . .	59
6.2	Anwendung der variablen Gross Error Schwelle im Jahr 2010 . . . . .	60
<b>7</b>	<b>Statistische Auswertung der Stationsmeldungen aus dem Jahr 2010</b>	<b>63</b>
7.1	Durchschnittliche Anzahl an Stationsmeldungen . . . . .	64
7.2	Beurteilung der Qualität der Stationsmeldungen durch die VERA-QC . . . . .	66
7.2.1	Vergleich mit den Ergebnissen des Projektes ALPEX . . . . .	70
<b>8</b>	<b>Statistische Auswertung der Deviations aus dem Jahr 2010</b>	<b>75</b>
8.1	Erkennung falscher SYNOP-Verschlüsselungen . . . . .	75
8.2	Erkennung falscher Einträge in der Stationsliste . . . . .	77
8.3	Erkennung mikrometeorologischer Phänomene . . . . .	77
8.4	Analyse auf Periodizitäten . . . . .	79
8.5	Meldungen der österreichischen internationalen Flughäfen . . . . .	86
<b>9</b>	<b>Conclusio</b>	<b>91</b>
	<b>Literaturverzeichnis</b>	<b>95</b>
<b>A</b>	<b>Konfigurationsdatei</b>	<b>99</b>
A.1	Konfigurationseinstellungen der VERA-QC . . . . .	100
<b>B</b>	<b>Statistische Auswertungen</b>	<b>103</b>
B.1	Statistische Auswertungen der Stationsmeldungen . . . . .	104
B.2	Anteil von akzeptierten, korrigierten und verworfenen Meldungen . . . . .	107
B.3	Anteil der Gross Errors . . . . .	110
<b>C</b>	<b>Lebenslauf</b>	<b>113</b>



# Abkürzungsverzeichnis

ALPEX	Alpine Experiment
BQC	Bayesian Quality Control
COPS	Convective and Orographically induced Precipitation Study
CQC	Complex Quality Control
DMA	Decision Making Algorithm
ECAM	European Conference on Application of Meteorology
ECMWF	European Centre for Medium-Range Weather Forecasts
GARP	Global Atmospheric Research Programme
GAW	Global Atmosphere Watch
GE	Gross Error
GFS	Global Forecast System
GTS	Global Telecommunications System
ICAO	International Civil Aviation Organization
ID	Inverse Distance
IQR	Interquartile Range
ISO	International Organization for Standardization
NCDC	National Climatic Data Center
NWP	Numerical Weather Prediction
OI	Optimum Interpolation
Q1	unteres (1.) Quartil
Q3	oberes (3.) Quartil
QA	Quality Assurance
QC	Quality Control
RMSE	Root Mean Square Error
SR	Spatial Regression
SYNOP	Synoptic Observations
UTC	Universal Time Coordinated
VarQC	Variational Quality Control
VERA	Vienna Enhanced Resolution Analysis
WMO	World Meteorological Organization

# Abbildungsverzeichnis

Abb. 1.1:	Gegenüberstellung von Mittelwert und wahren Wert normalverteilter Messwerte . . . . .	2
Abb. 2.1:	Veranschaulichung des Prinzips der „natürlichen Nachbarfindung“ für 9 Stationen . . . . .	12
Abb. 2.2:	Stationsverteilung mit 9 Stationen und den zur Diskretisierung benötigten Stützstellen . . . . .	17
Abb. 3.1:	Lage der britischen Station, mit der die Bias-Korrektur erklärt wird . . .	24
Abb. 3.2:	Erklärung der Bias-Korrektur in mehreren Schritten anhand des Parameters reduzierter Druck . . . . .	25
Abb. 3.3:	Zeitreihe der ungewichteten Deviations des reduzierten Drucks und Bias-Korrektur . . . . .	27
Abb. 3.4:	Bias-Korrekturen des reduzierten Drucks im Jänner, März, Mai und Juli 2010 . . . . .	28
Abb. 3.5:	Klassische und modifizierte Bias-Korrektur der äquivalentpotentiellen Temperatur im Oktober 2010 . . . . .	29
Abb. 3.6:	Schematische Darstellung der Möglichkeiten, Deviations für die Bias-Berechnung auszuwählen . . . . .	30
Abb. 3.7:	Bias-Korrektur der Windgeschwindigkeit für April 2010 . . . . .	32
Abb. 4.1:	Messzeitreihen und Häufigkeitsverteilungen für den reduzierten Bodendruck und die 6-stündigen Niederschlagssummen der Station Wien - Hohe Warte . . . . .	36
Abb. 4.2:	Schematische Darstellung der Normal Score Transformation . . . . .	38
Abb. 4.3:	Mittels Interquartilsabstand definierte klimatologische Grenzwerte bezüglich einer Standardnormalverteilung . . . . .	38
Abb. 4.4:	Berechnung klimatologischer Limits für 6-h Niederschlagsmesswerte mit Hilfe einer Normal Score Transformation . . . . .	39
Abb. 4.5:	Messzeitreihe der potentiellen Temperatur der Station Wien - Hohe Warte mit klimatologischen Limits . . . . .	40
Abb. 4.6:	Stationsspezifisches unteres klimatologisches Limit für den reduzierten Bodendruck . . . . .	42

Abb. 4.7:	Unteres klimatologisches Limit und Spannweite der Limits für den Stationsdruck . . . . .	42
Abb. 4.8:	Spannweite der klimatologischen Limits für die potentielle Temperatur .	43
Abb. 4.9:	Oberes klimatologisches Limit für die 6-stündigen Niederschlagssummen	44
Abb. 4.10:	Veranschaulichung einer alternativen Berechnung klimatologischer Limits anhand einer simulierten rechtsschiefen Verteilung . . . . .	45
Abb. 4.11:	Vergleich zweier Berechnungsmöglichkeiten von klimatologischen Grenzwerten . . . . .	47
Abb. 4.12:	Veranschaulichung der Problematik einer steilgipfligen Verteilung bei der Berechnung klimatologischer Grenzwerte . . . . .	48
Abb. 5.1:	Höhenabhängigkeit der Bias-Korrektur für die potentielle Temperatur .	51
Abb. 5.2:	Gegenüberstellung der Bias-Korrektur der potentiellen Temperatur im Winter und im Sommer . . . . .	52
Abb. 5.3:	Jahresgang der Bias-Korrektur der potentiellen Temperatur für Stationen in Spanien . . . . .	53
Abb. 5.4:	Bias-Korrektur des D-Wertes der potentiellen Temperatur und flächenmittelbereinigte Bias-Korrektur für Juli 2010 . . . . .	55
Abb. 5.5:	VERA-Analysen der potentiellen Temperatur vom 6. 9. 2010, nach drei unterschiedlichen QC-Varianten. . . . .	57
Abb. 6.1:	Gewichtete Deviations der äquivalentpotentiellen Temperatur und die daraus berechnete, variable Gross Error Schwelle . . . . .	59
Abb. 6.2:	Stationen, für welche die stationsspezifische variable Gross Error Schwelle Verwendung findet . . . . .	62
Abb. 7.1:	Stationen innerhalb der Europa-Domäne mit Verbindungslinien . . . . .	63
Abb. 7.2:	Durchschnittliche Anzahl der zur Verfügung stehenden Stationsmeldungen je Land . . . . .	65
Abb. 7.3:	Häufigkeitsverteilungen der berechneten Korrekturen für den reduzierten Druck und die potentielle Temperatur . . . . .	67
Abb. 7.4:	Anteile der akzeptierten, korrigierten und verworfenen Stationsmeldungen je Land . . . . .	69
Abb. 7.5:	Anteile der erkannten Gross Errors bezüglich aller Stationsmeldungen . .	70
Abb. 8.1:	Vergleich der gemeldeten Windgeschwindigkeiten von der Station Berlin - Alexanderplatz und Umgebung . . . . .	76
Abb. 8.2:	Darstellung der Deviations für die Station Wien - Innere Stadt und für umliegende Stationen . . . . .	78
Abb. 8.3:	Differenz zwischen den gemeldeten potentiellen Temperaturen der Station Wien - Innere Stadt und umliegenden Stationen . . . . .	78

---

Abb. 8.4:	Zeitreihe der gemeldeten potentiellen Temperaturen und der gewichteten Korrekturen der Station Innsbruck - Flughafen im Jahr 2010 . . . . .	79
Abb. 8.5:	Einseitiges diskretes Varianzspektrum für die potentielle Temperatur der Station Innsbruck - Flughafen im Jahr 2010 . . . . .	81
Abb. 8.6:	Einseitiges diskretes Varianzspektrum für die gewichteten Deviations des D-Wertes der potentiellen Temperatur der Station Innsbruck - Flughafen im Jahr 2010 . . . . .	82
Abb. 8.7:	Zeitreihe der potentiellen Temperatur und der zugehörigen gewichteten Deviations der Station Bischofshofen im Jahr 2010 . . . . .	83
Abb. 8.8:	Einseitiges diskretes Varianzspektrum der potentiellen Temperatur und der gewichteten Deviations der Station Bischofshofen im Jahr 2010 . . .	84
Abb. 8.9:	Fallstudie zur Erklärung der Periodizitäten in den Deviations der potentiellen Temperatur der Messstation Bischofshofen . . . . .	85
Abb. 8.10:	Gewichtete Deviations des reduzierten Bodendrucks, gemeldet von den österreichischen internationalen Flughäfen im Jahr 2010 . . . . .	89
Abb. 8.11:	Stationsverteilung in Österreich inklusive Stationen, die seit 16. 12. 2010 melden . . . . .	90
Abb. 8.12:	Gewichtete Deviations der Station LOWK für den Zeitraum Oktober 2010–März 2011 . . . . .	90

# Tabellenverzeichnis

Tab. 1.1:	Gegenüberstellung der Varianz benachbarter Stationen mit der Varianz der gesamten Europa-Domäne . . . . .	6
Tab. 6.1:	Statistik der Gross Error Erkennung mit und ohne der zusätzlichen stationsspezifischen IQR-Schwelle im Jahr 2010 . . . . .	61
Tab. 7.1:	Prozentuelle Aufschlüsselung der Stationsmeldungen nach Parameter und Beurteilung durch die VERA-QC . . . . .	66
Tab. 7.2:	Gegenüberstellung der im Zuge von ALPEX und dieser Diplomarbeit identifizierten Gross Errors für Länder in Europa . . . . .	74
Tab. 8.1:	Anzahl der fehlenden stündlichen Meldungen der Windkomponenten von österreichischen internationalen Flughäfen . . . . .	86
Tab. 8.2:	Statistik der Stationsmeldungen der österreichischen internationalen Flughäfen . . . . .	87
Tab. 8.3:	Bias-Korrekturen für die österreichischen internationalen Flughäfen . . . . .	88
Tab. B.1:	Durchschnittliche Anzahl der stündlichen Stationsmeldungen im Jahr 2010, aufgeschlüsselt nach Ländern . . . . .	104
Tab. B.2:	Durchschnittliche Anzahl der stündlichen Stationsmeldungen im Jahr 2010, bezogen auf die Fläche des jeweiligen Landes . . . . .	105
Tab. B.3:	Durchschnittliche Anzahl der stündlichen Stationsmeldungen im Jahr 2010, bezogen auf den jeweiligen Flächenanteil innerhalb der betrachteten Europadomäne . . . . .	106
Tab. B.4:	Beurteilung (akzeptieren/korrigieren/verwerfen) der Stationsmeldungen für den reduzierten Druck im Jahr 2010, sortiert nach Ländern . . . . .	107
Tab. B.5:	Beurteilung (akzeptieren/korrigieren/verwerfen) der Stationsmeldungen für die potentielle Temperatur im Jahr 2010, sortiert nach Ländern . . . . .	108
Tab. B.6:	Beurteilung (akzeptieren/korrigieren/verwerfen) der Stationsmeldungen für die Windgeschwindigkeit im Jahr 2010, sortiert nach Ländern . . . . .	109
Tab. B.7:	Anteil der erkannten Gross Errors an den Werten für den reduzierten Druck, die im Jahr 2010 zur Verfügung stehen (aufgeschlüsselt nach Ländern) . . . . .	110

Tab. B.8: Anteil der erkannten Gross Errors an den Werten für die potentielle Temperatur, die im Jahr 2010 zur Verfügung stehen (aufgeschlüsselt nach Ländern) . . . . .	111
Tab. B.9: Anteil der erkannten Gross Errors an den Werten für die äquivalentpotentielle Temperatur, die im Jahr 2010 zur Verfügung stehen (aufgeschlüsselt nach Ländern) . . . . .	112





# Kapitel 1

## Einleitung

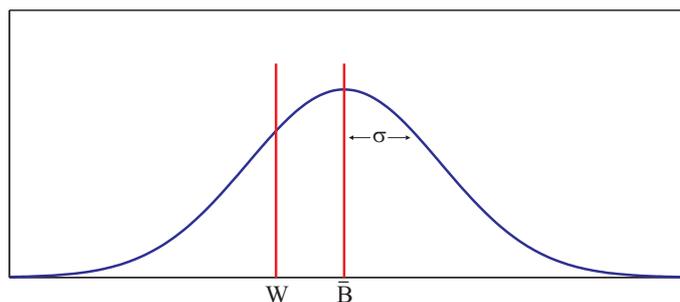
### 1.1 Meteorologische Messungen und mögliche Fehler

Beobachtungen der Natur stellen eine wichtige Grundlage jeder Naturwissenschaft dar und man versucht sie durch Messungen zu quantifizieren. In der Meteorologie handelt es sich beim beobachteten System meist um die Atmosphäre. Ziel einer meteorologischen Messung ist es somit, Erkenntnisse über den Zustand der Atmosphäre und der darin ablaufenden Prozesse zu erlangen. Diese werden durch die gemessenen Zustandsvariablen und ihre zeitlichen Änderungen beschrieben. Bei einer Messung handelt es sich um einen experimentellen Vorgang, der eine ausgewählte Zustandsvariable des betrachteten Systems mit einer zugehörigen, vordefinierten Referenz-Skala vergleicht (z.B.: die Zustandsvariable Temperatur mit der Kelvin-Skala). Das Ergebnis ist der Messwert, der sich aus dem bestimmten Zahlenwert und der Einheit der gemessenen Variable zusammensetzt.

Grundsätzlich kann man zwischen direkten und indirekten Messmethoden unterscheiden:

- Direkte Messmethoden basieren auf einem direkten Vergleich mit der Referenzskala, wie es bei der Messung von Länge, Masse und Zeit der Fall ist.
- Indirekte Messmethoden beruhen auf der Beobachtung, welche Auswirkung eine Zustandsvariable auf den verwendeten Sensor hat. Durch die Variationen des Messgerätes kann dann indirekt auf die Veränderungen der betrachteten Variable geschlossen werden.

Man muss beachten, dass vor allem indirekt erlangte Messwerte nur eine Annäherung an den wahren Wert darstellen und zusätzlich jede Messung mit Fehlern behaftet ist. In WMO (2008), herausgegeben von der World Meteorological Organization



**Abbildung 1.1:** Ergebnis (blau) einer mit einem Messinstrument durchgeführten Messreihe. Betrachtet wurde eine normalverteilte Variable mit dem wahren Wert  $W$  und dem Mittelwert der beobachteten Werte  $\bar{B}$ .  $\sigma$  bezeichnet die Standardabweichung.

(WMO), ist der wahre Wert einer Variable, der sich bei einer perfekten (jedoch undurchführbaren) Messung ergeben würde, definiert als der gemessene Wert  $\pm$  einem Unsicherheitsfaktor. Die unsichere Abweichung des gemessenen Wertes zum wahren Wert wird durch eine Vielzahl von Messfehlern verursacht, die mit Hilfe der Abbildung 1.1 wie folgt klassifiziert werden können:

- **Systematische Fehler** werden oft auch Bias genannt. Die Ursache dieser Fehler liegt in der gleichbleibenden Unvollkommenheit des Messverfahrens bzw. des Messgerätes. Beispiele dafür sind eine schlechte Kalibrierung, eine falsche Messanordnung, falsche Bedienung des Gerätes oder falsche Datenverarbeitung, aber auch die Rückwirkung des Messgerätes auf die zu messende Variable. Diese Fehler sind zeitlich konstant und können durch eine Bias-Korrektur reduziert werden. Mit Hilfe der Abbildung 1.1 kann der systematische Fehler einer normalverteilten Variable als Differenz des Mittelwertes der Messung und dem wahren Wert ( $\bar{B} - W$ ) berechnet werden.
- **Zufällige Fehler** entstehen durch zufällige Prozesse während der Messung und betreffen das Messgerät selbst, die direkte Umgebung oder den Beobachter. Ihr Vorzeichen variiert und sie sind räumlich und zeitlich nicht korreliert. Zufällige Fehler können, bei genügend oft wiederholter Messung, durch die Standardabweichung  $\sigma$  quantifiziert werden (vgl. Abbildung 1.1).
- Ein vorliegendes Messnetz kann nur einen gewissen räumlichen und zeitlichen Bereich meteorologischer Phänomene abdecken. Kleinskaligere Phänomene können nicht aufgelöst werden und führen zu **Repräsentativitätsfehlern**. In Bezug auf das operationelle europäische GTS-Messnetz (GTS, englisch: *global telecommunications system*) führen städtische Wärmeinseln, Küsteneffekte, kleinskalig orographisch bedingte Effekte oder auch subskalige Turbulenzen zu diesen mikrometeorologischen Fehlern. Abhängig von der Ursache

sind Repräsentativitätsfehler systematischer (Wärmeinseln, Küsteneffekte...) oder zufälliger Natur (Turbulenzen).

- Eine besondere Art von zufälligen Fehlern stellen die **Groben Fehler bzw. Gross Errors (GE)** dar. Sie sind charakterisiert durch ihr seltenes und zufälliges Auftreten (nicht normalverteilt), wie auch durch ihre große Amplitude, und werden meist von Fehlern bei der Beobachtung oder der Datenverarbeitung und -übertragung verursacht. Da Gross Errors ein Messergebnis stark verfälschen können, muss ihnen besondere Aufmerksamkeit geschenkt werden.

Je nachdem, welchem Zweck die Messdaten dienen, werden sie vor einer weiterführenden Analyse auf Fehler untersucht und wenn möglich auch korrigiert.

Die in diesem Kapitel angeführten Definitionen und Zusammenhänge können unter anderem in Emeis (2010) bzw. Strangeways (2003) (vor allem erstes Kapitel) oder in WMO (2008) nachgelesen werden. Einen guten Überblick über die Fehlerarten bieten Steinacker *et al.* (2000) oder Gandin (1988).

## 1.2 Begriffsdefinition: Qualitätssicherung und Qualitätskontrolle

Die beiden Begriffe Qualitätssicherung (englisch: *quality assurance*) und Qualitätskontrolle (englisch: *quality control*) finden in der Literatur unterschiedliche Verwendung. Im Sinne der WMO-Definition, wie sie in Zahumenský (2004) wiedergegeben wird, umfasst die Qualitätssicherung alle geplanten und systematisch durchgeführten Schritte zur Sicherstellung, dass die gesammelten Daten konsistent sind, die verlangten Qualitätskriterien erfüllt werden und alle Schritte durch eine umfassende Dokumentation nachvollziehbar sind. Die Qualitätskontrolle hingegen bezeichnet alle operationell durchgeführten Methoden, um das Fehlen von Daten zu erkennen, fehlerhafte Messungen auszusortieren und nach Möglichkeit auch zu korrigieren. Dazu zählt auch die menschliche Qualitätskontrolle, welche regelmäßige Besichtigungen der Messstationen oder regelmäßige Kontrollen der Messdaten durch den Synoptiker umfasst.

Der Begriff Qualitätskontrolle wird in dieser Arbeit mit QC (nach *quality control*) abgekürzt und bezieht sich auf die automatische QC im engeren Sinne, umfasst also die computergestützten Methoden zur Kontrolle bereits erhobener Daten.

### 1.2.1 Automatische Qualitätskontrolle

Mit der heutzutage immer größer werdenden Anzahl an automatischen Messstationen und der damit auch steigenden Datenmenge ist es immer wichtiger geworden, über die Qualität der erhobenen Messwerte Bescheid zu wissen und diese auch zu steigern. Nicht nur die numerische Wettervorhersage benötigt qualitativ hochwertige Anfangsbedingungen (Lorenc und Hammon 1988, Ingleby und Lorenc 1993), auch klimatologische Daten werden geprüft und ihre Qualität wird verbessert (Feng *et al.* 2004). Letztere werden in weiterer Folge für Klimastudien herangezogen und dienen als Grundlage zur Beurteilung von Klimamodellen. Die Verifikation numerischer Vorhersagemodelle stellt ein zusätzliches Anwendungsgebiet qualitativ hochwertiger Messdaten dar und auch Mesonetze (Messnetze, die der Erforschung mesoskaliger Phänomene dienen) oder Messkampagnen profitieren von einer sorgfältigen Qualitätskontrolle ihrer Messwerte. Vor allem wenn Daten von verschiedenen Organisationen erhoben und ausgetauscht werden, ist ein einheitlicher Qualitätsstandard unerlässlich.

Dieses Problem wurde bereits früh erkannt und so hat die World Meteorological Organization (WMO) 1980 gewisse Qualitätskontrollmethoden für den internationalen Austausch von meteorologischen Daten im „*WMO No. 544: Manual on the Global Observing System*“ vorgeschrieben und dieses seitdem immer wieder erweitert (WMO 2003). Zusätzlich werden auch die Serien „*Guide on the global data-processing system*“, „*Guide on the global observing system*“ und „*Guide to Meteorological Instruments and Methods of Observation*“ (WMO 1993, 2007, 2008) geführt, die ebenfalls Empfehlungen und Richtlinien bezüglich der Qualität der erhobenen Daten und der verwendeten Instrumente enthalten. In Salvati und Brambilla (2008) findet man einen guten Überblick über die erwähnten Richtlinien und Empfehlungen der WMO.

Mit dem wachsenden Bewusstsein bezüglich der Unerlässlichkeit hinreichender Qualitätskontrollen wurde in den letzten Jahrzehnten eine Vielzahl von automatischen Qualitätskontrollmethoden entwickelt, die Fehler erkennen und ausbessern. Die Geläufigsten werden im folgenden Abschnitt zusammengefasst.

## 1.3 Gängige Qualitätskontrollmethoden

Die Aufgabe einer Qualitätskontrolle ist es, fehlerbehaftete Meldungen zu identifizieren und nach Möglichkeit auch zu korrigieren. Dazu können die Meldungen der Stationen getrennt voneinander geprüft werden, oder man macht von der Kontinuität der betrachteten Variable Gebrauch und vergleicht sie mit Nachbarstationen.

Einige Qualitätskontrollmethoden verwenden zur Fehlererkennung auch zusätzliche Informationen, wie statistische Grenzwerte oder Hintergrundfelder (englisch: *background fields*) und sind entweder in einem Vorhersagealgorithmus eingebettet, wie es bei vielen numerischen Vorhersagemodellen (englisch: *numerical weather prediction, NWP*) der Fall ist, oder sind als eigenständiges Programm entwickelt. Die Methoden variieren stark in ihrer Komplexität und können eine oder mehrere der folgenden Qualitätskontrollen umfassen.

- **Limit Check (auch: Range Test):** Diese einfache Kontrolle vergleicht jeden Messwert einzeln mit möglichen physikalischen und klimatologischen Grenzwerten und dient vor allem dazu, Gross Errors zu erkennen. Sinnvolle Wertebereiche der wichtigsten meteorologischen Messgrößen sind in Zahumenský (2004) bzw. in WMO (1993) angeführt. In Feng *et al.* (2004) wird der Limit Check zur Erstellung einer klimatologischen Datenbank verwendet und in Baker (1992) werden Radiosondendaten gegen klimatologische Limits geprüft.
- **Temporal Consistency Check (auch: Step Test, Persistence Test):** Hierbei wird überprüft, ob die Tendenzen in den Messwerten einer Station zeitlich realistisch sind, oder ob zu große Sprünge vorliegen (Step Test). Beim Persistence Test wird kontrolliert, ob die Variabilität der Messwerte einer Station hinreichend groß ist. Sonst besteht die Vermutung, dass das Messgerät an einem bestimmten Wert „hängengeblieben“ ist. Shafer *et al.* (2000) sowie Fiebrich und Crawford (2001) beschreiben, wie diese Methoden in die Qualitätskontrolle für das Oklahoma Mesonet eingebaut sind und geben in Tabellen die maximal erlaubten Schwankungen bzw. minimal verlangte Variabilität für unterschiedliche Parameter an.
- **Internal Consistency Check:** Im Vergleich zur Kontrolle auf zeitliche und räumliche Korrelation der Messwerte eines Parameters wird beim Internal Consistency Check geprüft, ob die Messungen unterschiedlicher Parameter an einer Station untereinander konsistent sind. Reek *et al.* (1992) beschreibt die QC eines Datensatzes des National Climatic Data Centers (NCDC), wobei auch auf die beschriebene interne Konsistenz der Messdaten Wert gelegt wird (z.B.:  $T_{min} \leq T_{max}$ ). Komplexere Internal Consistency Checks machen unter anderem von der hydrostatischen oder geostrophischen Relation Gebrauch (Baker 1992).
- **Spatial Consistency Check (auch: Buddy Check):**  
Wie bereits erwähnt, kann ein bestehendes Messnetz aufgrund seiner Dichte und Anordnung nur bestimmte meteorologische Phänomene erfassen. Dabei

**Tabelle 1.1:** Gegenüberstellung der Varianz benachbarter Stationen mit der Varianz der gesamten Europa-Domäne für die Parameter potentielle und äquivalentpotentielle Temperatur, reduzierter Bodendruck und 6-stündiger Niederschlag zum Haupttermin 24. August 2011, 12:00 UTC. Für die ersten drei Parameter wurden Stationen in nächster Umgebung der Station Wien - Hohe Warte als „Nachbarstationen“ betrachtet, für den Parameter Niederschlag wurden Stationen im Nordosten Europas, die teilweise vom selben Regenschauer getroffen wurden für den Vergleich als „Nachbarstationen“ ausgewählt. Die Berechnung der Varianz der Parameter potentielle und äquivalentpotentielle Temperatur erfolgte, nachdem das Feld einer Gradientenbereinigung unterzogen wurde.

Parameter	Varianz	
	Nachbarstationen	gesamtes Feld
Potentielle Temperatur	0.70 °C <sup>2</sup>	< 15.99 °C <sup>2</sup>
Äquivalentpotentielle Temperatur	11.91 °C <sup>2</sup>	< 68.79 °C <sup>2</sup>
Reduzierter Bodendruck	0.33 hPa <sup>2</sup>	< 9.48 hPa <sup>2</sup>
Niederschlag (alle Werte)	8.18 mm <sup>2</sup>	> 1.85 mm <sup>2</sup>
Niederschlag (Werte > 0 mm)	13.51 mm <sup>2</sup>	> 7.57 mm <sup>2</sup>

gilt, dass die räumliche und zeitliche Auflösung der Beobachtungen feiner (höher) sein muss als jene des betrachteten Vorgangs. So können Feldexperimente wie COPS (Convective and Orographically-induced Precipitation Study, Wulfmeyer *et al.* 2011) Daten liefern, welche die Untersuchung eines Regenschauers ermöglichen, der vom operationellen GTS-Messnetz nicht vollständig aufgelöst oder gar nicht erkannt wird. Geht man jedoch von einem operationellen synoptischen Messnetz aus und betrachtet Parameter wie den reduzierten Bodendruck, potentielle oder äquivalentpotentielle Temperatur und die damit verbundenen meteorologischen Phänomene wie extratropische Drucksysteme, so können die Feldverteilungen der betrachteten Parameter als „glatt“ angesehen werden. Um diese Eigenschaft auch quantitativ zu betrachten, sind in Tabelle 1.1 die berechneten Varianzen (jeweils an Nachbarstationen und im Vergleich dazu für die gesamte Domäne) eines zufällig ausgewählten Haupttermins (zu dem auch verbreitet konvektiver Niederschlag gemessen wurde: 24. August 2011, 12:00 UTC) für die Parameter potentielle Temperatur, äquivalentpotentielle Temperatur, reduzierter Bodendruck und 6-stündiger Niederschlag angegeben. Als Datengrundlage dienen die dem Institut für Meteorologie und Geophysik zur Verfügung stehenden europaweiten GTS-Messwerte. Für die ersten drei Parameter wurden Nachbarstationen der Station Wien - Hohe Warte betrachtet. Für den Niederschlag wurden benachbarte Stationen im Nordosten Europas herangezogen, die teilweise vom selben Regenschauer getroffen wurden. Aus der Tabelle 1.1 ist abzulesen, dass die Varianz der Messwerte an benachbarten Stationen (Stationen der sekundären Nachbar-

schaft, vergl. Kapitel 2) für die Parameter potentielle und äquivalentpotentielle Temperatur sowie für den reduzierten Bodendruck klein ist im Vergleich zur jeweiligen Varianz der gesamten Europa-Domäne. Somit wird das Feld hier als „glatt“ bezeichnet, und es können Messwerte von benachbarten Stationen in einem Spatial Consistency Check verwendet werden um die Qualität einer Beobachtung zu beurteilen. Betrachtet man hingegen den Parameter Niederschlag, so übertrifft (vor allem bei konvektiven Niederschlagsereignissen) die Varianz an benachbarten Stationen jene des gesamten Niederschlagsfeldes, wodurch dieses nicht als „glatt“ bezeichnet wird.

Die vielen möglichen Spatial Consistency Checks basieren alle auf dem Prinzip, dass für einen bestimmten Messwert auch ein Schätzwert aus den Daten der umliegenden Stationen berechnet wird. Weicht der Messwert zu sehr vom Schätzwert ab, wird die Meldung als verdächtig erkannt und entweder adaptiert oder verworfen. Die einzelnen Buddy Checks unterscheiden sich hauptsächlich in der Interpolationsmethode mit welcher der Schätzwert ermittelt wird, wobei die folgenden häufig Verwendung finden:

- **Inverse Distance Interpolation (ID):** Für die Interpolation werden alle umliegenden Stationen innerhalb eines bestimmten Radius (Nachbarstationen) herangezogen. Die betrachtete Station, für welche die Interpolation durchgeführt wird (Zentralstation), geht nicht ein. Der Schätzwert wird aus der Summe der gewichteten Messwerte der Nachbarstationen berechnet, wobei die Gewichtung mit dem Kehrwert der Entfernung der jeweiligen Nachbarstation zur betrachteten Zentralstation erfolgt. Dabei kann der Radius mit unterschiedlicher Potenz oder in ausgeklügelten Funktionen eingehen. Diese bekannte Methode ist unter anderem in Wade (1987) beschrieben, ein etwas fortgeschrittenerer Ansatz wird in Barnes (1964) vorgestellt.
- **Spatial Regression Interpolation (SR):** Ähnlich wie bei der ID Interpolation wird bei der SR ein fixer Einflussbereich angegeben. Zusätzlich müssen die darin liegenden Stationen aber auch eine geforderte Mindestkorrelation mit der Zentralstation aufweisen, um in die Interpolation einzugehen. Der Schätzwert wird ermittelt, indem man von jeder Nachbarstation ausgehend eine lineare Regression zur Zentralstation berechnet. Die einzelnen Schätzungen werden mit dem Root Mean Square Error (RMSE), der sich aus früheren Schätzungen ergibt, gewichtet. Mithilfe des aus der Statistik ermittelten RMSE und des berechneten Schätzwertes kann auch ein Konfidenzintervall bestimmt werden. Liegt

ein Messwert außerhalb dieser Grenzen, besteht der Verdacht eines Gross Errors. Eine detaillierte Beschreibung der SR Interpolation in Hubbard und You (2005) zeigt auch, dass die SR der ID Interpolation unter meteorologisch interessanten Bedingungen überlegen ist.

- **Polynomial Interpolation:** Als Erweiterung der SR Interpolation, die auf einer linearen Regression basiert, kann auch eine Polynominterpolation verwendet werden. Dabei wird ein Polynom n-ten Grades in die umliegenden Beobachtungen gelegt und so der Schätzwert ermittelt. Heutzutage werden neben Polynomen höherer Ordnung auch Splines verwendet, die stückweise aus Polynomen zusammengesetzte Funktionen darstellen. Steinacker *et al.* (2000) beschreibt, wie man eine Spline-Interpolation auch als Optimierungsproblem formulieren kann, um die Krümmung des betrachteten Feldes zu minimieren.
- **Optimum Interpolation (OI):** Eine Qualitätskontrolle, welcher die OI als Interpolationsmethode zugrunde liegt, basiert ebenfalls auf dem Vergleich zwischen Schätz- und Messwert. Die OI Methode ist modellabhängig und benötigt ein Hintergrundfeld (englisch: *background field*), dem das gewichtete Beobachtungsincrement (Differenz aus Beobachtungen und Schätzung der Beobachtungen, englisch: *first guess*) aufaddiert wird. Die Gewichtung wird, wie bei der SR, aus der Statistik ermittelt. Zusätzlich werden auch die Fehler, die das Beobachtungs- und Hintergrundfeld behaften, in Form von symmetrischen und positiv definiten Kovarianzmatrizen berücksichtigt. Ein Vorteil gegenüber den erwähnten Spatial Regression Checks ist die Möglichkeit der OI auch isolierte Stationen betrachten zu können. Für die Berechnung des Hintergrundfeldes können Nachbarstationen verwendet werden, sind dafür aber nicht zwingend notwendig. In Kalnay (2003) findet man im Kapitel 5 eine Zusammenfassung vieler gängiger Datenassimilationsmethoden, unter anderem wird auch die OI behandelt. Lorenc (1981) beschreibt eine etwas abgeänderte Form der OI.
- **Bayesian QC (BQC):** Im Gegensatz zu allen bereits erwähnten Qualitätskontrollmethoden weist die BQC einen großen Vorteil auf. Sie kann die Wahrscheinlichkeit angeben, mit der eine Messung einen Gross Error darstellt. Dafür benötigt sie die Angabe eines Hintergrund- und eines Beobachtungsfeldes sowie deren Fehlerverteilungen und eine A-priori-Schätzung der Gross Error Wahrscheinlichkeit. Prinzipiell gibt es zwei Möglichkeiten, die A-posteriori-Wahrscheinlichkeit eines Gross Errors zu bestimmen. Die Abschätzung, ob ein Gross Error vorliegt, kann entweder für eine Station nach der anderen einzeln erfolgen, oder man betrachtet verschiedene Gross Error Kombinationen

(eine Auswahl an Stationsmeldungen, die mit einem Gross Error behaftet sind) und ermittelt die wahrscheinlichste. Lorenc und Hammon (1988) oder Ingleby und Lorenc (1993) geben einen guten Überblick über den Formalismus der BQC und beschreiben auch die unterschiedlichen Wege, die A-posteriori-Wahrscheinlichkeit eines Gross Errors zu ermitteln.

- **Variational QC (VarQC):** In der numerischen Wettervorhersage wird aktuell von den großen Vorhersagezentren ein variationeller Ansatz (meist 4D-VAR) für die Datenassimilation und Vorhersage verwendet. Dabei wird eine Kostenfunktion minimiert, die sich im Grunde aus dem Hintergrund- und dem Beobachtungsfeld sowie den Kovarianzmatrizen der Fehler dieser Felder zusammensetzt. Die Qualitätskontrolle ist in diesem Ansatz direkt eingebaut, indem der Term des Beobachtungsfeldes in der Kostenfunktion mit der Gross Error Wahrscheinlichkeit gewichtet wird. Die A-priori-Wahrscheinlichkeit für einen Gross Error erhält man statistisch aus historischen Daten mit Hilfe des Bayes-Theorems. Andersson und Järvinen (1999) beschreiben die variationelle QC, wie sie am European Centre for Medium-Range Weather Forecasts (ECMWF) implementiert ist.
- **Complex QC (CQC):** Die CQC wird erstmals von Gandin (1988) erwähnt. Da sich viele der aufgezählten QC Methoden gegenseitig nicht ausschließen, können sie nacheinander ausgeführt werden und man erhält mehrere Vorschläge zur Qualität einer Messung. Die Beobachtungen erhalten Flags, welche die Beurteilungen durch die einzelnen QC Methoden sammeln. Anschließend wird mit Hilfe eines „Decision Making Algorithm“ (DMA) entschieden, wie die einzelnen Beobachtungen aufgrund ihrer Flags behandelt werden sollen.



# Kapitel 2

## Die Qualitätskontrolle für VERA: VERA-QC

### 2.1 Problemstellung und mathematische Formulierung

Europaweit werden von Wetterdiensten im Allgemeinen stündlich SYNOP-Meldungen abgegeben. Zu den Haupt- und Nebenterminen stehen dem Institut für Meteorologie und Geophysik durchschnittlich 1800 bzw. 1300 Meldungen (vgl. Kapitel 7) zur Verfügung, die von dem am Institut entwickelten Analyseverfahren VERA (Vienna Enhanced Resolution Analysis, Pöttschacher *et al.* 1996) auf ein regelmäßiges Gitter analysiert werden. VERA bietet die Möglichkeit, neben den Messwerten auch Zusatzinformationen in die Analyse einzubinden, sogenannte Fingerprints. Dabei handelt es sich um von Messungen unabhängiges Vorwissen. Die Fingerprints beschreiben, wie komplexe Topographie ihre Umgebung unter idealisierten Bedingungen thermisch (reduziertes Luftvolumen in Tälern, gehobene Heiz- oder Kühlflächen) und dynamisch (Stau am Gebirge) beeinflussen kann (Steinacker *et al.* 2006). Zuvor sollen diese Daten jedoch auf ihre Qualität geprüft und gegebenenfalls korrigiert oder von der Analyse ausgeschlossen werden. Diese Aufgaben übernimmt die VERA-QC, die vor allem zwei Punkte berücksichtigen muss:

- Die betrachteten meteorologischen Parameter, wie zum Beispiel Druck oder potentielle Temperatur, weisen bezüglich der Auflösung des Messnetzes eine hohe Autokorrelation auf. Die räumliche und zeitliche Konsistenz der Messwerte ist eine Eigenschaft, die auch das Analysefeld beibehalten soll. Es kann also verlangt werden, dass die Krümmung bzw. Rauigkeit des Analysefeldes minimal sein soll, eine Bedingung die mathematisch gut formuliert werden kann.

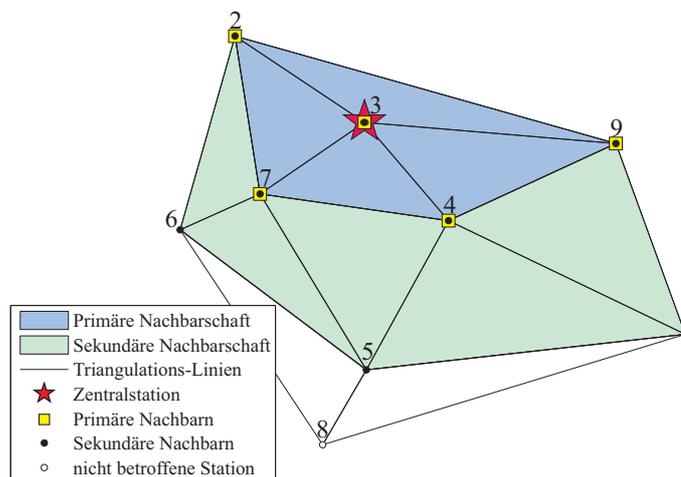
- Für die Überprüfung der räumlichen Konsistenz der Messwerte ist es notwendig, Nachbarstationen zu definieren. Dieser Schritt wird durch eine meist inhomogene Stationsverteilung erschwert. Betrachtet man die Alpenländer Österreich und Schweiz, so findet man eine sehr hohe Stationsdichte vor. In Italien oder Spanien wird hingegen der mittlere Stationsabstand bereits größer und in Küstengebieten liegt eine sehr einseitige Verteilung der Messstationen vor. Für die VERA-QC wurde aus diesem Grund die Methode der „natürlichen Nachbarfindung“ entwickelt, die sich auf die unterschiedlichen Stationsdichten anpassen kann.

In diesem Kapitel werden die mathematische Formulierung des Spatial Consistency Checks der VERA-QC und deren Vorteile vorgestellt.

### 2.1.1 Das Konzept der natürlichen Nachbarfindung

Bei der Konsistenzprüfung der Messwerte müssen für jede Station benachbarte Stationen gefunden werden, sodass deren Werte untereinander vergleichbar sind. Häufig erfolgt dies über die Definition eines fixen Einflussradius und oft auch unter Einbezug von statistischem Vorwissen wie Kovarianzen. Diese Methoden haben jedoch den Nachteil, dass sie sich nur sehr schlecht an eine inhomogene Stationsdichte anpassen können.

Um diesen Nachteil zu beheben wurde für die VERA-QC mit Hilfe der sogenannten



**Abbildung 2.1:** Veranschaulichung der „natürlichen Nachbarfindung“ für eine zufällige Verteilung von 9 Stationen. Station 3 ist in diesem Beispiel die Zentralstation. Die Stationen 2,3,4,7 und 9 sind primäre Stationen und die sekundäre Nachbarschaft wird von den Stationen 1,2,3,4,5,6,7 und 9 gebildet. Man beachte, dass die Zentralstation in der primären Nachbarschaft enthalten ist und die primäre Nachbarschaft wiederum als eine Teilmenge der sekundären Nachbarschaft definiert ist.

„Delaunay Triangulierung“ (Barber *et al.* 1996) die Methode der „natürlichen Nachbarfindung“ entwickelt. Bei der Triangulierung werden (in zwei Dimensionen) jeweils drei Stationen zu einem Dreieck derart verbunden, sodass sich im Umkreis, der das gebildete Dreieck umschließt, keine weitere Station befindet. Auch für höhere Dimensionen kann diese Triangulierung ausgeführt werden, indem man statt Dreiecke Tetraeder und statt Umkreise Umkugeln bzw. entsprechend höher dimensionale Körper bildet.

Abbildung 2.1 zeigt, wie neun zufällig verteilte Stationen durch die Triangulierung miteinander verbunden sind (schwarze Linien). Mit diesen Verbindungen können jetzt folgende Stationstypen definiert werden:

- Zentralstation ( $\mathcal{M}$ ): Diese Station bildet das Zentrum der betrachteten lokalen Nachbarschaft (Station Nr. 3 in Abbildung 2.1).
- Primäre Nachbarn ( $\mathcal{P}_p(\mathcal{M})$ ): Alle  $p = 1, 2, \dots, N(\mathcal{P})$  Stationen die direkt mit der Zentralstation verbunden sind, sowie die Zentralstation selbst, bilden die primäre Nachbarschaft (Stationen Nr. 2,3,4,7 und 9 in Abbildung 2.1).
- Sekundäre Nachbarn ( $\mathcal{S}_s(\mathcal{P}_p)$ ): Die sekundäre Nachbarschaft wird von allen  $s = 1, 2, \dots, N(\mathcal{S}(\mathcal{P}))$  natürlichen Nachbarn der einzelnen Primärstationen  $\mathcal{P}_p$  und von der primären Nachbarschaft inklusive Zentralstation gebildet (Stationen Nr. 1,2,3,4,5,6,7 und 9 in Abbildung 2.1). Man beachte, dass die primäre Nachbarschaft (inklusive Zentralstation) eine Teilmenge der sekundären Nachbarschaft ist. Diese Definition ermöglicht eine übersichtliche Darstellung des verwendeten mathematischen Formalismus, der in den folgenden Unterabschnitten präsentiert wird.

Der Grund für die Einteilung der Stationen in zwei Nachbarschaften wird im folgenden Abschnitt bei der Aufstellung der Kostenfunktion ersichtlich (vgl. Abschnitt 2.1.2, Gleichung (2.5), Index  $E$ ).

### 2.1.2 Aufstellung der Kostenfunktion

In der Meteorologie steht man häufig vor dem Problem, dass man räumlich oder auch zeitlich inhomogen verteilte Messwerte untereinander in Beziehung setzen möchte. Vor allem für die numerische Bearbeitung der Daten ist es oft erforderlich, die Beobachtungen auf ein regelmäßiges Gitter zu interpolieren. Dazu sucht man eine Funktion, die das Beobachtungsfeld bestmöglich beschreibt und zusätzlich gestellte Bedingungen erfüllt.

Es ist naheliegend zu fordern, dass die Abweichung der Funktion  $f(x_i)$  zu den

Messwerten  $y_i$  an den  $n$  Beobachtungspunkten (mit  $i = 1, \dots, n$ ) nur gering sein darf (vgl. Gleichung (2.1), erster Term - „Kostenfunktion“). Weisen zudem die Messwerte des betrachteten meteorologischen Parameters eine hohe Autokorrelation auf, so kann man fordern, dass der Verlauf der Funktion  $f(x_i)$  glatt, bzw. ihre Krümmung minimal sein soll (vgl. Gleichung (2.1), zweiter Term - „Penalty-Funktion“). Diese zwei Bedingungen, die auch unterschiedlich stark gewichtet werden können, resultieren in einem Minimierungsproblem, dessen Lösungen von sogenannten „Thin Plate Spline Funktionen“ getragen werden. Für zwei Dimensionen lautet die mathematische Formulierung dieser Bedingungen:

$$\frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2 + \mathcal{K} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (f_{x_1x_1}^2 + 2f_{x_1x_2}^2 + f_{x_2x_2}^2) dx_1 dx_2 \rightarrow Min \quad (2.1)$$

Die Indizes  $x_1x_1$ ,  $x_1x_2$  und  $x_2x_2$  geben an, wonach die Funktion  $f$  abgeleitet wird und  $\mathcal{K}$  bezeichnet einen Gewichtungsfaktor. Wahba (1990) beschreibt die präsentierten Zusammenhänge in angeführter Notation und gibt einen guten Überblick, wie man mit Hilfe von Splines Beobachtungsdaten verknüpft.

Auch die VERA-QC macht von dieser Methodik Gebrauch. Als erste Bedingung wird festgelegt, dass die gesuchte Funktion an den Beobachtungspunkten durch die Messwerte hindurchgehen, bzw. die Differenz  $y_i - f(x_i) = 0$  betragen soll. Somit verschwindet der erste Term in Gleichung (2.1) und es bleibt die Bedingung der minimalen Krümmung des Analysefeldes, die in einer sogenannten Kostenfunktion  $J$  formuliert wird.

Im Folgenden sei ein Beobachtungsfeld mit  $\Psi_o$  und ein fehlerfreies Analysefeld mit  $\Psi_a$  definiert. Die Krümmung des Analysefeldes wird mit  $\mathcal{C}_{\Psi_a}$  bezeichnet.

Die diskretisierte Form der Kostenfunktion des fehlerfreien Analysefeldes  $J(\Psi_a)$  wird für die VERA-QC wie folgt aufgestellt:

$$J(\Psi_a) = \sum_n^{N(G)} (\mathcal{C}_{\Psi_a})_n^2 \quad (2.2)$$

Man beachte, dass die Krümmungen aufgrund von beschränkter Computerleistung nicht wie üblich an einem regelmäßigen Gitter, sondern an den  $n = 1, \dots, N(G)$  Stationspositionen abgegriffen werden.

Die Krümmung  $\mathcal{C}_{\Psi_a}$  kann für zwei Dimensionen  $(x, y)$  folgend dargestellt werden:

$$(\mathcal{C}_{\Psi_a})_n^2 = \left[ \left( \frac{\partial^2 \Psi_a}{\partial x^2} \right)^2 + 2 \left( \frac{\partial^2 \Psi_a}{\partial x \partial y} \right)^2 + \left( \frac{\partial^2 \Psi_a}{\partial y^2} \right)^2 \right]_n \quad (2.3)$$

Für  $D$  Dimensionen wird Gleichung (2.3) zu:

$$(\mathcal{C}_{\Psi_a})_n^2 = \sum_{d_1=1}^D \sum_{d_2=1}^D \left( \frac{\partial^2 \Psi_a}{\partial d_1 \partial d_2} \right)_n^2 \quad (2.4)$$

$d_1$  und  $d_2$  bezeichnen räumliche Koordinaten oder die Zeit (z.B.: für vier Dimensionen gilt:  $d_1 = x, y, z, t$  und  $d_2 = x, y, z, t$ ).

Bevor Gleichung (2.2) minimiert werden kann, muss die unbekannte Krümmung des Analysefeldes  $\Psi_a$  durch jene des Beobachtungsfeldes  $\Psi_o$  angenähert werden. Dafür wird eine von der Krümmung des Beobachtungsfeldes ausgehende Taylorreihenentwicklung erster Ordnung durchgeführt:

$$\sum_{d_1=1}^D \sum_{d_2=1}^D \left( \frac{\partial^2 \Psi_a}{\partial d_1 \partial d_2} \right)_n^2 = \sum_{d_1=1}^D \sum_{d_2=1}^D \left[ \frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} + \sum_E \frac{\partial}{\partial \Psi_E} \left( \frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} \right) (\Psi_a - \Psi_o)_E \right]_n^2 \quad (2.5)$$

Die Abweichung des Analysefeldes vom Beobachtungsfeld ( $\Psi_a - \Psi_o$ ) in Gleichung (2.5) wird als Deviation ( $\Delta\Psi$ ) bezeichnet und stellt die einzige unbekannte Größe dar.

An dieser Stelle muss noch dem Index  $E$  besondere Aufmerksamkeit geschenkt werden. Er bezeichnet nicht nur die betrachtete Station  $n$ , sondern auch ihre Nachbarstationen. Ihnen wird die Möglichkeit zugesprochen, fehlerhaft zu sein und somit werden auch sie zur Variation freigegeben. In diesem Punkt unterscheidet sich die VERA-QC von den meisten herkömmlichen Qualitätskontrollmethoden, die einen Fehler nur bei der jeweils betrachteten Station zulassen.

Setzt man Gleichung (2.5) in Gleichung (2.2) ein und verwendet die Definitionen der primären und sekundären Nachbarschaften aus Abschnitt 2.1.1, so schreibt sich die Kostenfunktion als:

$$J = \sum_m \sum_s \sum_{d_1} \sum_{d_2} \left[ \frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} + \sum_p \frac{\partial}{\partial \Psi_p} \left( \frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} \right) \Delta\Psi_p \right]_s^2 \quad (2.6)$$

Die erste Summe geht über alle  $m = 1, \dots, N(\mathcal{M})$  Stationen in der gesamten Domäne, wobei eine Station nach der anderen als Zentralstation  $\mathcal{M}$  betrachtet wird. Zu jeder Zentralstation gehört auch eine lokale, sekundäre Nachbarschaft, über die sich die zweite Summe erstreckt. Die Krümmungen des Beobachtungsfeldes werden somit an allen sekundären Nachbarstationen abgegriffen. Die mit den Deviations multiplizierten Ableitungen der Krümmungen nach den Stationswerten

werden jedoch nur an den primären Stationen ausgewertet. Anschaulich gesprochen werden die Informationen des Beobachtungsfeldes an allen Stationen der sekundären Nachbarschaft einer Zentralstation abgefragt. Zur Variation freigegeben bzw. fehlerbehaftet können jedoch nur Stationen der primären Nachbarschaft sein.

### 2.1.3 Minimierung der Kostenfunktion

Um die unbekannt Deviations ( $\Delta\Psi = \Psi_a - \Psi_o$ ) zu ermitteln, muss die Kostenfunktion  $J$  noch nach den Stationswerten  $\Psi_m$  abgeleitet werden:

$$\frac{\partial J}{\partial \Psi_m} = 2 \sum_m \sum_s \sum_{d_1} \sum_{d_2} \left\{ \frac{\partial}{\partial \Psi_m} \left( \frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} \right) \cdot \left[ \frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} + \sum_p \frac{\partial}{\partial \Psi_p} \left( \frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} \right) \Delta \Psi_p \right]_s \right\} = 0 \quad (2.7)$$

Man erhält ein Gleichungssystem mit  $N(\mathcal{M})$  Zeilen, das nach den Deviations aufgelöst werden kann. Die so ermittelten Korrekturvorschläge minimieren die Kostenfunktion bzw. die Krümmung des Analysefeldes. Der VERA-QC bleibt nur noch die Entscheidung zu treffen, ob die Messwerte damit korrigiert werden oder ob gar ein Gross Error vorliegt.

### 2.1.4 Diskretisierung

Um die Ableitungen der Kostenfunktion auch numerisch ausführen zu können, müssen die Krümmungen und deren Ableitungen aus Gleichung (2.7) diskretisiert werden. Hierfür wird die bewährte Diskretisierungsmethode der finiten Differenzen gewählt. Wie bereits der Name verrät, werden bei dieser Methode die Ableitungen durch endliche Differenzen ersetzt.

Betrachtet man beispielsweise eine beliebige Funktion  $f(x)$ , so sollen deren Funktionswerte an  $x_i$  Stützstellen bekannt sein. In der Umgebung der Stützstelle  $x_0$  kann der Funktionswert  $f(x)$  durch eine Taylorreihe approximiert werden:

$$f(x) = f(x_0) + \frac{df}{dx}(x_0) \underbrace{(x - x_0)}_{\Delta x} + \frac{1}{2!} \frac{d^2 f}{dx^2}(x_0) (x - x_0)^2 + \dots \quad (2.8)$$

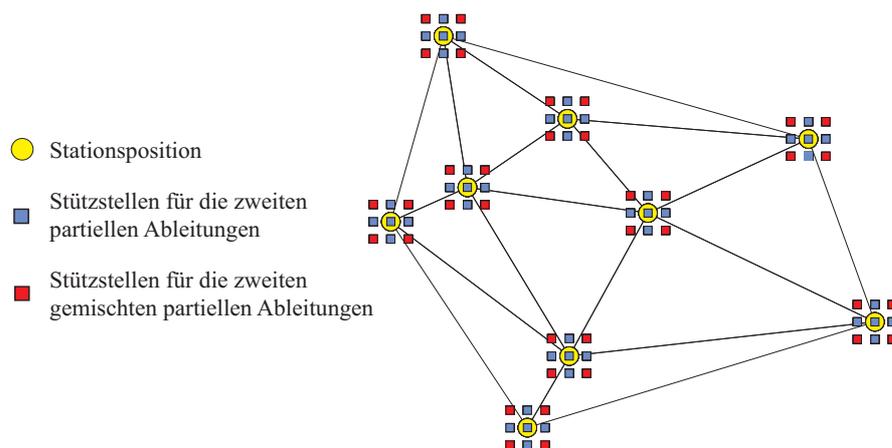
Eine Umformung nach  $\frac{df}{dx}(x_0)$  mit  $x = x_0 + \Delta x$  ( $\Delta x$  gibt die Entfernung zur Stützstelle an) ergibt:

$$\left. \frac{df}{dx} \right|_{x_0} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} - \underbrace{\frac{\Delta x}{2} \frac{d^2 f}{dx^2}}_{O(\Delta x)} - \dots \quad (2.9)$$

Die Terme auf der rechten Seite, die noch Ableitungen enthalten, werden zum sogenannten Diskretisierungsfehler  $O(\Delta x)$  zusammengefasst und müssen vernachlässigt werden. Gleichung (2.9) stellt eine Diskretisierung 1. Ordnung dar (der Fehler ist proportional zur ersten Potenz von  $\Delta x$ ) und es handelt sich um sogenannte Vorwärts-Differenzen (neben dem betrachteten Punkt  $x_0$  wird der in positiver x-Richtung gelegene Punkt  $x_0 + \Delta x$  betrachtet). Für eine zweifache Ableitung nach  $x$  ergibt sich folgende Diskretisierung 2. Ordnung:

$$\left. \frac{d^2 f}{dx^2} \right|_{x_0} = \frac{f(x_0 + \Delta x) - 2f(x_0) + f(x_0 - \Delta x)}{\Delta x^2} + O(\Delta x^2) \quad (2.10)$$

Normalerweise befinden sich die benötigten Stützstellen auf einem regelmäßigen Gitter, das sich über die gesamte betrachtete Domäne erstreckt. Um jedoch die Anzahl der Gitterpunkte und somit den Rechenaufwand gering zu halten, wird für die VERA-QC um jede Station ein kleineres lokales Gitter gelegt, das nur aus den benötigten Stützstellen besteht. Dadurch wird die Dichte der Stützstellen auch automatisch an die inhomogene Stationsverteilung angepasst. Der Abstand der Gitterpunkte zu den Stationen kann in der Konfigurationsdatei variiert werden. Abbildung 2.2 veranschaulicht diesen Schritt für die bereits in Abbildung 2.1 dargestellte, zweidimensionale Verteilung von 9 Stationen.



**Abbildung 2.2:** Darstellung der zufälligen Stationsverteilung (vgl. Abbildung 2.1) von 9 Stationen (gelbe Kreise). Zusätzlich ist um jede Station das lokale Gitter mit den Stützstellen, die zur Diskretisierung im zweidimensionalen Fall benötigt werden, eingezeichnet.

Die Interpolation der Stationswerte auf die umliegenden Gitterpunkte geschieht mit einer Inverse Distance Interpolation:

$$\Psi_n \simeq \sum_s \Psi_s \frac{1/(\beta + d_{n,s}^\alpha)}{\sum_s [1/(\beta + d_{n,s}^\alpha)]} \quad (2.11)$$

In Gleichung (2.11) bezeichnet  $\Psi_n$  die  $n$  Gitterpunktwerte, die jeweils für eine aktuell betrachtete sekundäre Nachbarschaft berechnet werden. Die Stationswerte der  $s$  Stationen in dieser lokalen Nachbarschaft werden mit  $\Psi_s$  angegeben. Der Bruch auf der rechten Seite stellt die normierte Gewichtung dar, mit der die einzelnen Stationswerte in die Interpolation eingehen. Diese Gewichtung ist von den, mit dem benutzerdefinierten Parameter  $\alpha$  potenzierten Abständen zwischen Gitterpunkten und Stationen ( $d_{n,s}$ ) und vom Parameter  $\beta$ , der den Grad der Glättung regelt, abhängig. Die Verwendung dieser simplen Interpolationsmethode ist gerechtfertigt, da in erster Linie die Änderungen der Krümmungen betrachtet werden und somit der Absolutwert an den Gitterpunkten von zweitrangiger Bedeutung ist.

Mit den auf die Gitterpunkte interpolierten Messwerten können nun die Krümmungen und deren Ableitungen aus Gleichung (2.7) mit dem Formalismus (2.9) und (2.10) diskretisiert werden. Betrachtet man eine beliebige Station  $s$  mit den Koordinaten  $[x_s, y_s]$ , so lässt sich eine zweifache Ableitung nach  $y$  wie folgt ausdrücken:

$$\left( \frac{\partial^2 \Psi_o}{\partial y^2} \right)_s \simeq \frac{\Psi(x_s, y_s + \Delta y) - 2\Psi(x_s, y_s) + \Psi(x_s, y_s - \Delta y)}{(\Delta y)^2} \quad (2.12)$$

$\Delta y$  bezeichnet dabei den Abstand zwischen den Gitterpunkten in  $y$ -Richtung. Die Diskretisierung der Ableitung der Krümmung nach dem Stationswert  $\Psi_p$  lautet unter der Verwendung von Vorwärts-Differenzen:

$$\frac{\partial}{\partial \Psi_p} \left( \frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} \right) \simeq \frac{\frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} (\dots, \Psi_{p-1}, \Psi_p + \Delta \Psi_p, \Psi_{p+1}, \dots)}{\Delta \Psi_p} - \frac{\frac{\partial^2 \Psi_o}{\partial d_1 \partial d_2} (\dots, \Psi_{p-1}, \Psi_p, \Psi_{p+1}, \dots)}{\Delta \Psi_p} \quad (2.13)$$

Man beachte, dass dem Messwert der betrachteten Station  $p$  im ersten Term der rechten Seite eine Erhöhung von  $\Delta \Psi_p$ , welche für die Diskretisierung mittels Vorwärts-Differenzen notwendig ist, hinzugefügt wird.

## 2.1.5 Kostenfunktionsreduktion

Nachdem die Gleichungen (2.7) diskretisiert und das so entstandene Gleichungssystem nach den Deviations  $\Delta \Psi$  aufgelöst wurde, muss die VERA-QC entscheiden,

wie sie mit den Stationsmeldungen und den zugehörigen Korrekturvorschlägen umgeht. Dabei wird jeweils nicht nur der Absolutwert der Deviation berücksichtigt, sondern auch welche Auswirkung eine Korrektur des betrachteten Messwertes auf die Kostenfunktion bzw. Krümmung des Feldes hätte.

Für die folgenden Überlegungen soll eine Stationsverteilung betrachtet werden, in der die Messung einer Station einen großen Fehler aufweist. Die Messungen der umliegenden Stationen sollen jedoch von keinen Fehlern gestört sein. Durch den Spline-Ansatz der VERA-QC kann es bei der Qualitätskontrolle zu einer Übertragung des Fehlers auf umliegende Stationen kommen. Die Krümmung des Beobachtungsfeldes wird zwar hauptsächlich durch die Reduktion des großen Fehlers verringert, zusätzlich kann dies aber auch durch das Anheben der umliegenden Stationswerte in Richtung der fehlerbehafteten Meldung geschehen. Somit werden auch für diese Stationen „falsche“ Korrekturen berechnet.

Um zu unterscheiden, welche Korrektur angebracht werden soll, wird das Ausmaß der Kostenfunktionsreduktion betrachtet. Dabei vergleicht man die Krümmung des ursprünglichen Beobachtungsfeldes mit der Krümmung, die sich nach Anbringen eines einzelnen Korrekturvorschlags ergibt. Würde man (in zwei Dimensionen) nur die Korrektur der Station mit dem großen Fehler anbringen, würde sich eine Kostenfunktionsreduktion von 100% ergeben. Das alleinige Anbringen der „falschen“ Korrektur einer umliegenden Station würde nur zu einer sehr schwachen Reduktion der Kostenfunktion führen. Gewichtet man die berechneten Deviations mit dieser Kostenfunktionsreduktion, so verschwindet der unerwünschte Effekt der Fehlerübertragung weitgehend. Die Kostenfunktionsreduktion  $\mathcal{W}$  lässt sich für eine betrachtete Station  $\mathcal{M}$  und der für diese Station berechneten Korrektur  $\Delta\Psi_{\mathcal{M}}$  folgendermaßen ausdrücken:

$$\mathcal{W}_m = \frac{\mathcal{J}(\Psi_{o,\mathcal{S}\neq\mathcal{M}}, \Psi_{o,\mathcal{S}=\mathcal{M}}) - \mathcal{J}(\Psi_{o,\mathcal{S}\neq\mathcal{M}}, \Psi_{o,\mathcal{S}=\mathcal{M}} + \Delta\Psi_{\mathcal{M}})}{\mathcal{J}(\Psi_{o,\mathcal{S}\neq\mathcal{M}}, \Psi_{o,\mathcal{S}=\mathcal{M}})} \quad (2.14)$$

Die gewichteten Korrekturen ergeben sich aus  $\Delta\Psi_w = \mathcal{W} \Delta\Psi$ . Durch den Index  $\mathcal{S}$  wird angegeben, dass für die Berechnung der Kostenfunktionsreduktion  $\mathcal{W}_m$  nur die Krümmungen innerhalb der sekundären Nachbarschaft der Station  $\mathcal{M}$  herangezogen werden. Eine Berücksichtigung der Krümmung des gesamten Feldes würde zu lange Rechenzeiten verursachen und die Betrachtung der unmittelbaren Nachbarschaft ist meist ausreichend. Dennoch muss man in Kauf nehmen, dass global berechnete Deviations angebracht werden, um lokal die Kostenfunktionsreduktion zu berechnen (vgl. Kapitel 7, Abschnitt 7.2.1).

### 2.1.6 Beurteilung der Stationsmeldungen (akzeptieren, korrigieren oder verwerfen)

Die Kostenfunktionsreduktion dient nicht nur der Gewichtung der Deviations, sondern ermöglicht auch die Entscheidung, ob eine Stationsmeldung akzeptiert, korrigiert oder verworfen werden soll. Die VERA-QC entscheidet zunächst im Rahmen der GE-Erkennung, ob ein Gross Error vorliegt:

- Ein Gross Error liegt vor, wenn die Kostenfunktionsreduktion einen benutzerdefinierten Schwellwert überschreitet. Die Station, welche die fehlerbehaftete Meldung abgegeben hat, wird als GE-Station bezeichnet. Um zu verhindern, dass betragsmäßig kleine Korrekturvorschläge zu Gross Errors führen, wird zusätzlich noch eine Bedingung gestellt: Die gewichtete Korrektur der vermeintlichen GE-Station muss ein benutzerdefiniertes Vielfaches des Medians aller Korrekturvorschläge überschreiten. Im Kapitel 6 wird ein zusätzliches Kriterium zur Gross Error Erkennung vorgestellt, das von stationspezifischen statistischen Informationen Gebrauch macht und in der operationellen VERA-QC verwendet wird.

Werden Stationsmeldungen als Gross Errors identifiziert, so werden sie nicht weiter berücksichtigt und die Qualitätskontrolle wird erneut für die restlichen Stationen durchgeführt. Es bleibt die Entscheidung zu treffen, ob eine Meldung mit den neu berechneten, gewichteten Deviations korrigiert werden soll, oder ohne Korrektur akzeptiert wird:

- Eine gewichtete Korrektur wird angebracht, sobald ihr Betrag einen benutzerdefinierten Schwellwert überschreitet.

### 2.1.7 Behandlung von Stationscluster

Einen Sonderfall stellen sehr eng benachbarte Stationen dar, bei denen es trotz Gewichtung mit der Kostenfunktionsreduktion zur Übertragung eines Fehlers auf eine/mehrere nahe gelegene Nachbarstation/en kommen kann. Aus diesem Grund wurde die sogenannte Cluster-Behandlung entwickelt:

Stationen werden als Teil eines Clusters betrachtet, wenn der Abstand zwischen jeweils zwei Stationen kleiner ist als ein benutzerdefinierter Prozentsatz des Medians der Abstände in der betrachteten Subdomäne (sekundäre Nachbarschaft). Diese Stationen werden zu einer fiktiven Cluster-Station zusammengefasst, deren Koordinaten als Mittelwert der ursprünglichen Koordinaten hervorgehen. Der fiktive Stationswert

setzt sich aus den einzelnen Stationswerten zusammen, die mit der Anzahl ihrer jeweiligen primären Nachbarn gewichtet werden.

### 2.1.8 Ablauf der VERA-QC mit Cluster-Behandlung und Gross Error Erkennung

Bevor die VERA-QC zum ersten Mal durchgeführt wird, erfolgt die Überprüfung, ob Stationen zu Stationsclustern zusammengefügt werden müssen.

- Die QC wird auf die modifizierte Stationsverteilung mit fiktiven Cluster-Stationen angewendet und berechnet für jede Station mit Hilfe der Kostenfunktionsreduktion gewichtete Deviations. Diese werden dann wieder auf die einzelnen Cluster-Stationen aufgeteilt und deren ursprünglichen Meldungen aufgeprägt. Mit den gewichteten Deviations wird auch die GE-Erkennung durchgeführt und Stationen, deren Meldungen als Gross Errors identifiziert wurden, werden für den betrachteten Parameter gelöscht.
- Erneut wird die VERA-QC ausgeführt, jedoch diesmal für die reduzierte ursprüngliche Stationsverteilung. Die Cluster-Stationen werden wieder individuell, aber mit den bereits korrigierten Messwerten betrachtet, Stationen mit GE wurden nicht mehr berücksichtigt.

Die Korrekturen der Cluster-Stationen setzen sich somit aus den gewichteten Deviations beider QC-Läufe zusammen, die restlichen Stationen erhalten die gewichteten Deviations des zweiten Laufs. Sie stellen die Basis für die Entscheidung dar, ob ein Messwert korrigiert oder unkorrigiert akzeptiert wird.



# Kapitel 3

## Die Bias-Korrektur

### 3.1 Begriffsdefinition

Im Zuge des „*Global Atmosphere Watch*“ (GAW) Programms der WMO wird in Hinblick auf Standardisierung ein Glossar (Klausen *et al.* 2010) geführt, das Begriffe der Qualitätskontrolle und -sicherstellung (QC/QA) im meteorologischen Sinne enthält. Unter anderem bedient es sich dabei der Definitionen aus dem „*Internationalen Wörterbuch der Metrologie*“ (Working Group 2 of the Joint Committee for Guides in Metrology 2008). Die Begriffe „systematischer Fehler“ und „Bias“ werden im Letzteren (Seite 22–23) wie folgt definiert:

*„systematic error: component of measurement error that in replicate measurements remains constant or varies in a predictable manner“*

*„bias: estimate of a systematic measurement error“*

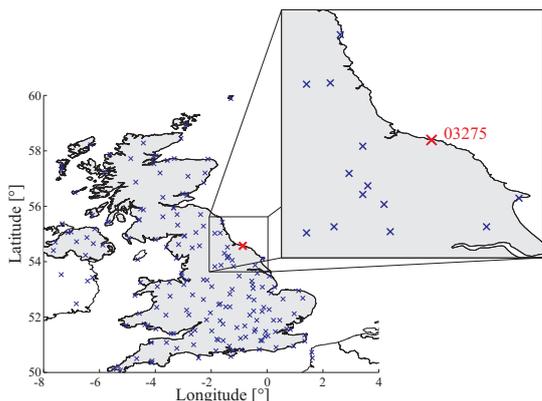
Sie werden in dieser Arbeit auch in diesem Sinne verwendet.

### 3.2 Das Problem systematischer Fehler und deren Lösung: die Bias-Korrektur

Jede Messung ist mit Fehlern behaftet, unter anderem auch mit systematischen Fehlern. Diese sind durch eine asymmetrische Verteilung charakterisiert und heben sich somit bei wiederholten Messungen im Mittel nicht auf. Sie führen dazu, dass Messungen systematisch zu hohe oder zu niedrige Werte liefern, was häufig auf schlecht kalibrierte Messgeräte zurückzuführen ist.

Eine Möglichkeit, derartige Fehler zu berücksichtigen und auch operationell auszubessern, ist die Zuhilfenahme der Fehlerstatistik. Jeder QC-Lauf liefert gewichtete und ungewichtete Korrekturvorschläge (Deviations), die sich zeitlich

betrachtet in Summe für systematische Fehler nicht aufheben. Sammelt man die ungewichteten Deviations über einen längeren Zeitraum, so kann man daraus den systematischen Fehler abschätzen und eine sogenannte Bias-Korrektur berechnen, die den Messwerten zu Beginn jedes QC-Laufs aufgeprägt wird. Das Prinzip der Bias-Korrektur wird anhand des Beispiels der Druckmeldungen einer britischen Station mit der WMO-Stationsnummer 03275 („Station 03275“) und mit den Abbildungen 3.1 und 3.2 im Folgenden genauer erklärt.



**Abbildung 3.1:** GTS-Stationen in Großbritannien (blaue Kreuze), die regelmäßig Meldungen für den Parameter Druck abgeben. Das rote Kreuz markiert die fehlerbehaftete Station (WMO-Nr: 03275).

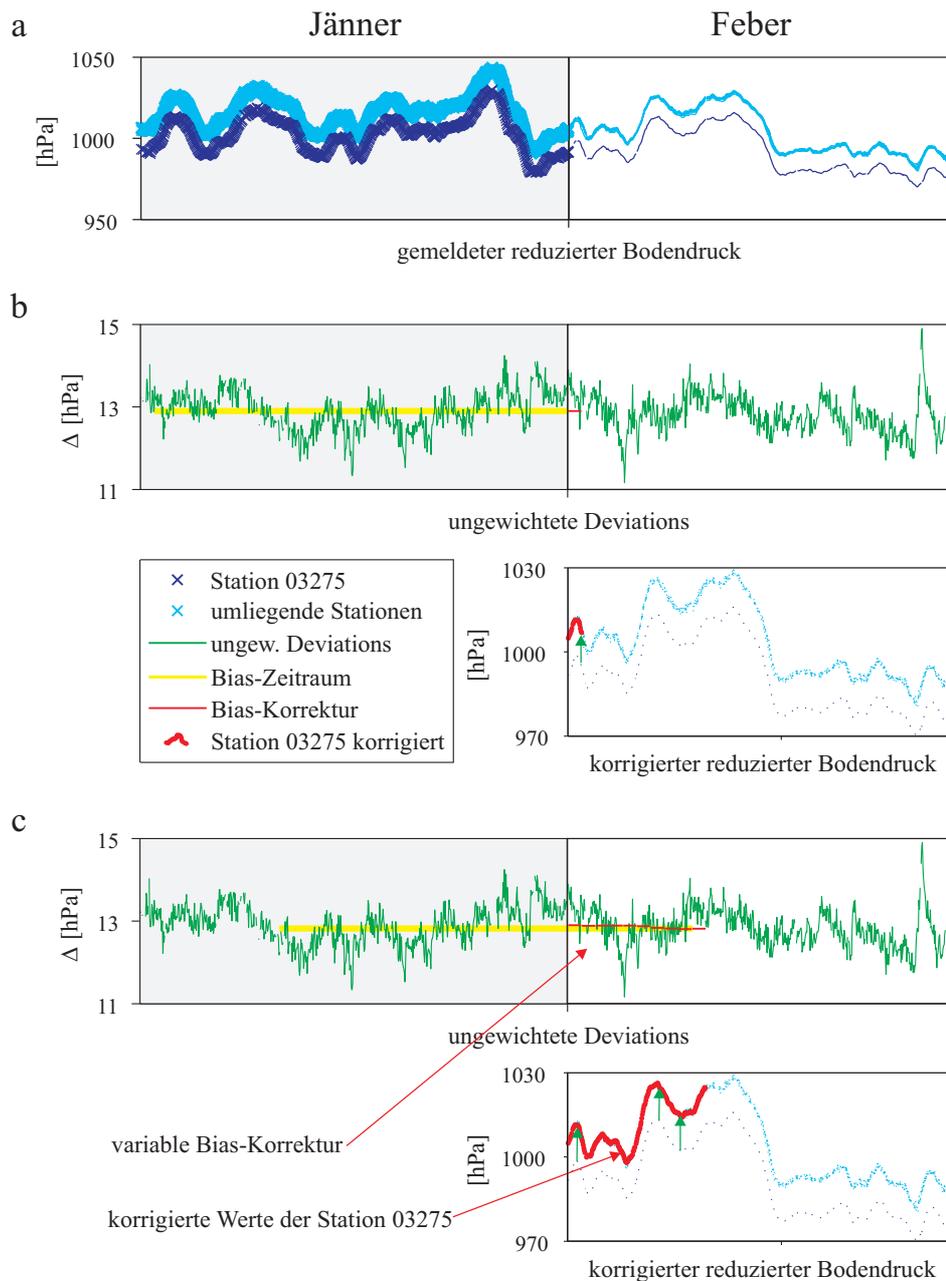
Stationen näher betrachtet und miteinander verglichen.

Dazu sind in Abbildung 3.2a die reduzierten Druckwerte der umliegenden Stationen in Hellblau und jene der fehlerbehafteten Station in Dunkelblau für die Monate Jänner und Feber 2010 dargestellt. Man erkennt, dass für die umliegenden Stationen durchwegs höhere Werte des reduzierten Bodendrucks berechnet werden als für die Station 03275. In Abbildung 3.2b sind die ungewichteten Deviations für die fehlerbehaftete Station in Grün dargestellt. Die Qualitätskontrolle korrigiert die Werte der Station 03275 zu jedem Termin um ca. 13 hPa nach oben und passt sie somit den Meldungen der umliegenden Stationen an. Eine Stationsmeldung mit derartig hoher ungewichteter Korrektur wird in den meisten Fällen als Gross Error identifiziert und entfernt, obwohl sie durchaus einen Informationsgehalt vorweisen kann.

Um dies zu verhindern, wird aus den gesammelten Deviations eine Bias-Korrektur berechnet, die den systematischen Fehler der betrachteten Station zu Beginn der QC korrigiert und somit die Stationsmeldung für eine weiterführende Analyse verfügbar macht. Die Bias-Korrektur wird definiert als:

*der Median der ungewichteten Deviations, die im Bias-Zeitraum von der QC berechnet und gesammelt wurden.*

In Abbildung 3.1 sind die Stationen Großbritanniens, von denen regelmäßig (bestenfalls stündlich) Meldungen für den Parameter Druck zur Verfügung stehen, durch blaue Kreuze symbolisiert. Die Station 03275 wurde rot hervorgehoben, da bei einem Vergleich mit den umliegenden Stationen erkannt wurde, dass ihre Druckmeldungen mit einem systematischen Fehler behaftet sind. Im vorliegenden Abschnitt werden die im Unterausschnitt hervorgehobenen



**Abbildung 3.2:** Anhand des Beispiels der Druckmeldungen jener Stationen, die in Abbildung 3.1 hervorgehoben sind, wird das Prinzip der Bias-Korrektur erklärt. In (a) ist der reduzierte, gemeldete Stationsdruck der umliegenden Stationen (höhere Werte, hellblau) und der fehlerbehafteten Station 03275 (tiefere Werte, dunkelblau) für Jänner und Feber 2010 dargestellt. In (b) und (c) sind die ungewichteten Deviations der Station 03275 jeweils in Grün gekennzeichnet, der Berechnungszeitraum der Bias-Korrektur ist als gelber Balken und die Bias-Korrektur für den 1. Feber 2010 als roter Balken dargestellt. Die roten Punkte kennzeichnen die Meldungen der Station 03275, die durch die Bias-Korrektur den umliegenden Stationsmeldungen angepasst wurden. Man beachte, dass sich der Berechnungszeitraum für die Bias-Korrektur (gelber Balken) von (b) auf (c) verschoben hat.

Der benutzerdefinierte (in der Konfigurationsdatei einstellbare) Bias-Zeitraum ist für den Parameter Druck mit 30 Tagen vordefiniert. Pro Station entspricht dies 720 möglichen stündlichen Meldungen bzw. Deviations. Natürlich muss für die Bias-Berechnung der jeweils betrachteten Station eine benutzerdefinierte Mindestanzahl an Deviations (empfohlen: 30) in diesem Zeitraum vorliegen.

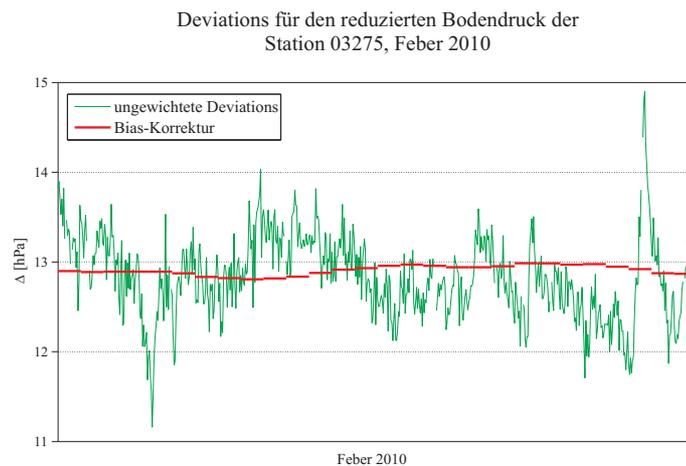
In Abbildung 3.2b ist die Berechnung der Bias-Korrektur für die Station 03275 für den 1. Feber 2010 dargestellt. Der gelbe Balken kennzeichnet den Bias-Berechnungszeitraum und ist auf Höhe des Medians der ungewichteten Deviations aus diesem Zeitfenster eingezeichnet. Die so errechnete Bias-Korrektur (roter Balken) ist für alle Termine am 1. Feber gültig und wird am Ende eines jeden Tages für den folgenden neu berechnet.

Wurde zu Beginn der QC eine Bias-Korrektur angebracht, so ergeben sich auch geringere Deviations. Um die Berechnung der Bias-Korrektur nach dem beschriebenen Kalkül fortführen zu können, wird bei einer QC mit Bias-Korrektur am Ende der QC die Bias-Korrektur wieder den berechneten Deviations aufaddiert, um die gesamte angebrachte Korrektur darzustellen.

Im unteren rechten Eck der Abbildung 3.2b sind die korrigierten Werte des reduzierten Bodendrucks der umliegenden Stationen (höhere Werte, hellblau) und die gemeldeten Werte der fehlerbehafteten Station (tiefere Werte, dunkelblaue Punkte) dargestellt. Die roten Punkte beschreiben die korrigierten Druckmeldungen der Station 03275, die durch eine Qualitätskontrolle mit Bias-Korrektur angehoben und somit den umliegenden Beobachtungen angepasst wurden.

Abbildung 3.2c stellt die Berechnung der Bias-Korrektur und deren positive Auswirkung auf die fehlerbehafteten Stationsmeldungen für den 1. bis 10. Feber 2010 dar. Man beachte, wie der Bias-Berechnungszeitraum (gelber Balken, hier für den 10. Feber) mitwandert.

Die Schwankungen der täglich neu berechneten Bias-Korrekturen sind in Abbildung 3.3 besser erkennbar. Es ist ersichtlich, dass im Falle einer Behebung des systematischen Fehlers (z.B.: Kalibrierung oder Austausch eines Messgerätes) die Bias-Korrektur den halben Bias-Zeitraum benötigen würde, um sich auf die nunmehr „richtigen“ Meldungen einzustellen. In diesem Zeitraum würde eine, durch die angebrachte „falsche“ Bias-Korrektur verfälschte Meldung als Gross Error erkannt oder bestenfalls wieder in die der Bias-Korrektur entgegengesetzte Richtung korrigiert werden.



**Abbildung 3.3:** Grün: ungewichtete Deviations des reduzierten Drucks für die Station 03275 (Feber 2010). Rot: Bias-Korrektur, berechnet als Median der ungewichteten Deviations der jeweils letzten 30 Tage.

#### **Bemerkung:**

Betrachtet man die Werte des reduzierten Bodendrucks in Abbildung 3.2a, so lautet die erste Vermutung, dass hinter diesem systematischen Fehler ein Reduktionsfehler steckt. Um sicher zu gehen, dass alle Druckmeldungen nach dem gleichen Formalismus auf Meeresebene reduziert werden, erfolgt vor der Qualitätskontrolle im Rahmen der Dekodierung eine Reduktion des gemeldeten, noch nicht reduzierten Stationsdrucks. Es wird somit der berechnete und nicht der gemeldete reduzierte Druck geprüft. Die Reduktion wird mit Hilfe einer aktuellen WMO-Stationsliste vorgenommen, in der, wie es sich herausstellte, die Höhenangabe bezüglich der Station 03275 fehlerhaft ist. Somit können die Bias-Korrekturen auch dazu dienen, die verwendete Stationsliste zu überprüfen und gegebenenfalls zu korrigieren.

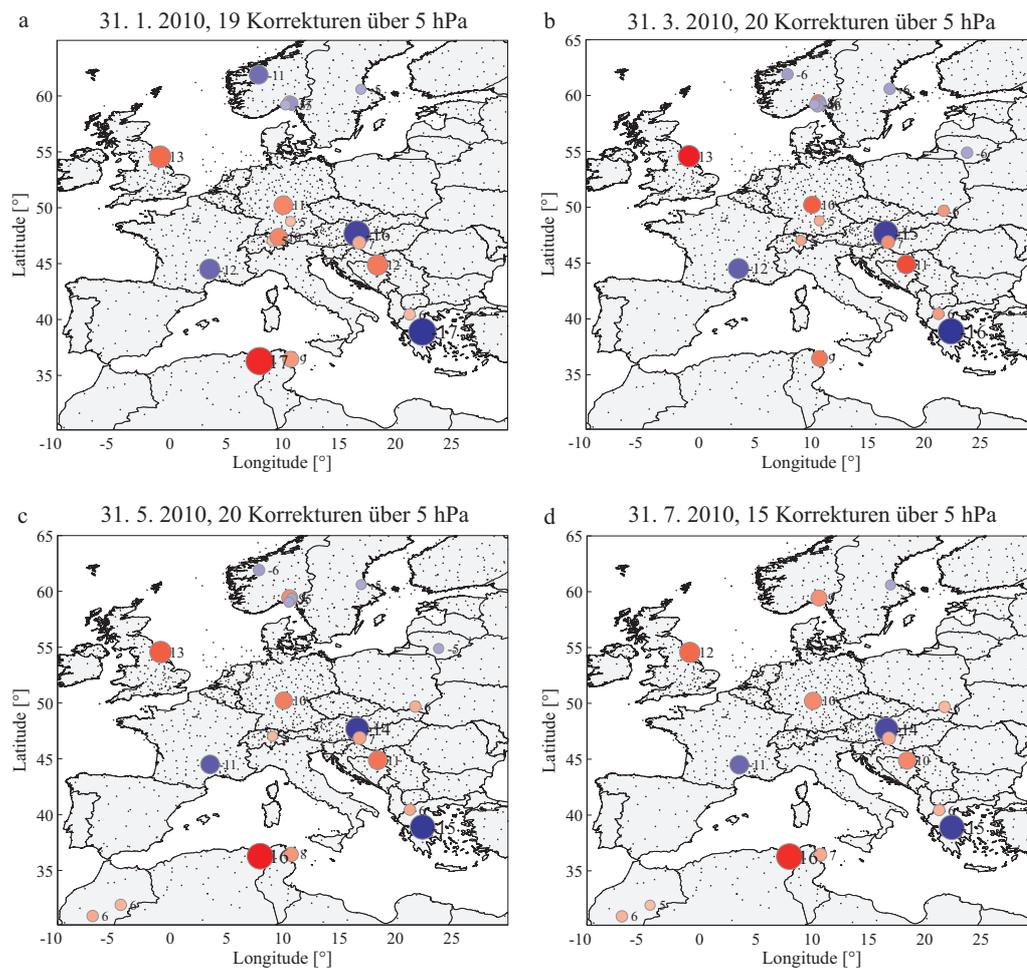
## **3.3 Auswertung der Bias-Korrekturen**

### **3.3.1 Reduzierter Bodendruck**

In Abbildung 3.4 sind für GTS-Stationen in Europa Bias-Korrekturen über 5 hPa für den reduzierten Bodendruck in den Monaten Jänner, März, Mai und Juli dargestellt. Positive Bias-Korrekturen sind rot, negative blau markiert. Betrachtet man die Bias-Korrekturen im zeitlichen Verlauf (linke obere bis rechte untere Abbildung), so erkennt man, dass sich zwar der Betrag der Korrekturen geringfügig ändern kann, das Vorzeichen jedoch, wie bei einem systematischen Fehler zu erwarten ist, erhalten bleibt. Weiters fällt auf, dass für einige Stationen in einem Monat Bias-Korrekturen eingezeichnet sind, im nächsten jedoch nicht mehr oder umgekehrt. Das beruht

darauf, dass die jeweilige Korrektur betragsmäßig die hier, für eine übersichtlichere Darstellung gewählte Schwelle von 5 hPa nicht überschreitet, oder, dass die, für die Berechnung der Bias-Korrektur geforderte Mindestanzahl an Meldungen bzw. Deviations (hier: 30) im Bias-Berechnungszeitraum (hier: 30 Tage bzw. maximal 720 stündliche Meldungen) nicht vorliegt.

### Bias-Korrekturen für den reduzierten Bodendruck [hPa]



**Abbildung 3.4:** Positive (rot) und negative (blau) Bias-Korrekturen des reduzierten Bodendrucks für GTS-Stationen in Europa, jeweils für den letzten Tag der Monate Jänner (a), März (b), Mai (c) und Juli (d) 2010. Alle verfügbaren Stationen sind zusätzlich als schwarze Punkte eingetragen.

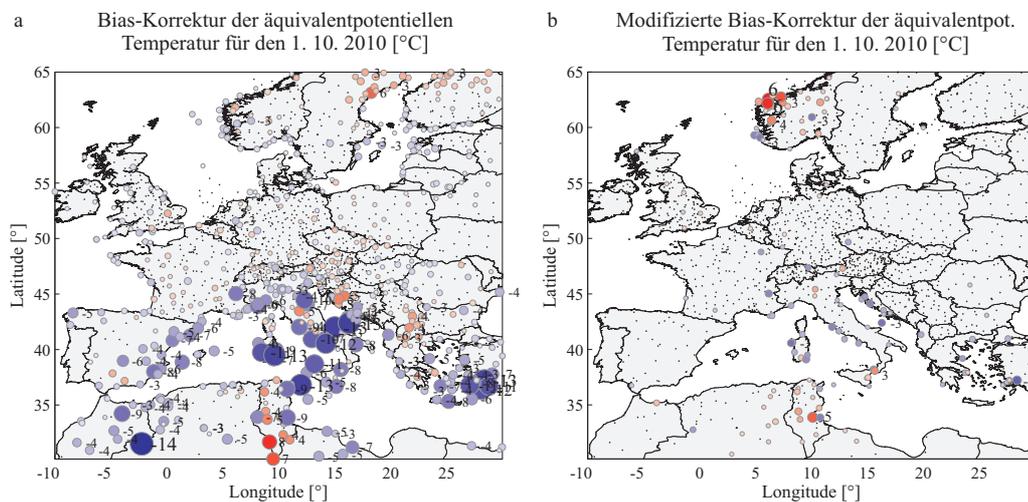
Abschließend sei erwähnt, dass die Bias-Korrektur für den reduzierten Bodendruck im operationellen Einsatz sehr zufriedenstellende Ergebnisse liefert und sich seit Beginn ihrer Anbringung mit 1. Jänner 2010 bewährt hat.

### 3.3.2 Potentielle Temperatur

Die Problematik der potentiellen Temperatur und deren Bias-Korrektur ist in einem eigenen Kapitel genauer beschrieben. An dieser Stelle sei nur der Verweis auf Kapitel 5 gegeben.

### 3.3.3 Äquivalentpotentielle Temperatur

Die gerade vorgestellte, simple Bias-Korrektur verlangt für die äquivalentpotentielle Temperatur, wie für die potentielle Temperatur, nach Modifikationen. In Abbildung 3.5a ist die Bias-Korrektur, berechnet als Median der ungewichteten Deviations der letzten 30 Tage, dargestellt.



**Abbildung 3.5:** Gegenüberstellung der Bias-Korrekturen für den Parameter äquivalentpotentielle Temperatur für den 1. 10. 2010. Dargestellt sind Korrekturen ab 1 °C, die Beschriftung erfolgt ab 3 °C. (a): Bias-Korrektur berechnet als Median der ungewichteten Deviations der letzten 30 Tage. (b): Bias-Korrektur berechnet als Median der ungewichteten, flächenmittelbereinigten Deviations der Monate September und März

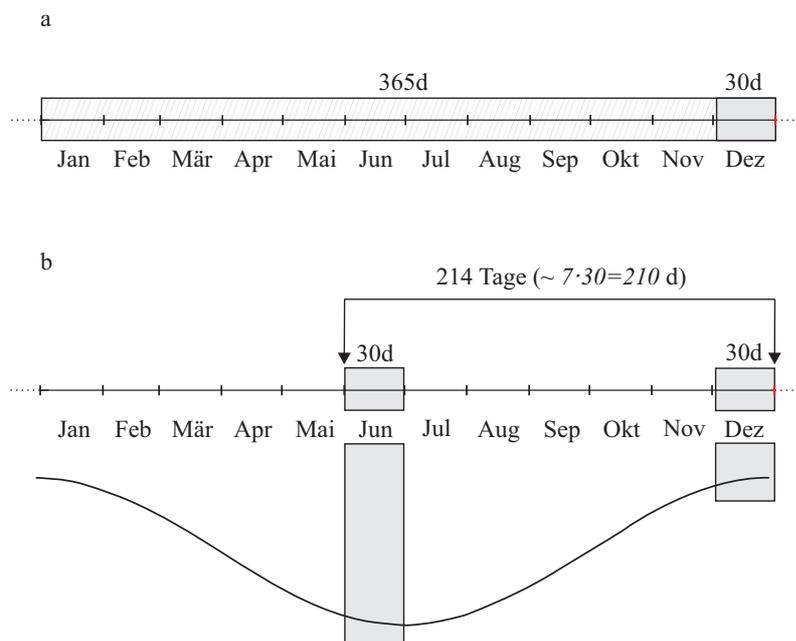
Es fällt auf, dass vor allem Küstenstationen starke negative Korrekturen erhalten, die auch einem Jahresgang unterliegen. In den Sommermonaten ist der Mittelmeerraum von höheren negativen Korrekturen bis zu -17 °C betroffen, welche sich in den Wintermonaten abschwächen und eher nach Norden, vor allem an die Küstengebiete Norwegens und der Ostsee, verlagern.

Um die regionale Abhängigkeit der Bias-Korrekturen zu beheben, wird, wie es sich auch für die potentielle Temperatur bewährt hat, das Flächenmittel der Korrekturen der jeweils primären Nachbarstationen von den Deviations abgezogen. Die Zentralstation selbst geht in dieser Mittelung nicht ein. Die jahreszeitliche

Abhängigkeit wird reduziert, indem man den Bias-Berechnungszeitraum auf 365 Tage erweitert. Durch diese Modifikationen kann die Korrektur ganzer Landstriche in eine Richtung verhindert und die Bias-Korrektur von einer jahreszeitlichen Abhängigkeit befreit werden.

Um Rechenzeit zu sparen wurde eine weitere Berechnungsmöglichkeit der Bias-Korrektur implementiert, die ebenfalls erlaubt, eine jahreszeitliche Abhängigkeit der Messwerte herauszufiltern. Die ursprüngliche Bias-Korrektur basiert auf der Definition eines benutzerdefinierten Bias-Berechnungszeitraums (z.B.: 30 Tage oder 365 Tage), der ununterbrochen in die Vergangenheit reicht (vgl. Abbildung 3.6a für den 31. Dezember). Alternativ dazu können auch nur einzelne Intervalle am Anfang und am Ende des Bias-Berechnungszeitraums herangezogen werden, wie in Abbildung 3.6b veranschaulicht wird. Betrachtet man den 31. Dezember und definiert den Bias-Berechnungszeitraum mit 214 Tagen, wird die Bias-Korrektur

Verschiedene Möglichkeiten, eine Bias-Korrektur für den 31. Dezember zu berechnen



**Abbildung 3.6:** Schematische Darstellung der Möglichkeiten, Deviations für die Bias-Berechnung auszuwählen. (a): Bei der simplen Berechnung der Bias-Korrektur werden die Deviations der letzten 30 Tage (grau hinterlegt) oder 365 Tage (schraffiert) herangezogen. (b): Zusätzlich gibt es die Möglichkeit, Deviations des letzten Monats und eines weiter zurückliegenden Monats (zum Beispiel ein halbes Jahr zurückliegend) zu verwenden. Unterhalb der Zeitachse ist ein cosinusförmiger Jahresgang der Messwerte simuliert. Die dargestellten Zeiträume sind variabel und können in der Konfigurationsdatei definiert werden.

aus den Deviations der Monate Dezember und Juni berechnet. Die Breite der beiden betrachteten Intervalle (hier: 30 Tage) ist selbstverständlich auch in der Konfigurationsdatei einstellbar. Unterhalb der Zeitachse ist ein idealisierter Jahresgang (zum Beispiel der Deviations der potentiellen Temperatur) dargestellt, der veranschaulichen soll, wie eine jahreszeitliche Abhängigkeit durch die Betrachtung von Monaten, die genau ein halbes Jahr auseinander liegen, reduziert wird. Die Mediane, berechnet aus den Werten im Dezember und Juni oder Feber und August, sollten kaum vom Median, der aus allen 12 Monaten berechnet wurde, abweichen.

Vergleicht man Bias-Korrekturen berechnet aus Deviations der letzten 365 Tage mit jenen, berechnet aus den Deviations des letzten und aus einem sechs Monate zurückliegenden Monat, kann man kaum einen Unterschied erkennen. Dennoch ist die zweite Variante vorzuziehen, da sie nur ein halbes Jahr gesammelter Korrekturvorschläge benötigt. Sind anfangs die sechs Monate noch nicht erreicht, wird die Bias-Korrektur nur mit Hilfe der Deviations des vorliegenden Zeitintervalls (zum Beispiel der letzten 30 Tage) durchgeführt.

In Abbildung 3.5 sind die ursprünglichen (a) und die modifizierten (b) Bias-Korrekturen für den 1. Oktober 2010 gegenüberstellt und die Verbesserung aufgrund der beschriebenen Erweiterungen ist deutlich erkennbar.

### 3.3.4 Windgeschwindigkeit

Bisher wurden die beiden Komponenten des horizontalen Windvektors ( $u$  und  $v$ ) getrennt behandelt und die beiden Felder mussten unabhängig voneinander der Glattheitsbedingung der VERA-QC genügen. Dies konnte zu einer Betrags- und Richtungsänderung der gemeldeten Windvektoren führen. Aus meteorologischer Sicht ist es jedoch sinnvoller, die Windrichtung und -geschwindigkeit getrennt voneinander zu kontrollieren.

Als erster Schritt in diese Richtung wurde im Zuge der vorliegenden Diplomarbeit die Möglichkeit implementiert, zusätzlich zu den einzelnen Windkomponenten auch die Windgeschwindigkeit zu prüfen.

Betrachtet man den horizontalen Windvektor  $\vec{v}$ , so kann dieser (insofern  $\vec{v} \neq \vec{0}$  gilt) folgendermaßen dargestellt werden:

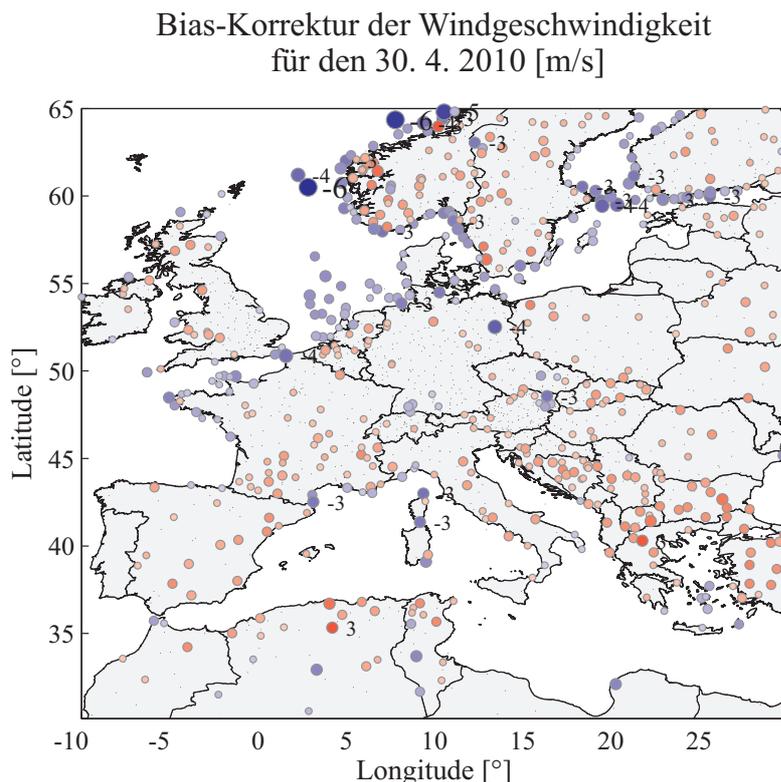
$$\vec{v} = |\vec{v}| \cdot \vec{v}_{norm} \quad \text{mit} \quad \vec{v}_{norm} = \frac{1}{|\vec{v}|} \cdot \vec{v} \quad (3.1)$$

wobei  $|\vec{v}|$  den Betrag von  $\vec{v}$  und  $\vec{v}_{norm}$  den Einheitsvektor von  $\vec{v}$  bezeichnet. Von der Qualitätskontrolle wird nun die Windgeschwindigkeit, also  $|\vec{v}|$  geprüft, die Wind-

richtung bleibt unverändert. Als Ergebnis erhält man die korrigierte Windgeschwindigkeit  $|\vec{v}|_{korr}$  und über

$$\vec{v}_{korr} = |\vec{v}|_{korr} \cdot \vec{v}_{norm}, \quad (3.2)$$

mit  $\vec{v}_{norm}$  aus Gleichung (3.1), kann auf den korrigierten Windvektor  $\vec{v}_{korr}$  und somit auch auf die korrigierten Komponenten des Windvektors  $u_{korr}$  und  $v_{korr}$  geschlossen werden. Die gemeldete Windrichtung bleibt bei dieser Methode unverändert.



**Abbildung 3.7:** Bias-Korrektur [m/s] der Windgeschwindigkeit für den 30. 4. 2010, berechnet als Median der ungewichteten Deviations der letzten 30 Tage. Hervorgehoben sind in Rot positive und in Blau negative Korrekturen über 1 m/s, die Beschriftung erfolgt ab 3 m/s. Man kann erkennen, dass vor allem im Norden der Domäne die Windgeschwindigkeiten der Küstenstationen nach unten und jene der Inlandstationen nach oben korrigiert werden. Ebenso erhalten im Vergleich zur Umgebung Messungen von höher gelegenen Stationen negative Korrekturen.

Auch für die Windgeschwindigkeit wird laufend eine Bias-Korrektur berechnet und angebracht. In Abbildung 3.7 ist diese für den 30. April 2010 dargestellt. Es stechen vor allem zwei Auffälligkeiten hervor:

- Für Stationen im Bereich der nördlichen Küstengebiete Europas werden hauptsächlich negative Korrekturen berechnet, Messwerte von Inlandstationen hingegen werden nach oben korrigiert.

- Stationen, die im Vergleich zur Umgebung höher gelegen sind, erhalten tendenziell negative Korrekturen. So beträgt beispielsweise die Bias-Korrektur im Sommer 2010 für die niederösterreichische Messstation Leiser Berge (WMO-Nr.: 11083, Höhe 457 Meter) ca. -4 m/s. Die umliegenden Stationen weisen eine durchschnittliche Stationshöhe von nur 200 Metern auf.

Die beiden Gegebenheiten, dass in Küstengebieten und an erhöhten, exponierten Stationen im Mittel höhere Windgeschwindigkeiten als im Landesinneren oder an Flachlandstationen vorherrschen, werden von der VERA-QC richtig erkannt. Durch die Bedingung der minimalen Krümmung des Analysefeldes erfolgen an diesen Stellen negative Korrekturen. Daraus resultiert eine Bias-Korrektur, die diesen systematischen Repräsentationsfehler behebt. Eine Bias-Korrektur sollte sich jedoch nur auf systematische Fehler einzelner Stationen auswirken und nicht flächendeckend die Windgeschwindigkeiten ganzer Landstriche in eine Richtung korrigieren. Es wurde versucht, diesen Nachteil, ähnlich wie bei der potentiellen und äquivalentpotentiellen Temperatur, durch das Abziehen des Flächenmittels der Deviations der jeweiligen primären Nachbarn auszubessern. Nach einigen Monaten kam es jedoch dazu, dass sich Bias-Korrekturen und Deviations gegenseitig immer stärker aufschaukelten und betragsmäßig immer größere negative Deviations für Stationen in Süd-Schweden berechnet wurden. Die Ursache liegt vermutlich darin, dass in Bereichen mit besonders vielen Küstenstationen, die systematisch hohe Windgeschwindigkeiten melden, nicht nur die Deviations (und damit auch Messwerte) dieser Stationen sondern auch jene der umliegenden Stationen durch die Korrektur mit dem Flächenmittel angehoben werden. Geschieht dies systematisch, müssen auch systematisch betragsmäßig immer höhere negative Korrekturen angebracht werden, um die Werte dem restlichen Feld anzupassen. Diese Methode musste somit verworfen werden und zur Zeit wird die Bias-Korrektur für die Windgeschwindigkeit ohne spezielle Modifikation berechnet, sie korrigiert also auch systematische Repräsentativitätsfehler in Küstengebieten.



# Kapitel 4

## Der Limit-Check

### 4.1 Einführung klimatologischer Limits

Die VERA-QC sieht bereits eine Überprüfung auf Einhaltung physikalischer Grenzwerte vor und wird im Rahmen dieser Diplomarbeit um eine Kontrolle auf Einhaltung klimatologischer Limits erweitert. Im Vordergrund steht die Implementierung des sogenannten Limit-Checks und die Bereitstellung eines Programms zur Berechnung klimatologischer Grenzwerte. Im Detail werden die Parameter Niederschlag, Windgeschwindigkeitskomponenten  $u$  und  $v$ , potentielle und äquivalentpotentielle Temperatur, sowie der reduzierte Bodendruck und der Stationsdruck betrachtet. Es wird darauf hingewiesen, dass die bereitgestellten Methoden und Programme zur Berechnung stationsspezifischer klimatologischer Limits, sofern möglich, nur auf Messzeitreihen angewendet werden sollten, die den von der WMO definierten Klima-Referenzzeitraum von 30 Jahren umfassen. In diesem Kapitel werden die beschriebenen Methoden jedoch auf die Daten von Jänner bis Juni 2010 angewendet, um ihre Wirkungsweise auch graphisch ohne großen Rechenaufwand veranschaulichen zu können.

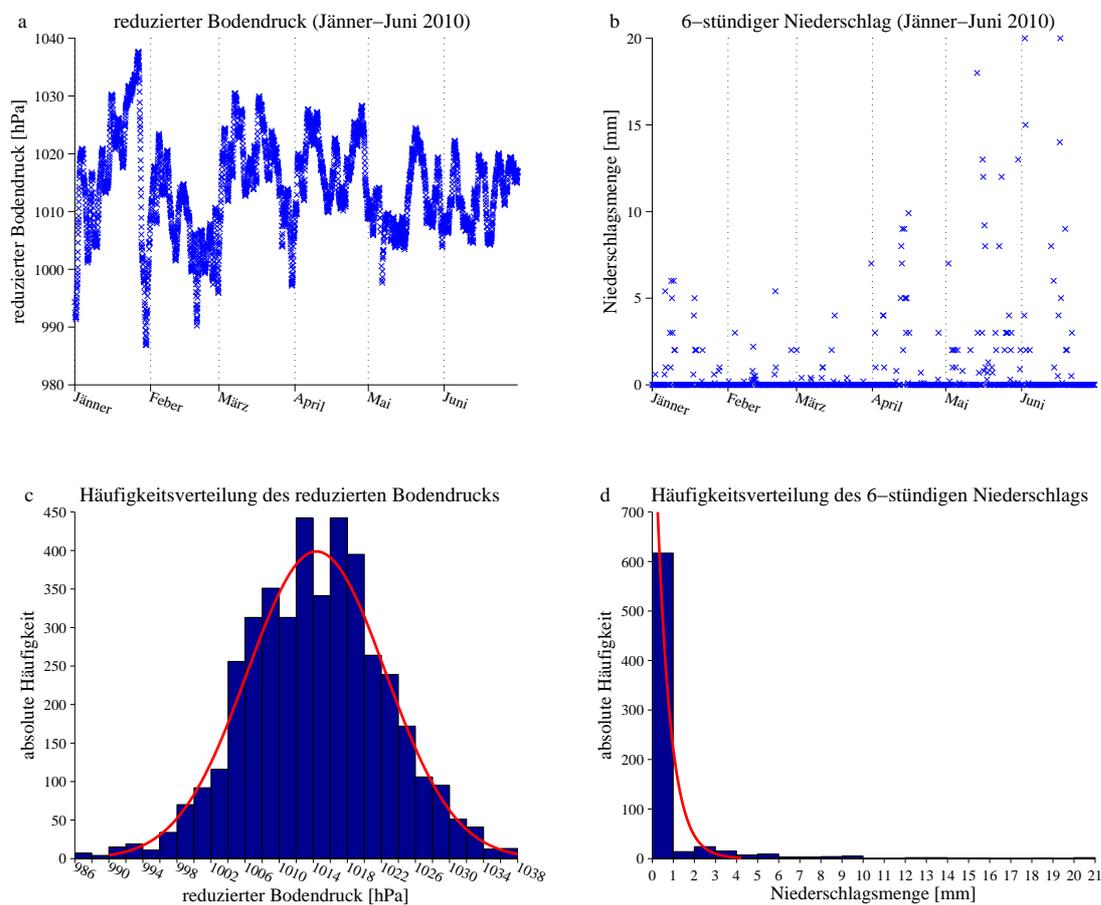
#### 4.1.1 Limit-Check

Zu den einfachsten Qualitätskontrollen gehört die Überprüfung auf Einhaltung physikalischer und klimatologischer Grenzwerte, der sogenannte Limit-Check. Jede Stationsmeldung wird zuerst auf ihre physikalische Plausibilität überprüft, indem sie benutzerdefinierten physikalisch möglichen Minima und Maxima gegenübergestellt wird. Befindet sich die Messung im physikalisch sinnvollen Bereich, erfolgt die Überprüfung auf Einhaltung stationsspezifischer klimatologischer Limits. Die Berechnung dieser Grenzwerte setzt das Vorhandensein eines Archivs langer Messzeitreihen voraus und ist im Abschnitt 4.1.2 genauer beschrieben. Bestehen Messwerte eine dieser

beiden Kontrollen nicht, werden sie eliminiert und durch einen entsprechenden Flag gekennzeichnet. Meldungen von Stationen bzw. Parametern, deren Aufzeichnungen für eine klimatologische Auswertung nicht ausreichen, können nur auf ihre physikalische Plausibilität überprüft werden.

### 4.1.2 Klimatologische Grenzwerte

Liegt für eine Station und einen betrachteten Parameter eine hinreichend lange Messzeitreihe vor, können daraus stationsspezifische, klimatologisch mögliche Höchst- und Tiefstwerte berechnet werden. Bei diesem Vorgehen muss darauf geachtet werden, dass nicht alle Parameter die gleichen statistischen Eigenschaften aufweisen. Dies wird offensichtlich, wenn man beispielsweise eine Messreihe



**Abbildung 4.1:** Gegenüberstellung der Messzeitreihen und der Häufigkeitsverteilungen des reduzierten Bodendrucks (a, c) und der sechsstündigen Niederschlagssummen (b, d) für den Zeitraum von Jänner bis Juni 2010, gemessen an der Station Wien - Hohe Warte (WMO-Nr.: 11035). Die rote Linie in (c) stellt eine nicht normierte Wahrscheinlichkeitsdichtefunktion der Normalverteilung ( $\mu = 1014.38$  hPa,  $\sigma = 8.12$  hPa) dar und in (d) jene einer Exponentialverteilung ( $\mu = 0.64$  mm).

des reduzierten Bodendrucks jener von sechsstündigen Niederschlagssummen gegenüberstellt (Abbildung 4.1).

In Abbildung 4.1a ist die Messzeitreihe des reduzierten Bodendrucks der Station Wien - Hohe Warte von Jänner bis Juni 2010 dargestellt. Bildet man mit diesen Daten ein Histogramm (Abbildung 4.1c), so ist zu erkennen, dass diese annähernd einer Normalverteilung folgen. Betrachtet man dazu im Vergleich die Messzeitreihe der sechsstündigen Niederschlagssummen (Abbildung 4.1b), lässt sich sofort ein signifikanter Unterschied erkennen: Die Messwerte häufen sich bei sehr geringen Werten und der Messwert Null tritt am häufigsten auf – es handelt sich um eine positive, rechtsschiefe Verteilung. Diese Eigenschaften gehen auch aus der Häufigkeitsverteilung der Niederschlagssummen (Abbildung 4.1d) hervor, an die eine Exponentialverteilung angepasst wurde (rote Linie).

Im Folgenden werden die Messreihen der Windkomponenten, der potentiellen und der äquivalentpotentiellen Temperatur und jene des reduzierten Bodendrucks als annähernd normalverteilt angenommen.

Um die klimatologischen Grenzwerte für alle Parameter nach einem einheitlichen Kalkül berechnen zu können, müssen die nicht normalverteilten Niederschlagsdaten zuerst mit einer Transformation einer Normalverteilung angepasst werden. Dafür wird eine Normal Score Transformation gewählt, wie sie Goovaerts (1997) beschreibt.

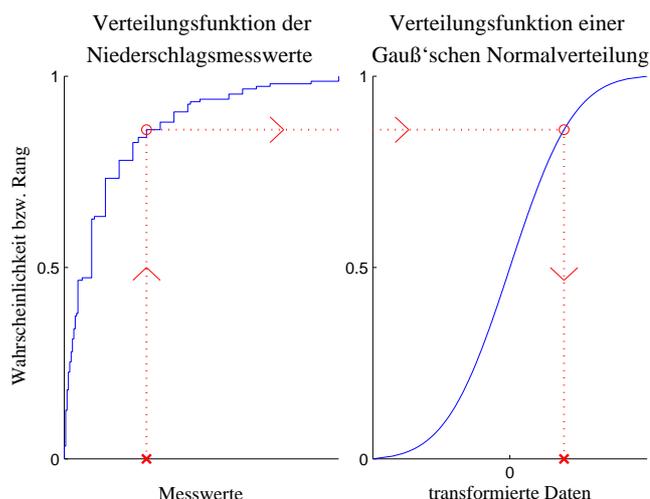
Die Messdaten werden zuerst aufsteigend gereiht und bekommen gemäß ihrer Position einen Rang zugewiesen. Diesem wird der jeweils äquivalente Rang einer Standardnormalverteilung gegenübergestellt und die Verteilungswerte dieser äquivalenten Ränge bilden den transformierten Datensatz.

Wie in Abbildung 4.2 ersichtlich, erfolgt die Vergabe der Ränge durch die Berechnung der empirischen kumulativen Verteilungsfunktion der Messwerte, welcher die kumulative Verteilungsfunktion einer Normalverteilung gegenübergestellt wird. Formal lässt sich diese Transformation wie folgt ausdrücken:

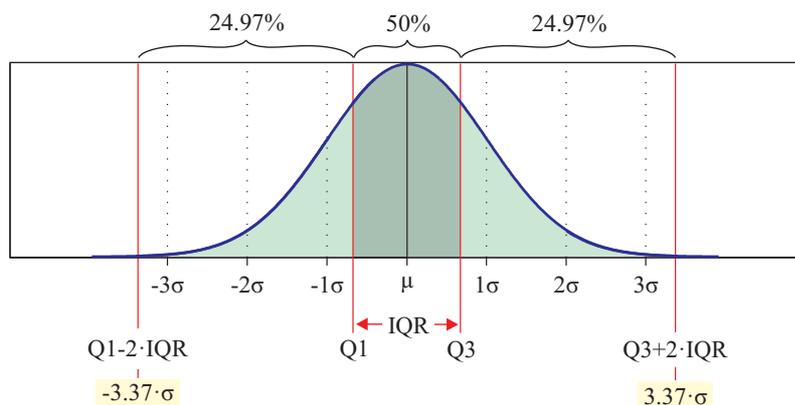
$$\phi(z) = G^{-1}[F_z] \quad (4.1)$$

wobei  $\phi(z)$  die transformierten Daten,  $F_z$  die kumulative Verteilungsfunktion der Messwerte  $z$  und  $G^{-1}$  die inverse Gauß'sche Verteilungsfunktion bezeichnet.

Können alle Messreihen als annähernd normalverteilt betrachtet werden, bzw. wurden sie durch eine Transformation einer Normalverteilung angepasst, so erfolgt die



**Abbildung 4.2:** Schematische Darstellung der Normal Score Transformation, wobei die empirische Verteilungsfunktion der Messwerte (blaue Linie, links) jener einer Gauß'schen Normalverteilung (blaue Linie, rechts) gegenübergestellt wird. Die punktierte rote Linie mit den Pfeilen markiert die einzelnen Schritte der Transformation: von den Messdaten (rotes Kreuz, links) bis zu den transformierten Daten (rotes Kreuz, rechts).



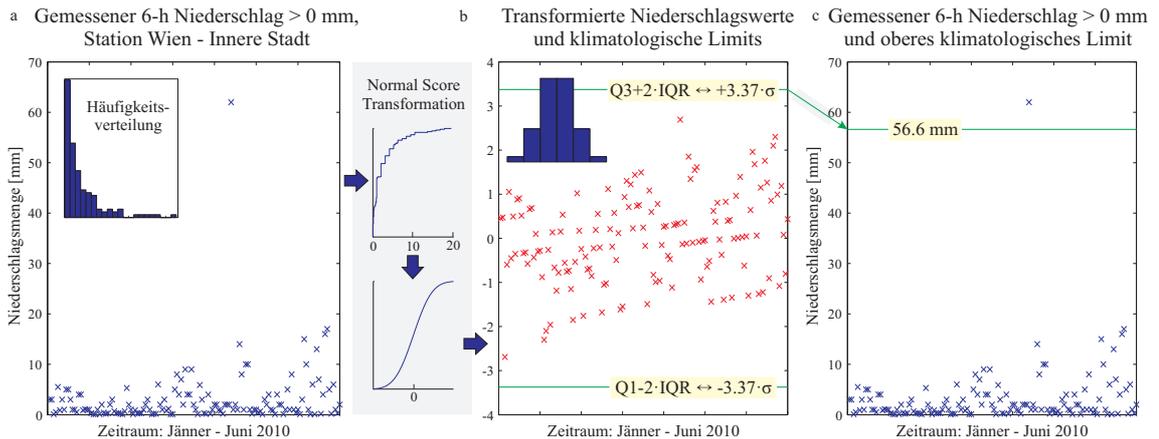
**Abbildung 4.3:** Die in Gleichung 4.2 definierten klimatologischen Grenzwerte sowie das untere ( $Q_1$ ) und obere ( $Q_3$ ) Quartil sind für die in Blau dargestellte Normalverteilung durch rote vertikale Linien gekennzeichnet. Zusätzlich sind noch der Mittelwert ( $\mu$ ) und die beidseitig vom Mittelwert ausgehenden Intervalle bis zur ersten, zweiten und dritten Standardabweichung ( $\sigma$ ) durch vertikale schwarze Linien markiert. Der Interquartilsabstand ( $IQR = Q_3 - Q_1$ ) beinhaltet bezüglich einer Standardnormalverteilung 50 % aller Messwerte, das Intervall zwischen den beiden klimatologischen Grenzwerten 99,94 % der Messwerte.

Berechnung klimatologischer Limits. Die stationsspezifischen Unter- bzw. Obergrenzen ( $Lim_1$  bzw.  $Lim_2$ ) werden definiert als:

$$Lim_1 = Q_1 - 2 \cdot IQR \quad \text{bzw.} \quad Lim_2 = Q_3 + 2 \cdot IQR \quad (4.2)$$

wobei  $Q_1$  bzw.  $Q_3$  das untere bzw. obere Quartil und  $IQR = Q_3 - Q_1$  den

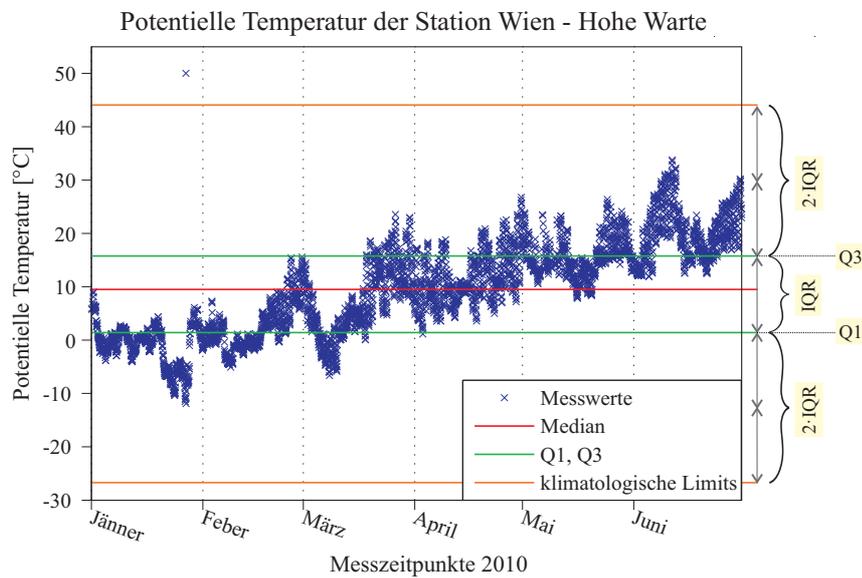
Interquartilsabstand bezeichnen. Wurden die Grenzwerte aus transformierten Daten berechnet, müssen diese Limits anschließend mit einer geeigneten Rücktransformation wieder den ursprünglichen Daten angepasst werden. In Abbildung 4.3 sind diese Grenzwerte für eine Standardnormalverteilung dargestellt und können für normalverteilte Daten mit der Standardabweichung  $\sigma$  in die Limits  $\pm 3.37 \cdot \sigma$  umgerechnet werden.



**Abbildung 4.4:** Schrittweise Berechnung der klimatologischen Limits für die 6-stündigen Niederschlagssummen, gemessen im Zeitraum Jänner - Juni 2010 an der Station Wien - Innere Stadt. (a) zeigt die Messwerte (blaue Kreuze) und die daraus berechnete Häufigkeitsverteilung (linkes oberes Eck). Diese Daten werden mit Hilfe einer Normal Score Transformation (grau hinterlegter Bereich) einer Normalverteilung angepasst. Die transformierten Daten sind in (b) durch rote Kreuze gekennzeichnet, mit zugehöriger Häufigkeitsverteilung im linken oberen Eck. Die berechneten Grenzwerte sind durch grüne horizontale Linien dargestellt, der obere Schwellwert wird mit einer inversen Normal Score Transformation wieder den Messwerten angepasst (c). In diesem Beispiel würde der Messwert von 62 mm in 6 Stunden verworfen werden.

Die Berechnung der klimatologischen Limits für die 6-stündigen Niederschlagsmesswerte (Jänner - Juni 2010) der Station Wien - Innere Stadt wird Schritt für Schritt in der Abbildung 4.4 veranschaulicht. In (a) ist der gemessene 6-h Niederschlag durch blaue Kreuze gekennzeichnet, im linken oberen Eck ist die daraus ermittelte Häufigkeitsverteilung dargestellt. Auf den ersten Blick ist zu erkennen, dass diese stark von einer Normalverteilung abweicht und somit müssen die Daten zuerst mit einer Normal Score Transformation einer Normalverteilung angepasst werden, bevor die in Gleichung 4.2 definierten klimatologischen Grenzwerte berechnet werden können. Die transformierten Daten sind in (b) durch rote Kreuze gekennzeichnet, im linken oberen Eck ist wieder die Häufigkeitsverteilung der transformierten Daten dargestellt. Mit den transformierten Daten kann die Berechnung der klimatologischen Limits (grüne horizontale Linien) erfolgen. Da

der untere Grenzwert für den Parameter Niederschlag naturgemäß Null beträgt, muss im nächsten Schritt nur das obere klimatologische Limit mit einer inversen Normal Score Transformation den Messdaten wieder angepasst werden (grüner Pfeil von (b) nach (c)). In (c) erkennt man, dass ein Messwert (62 mm in 6 Stunden, prinzipiell als korrekt einzustufen) über dem oberen klimatologischen Limit (grüne horizontale Linie) liegt und somit verworfen werden würde. Dies scheint aufgrund der vorliegenden Messwerte (blaue Kreuze in (a) und (c)) jedoch vernünftig. An dieser Stelle sei nochmals erwähnt, dass für die Berechnung der klimatologischen Limits längere Messzeitreihen herangezogen werden sollten.



**Abbildung 4.5:** Messzeitreihe der potentiellen Temperatur (blaue Kreuze), gemessen an der Station Wien - Hohe Warte (Jänner–Juni 2010). Der Median der Messwerte wird durch die rote Linie markiert, das untere Quartil (Q1) durch die untere und das obere Quartil (Q3) durch die obere grüne Linie und die beiden klimatologischen Grenzwerte werden durch die orangenen Linien gekennzeichnet.

In Abbildung 4.5 ist die Auswertung der klimatologischen Grenzwerte für eine Messzeitreihe der potentiellen Temperatur der Station Wien - Hohe Warte veranschaulicht. Eine Normal Score Transformation musste in diesem Fall nicht durchgeführt werden. Neben den Messwerten (blaue Kreuze) und den klimatologischen Limits (orange Linien) sind auch das obere und untere Quartil (grüne Linien), wie der Median der Messwerte (rote Linie) eingezeichnet. Die Messdaten wurden dahingehend manipuliert, dass Ende Jänner ein sogenannter Kommafehler eingebaut wurde, der einen Gross Error darstellen soll. Es wird der Fall simuliert, dass aufgrund eines Problems innerhalb des Prozesses der Datenübertragung anstatt der gemessenen

5 °C, 50 °C ausgelesen werden.

Wie in der Abbildung 4.5 erkennbar ist, kann ein derartiger Fehler bereits von diesem simplen Check auf Einhaltung klimatologisch möglicher Limits erkannt werden.

### 4.1.3 Auswertung der klimatologischen Grenzwerte

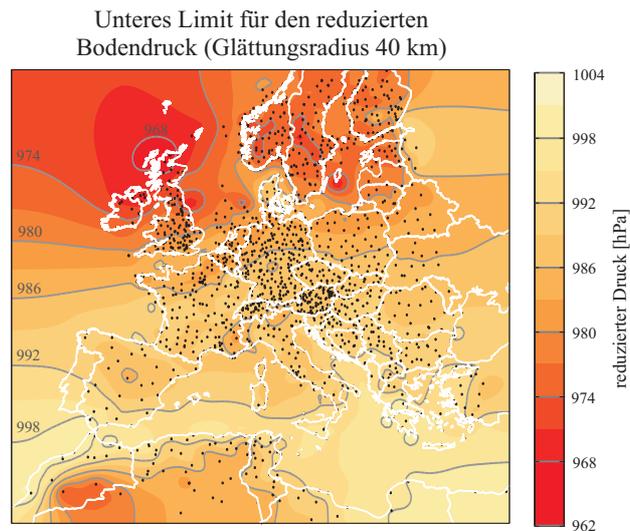
Die berechneten klimatologisch möglichen Limits wurden nicht nur für einzelne Stationen auf ihre Plausibilität überprüft, sondern auch flächig dargestellt. Dazu mussten die Grenzwerte, welche an den unregelmäßig verteilten Stationspositionen vorliegen, auf ein regelmäßiges Gitter interpoliert werden. Dieser Schritt erfolgte mittels VERA.

#### Reduzierter Bodendruck

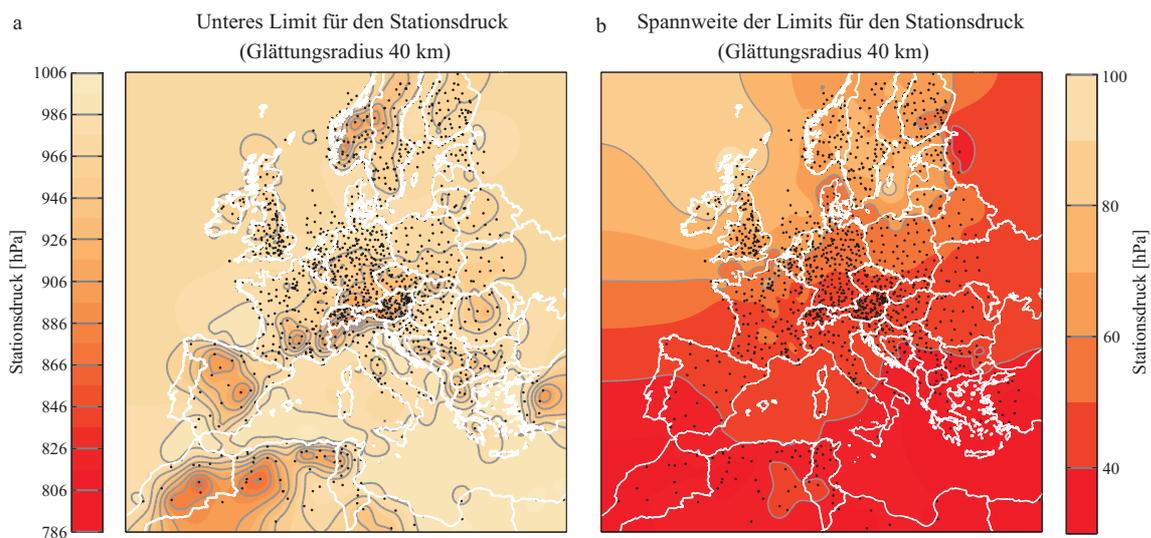
Bei der Berechnung klimatologischer Grenzwerte für den reduzierten Bodendruck wird empfohlen, keine Normal Score Transformation anzuwenden. Zum einen können die Messwerte des reduzierten Bodendrucks bereits als annähernd normalverteilt angenommen werden (siehe Abbildung 4.1c), zum anderen werden Quartile für die Berechnung der klimatologischen Limits herangezogen, welche nicht von Extremwerten beeinflusst werden.

In unserem Fall stellen Gross Errors solche Extremwerte dar, die bei der Berechnung der Grenzwerte keine Bedeutung haben sollten. Durch die Transformation in normalverteilte Daten werden die Gross Errors jedoch gleich gewichtet wie die restlichen Messdaten und führen so zu höheren bzw. tieferen klimatologisch möglichen Höchst- und Tiefstwerten. Der Limit Check würde dadurch unnötiger Weise zu „großzügig“ ausfallen. Möchte man dennoch alle Daten vor der Berechnung der klimatologischen Limits transformieren, so muss bei der Normal Score Transformation noch eine Änderung vorgenommen werden, die den Rändern der Verteilung eine geringere Gewichtung zuweist. Wird das präsentierte Kalkül zur Berechnung klimatologischer Limits auf längere Zeitreihen angewandt, so sollte auch überprüft werden, ob eventuell die Definition der Grenzwerte für einzelne Parameter auf  $Lim_1 = Q1 - 3 \cdot IQR$  bzw.  $Lim_2 = Q3 + 3 \cdot IQR$  erweitert werden soll.

Betrachtet man die Grenzwerte für den reduzierten Bodendruck, so erkennt man, dass in den nördlichen Regionen nicht nur, wie in Abbildung 4.6 ersichtlich, das untere klimatologische Limit am tiefsten liegt, sondern auch das obere am höchsten. Da die Variabilität des reduzierten Bodendrucks in den nördlichen Gebieten am höchsten ist, ergibt sich in diesem Bereich auch eine größere Spanne der berechneten Grenzwerte. Um zu überprüfen, ob die klimatologischen Limits auch wirklich die



**Abbildung 4.6:** Mittels VERA auf ein 40 km Gitter interpoliertes unteres klimatologisches Limit für den reduzierten Bodendruck. Die Grenzwerte wurden aus den Stationsmeldungen des Zeitraums Jänner–Juni 2010 berechnet. Eine Glättung des Feldes erfolgte mittels Einflussradius von 40 km.



**Abbildung 4.7:** Unteres klimatologisches Limit (a), bzw. Spannweite der Limits (b) für den Stationsdruck, die mit Hilfe von VERA auf ein regelmäßiges 40 km Gitter interpoliert und anschließend geglättet wurden. Die Berechnungen basieren auf den Daten des Zeitraums Jänner–Juni 2010.

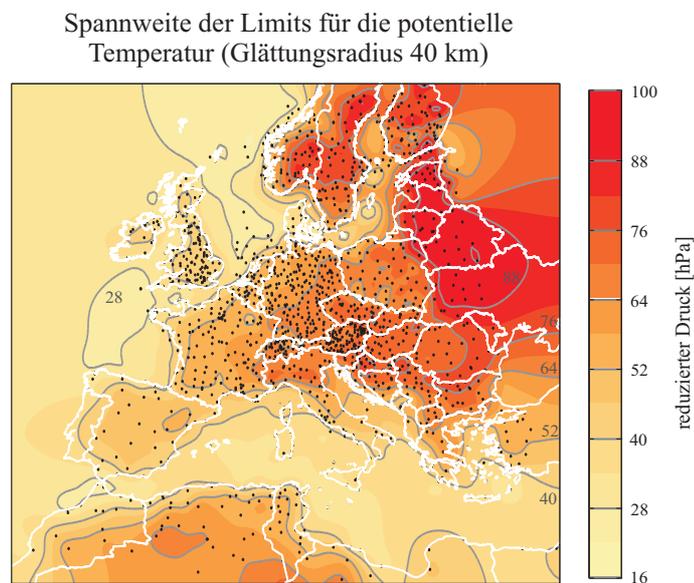
Eigenschaften des betrachteten Parameters widerspiegeln können, wurden auch für den Stationsdruck Grenzwerte berechnet (Abbildung 4.7).

Untersucht man das untere klimatologische Limit für den Stationsdruck (Abbildung 4.7a), so ist erkennbar, dass sich dessen Höhenabhängigkeit auch gut in den berechneten Limits widerspiegelt. Theoretisch sollte diese Eigenschaft in den klimatologischen Grenzwerten für den reduzierten Bodendruck nicht mehr

erkennbar sein. Trotzdem macht sich bei höher gelegenen Stationen ein geringer Reduktionseinfluss bemerkbar, der, wie zum Beispiel an zwei Stationen im Südosten Marokkos (Abbildung 4.6), durch die geringe Stationsdichte und die Randlage bei der Interpolation zusätzlich hervorgehoben wurde.

In Abbildung 4.7b wird die Spannweite der Limits für den Stationsdruck (Differenz aus oberem und unterem Limit) dargestellt. Wie bereits erwähnt, ergibt sich bei höherer Variabilität der Messwerte eine breitere IQR und somit tiefere und höhere Grenzwerte bzw. eine größere Spannweite der Limits, welche im Norden größer als im Süden ist.

### Potentielle und äquivalentpotentielle Temperatur



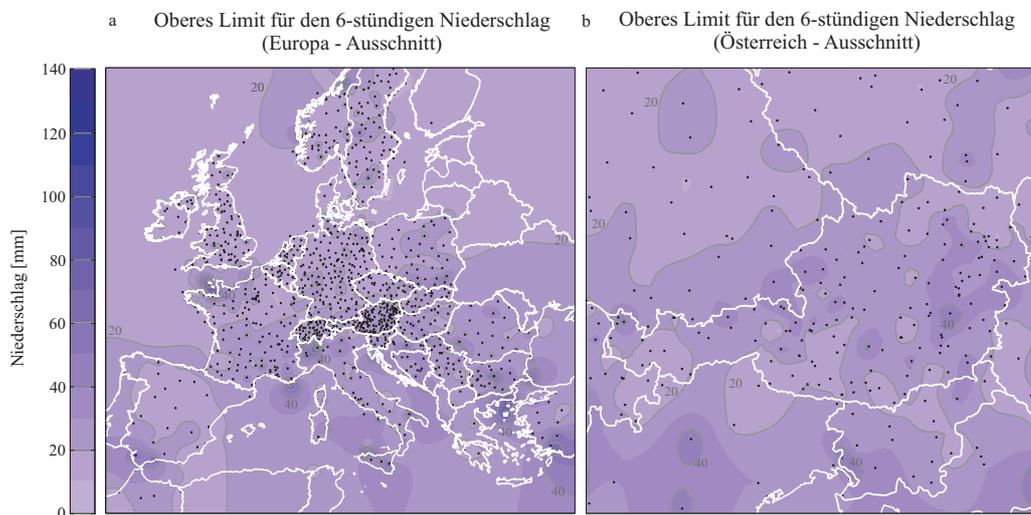
**Abbildung 4.8:** Spannweite (Differenz aus oberem und unterem Limit) der stationsspezifischen klimatologischen Grenzwerte für die potentielle Temperatur. Als Datengrundlage dienen die Stationsmeldungen des Zeitraums Jänner–Juni 2010. Die Interpolation auf ein regelmäßiges 40 km Gitter erfolgte mittels VERA und zusätzlich wurde eine Glättung der Gitterpunktwerte durchgeführt.

Für die potentielle Temperatur ist in Abbildung 4.8 die Spannweite der Limits (oberes minus unteres Limit) dargestellt. Es ist eindeutig zu erkennen, dass die Variabilität der Messwerte bzw. die Spannweite der klimatologischen Grenzwerte der potentiellen Temperatur über den Kontinenten deutlich höher ist als jene über dem Mittelmeer bzw. dem Atlantik. Im Bereich von der Ukraine bis zum Baltikum nimmt diese ihr Maximum ein. Die flächige Darstellung der äquivalentpotentiellen

Temperatur bietet ein vergleichbares Bild, wobei der Unterschied zwischen Kontinent und Meeresoberfläche etwas geringer ausfällt.

Die Messwerte der potentiellen und vor allem jene der äquivalentpotentiellen Temperatur enthalten einen ausgeprägten Jahresgang, der die Variabilität der Messwerte erhöht. Möchte man den zusätzlichen Aufwand vermeiden, diesen Jahresgang von den Messdaten abzuziehen, wäre eine vorangehende Normal Score Transformation der Messdaten von Vorteil. Dabei muss aber, wie bereits bei der Betrachtung des reduzierten Bodendrucks erwähnt, auf den Einfluss der Gross Errors Rücksicht genommen werden.

### Windkomponenten $u$ und $v$ sowie der 6-stündige Niederschlag



**Abbildung 4.9:** Flächenige Darstellung des stationenspezifischen oberen klimatologischen Limits für die 6-stündigen Niederschlagssummen, berechnet aus den Stationsmeldungen des Zeitraums Jänner–Juni 2010. Das obere Limit wurde mittels VERA auf ein Gitter mit  $\Delta x = \Delta y = 11.6$  km interpoliert.

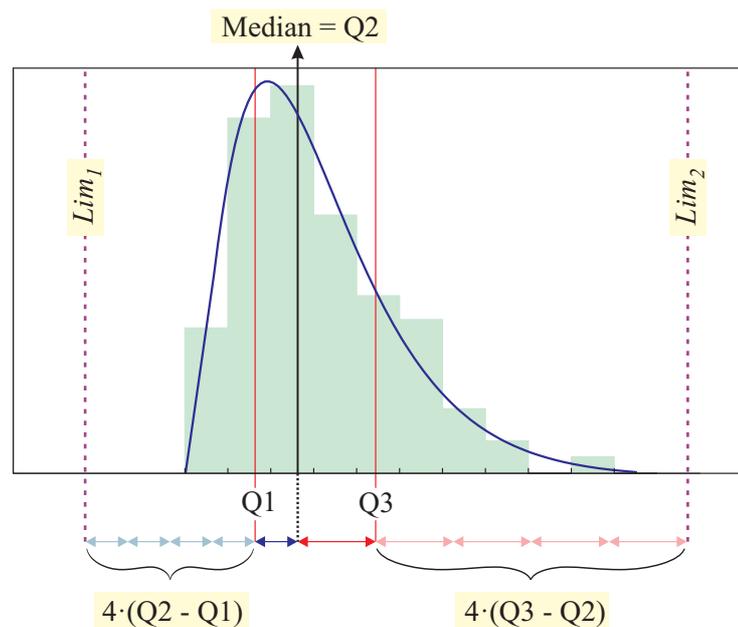
Die Windkomponenten  $u$  und  $v$  und vor allem die 6-stündigen Niederschlagssummen sind stark von den im betrachteten Halbjahr aufgetretenen Wetterlagen abhängig. So kommt es, dass die berechneten oberen Limits für den 6-stündigen Niederschlag (Abbildung 4.9a) lokale Maxima aufweisen, die sich sehr stark von ihrer Umgebung abheben (beispielsweise an der Küste Süd- oder Nordfrankreichs oder im westlichen Oberitalien). Das kann damit erklärt werden, dass ergiebige Niederschlagsereignisse örtlich sehr begrenzt sind und oftmals nur wenige Stationen betreffen.

Aus diesem Grund sollten für die Berechnung der klimatologischen Grenzwerte der Windgeschwindigkeitskomponenten und der 6-stündigen Niederschlagssummen

unbedingt längere Zeitreihen herangezogen werden. Betrachtet man den Österreich - Ausschnitt (Abbildung 4.9b), so erreichen die oberen Limits 20–40 mm in sechs Stunden, ein Wert, der leicht übertroffen werden kann. Auch wenn in dieser Darstellung lokale Höchstwerte durch die Interpolation abgeschwächt wurden, wird deutlich, dass Messzeitreihen von einem halben Jahr zu kurz sind, um klimatologisch mögliche Extremwerte für den Niederschlag zu ermitteln.

Ebenso wie beim Niederschlag sollte man auch bei der Berechnung der klimatologischen Grenzwerte der Windkomponenten  $u$  und  $v$  auf transformierte Daten zurückgreifen. Vor allem in komplexem Gelände kann es zu einer ausgeprägten Richtungsabhängigkeit der Windmesswerte kommen. In Folge davon kann auch die Häufigkeitsverteilung für eine Windkomponente, ähnlich jener der Niederschlagsmessungen, eine sehr einseitige Verteilung darstellen.

#### 4.1.4 Alternative Berechnung klimatologischer Limits



**Abbildung 4.10:** Veranschaulichung einer alternativen Möglichkeit der Berechnung klimatologischer Limits, welche die Schiefe einer Verteilung berücksichtigt. Die grünen Balken repräsentieren eine simulierte rechtsschiefe Häufigkeitsverteilung, die blaue Linie ist die zugehörige Dichtefunktion. Der Median ist durch die schwarze vertikale Linie und das untere bzw. obere Quartil durch rote vertikale Linien dargestellt. Mit Hilfe der Abstände zwischen Median und unterem bzw. oberem Quartil werden das untere ( $Lim_1$ ) und das obere ( $Lim_2$ ) klimatologische Limit (strichlierte vertikale Linien) berechnet.

Eine Möglichkeit die Berechnung der klimatologischen Limits zu verbessern ist eine mögliche Schiefe in den Häufigkeitsverteilungen der Messwerte zu berücksichtigen.

In Abbildung 4.10 ist zur Veranschaulichung eine rechtsschiefe Verteilung simuliert. Schiefe Verteilungen sind unter anderem dadurch charakterisiert, dass sich der Abstand zwischen dem unteren Quartil ( $Q_1$ ) und dem Median ( $Q_2$ ) vom Abstand zwischen dem Median ( $Q_2$ ) und dem oberen Quartil ( $Q_3$ ) sehr unterscheidet. Diese beiden Abstände (vom Median zum unteren bzw. zum oberen Quartil, in Abbildung 4.10 durch einen blauen bzw. roten Doppelpfeil gekennzeichnet) können zur Definition der klimatologischen Limits herangezogen werden:

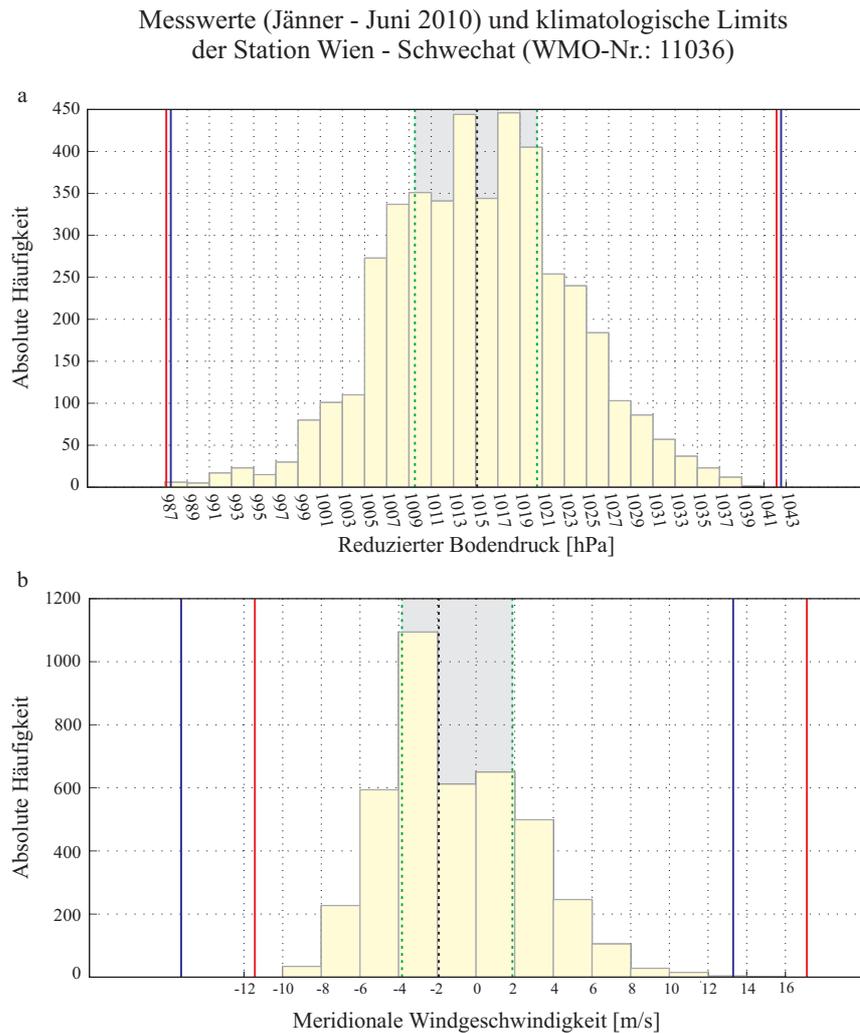
$$Lim_1 = Q_1 - 4 \cdot (Q_2 - Q_1) \quad \text{bzw.} \quad Lim_2 = Q_3 + 4 \cdot (Q_3 - Q_2) \quad (4.3)$$

Mit dieser Definition fällt das untere klimatologische Limit in Abbildung 4.10 nicht „unnötigerweise“ zu großzügig aus, wie es bei der Verwendung des Interquartilsabstands zur Berechnung der Grenzwerte der Fall wäre.

Im Folgenden sollen die klimatologischen Limits mit beiden vorgestellten Berechnungsmöglichkeiten für die Messzeitreihen der Station Wien - Schwechat (Jänner bis Juni 2010) ermittelt und einander gegenübergestellt werden.

Für die Parameter reduzierter Bodendruck und Windkomponente  $v$  sind die Ergebnisse in Abbildung 4.11a und 4.11b dargestellt. Die mit Gleichung 4.2 berechneten klimatologischen Grenzwerte sind durch vertikale blaue Linien gekennzeichnet, jene berechnet mit Gleichung 4.3 sind als rote vertikale Linien abgebildet. Werden die klimatologischen Limits für annähernd normalverteilte Daten berechnet, macht es aufgrund der zentralen Lage des Medians zwischen dem unteren und oberen Quartil keinen großen Unterschied mit welcher Berechnungsmethode die Grenzwerte ermittelt werden. Dieser Fall ist in Abbildung 4.11a für den reduzierten Bodendruck dargestellt. Betrachtet man jedoch die meridionale Windgeschwindigkeit in Abbildung 4.11b, die abhängig von den lokalen Gegebenheiten sehr stark richtungsabhängig sein kann und dadurch eine schiefe Verteilungsfunktion der Messwerte aufweist, ist der Vorteil der letzteren Berechnungsmethode (Gleichung 4.3) erkennbar. Durch die Berücksichtigung der Schiefe der Häufigkeitsverteilung der Messwerte ist das untere Limit nicht zu großzügig und das obere Limit nicht zu streng.

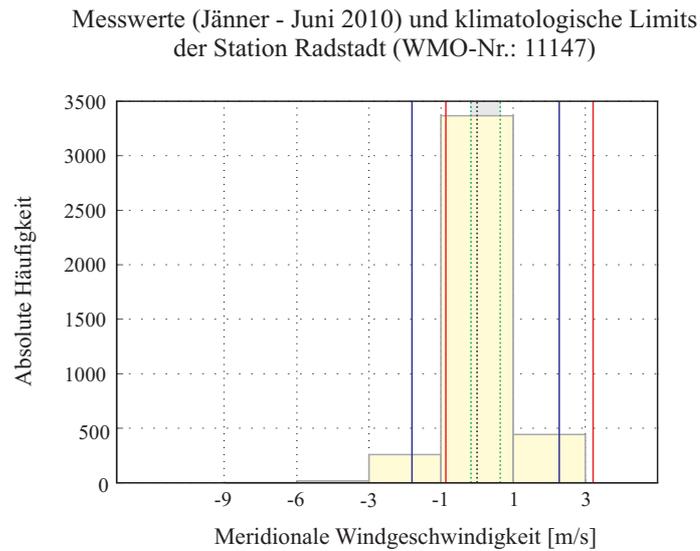
An dieser Stelle sei noch erwähnt, dass beide Berechnungsmöglichkeiten auch eine Schwachstelle aufweisen. Bei sehr steilgipfligen Verteilungen liegen das obere und untere Quartil sehr eng beieinander, woraus zu restriktive klimatologische Grenzwerte abgeleitet werden. Dieses Problem tritt vor allem für Windmessungen in komplexem Gelände auf. Zur Veranschaulichung ist in Abbildung 4.12 die Berechnung der klimatologischen Limits für die meridionale Windgeschwindigkeit an der Station



**Abbildung 4.11:** Anhand der Messzeitreihen der Station Wien - Schwechat werden klimatologische Grenzwerte für den reduzierten Bodendruck (a) und die meridionale Windgeschwindigkeit (b) nach den beiden vorgestellten Methoden berechnet und gegenübergestellt. Die blauen vertikalen Linien kennzeichnen die klimatologischen Limits nach der Berechnungsmethode, die den Interquartilsabstand verwendet (Gleichung 4.2) und die roten repräsentieren die Grenzwerte nach der Berechnungsmethode, welche die Schiefe der Verteilung berücksichtigt (Gleichung 4.3). Zusätzlich sind in Grün das untere und obere Quartil, sowie in Schwarz der Median eingezeichnet. Der grau hinterlegte Bereich enthält 50 Prozent der Messdaten.

Radstadt dargestellt. Beide Berechnungsmöglichkeiten ergeben sehr strenge klimatologische Grenzwerte. Ähnliche Probleme treten bei zu flachgipfligen Verteilungen auf, welche sehr tiefe bzw. hohe Limits ergeben.

Man beachte jedoch, dass in Abbildung 4.12 auch Windgeschwindigkeiten von 0 m/s berücksichtigt wurden, welche eine steilgipflige Häufigkeitsverteilung bedingen. Bei der Berechnung klimatologischer Grenzwerte sollte man somit Kalmen ausschließen. Mit der Definition von minimalen bzw. maximalen Schwellwerten könnten zu



**Abbildung 4.12:** Veranschaulichung der Problematik von steilgipflich verteilten Messwerten bei der Berechnung klimatologischer Limits anhand der meridionalen Windgeschwindigkeitsmesswerte an der Station Radstadt. Wie in Abbildung 4.11 kennzeichnen die blauen vertikalen Linien die klimatologischen Limits nach Gleichung 4.2 und die roten jene nach Gleichung 4.3. In Grün sind zusätzlich das untere und obere Quartil, sowie in Schwarz der Median eingezeichnet. Der grau hinterlegte Bereich enthält 50 Prozent der Messdaten.

restriktive bzw. zu weite klimatologische Limits verhindert werden. Eine weitere Möglichkeit wäre, den Faktor mit dem der Abstand zwischen dem oberen bzw. unterem Quartil und dem Median gewichtet wird nicht nur vom Parameter sondern auch von der Wölbung der stationsspezifischen Verteilung abhängig zu machen. Das hier präsentierte Kalkül soll Anregung zu weiteren Untersuchungen geben die, um ein breiteres Spektrum an Extremwerten zu berücksichtigen, mit längeren Messzeitreihen durchgeführt werden sollten.

# Kapitel 5

## Qualitätskontrolle der potentiellen Temperatur

### 5.1 Die potentielle Temperatur

Synoptisch-skalige Bewegungen in der freien Atmosphäre laufen in guter Näherung isentrop ab, wobei die Entropie des betrachteten Luftpakets definitionsgemäß konstant bleibt. Adiabatisch reversible Prozesse, also jene ohne Entropietransfer über die Systemgrenzen hinaus und ohne Entropieproduktion, stellen isentrope Prozesse dar. Eine Größe, die auf isentropen Flächen bzw. bei isentropen Prozessen konstant bleibt, ist die potentielle Temperatur  $\theta$ :

$$\theta = T \left( \frac{p_0}{p} \right)^\kappa ; \quad \kappa = \frac{R}{c_p} \quad p_0 = 1000 \text{hPa} \quad (5.1)$$

mit  $T$  der Temperatur,  $p$  dem lokalen Druck,  $p_0$  dem Referenzdruck,  $R=287.04 \text{ J kg}^{-1}\text{K}^{-1}$  der Gaskonstante für trockene Luft und  $c_p = 1005 \text{ J kg}^{-1}\text{K}^{-1}$  der spezifischen Wärmekapazität bei konstantem Druck für trockene Luft. Formal lässt sich der oben erklärte Zusammenhang mit Hilfe folgender Formeln ausdrücken:

$$dh = \alpha dp + T ds \quad (5.2)$$

$$dh = c_p dT \quad (5.3)$$

$$T ds = c_p dT - \alpha dp \quad \text{bzw.} \quad ds = c_p \left( \frac{dT}{T} - \kappa \frac{dp}{p} \right) \quad (5.4)$$

$h$  bezeichnet die spezifische Enthalpie,  $\alpha$  das spezifische Volumen und  $s$  die spezifische Entropie. Die Gibbsche Gleichung (5.2) kann mit der Beziehung (5.3) zu

Gleichung (5.4) umgeformt werden, die wiederum mit der Definition der potentiellen Temperatur (Gleichung (5.1)) folgenden Zusammenhang ergibt:

$$ds = c_p \frac{d\theta}{\theta} \quad (5.5)$$

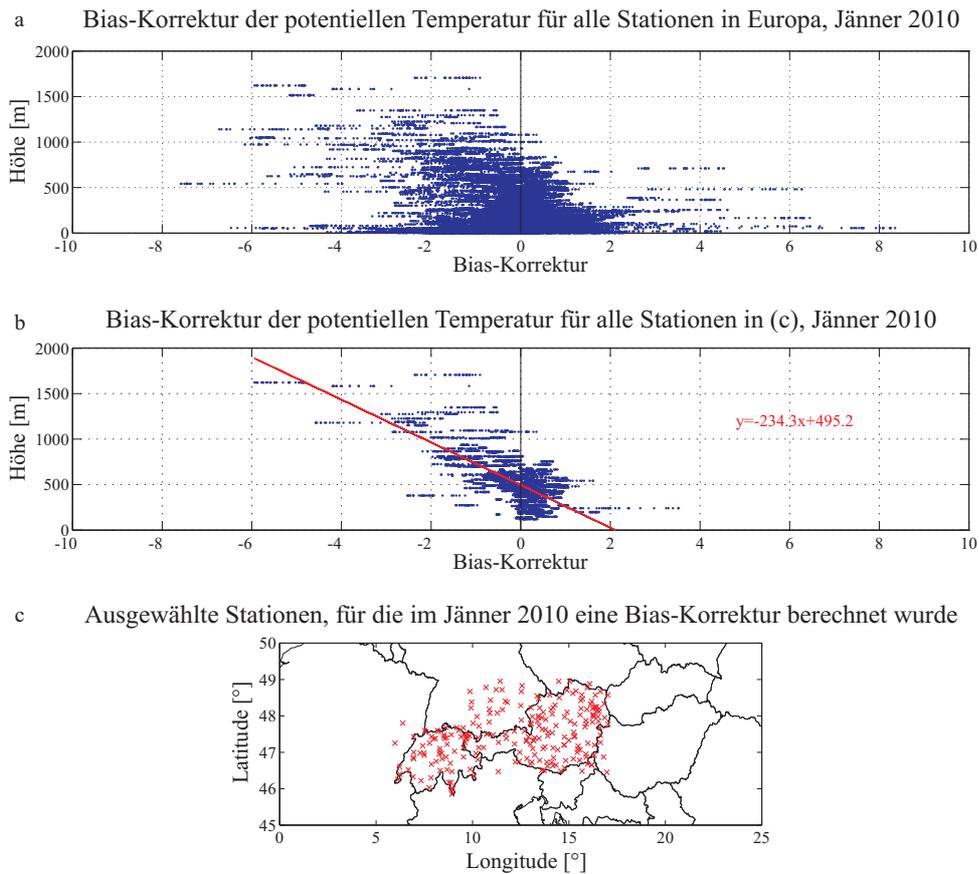
Wie erwähnt, bleibt bei isentropen Prozessen die Entropie konstant,  $ds = 0$ , womit sich aus Gleichung (5.5) auch  $d\theta = 0$  bzw.  $\theta = \text{const}$  ableiten lässt. Die präsentierten Zusammenhänge und Definitionen können im 7. Band der Reihe *Bergmann-Schaefer: Lehrbuch der Experimentalphysik* im Kapitel 3 (Bauer et al. 2001) nachgelesen werden.

## 5.2 Problemstellung

Anschaulich ausgedrückt, beschreibt die potentielle Temperatur  $\theta$  jene Temperatur, die ein Luftpaket annimmt, wenn es isentrop auf den Referenzdruck  $p_0 = 1000$  hPa gebracht wird. Somit eignet sich die potentielle Temperatur dazu, Temperaturen aus verschiedenen Höhenlagen vergleichbar zu machen und Luftmassen zu unterscheiden, sofern in der Atmosphäre trockenisentropen Verhältnisse vorherrschen. Da diese Bedingung unter realen Verhältnissen nicht immer erfüllt ist und der Unterschied zwischen dem geometrischen (realen) und dem trockenisentropen Temperaturgradienten ( $\partial T/\partial z = -0.98$  K/100m) regional schwankt, sind die aus den lokal gemessenen Temperatur- und Druckwerten berechneten potentiellen Temperaturen einer höher gelegenen Station und jene Temperaturen einer Station im Flachland nicht uneingeschränkt vergleichbar.

Geht man von einer vertikalen Temperaturverteilung aus, die im Mittel stabile Verhältnisse präsentiert, so ist der geometrische Temperaturgradient geringer als jener bei trockenisentropen Verhältnissen. Das hat zur Folge, dass die gemessene Temperatur einer höher gelegenen Station, wie auch die daraus berechnete potentielle Temperatur, höhere Werte annehmen als es bei trockenisentropen Verhältnissen der Fall wäre. Für eine tiefer gelegene Station fällt dieser Effekt natürlich geringer aus. Der Einfluss von kleinen Fehlern bezüglich des Stationsdrucks kann vernachlässigt werden.

Die Qualitätskontrolle vergleicht nun die potentiellen Temperaturen räumlich (horizontal) und korrigiert die höheren Werte der hochgelegenen Stationen, die in



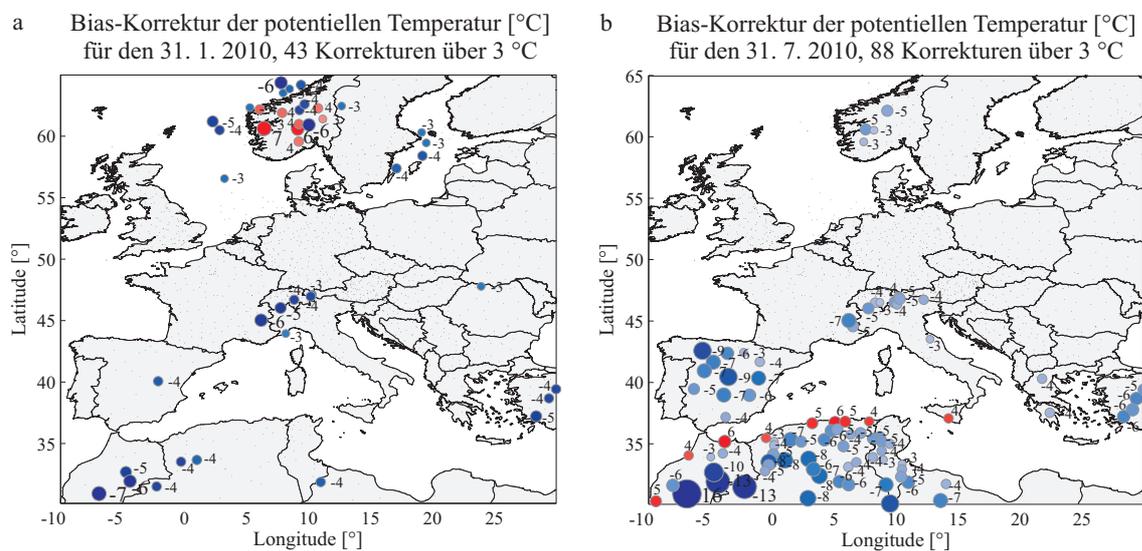
**Abbildung 5.1:** Darstellung der Abhängigkeit der Bias-Korrektur für die potentielle Temperatur von der Höhe für alle Stationen in Europa (a) und alle Stationen im Großraum Österreich und Schweiz (b). Letztere Stationen sind in (c) durch rote Kreuze gekennzeichnet.

geringerer Anzahl vorkommen, nach unten. Aus den systematisch negativen Deviations für die potentielle Temperatur resultiert auch ein negativer Bias, der eine starke Höhenabhängigkeit aufweist. Dieser Zusammenhang ist in den Abbildungen 5.1a–c dargestellt, die sich auf die Bias-Korrekturen vom 1. bis 31. Jänner 2010 beziehen. In Abbildung 5.1a sind die Bias-Korrekturen aller Stationen in Europa gegen die Höhe aufgetragen und lassen bereits die erwähnte Tendenz höher gelegener Stationen zu negativen Bias-Korrekturen erkennen. Betrachtet man nur Stationen im Großraum Österreich und Schweiz (rote Kreuze in Abbildung 5.1c), so nimmt deren Anteil an höher gelegenen Stationen zu und die Höhenabhängigkeit der Bias-Korrekturen (Abbildung 5.1b) wird deutlicher erkennbar. In den Sommermonaten werden für die hochgelegenen Stationen betragsmäßig noch größere negative Bias-Korrekturen berechnet.

Für eine hochgelegene Talstation in den Alpen kommt es vor allem im Sommer mit erhöhter Einstrahlung durch das reduzierte Talvolumen zu einer stärkeren Erwärmung, als dies im Flachland der Fall wäre. Ebenso sind Stationen auf einer

Hochebene in den Sommermonaten von einer stärkeren Erwärmung betroffen (wie zum Beispiel am spanischen Hochplateau). Diese lokale Überhitzung weist einen starken Jahresgang auf, der sich auch in den Bias-Korrekturen widerspiegelt.

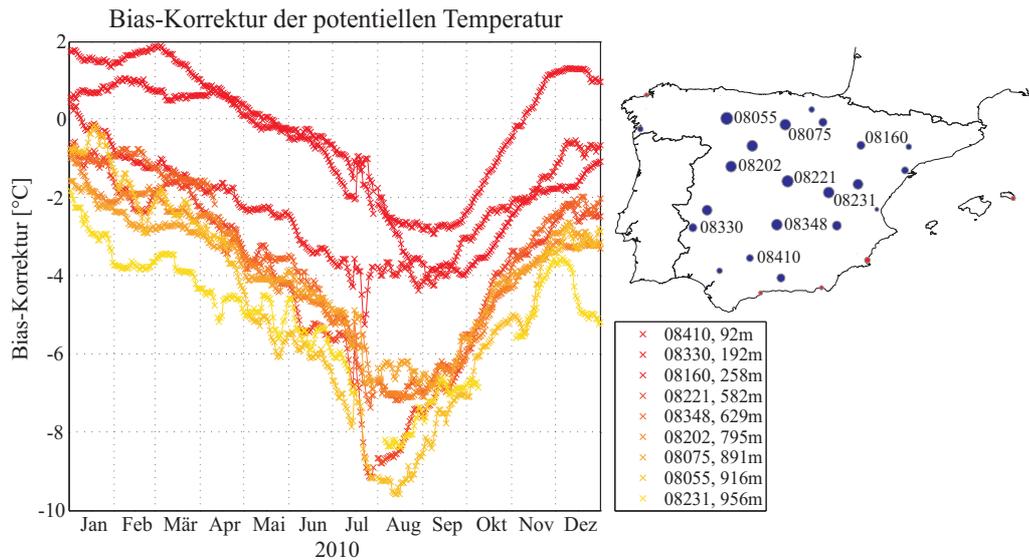
An dieser Stelle sei erwähnt, dass nicht alle zur Verfügung stehenden Stationen für die VERA-QC und eine anschließende VERA-Analyse in Betracht gezogen werden. Ein vorgeschalteter Stationsauswahlalgorithmus sorgt dafür, dass nur vergleichbare Stationen nahe der Minimumtopographie (Bica *et al.* 2007) in die VERA-QC und in die anschließende VERA-Analyse eingehen. Somit wurden hochgelegene Bergstationen, die keiner lokalen Überhitzung mehr unterliegen, in der oben beschriebene Problemstellung nicht eingeschlossen. Abbildung 5.1, welche maximal Stationen in hochgelegenen Tälern und auf Hochplateaus berücksichtigt, würde für die freie Atmosphäre bzw. für Bergstationen ein anderes Bild zeigen.



**Abbildung 5.2:** Gegenüberstellung der Bias-Korrekturen für die potentielle Temperatur (a) im Winter (Bias-Korrektur für 31. Jänner 2010) und (b) im Sommer (Bias-Korrektur für 31. Juli 2010). Für eine bessere Übersichtlichkeit sind nur Bias-Korrekturen über 3 °C dargestellt.

Zur Veranschaulichung des Problems der Höhenabhängigkeit der Bias-Korrekturen sind in Abbildung 5.2 die Bias-Korrekturen der potentiellen Temperatur für den 31. Jänner und den 31. Juli 2010 gegenübergestellt und Abbildung 5.3 präsentiert den Jahresgang der Bias-Korrekturen für spanische Stationen in unterschiedlichen Höhenlagen.

Am Ende dieses Kapitels zeigt Abbildung 5.5a die VERA-Analyse vom 6. September 2010 um 06:00 UTC (zu diesem Termin konnte man föhnige Tendenzen erkennen), nachdem die Werte der potentiellen Temperatur von der Qualitätskontrolle mit



**Abbildung 5.3:** Jahresgang der Bias-Korrektur für die potentielle Temperatur ausgewählter Stationen in Spanien (durch Stationsnummern gekennzeichnet, Punktgröße proportional zur Größe der Bias-Korrektur eines Sommertags). In der Legende sind neben den Stationsnummern auch die Stationshöhen in Meter eingetragen, die Farbtöne von rot bis gelb kennzeichnen niedrige bis hochgelegene Stationen. Neben dem ausgeprägten Jahresgang der Bias-Korrekturen ist auch zu erkennen, dass höher gelegene Stationen (gelbe Farbtöne) eher negative Bias-Korrekturen erhalten.

Bias-Korrektur (Bias-Berechnungszeitraum: 30 Tage) überprüft und korrigiert wurden. Als Resultat der vor allem in den Sommermonaten negativen Bias-Korrekturen erhält man ein stark geglättetes  $\theta$ -Feld, in dem der Einfluss der Gebirge kaum mehr erkennbar ist.

### 5.3 Lösungsvorschlag

Alternativ zur potentiellen Temperatur kann auch der D-Wert kontrolliert werden, welcher als Differenz zwischen der potentiellen Temperatur gemäß Standardatmosphäre und der berechneten potentiellen Temperatur definiert ist. Dies hat zur Folge, dass der aktuelle Temperaturgradient nicht mehr durch den trockenisentropen, sondern durch den wahrheitsgetreueren Temperaturgradienten der Standardatmosphäre angenähert wird.

Um  $\theta$  der Standardatmosphäre zu berechnen, benötigt man  $\partial\theta/\partial z$ , den man durch Ableiten der Definition (5.1) nach der Vertikalkoordinate  $z$  unter Zuhilfenahme der statischen Grundgleichung und der Zustandsgleichung für ideale Gase erhält:

$$\frac{\partial\theta}{\partial z} = \frac{\theta}{T} \left( \frac{\partial T}{\partial z} + \frac{g}{c_p} \right) \quad (5.6)$$

In Gleichung (5.6) beschreibt  $g$  die Erdbeschleunigung und  $g/c_p$  den isentropen Temperaturgradienten:

$$\frac{g}{c_p} = -\frac{\partial T}{\partial z} \equiv \Gamma = 0.98 \text{ K/100m} \quad (5.7)$$

In guter Näherung gilt in der unteren Troposphäre  $\theta/T \approx 1$  und Gleichung (5.6) vereinfacht sich zu:

$$\frac{\partial \theta}{\partial z} \approx \frac{\partial T}{\partial z} + \Gamma \quad (5.8)$$

Für die Standardatmosphäre ist in der Troposphäre  $\partial T/\partial z = -0.65 \text{ K/100m}$  festgelegt, womit

$$\frac{\partial \theta}{\partial z} \approx 0.35 \text{ K/100m} \quad (5.9)$$

resultiert.

Die potentielle Temperatur gemäß Standardatmosphäre im Stationsniveau  $\theta_{stand}(z)$  wird über

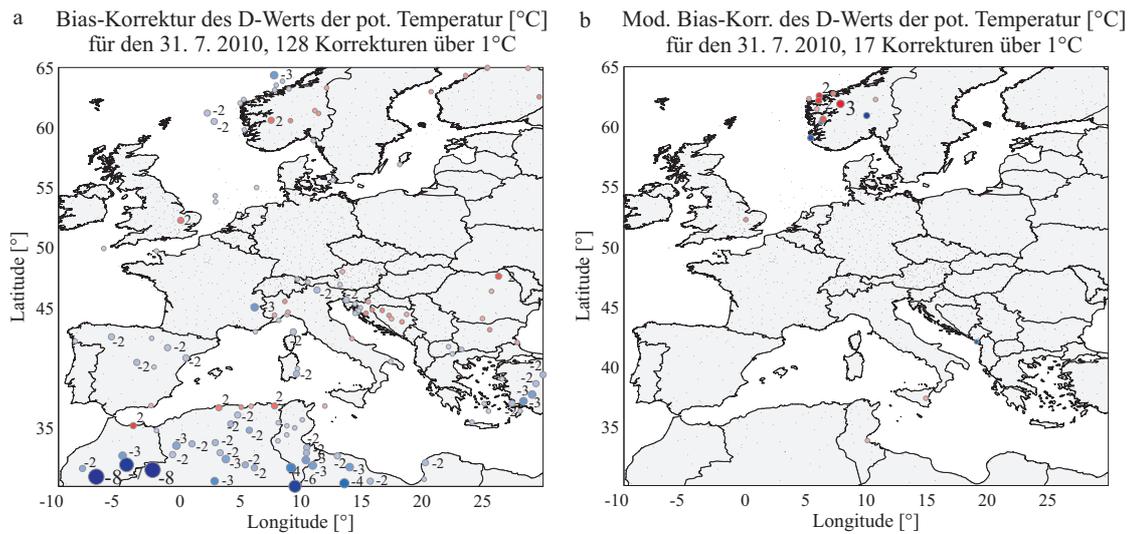
$$\theta_{stand}(z) = \theta_{stand}(z_0) + \left(\frac{\partial \theta}{\partial z}\right) h \quad (5.10)$$

berechnet, mit  $\theta_{stand}(z_0) = 15 \text{ °C}$ , der potentiellen Temperatur der Standardatmosphäre im Meeresniveau und  $h$  der Stationshöhe in Meter.

Nachdem die D-Werte von der QC überprüft und die berechneten gewichteten Deviations angebracht wurden, werden die Werte der Standardatmosphäre wieder den D-Werten aufaddiert und ergeben somit die korrigierten potentiellen Temperaturen. Die daraus resultierende VERA-Analyse für den Fall vom 6. September 2010 um 06:00 UTC ist in Abbildung 5.5b dargestellt und liefert für den Alpenraum ein zufriedenstellendes Resultat.

Die Kontrolle der Abweichung zur Standardatmosphäre führt zu einer Reduktion der negativen Korrekturen für höher gelegene Stationen. Lediglich in den Gebieten, wo die vertikale Temperaturabnahme aufgrund von lokaler Überhitzung noch geringer ausfällt als durch den Temperaturgradienten der Standardatmosphäre vorgegeben, ergeben sich noch immer negative Deviations und in Folge daraus auch negative Bias-Korrekturen. Dies ist vor allem im Bereich der spanischen Hochebene, in Nordafrika und auch in den westlichen Gebirgen der Türkei erkennbar (siehe Abbildung 5.4a).

Um dem entgegenzuwirken wird den Deviations das Flächenmittel der Korrekturen der jeweiligen primären Nachbarstationen abgezogen und somit eine durchwegs negative (bzw. gegebenenfalls positive) Korrektur ganzer Landstriche verhindert. Zusätzlich werden bei der Berechnung der Bias-Korrektur für den D-Wert der



**Abbildung 5.4:** (a): Bias-Korrekturen der D-Werte der potentiellen Temperaturen für den 31. 7. 2010, berechnet als Median der ungewichteten Deviations der vorangegangenen 243 Tage. (b): Den Deviations wurde zusätzlich das Flächenmittel über die Korrekturen der jeweils primären Nachbarstationen abgezogen und die Berechnung erfolgte aus den Deviations der Monate Juli und Jänner. Hervorgehoben sind in Blau negative und in Rot positive Korrekturen über 1 °C, die Beschriftung erfolgt für Bias-Korrekturen ab 2 °C (Werte gerundet).

potentiellen Temperatur die Deviations der letzten 365 Tage herangezogen, um den Jahresgang des vertikalen Temperaturgradienten zu berücksichtigen. Um Rechenzeit zu sparen und die Vorlaufzeit zu verkürzen, können alternativ auch Deviations des letzten und eines, ein halbes Jahr zurückliegenden Monats als Berechnungsgrundlage dienen (vgl. Kapitel 3, Abschnitt 3.3.3).

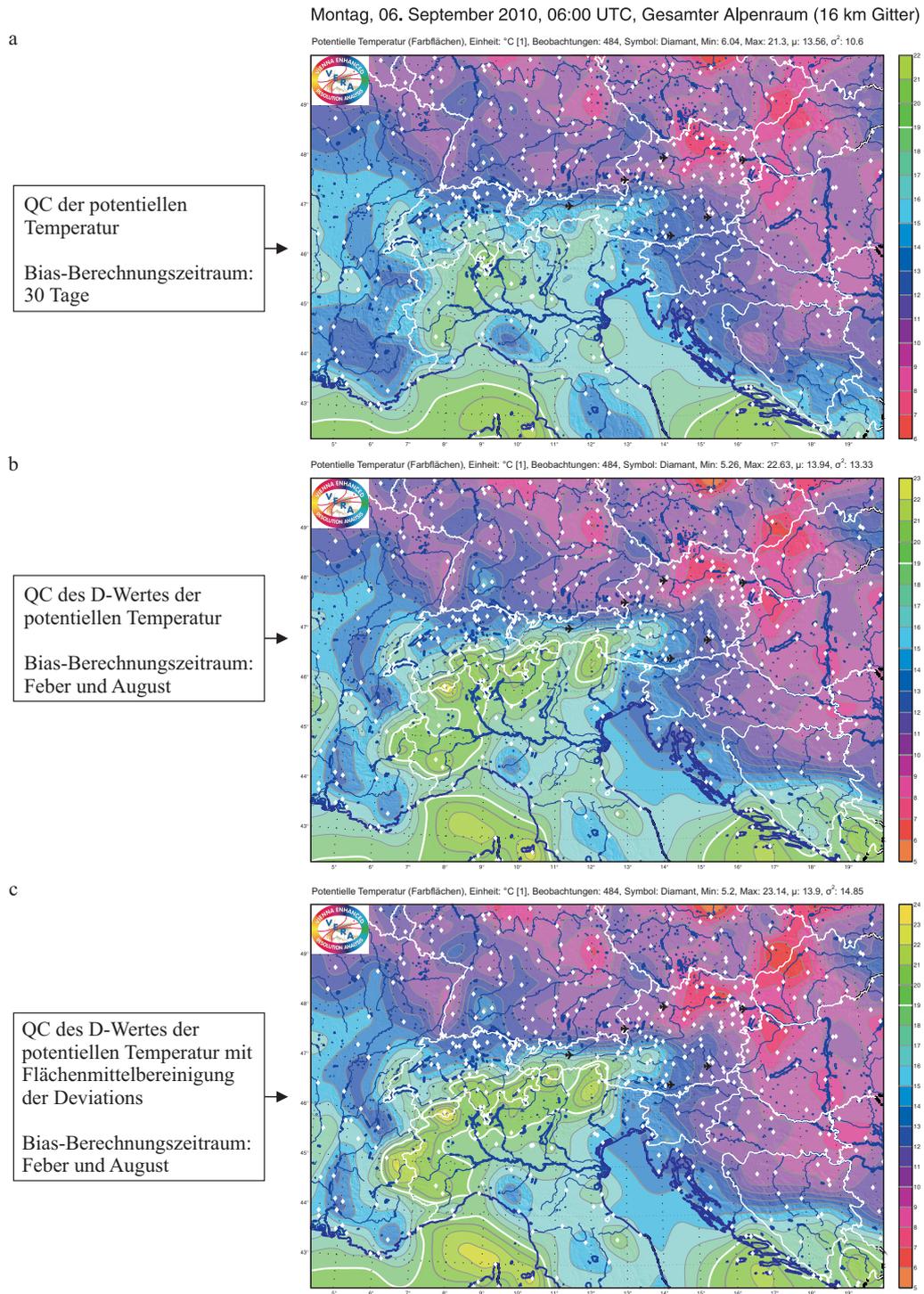
Abbildung 5.4 vergleicht die Bias-Korrekturen, berechnet aus den Deviations für die D-Werte der potentiellen Temperatur der letzten 243 Tage (Deviations werden erst seit Dezember 2009 gesammelt) in (a), und jene, berechnet aus den flächenmittelbereinigten Deviations der D-Werte der Monate Juli und Jänner in (b). Gut erkennbar ist der positive Einfluss der Flächenmittelbereinigung, sodass die modifizierten Bias-Korrekturen (b) betragsmäßig kaum über 1 °C reichen und sich lediglich in Norwegen bemerkbar machen. Außerdem weist die modifizierte Bias-Korrektur keinen Jahresgang mehr auf. Die Analyse der potentiellen Temperatur vom 6. September 2010 um 06:00 UTC mit vorangegangener modifizierter Qualitätskontrolle (Bias-Korrektur bezüglich zweier 30-Tage Intervalle ( $\sim$  Feber und August), Betrachtung des D-Werts und Flächenmittelbereinigung der Deviations) ist in Abbildung 5.5c dargestellt.

Die QC des D-Wertes an Stelle der potentiellen Temperatur führt in den Alpen

bereits zu einer deutlichen Verbesserung. Die Flächenmittelbereinigung der Deviations hingegen verhindert auch für die bereits oben erwähnten Gebiete mit lokaler Überhitzung eine verfälschende negative Korrektur.

**Bemerkung:**

Wird bei der QC der potentiellen Temperatur der D-Wert kontrolliert, so ergeben sich für höher gelegene Stationen höhere  $\theta$ -Werte. Es ist zu beachten, dass sich diese Änderung auf alle von der potentiellen Temperatur abgeleiteten Größen auswirkt. So zum Beispiel auch auf die Berechnung der Temperaturen der Täler und Niederungen, in der neben dem reduzierten Bodendruck und einem vorgegebenen vertikalen Temperaturgradienten auch die potentielle Temperatur eingeht. Somit wird auch die Temperatur der Täler und Niederungen an den Positionen höher gelegener Stationen angehoben. Je nach beabsichtigter Anwendung kann die QC des D-Werts der potentiellen Temperatur von Vor- oder Nachteil sein und muss auf den Benutzer abgestimmt werden. Um Luftmassengrenzen zu erkennen, könnte eine Analyse nach der QC der potentiellen Temperatur besser geeignet sein (Abbildung 5.5a). Im synoptischen Dienst kann es jedoch von Vorteil sein, Werte der potentielle Temperatur zur Verfügung gestellt zu bekommen, die sich näher an den gemessenen Werten befinden. Betrachtet man beispielsweise eine Föhn-Wetterlage, so würden die vom Föhn verursachten hohen Temperaturen der höher gelegenen Talstationen von einer QC der potentiellen Temperatur stark nach unten korrigiert werden und der Einfluss des Föhns würde geringer eingeschätzt werden. In diesem Fall wäre eine Analyse nach QC des D-Wertes der potentiellen Temperatur vorteilhafter (Abbildung 5.5b, bzw. Abbildung 5.5c).



**Abbildung 5.5:** VERA-Analyse der qualitätskontrollierten potentiellen Temperatur vom 6. September 2010, 06:00 UTC, für den Alpenraum-Ausschnitt. Eingehende Stationen sind durch weiße Diamant-Symbole gekennzeichnet. (a), (b) und (c) zeigen die Analysen, nachdem die Messwerte von einer VERA-QC mit jeweils unterschiedlichen Einstellungen kontrolliert wurden. Man erkennt, vor allem in (c), dass die höheren potentiellen Temperaturwerte im Gebirge (zu diesem Termin konnte man föhnige Tendenzen erkennen) erhalten bleiben.

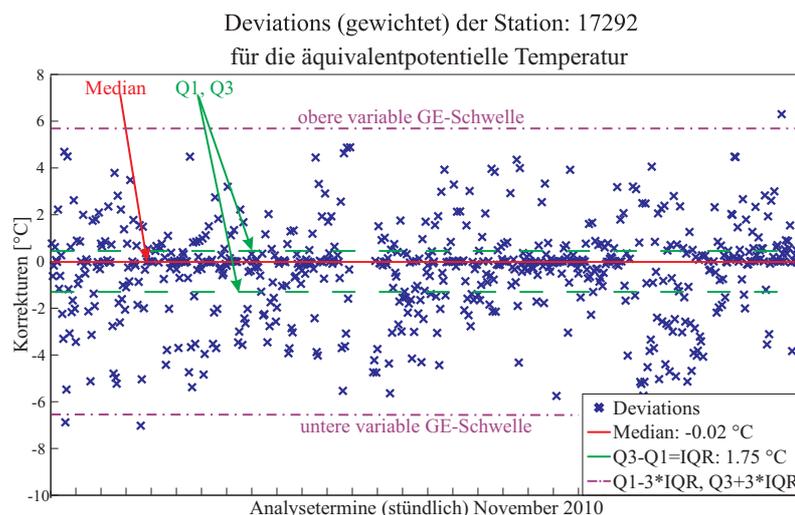


# Kapitel 6

## Die variable Gross Error Schwelle

### 6.1 Definition einer stationenspezifischen variablen Gross Error Schwelle

Die Gross Error Erkennung, wie sie im Abschnitt 2 beschrieben ist, hat sich sehr gut bewährt. Dennoch gibt es Fälle, in denen Gross Errors nicht erkannt werden und somit wurde die Einführung eines zusätzlichen Kriteriums notwendig. Um nicht nur wie bisher die räumliche Verteilung der Messwerte bei der Gross Error Erkennung zu berücksichtigen, wird auf die Fehlerstatistik der einzelnen Stationen zurückgegriffen. Die Methodik der variablen, stationenspezifischen GE-Schwelle wird im Folgenden mit Hilfe der Abbildung 6.1 veranschaulicht.



**Abbildung 6.1:** Gewichtete Deviations der Station 17292 für die äquivalentpotentielle Temperatur im November 2010 (blaue Kreuze). Daraus berechnet sind der Median in Rot, das obere und das untere Quartil in Grün und die variablen Schwellen zur Gross Error Erkennung in Violett dargestellt.

Darin sind in Blau die gewichteten Deviations der Station 17292 für die äquivalentpotentielle Temperatur der stündlichen Analysetermine im November 2010 dargestellt. Zusätzlich ist der Median der Deviations als rote Linie und das untere und das obere Quartil (Q1, Q3) jeweils als grün gestrichelte Linie eingezeichnet. Die variable untere und obere Schwelle zur Erkennung von Gross Errors ist definiert als:

$$Q1 - \text{Vielfaches} \cdot IQR \quad \text{bzw.} \quad Q3 + \text{Vielfaches} \cdot IQR \quad (6.1)$$

mit dem Interquartilsabstand  $IQR = Q3 - Q1$  und ist jeweils durch die violette strichpunktierte Linie gekennzeichnet. Als Vielfaches der IQR wurde der dreifache Interquartilsabstand gewählt, wobei dieser Faktor vom Benutzer in der Konfigurationsdatei einstellbar ist.

Über- oder unterschreitet nun eine gewichtete Korrektur diese stationsspezifischen Schwellwerte, so wird die Meldung als Gross Error erkannt und verworfen. Um zu verhindern, dass auf diese Weise zu kleine Korrekturen bereits als Gross Errors interpretiert werden, muss die gewichtete Korrektur zusätzlich noch eine benutzerdefinierte Schwelle (hier: von 4 °C) überschreiten. Dieses Problem tritt bei Deviations mit geringer Streubreite auf.

Die auf die beschriebene Art berechneten variablen Grenzwerte zur Gross Error Erkennung sind jeweils ein Monat gültig und werden am Ende des Monats aus den vorliegenden gewichteten Deviations des endenden Monats neu berechnet. Dabei wird die Bedingung gestellt, dass für die jeweils betrachtete Station eine Mindestanzahl an Deviations (hier: 50) im Berechnungszeitraum vorliegen muss. Die variablen Schwellwerte aus Abbildung 6.1 sind somit für die Station 17292 an allen Analyseterminen im Dezember 2010 gültig.

## 6.2 Anwendung der variablen Gross Error Schwelle im Jahr 2010

Die beschriebene IQR-Bedingung ist seit Jänner 2010 als Erweiterung der bisherigen Gross Error Erkennung im Einsatz und wird auf die Windgeschwindigkeitskomponenten ( $u$  und  $v$ ), die potentielle und äquivalentpotentielle Temperatur ( $\theta$  und  $\theta_e$ ), den reduzierten Bodendruck ( $p_{red}$ ) und die Windgeschwindigkeit ( $|\vec{v}|$ ) angewendet. Aus Tabelle 6.1 kann man entnehmen, wie viel Prozent der zur Verfügung stehenden Meldungen im Jahr 2010 mit und ohne der IQR-Bedingung als Gross Error erkannt wurden.

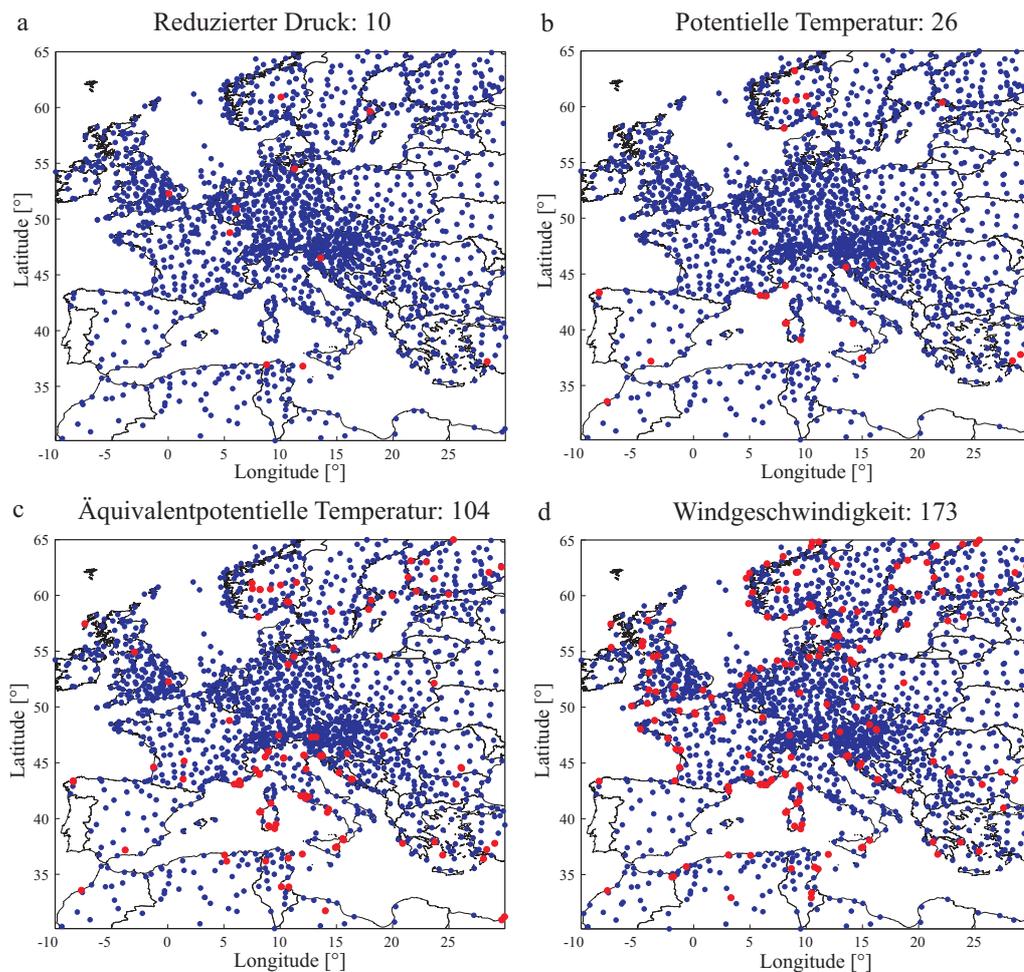
**Tabelle 6.1:** Vergleich der Gross Error Erkennung mit und ohne der zusätzlichen stationsspezifischen variablen IQR-Schwelle im Jahr 2010. In der ersten Zeile wird angegeben, wie viel Prozent der, für den jeweiligen Parameter zur Verfügung stehenden Stationsmeldungen von der Gross Error Erkennung ohne der IQR-Bedingung als GE identifiziert wurden. Die zweite Zeile enthält jene Gross Errors, die von der erweiterten Gross Error Erkennung gefunden wurden.

Erkannte Gross Errors	$u$	$v$	$\theta$	$\theta_e$	$p_{red}$	$ \vec{v} $
ohne IQR-Bedingung [%]	0.428	0.468	0.306	0.338	0.324	0.493
mit IQR-Bedingung [%]	0.551	0.592	0.369	0.648	0.334	0.586

Auch mit dem zusätzlichen Kriterium werden, wie per Definition von Gross Errors verlangt, nur wenige Meldungen verworfen. Lediglich bei der äquivalentpotentiellen Temperatur hat sich die Anzahl der erkannten Gross Errors beinahe verdoppelt. Der Grund dafür ist, dass die zusätzliche IQR-Bedingung nur verwendet wird, wenn der Absolutwert der variablen GE-Schwelle einen gewissen benutzerdefinierten Schwellwert (4 °C in der operationellen VERA-QC) überschreitet. Der Absolutwert der variablen GE-Schwelle ist umso größer, je größer die Varianz der gewichteten Deviations ist. Diese ist wiederum direkt proportional zur Varianz der Messwerte. Somit ist die effektive Anzahl der Stationen, für die dieses zusätzliche GE-Kriterium zum Einsatz kommt, nicht bei allen Parametern gleich.

In Abbildung 6.2 sind alle Stationen blau eingezeichnet, von denen im Oktober 2010 Meldungen vorliegen. Jene Stationen, an denen für den dargestellten Parameter die IQR-Bedingung zur GE-Erkennung verwendet werden kann, sind rot markiert. Bei der äquivalentpotentiellen Temperatur und der Windgeschwindigkeit kommt die zusätzliche GE-Erkennung aufgrund ihrer höheren Variabilität am öftesten zum Einsatz.

Stationen mit zusätzlicher GE-Bedingung  
Oktober 2010

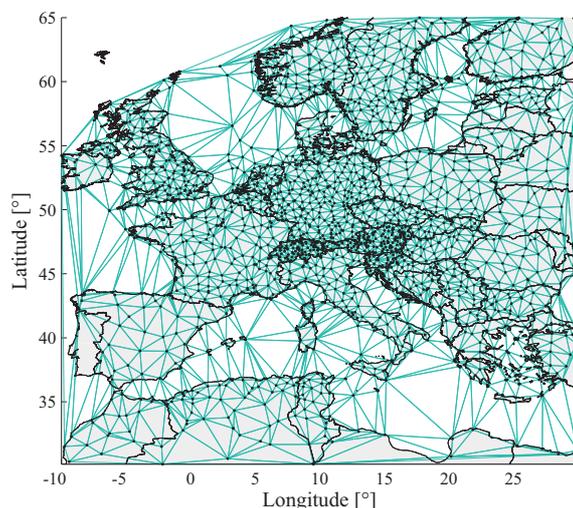


**Abbildung 6.2:** Blau: Stationen, für die im Oktober 2010 Werte für die Parameter reduzierter Druck (a), potentielle Temperatur (b), äquivalentpotentielle Temperatur (c) und Windgeschwindigkeit (d) zur Verfügung stehen. Rot: Stationen, für welche die zusätzliche IQR-Bedingung bei der Gross Error Erkennung zu tragen kommt.

# Kapitel 7

## Statistische Auswertung der Stationsmeldungen aus dem Jahr 2010

Im vorliegenden Kapitel werden die dem Institut für Meteorologie und Geophysik zur Verfügung stehenden Messwerte und deren Behandlung durch die VERA-QC genauer betrachtet und auf Auffälligkeiten hin untersucht. Die Ergebnisse sind selbstverständlich von den verwendeten Einstellungen abhängig, die zugehörige Konfigurationsdatei ist im Anhang A.1 angefügt.



**Abbildung 7.1:** Betrachtete Europa-Domäne mit Positionen der GTS-Stationen (schwarze Punkte), von denen am 13. 7. 2010, 06:00 UTC eine Meldung zur Verfügung steht. Die grünen Linien zeigen an, welche Stationen als natürliche Nachbarn verbunden worden sind.

In Abbildung 7.1 ist die betrachtete Europa-Domäne dargestellt und zusätzlich sind alle Positionen jener Stationen, von denen eine Meldung zu einem zufällig ausgewählten Analysetermin im Jahr 2010 zur Verfügung steht, mit schwarzen Punkten markiert. Die stark variierende Dichte an Stationsmeldungen stellt für jede Qualitätskontrolle eine große Herausforderung dar. Ein Vorteil der VERA-QC liegt darin, dass sie sich mit dem Konzept der „natürlichen Nachbarfindung“ (Kapitel 2, Abschnitt 2.1.1) an die lokale Stationsdichte anpassen kann und die Definition eines Einflussradius nicht nötig ist. Die grünen Linien in Abbildung 7.1 zeigen an, welche Statio-

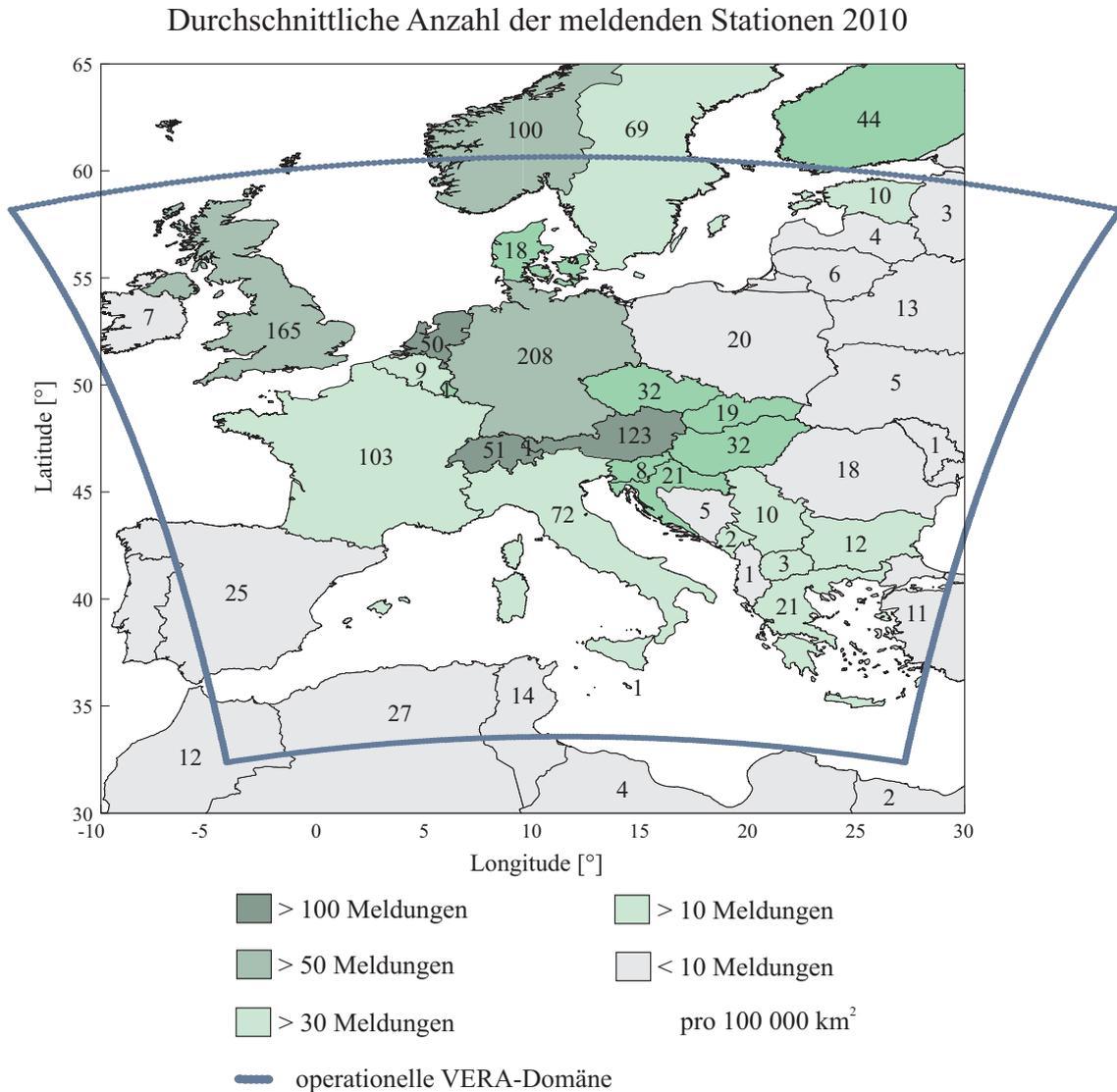
nen als Nachbarstationen miteinander verbunden worden sind. Man erkennt, dass die Stationsdichte von Land zu Land variiert und aus diesem Grund werden in den nächsten Abschnitten die vorhandenen Meldungen bezüglich länderspezifischer Unterschiede untersucht.

## 7.1 Durchschnittliche Anzahl an Stationsmeldungen

In Abbildung 7.2 ist die Anzahl der durchschnittlich vorhandenen, stündlichen Stationsmeldungen im Jahr 2010, aufgeschlüsselt nach Ländern, dargestellt (Zahlenwert innerhalb der jeweiligen Ländergrenzen). Die Zählung erfolgte, sobald von einer Station für einen Parameter eine Meldung vorliegt, ungeachtet der Qualität der Meldung.

Um das Problem unterschiedlicher Stations- bzw. Meldungsdichten innerhalb der betrachteten Europa-Domäne zu veranschaulichen (vgl. Abbildung 7.1), bezieht sich die Farbgebung auf die Anzahl der durchschnittlichen Meldungen je Land, bezogen auf eine Fläche von 100 000 km<sup>2</sup>. Für Länder, deren Grenzen über die Europa-Domäne hinausreichen, wurde der Flächenanteil innerhalb der betrachteten Domäne geschätzt. Die zugrundeliegenden Zahlenwerte können auch den Tabellen B.1, B.2 und B.3 im Anhang entnommen werden, in denen eine zusätzliche Unterscheidung zwischen Haupt- und Nebenterminen vorgenommen wird.

Aus Abbildung 7.2 geht hervor, dass die Meldungsdichte in Zentraleuropa, Großbritannien und Skandinavien im Vergleich zu den südlichen und östlichen Bereichen der betrachteten Europa-Domäne deutlich höher ist und in der Schweiz, den Niederlanden und in Österreich ein Maximum einnimmt. Auch Deutschland, Großbritannien und Norwegen weisen eine zufriedenstellende Anzahl an Meldungen auf. Die Qualität dieser Meldungen wird im folgenden Abschnitt 7.2.1 behandelt. Der durchschnittliche Stationsabstand (Median der Abstände zwischen den verbundenen Stationen) beträgt in Österreich und der Schweiz ca. 30 km, in Deutschland ca. 50 km, in Frankreich ca. 70 km, in Italien ca. 75 km und in Spanien bereits 126 km. Für die gesamte Domäne kann ein medianer Stationsabstand von 57 km errechnet werden.



**Abbildung 7.2:** Durchschnittliche Anzahl (gerundet) der Stationen, von denen im Jahr 2010 zu den stündlichen Terminen eine Meldung vorliegt. Die Zählung erfolgte, sobald für einen Parameter eine Meldung zur Verfügung steht, ungeachtet der Qualität der Meldung. Die Zahlenwerte innerhalb der Ländergrenzen geben die Anzahl der durchschnittlichen stündlichen Meldungen für das jeweilige Land an. Die Farbgebung bezieht sich auf die Anzahl der durchschnittlichen Meldungen in einem Land, bezogen auf eine Fläche von 100 000 km<sup>2</sup>. Für Länder, deren Grenzen über die Europa-Domäne hinausreichen, wurde der Flächenanteil innerhalb der betrachteten Domäne geschätzt (Tabelle B.3 enthält die zugrundeliegenden Zahlenwerte). Zusätzlich ist in Blau die operationelle Europa-Domäne von VERA eingezeichnet.

## 7.2 Beurteilung der Qualität der Stationsmeldungen durch die VERA-QC

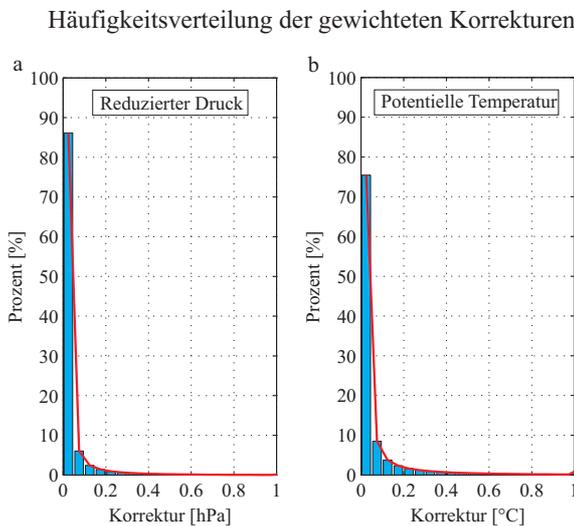
Eine ausgereifte Qualitätskontrolle kann Stationsmeldungen akzeptieren, verwerfen oder auch korrigieren. Welche Entscheidung für eine Meldung getroffen wird, hängt nicht nur von der Methode der Qualitätskontrolle ab, sondern auch von den benutzerdefinierten Einstellungen. Im Anhang A.1 sind die Konfigurationseinstellungen der operationellen VERA-QC angeführt, die auch zu den im Folgenden präsentierten Ergebnissen geführt haben.

**Tabelle 7.1:** Prozentuelle Aufschlüsselung der Stationsmeldungen nach Parameter und Beurteilung durch die VERA-QC. Die Prozentangaben beziehen sich auf die Anzahl aller, zur Verfügung stehender Meldungen je Parameter im Jahr 2010 (Niederschlag ( $NS$ ): 1 811 969, Windkomponente ( $u$ ): 10 830 688, Windkomponente ( $v$ ): 10 830 688, potentielle Temperatur ( $\theta$ ): 10 512 621, äquivalentpotentielle Temperatur ( $\theta_e$ ): 10 484 136, reduzierter Druck ( $p_{red}$ ): 10 210 115 und Windgeschwindigkeit ( $|\vec{v}|$ ): 10 830 688).

	$NS$	$u$	$v$	$\theta$	$\theta_e$	$p_{red}$	$ \vec{v} $
akzeptiert [%]	99.13	94.57	94.71	96.71	93.06	17.07	94.49
korrigiert [%]	$\sim 0$	4.76	4.57	2.87	6.23	82.24	4.80
verworfen [%]	0.87	0.67	0.72	0.42	0.70	0.69	0.71
Gross Error [%]	0.87	0.55	0.60	0.37	0.65	0.33	0.59
Cluster, gesamt [%]	2.51	4.66	4.66	4.73	4.73	4.59	4.66
Cluster, korrigiert [%]	0.00	0.85	0.82	0.49	1.19	4.59	0.89
Cluster, nicht korr. [%]	2.51	3.81	3.85	4.23	3.54	0.00	3.77
unphysikalisch [%]	$\sim 0$	$\sim 0$	$\sim 0$				
nicht verbunden [%]	$\sim 0$	$\sim 0$	$\sim 0$	0.00	0.00	0.00	$\sim 0$
melden zu selten [%]	0.00	0.12	0.12	0.05	0.05	0.35	0.12

In der Tabelle 7.1 sind die Entscheidungen, welche die VERA-QC für die Stationsmeldungen im Jahr 2010 getroffen hat, prozentuell aufgeschlüsselt und nach Parameter sortiert. Man erkennt, dass weit über 90 % aller Stationsmeldungen von der VERA-QC unkorrigiert akzeptiert wurden, mit Ausnahme der Meldungen für den reduzierten Bodendruck. Der Grund dafür liegt in einer benutzerdefinierten Einstellung. Korrekturen werden nur angebracht, wenn sie diese parameterabhängige fixe Schwelle überschreiten. Für den reduzierten Bodendruck liegt die Schwelle bei 0 hPa, für die potentielle und äquivalentpotentielle Temperatur und den Wind bei 1 °C bzw. 1 m/s und für den 6h-Niederschlag bei 50 mm (vgl. Anhang A.1).

Für den Niederschlag bewirkt diese Einstellung, dass die Qualitätskontrolle im Wesentlichen auf die Gross Error Erkennung reduziert wird. Obwohl für den reduzierten Druck alle Korrekturen unabhängig ihres Betrages angebracht werden, bleiben 17.07% unkorrigiert. Der Grund dafür liegt in der Gewichtung mit der Kostenfunktionsreduktion. Der erste Korrekturvorschlag bzw. die ungewichteten Deviations werden global berechnet und dann mit der Kostenfunktionsreduktion gewichtet. Diese wird lokal unter Anbringung der global ermittelten Korrekturen berechnet und somit besteht die Möglichkeit, dass negative Werte für die Kostenfunktionsreduktionen resultieren. Diese werden dann auf Null gesetzt und verursachen somit die 17.07% unkorrigierte Meldungen.



**Abbildung 7.3:** Häufigkeitsverteilung der berechneten Korrekturen für alle Stationsmeldungen im Jahr 2010 für die Parameter (a) reduzierter Druck und (b) potentielle Temperatur.

Auch wenn über 80% der Stationsmeldungen für den reduzierten Druck im Jahr 2010 korrigiert wurden, handelt es sich dabei lediglich um sehr geringe Korrekturen in der Größenordnung von 0.05 hPa, die der Unschärfe von Druckmeldungen entsprechen (Angaben von gemessenen Druckwerten werden auf 1/10 hPa gerundet). In Abbildung 7.3 sind die Häufigkeitsverteilungen der berechneten gewichteten Korrekturen für die Parameter reduzierter Druck und potentielle Temperatur dargestellt. Auffällig ist der hohe Anteil sehr geringer Korrekturen im Zehntel hPa- bzw. °C-Bereich.

Im Falle der potentiellen Temperatur wurden die dargestellten Korrekturen nicht angebracht. (Bemerkung: Es treten auch noch höhere Korrekturen als in Abbildung 7.3 dargestellt auf, deren Anzahl geht jedoch in dieser Darstellung verloren. Ebenso werden hier die Korrekturvorschläge für Gross Errors nicht berücksichtigt.)

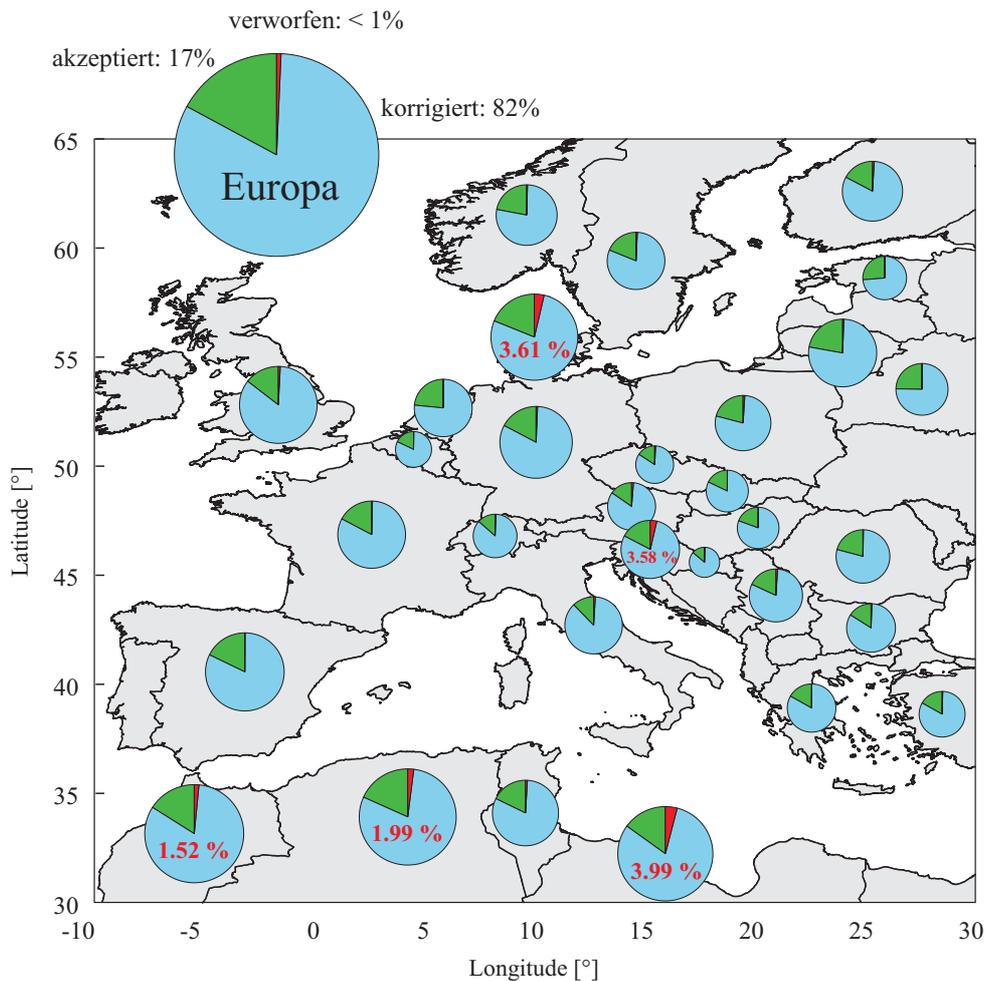
Der geringe Anteil an Stationsmeldungen der verworfen werden muss, beträgt erfreulicherweise für alle betrachteten Parameter unter 1%. Dabei handelt es sich zum einen um Meldungen die als Gross Error identifiziert wurden, zum anderen um Beobachtungen von Stationen, die das Kriterium der Mindestanzahl an Meldungen nicht erfüllen. Die Berechnung einer Bias-Korrektur ist an die

Bedingung gebunden, dass von einer Station im Bias-Berechnungszeitraum (z.B.: 30 Tage) eine Mindestanzahl an Meldungen (z.B.: 30 Meldungen) für den jeweiligen Parameter vorliegen muss. Kann dieses Kriterium nicht erfüllt werden, so wird keine Bias-Korrektur berechnet. Die betroffene Station wird solange von einer weiterführenden Analyse ausgeschlossen (von der VERA-QC verworfen), bis wieder regelmäßig genügend Meldungen vorliegen und eine statistische Beurteilung der Station, nämlich eine Bias-Korrektur, berechnet werden kann. Stationsmeldungen können aber auch verworfen werden, wenn unphysikalische Werte gemeldet werden, oder die Station nicht mit Nachbarn verbunden werden konnte. Die beiden letzten Gründe sind jedoch zweitrangig einzustufen. Im betrachteten Zeitraum sind nur 39 Druckwerte als unphysikalisch identifiziert worden. Der Fall, dass für eine Station keine Nachbarn gefunden wurden, ist nie aufgetreten.

Weiterführend wurde untersucht, ob länderspezifische Unterschiede bezüglich des Verhältnisses von akzeptierten, korrigierten und verworfenen Stationsmeldungen existieren. In Abbildung 7.4 sind zu diesem Zweck Kreisdiagramme, die dieses Verhältnis für den Parameter reduzierter Druck veranschaulichen, für Länder mit durchschnittlich mindestens 10 Stationsmeldungen pro Analysetermin dargestellt. Das Kreisdiagramm in der linken oberen Ecke präsentiert die Daten für die gesamte Europa-Domäne, wobei der blaue Kreisausschnitt jeweils den Anteil an korrigierten Meldungen, der grüne Ausschnitt jenen der akzeptierten und der rote den Anteil der verworfenen Meldungen symbolisiert.

Man kann erkennen, dass es kaum länderspezifische Unterschiede im Verhältnis von akzeptierten, korrigierten und verworfenen Stationsmeldungen gibt. Lediglich für die Stationsmeldungen von Dänemark, Slowenien, Marokko, Algerien und Libyen ist der Anteil an verworfenen Meldungen erhöht ( $>1.5\%$ ). Dies ist hauptsächlich auf zwei Ursachen zurückzuführen: In Dänemark, Slowenien und Libyen werden viele Meldungen verworfen, weil die betroffenen Stationen zu selten melden. In Marokko und Algerien hingegen werden vermehrt Gross Errors gefunden und die Stationsmeldungen aus diesem Grund verworfen. Die in Abbildung 7.4 dargestellten Daten sind im Anhang B.2 in Tabellenform beigelegt. Zusätzlich liegt auch dieselbe Auswertung für die Parameter potentielle Temperatur und Windgeschwindigkeit vor. Wie beim reduzierten Druck können auch bei den anderen Parametern keine länderspezifischen Unterschiede im Verhältnis der akzeptierten zu den korrigierten und zu den verworfenen Stationsmeldungen gefunden werden.

Den größten Anteil an den verworfenen Stationsmeldungen machen die erkannten Gross Errors aus (nur beim reduzierten Bodendruck gehen etwas mehr Meldungen

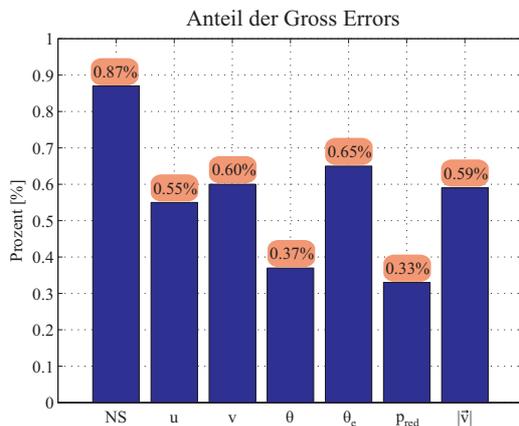


**Abbildung 7.4:** Kreisdiagramme der Anteile an akzeptierten (grün), korrigierten (blau) und verworfenen (rot) Stationsmeldungen des Parameters reduzierter Druck im Jahr 2010; Die Qualitätskontrolle wurde mit allen Erweiterungen (inklusive Bias-Korrektur) auf die vorliegenden Stationsmeldungen angewendet. Man beachte, dass es sich bei den angebrachten Korrekturen überwiegend um Werte der Größenordnung von 0.05 hPa handelt (vergleiche Abbildung 7.3). Die Darstellung erfolgt für Länder in der betrachteten Domäne, von denen im Zeitraum 2010 durchschnittlich mindestens 10 Stationsmeldungen pro Analysetermin vorliegen. Für Länder, in denen mehr als 1.5% der Meldungen verworfen werden müssen, ist dieser Anteil in roten Ziffern dem Kreisdiagramm beigelegt. Die Größe der Kreisdiagramme ist willkürlich gewählt um eine bessere Lesbarkeit zu gewährleisten.

verloren, weil Stationen zu selten melden, vergleiche Tabelle 7.1).

In Abbildung 7.5 sind die Anteile der erkannten Gross Errors an allen zur Verfügung stehenden Stationsmeldungen im Jahr 2010 nach Parameter aufgeschlüsselt dargestellt. Das Charakteristikum, dass Gross Errors nur selten auftreten, spiegelt sich in ihrem geringen Anteil < 1% wider. Die zweite Bedingung, dass Gross Errors zufällig auftreten, ist für die von der VERA-QC als Gross Errors identifizierten

Meldungen ebenso erfüllt. Seit der Implementierung der Bias-Korrektur werden zeitlich und räumlich persistente Korrekturvorschläge mit hoher Amplitude nicht mehr als Gross Errors markiert und verworfen, sondern durch die Bias-Korrektur ausgebessert.



**Abbildung 7.5:** Anteile der erkannten Gross Errors bezüglich aller Stationsmeldungen im Jahr 2010 für die Parameter Niederschlag ( $NS$ ), Windkomponenten ( $u$ ,  $v$ ), potentielle und äquivalentpotentielle Temperatur ( $\theta$ ,  $\theta_e$ ), reduzierter Bodendruck ( $p_{red}$ ) und Windgeschwindigkeit ( $|\vec{v}|$ ).

Im Anhang B.3 sind Tabellen angeführt, in denen die Anzahl und der prozentuelle Anteil der erkannten Gross Errors, aufgeschlüsselt nach Ländern, für die Parameter reduzierter Druck, potentielle- und äquivalentpotentielle Temperatur angegeben sind. Anhand dieser Angaben kann man erkennen, dass eine leichte lokale Abhängigkeit der erkannten Gross Errors besteht. In Ländern wie Algerien, Marokko oder Tunesien werden geringfügig mehr Gross Errors gefunden, was aber auf die Lage am Rand der betrachteten Domäne zurückgeführt werden kann und zusätzlich durch die dort vorherrschende geringere Stationsdichte verstärkt wird.

### 7.2.1 Vergleich mit den Ergebnissen des Projektes ALPEX

Bei ALPEX (Alpine Experiment) handelt es sich um ein Feldexperiment, das 1982 im Rahmen von GARP (Global Atmospheric Research Programme) durchgeführt wurde. Während der Monate Mai und April 1982 wurden in mehreren Ländern Europas intensive synoptische Messungen durchgeführt (Messungen mit einem Forschungsflugzeug in der Schweiz bzw. Schiffen im Mittelmeer sowie vermehrt manuelle und automatische synoptische Bodenbeobachtungen und Radiosondenaufstiege). Das Hauptaugenmerk der Messkampagne lag auf der Erfassung meteorologischer Phänomene im alpinen Bereich und bildete die Grundlage für viele wissenschaftliche Arbeiten und Erkenntnisse im Gebiet der Gebirgsmeteorologie.

Am Institut für Meteorologie und Geophysik der Universität Innsbruck beschäftigte

sich 1988 ein Projekt mit der Beurteilung und Korrektur der im Zuge von ALPEX erhobenen synoptischen Beobachtungen. Eine Beschreibung und eine Dokumentation der Arbeit, sowie eine Zusammenfassung der Ergebnisse, liegen am Institut für Meteorologie und Geophysik der Universität Wien unter dem Titel „*The Quality of ALPEX Level-II B Synop and Rawin Data - A Documentation of Errors*“ vor. Die Autoren R. Steinacker, W. Klötzer und A. Lanzinger konzentrierten sich auf zwei Fehlerarten, zufällig verteilte und systematische Fehler. Die während der intensiven Messperiode im Mai und April 1982 erhobenen Temperatur- und Bodendruckbeobachtungen wurden manuell auf ihre dreidimensionale ( $x$ ,  $y$  und  $t$ ) Konsistenz geprüft und gegebenenfalls auch korrigiert. Der große Vorteil einer solchen arbeitsintensiven manuellen Qualitätskontrolle gegenüber einer automatischen liegt darin, dass der Synoptiker seine jahrelange Erfahrung beim räumlich- und zeitlichen Vergleich der Messwerte einbringen kann. Es ist jedoch zu beachten, dass die Entscheidungen, die dieser Synoptiker bei der Kontrolle der Daten trifft, natürlich subjektiv gefällt werden und von der Tendenz des Synoptikers entweder „nie einer Messung zu vertrauen“ oder „nie eine Messung zu verwerfen“ beeinflusst sind.

Bei der Kontrolle der ALPEX-Daten (Mai - Juni 1982) ist man auf erhebliche länderspezifische Unterschiede in der Qualität der europaweit gesammelten Messwerte gestoßen. Im Vergleich dazu haben die im Zuge dieser Diplomarbeit durchgeführten Auswertungen der europaweiten Messwerte aus dem Jahr 2010 jedoch keine großen Abweichungen in der Qualität der Beobachtungen der einzelnen Länder ergeben.

In Tabelle 7.2 sind exemplarisch der prozentuelle Anteil der (mittels manueller QC) erkannten Gross Errors in den ALPEX-Daten und der Anteil der (mittels VERA-QC) in den Daten des Jahres 2010 erkannten Gross Errors gegenübergestellt. Man muss beachten, dass in dieser Tabelle unterschiedliche Datengrundlagen eingeflossen sind. So umfasst der ALPEX-Datensatz die dreistündigen synoptischen Beobachtungen im Jahr 1982 (total 496 mögliche Messungen pro Station), die Datengrundlage dieser Diplomarbeit bilden die stündlichen SYNOP-Meldungen des gesamten Jahres 2010 (total 8760 mögliche Messungen pro Station). Ebenso unterscheiden sich die beiden betrachteten Europadomänen (die ALPEX-Domäne ist vor allem in ihrer West-Ost-Erstreckung kleiner). Nur die Anzahl der zur Verfügung stehenden Messungen (ca. 1000 - 1100 Stationen) ist vergleichbar geblieben. Auch die gegenübergestellten Parameter weichen voneinander ab. Von der VERA-QC werden operationell die potentielle Temperatur  $\Theta$  und der reduzierte Bodendruck  $p_{red}$  kontrolliert, bei der Qualitätskontrolle der ALPEX-Daten betrachtete man

spezielle abgeleitete Größen der Temperatur- und Bodendruckbeobachtungen. Den sogenannten D-Wert berechnet aus dem gemessenen Bodendruck,  $D_p$ , ermittelt man als Differenz der Stationshöhe (in geopotentiellen Metern) und der Höhe der Druckfläche des gemessenen Stationsdrucks bei Annahme der Standardatmosphäre. Analog dazu ist der D-Wert der Temperatur definiert als Differenz zwischen der virtuellen Temperatur an der Station und der Temperatur der Standardatmosphäre auf Stationshöhe.

Obwohl für Tabelle 7.2 unterschiedliche Datengrundlagen herangezogen wurden, kann ein Vergleich der prozentuellen Anteile der gefundenen Gross Errors an den jeweils pro Land zur Verfügung stehenden Beobachtungen durchgeführt werden. Die Reihung der Länder erfolgte nach der Anzahl der, in den Druckmeldungen aus dem Jahr 1982 gefundenen Gross Errors.

Betrachtet man die mittleren Anteile der gefundenen groben Fehler (vorletzte Zeile in Tabelle 7.2), so ist vor allem für den Parameter Druck zu erkennen, dass sich die Qualität der Beobachtungen bzw. das Vorkommen von zufällig verteilten groben Fehlern in den letzten Jahrzehnten deutlich verbessert hat. Ebenso sind die Qualitätsunterschiede der Messungen aus verschiedenen Ländern deutlich geringer geworden, was sich auch in einer geringeren Standardabweichung der Gross Error Anteile widerspiegelt (letzte Zeile in Tabelle 7.2).

Lediglich für zwei Länder konnten im Jahr 2010 mehr grobe Fehler in den Druckbeobachtungen gefunden werden, zum einen für Großbritannien und zum anderen für die Schweiz (fett markiert). Aber auch die Qualität der Temperaturmeldungen im Jahr 2010 wurde für diese beiden Länder schlechter beurteilt. Im Falle Großbritanniens muss man beachten, dass der ALPEX-Datensatz nur Stationsmeldungen aus dem südlichsten Teil Großbritanniens (19 Stationen meldeten Druckbeobachtungen, 27 Stationen Temperaturwerte) berücksichtigt. Im Jahr 2010 wurden jedoch durchschnittlich 165 SYNOP-Meldungen aus dem gesamten Vereinigten Königreich kontrolliert. Es ist somit zu hinterfragen, in welchem Ausmaß die beiden Datensätze vergleichbar sind. Für die Schweiz hat sich die Anzahl der untersuchten Beobachtungen verdoppelt, doch signifikante Qualitätsunterschiede ergeben sich nur für die Temperaturwerte.

Auch in der mittleren Anzahl der gefundenen Gross Errors für die Temperaturmessungen ist beim Vergleich der Daten aus 1982 und jenen aus 2010 im Mittel nur eine geringe Verbesserung ersichtlich, einzelne Länder wurden 1982 sogar besser beurteilt. Dieser Unterschied kann eventuell mit den verwendeten Qualitätskontrollmethoden erklärt werden. Bei extremen Wetterlagen, wie zum Beispiel bei Föhn- oder Inversionswetterlagen, können in den Temperaturwerten an Stationen

im Gebirge und an Stationen in der Ebene oder an benachbarten Stationen in verschiedenen Höhenlagen sehr große Unterschiede entstehen. Auch eine ausgereifte automatische Qualitätskontrolle kann diese korrekten Temperaturunterschiede in komplexem Gelände nicht immer richtig einstufen und identifiziert fälschlicherweise Gross Errors. Ein Synoptiker hingegen kann mit seiner Erfahrung die Kriterien für eine Gross Error Erkennung je nach Wetterlage und Orographie anpassen und Stationsmeldungen besser beurteilen.

Zusammenfassend kann festgestellt werden, dass sich die Qualität der synoptischen Beobachtungen in Europa (beurteilt durch die Anzahl der identifizierten Gross Errors) von 1982 bis 2010 verbessert hat und die länderspezifischen Unterschiede in den Messwerten abgenommen haben. Diese zwei Aussagen spiegeln sich vor allem in der Betrachtung der Bodendruckmessungen wieder.

**Tabelle 7.2:** Gegenüberstellung der im Zuge von ALPEX (März - April 1982) und im Zuge dieser Diplomarbeit (2010) identifizierten Gross Errors für Länder in Europa. (weitere Details siehe Text)

Länder	Gross Error [%]			
	$D_p$ (1982)	$p_{pred}$ (2010)	$D_T$ (1982)	$\Theta$ (2010)
Rumänien (ROU)	5.61	0.26	0.57	0.59
Tschechoslowakei (1982)	5.26		0.27	
Tschechische Republik (CZE)		0.15		0.32
Slowakei (SVK)		0.06		0.15
Frankreich - Automaten (1982)	3.04		0.76	
Frankreich (FRA)		0.29		0.26
Jugoslawien (1982)	2.48		0.65	
Slowenien (SVN)		0.12		0.08
Kroatien (HRV)		0.38		0.58
Bosnien und Herzegowina (BIH)		0.27		0.18
Montenegro (MNE)		0.05		0.18
Serbien (SRB)		0.23		0.23
Mazedonien (MKD)		0.29		0.11
Albanien (ALB)	2.39	0.01	0.46	0.04
Bulgarien (BGR)	1.56	0.31	0.51	0.41
Italien (ITA)	1.56	0.42	0.42	0.63
Belgien (BEL)	1.54	0.24	0.07	0.30
Türkei (TUR)	1.51	0.36	0.45	0.33
Sowjetunion (1982)	1.33		0.46	
Russische Föderation (RUS)		0.08		0.13
Spanien (ESP)	1.26	0.05	0.32	0.34
Polen (POL)	1.24	0.30	0.25	0.28
Griechenland (GRC)	1.16	0.16	0.28	0.19
Österreich (AUT)	0.93	0.22	0.13	0.27
Ungarn (HUN)	0.82	0.03	0.19	0.04
Frankreich - manuell (1982)	0.81		0.01	
DDR (1982)	0.58		0.15	
Deutschland (DEU)		0.45		0.33
<b>Vereinigtes Königreich (GBR)</b>	<b>0.21</b>	<b>0.62</b>	<b>0.01</b>	<b>0.41</b>
Luxemburg (LUX)	0.21	0.21	~0	~0
Deutschland (1982)	0.19		0.03	
Niederlande (NLD)	0.18	0.04	0.05	0.02
<b>Schweiz (CHE)</b>	<b>0.06</b>	<b>0.15</b>	<b>0.03</b>	<b>0.42</b>
Mittelwert	1.54	0.21	0.29	0.26
Standardabweichung	1.49	0.16	0.22	0.17

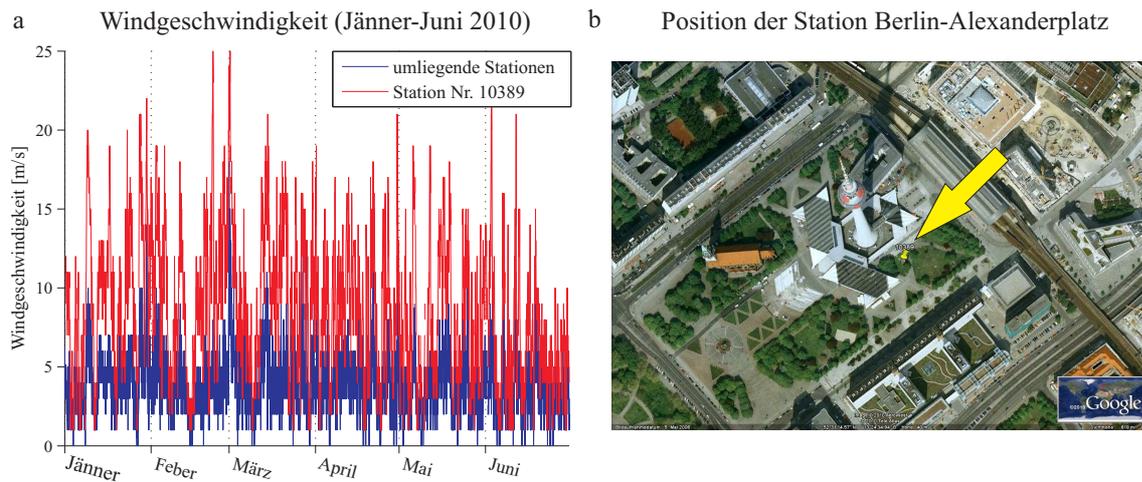
# Kapitel 8

## Statistische Auswertung der Deviations aus dem Jahr 2010

Sammelt man die von der VERA-QC berechneten Korrekturen, so können, wie bereits in Kapitel 4 beschrieben, systematische Fehler mit Hilfe einer Bias-Korrektur behoben werden. Die gesammelten Deviations bieten jedoch noch weiteren Informationsgehalt, der auf den ersten Blick nicht offensichtlich erscheint. Im vorliegenden Kapitel werden einige Ergebnisse aus statistischen Untersuchungen der gesammelten Deviations präsentiert.

### 8.1 Erkennung falscher SYNOP-Verschlüsselungen

In Kapitel 3 Abschnitt 3.3.4 wurde die Bias-Korrektur der Windgeschwindigkeit für Europa in Abbildung 3.7 dargestellt. Man erkennt eine Station im Nordosten Deutschlands, die durchwegs hohe negative Korrekturen erhält, woraus sich im April 2010 eine Bias-Korrektur von -4 m/s ergibt. Der Grund für die negativen Deviations könnte in den physikalischen Eigenschaften des gemessenen Parameters liegen. So werden Meldungen von Küstenstationen, die höhere Windgeschwindigkeiten aufweisen, von VERA-QC tendenziell nach unten korrigiert, Stationsmeldungen im Landesinneren hingegen nach oben. Ebenso erhalten Bergstationen, die aufgrund ihrer exponierten Lage stärkeren Windgeschwindigkeiten ausgesetzt sind, vermehrt negative Deviations. Da die betrachtete Station in Deutschland jedoch im Landesinneren auf 37 Meter Seehöhe liegt, kann keiner der beiden Gründe als Erklärung für die negativen Korrekturen in Frage kommen.



**Abbildung 8.1:** (a): gemeldete Windgeschwindigkeiten [m/s] der Station 10389 (rot) sowie für vier umliegende Stationen: 10381, 10382, 10384, 10385 (blau) im ersten Halbjahr 2010. (b): unmittelbare Umgebung der Station Berlin - Alexanderplatz (WMO-Nr.: 10389); [Quelle: Google Inc. (2011), Google Earth am 28. 1. 2011]

Es handelt sich bei der betrachteten Station um die Station Berlin - Alexanderplatz (WMO-Stationsnummer: 10389), deren Standort in Abbildung 8.1b dargestellt ist. Weiters vergleicht Abbildung 8.1a die gemeldeten Windgeschwindigkeiten der Station Berlin - Alexanderplatz (rot) und jene von vier in einem Umkreis von 17 km liegenden benachbarten Stationen (WMO-Nr.: 10381, 10382, 10384, 10385, blau). Die Station 10389 liefert durchwegs höhere Messwerte als die vier umliegenden, die sich auf freien Plätzen wie Wiesen oder Rollfeldern befinden. Eine beispielhaft angeführte SYNOP-Meldung der Station Berlin - Alexanderplatz vom 11. 4. 2010, 12:00 UTC zeigt, dass die Windgeschwindigkeit, wie auch die von den umliegenden Stationen, in der Einheit [m/s] verschlüsselt wird:

*AAXX 11121 10389 16 /// /0305 10071 20015 30163 40265 50002 60001*

(1: Meter pro Sekunde – mit Anemometer gemessen)

[Quelle: <http://www.ogimet.com/home.phtml.en>, am 10. 4. 2011]

Nimmt man jedoch an, dass die Station 10389 ihre Windgeschwindigkeit in der Einheit Knoten [kn] angibt und man die Messwerte mit dem Verhältnis  $1 \text{ kn} \approx 0.514 \text{ m/s}$  in eine Angabe der Einheit [m/s] umrechnet, wird eine sehr gute Übereinstimmung mit den umliegenden Stationsmeldungen erzielt. Die Deviations können somit helfen, Fehler bei der Verschlüsselung der Daten aufzudecken.

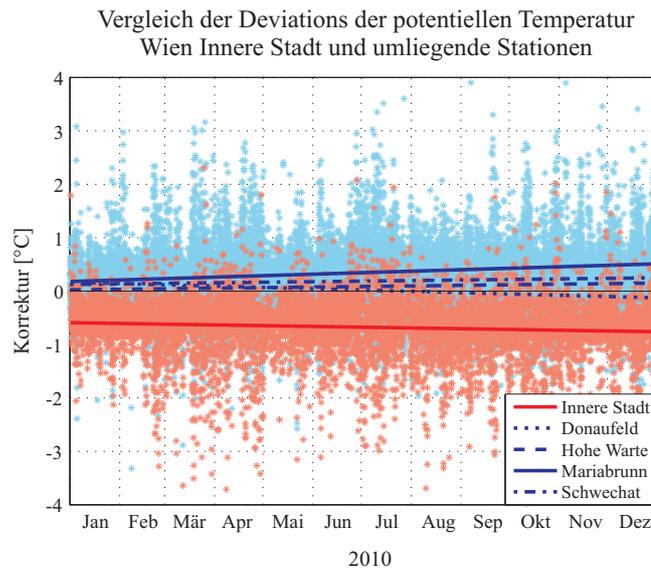
## 8.2 Erkennung falscher Einträge in der Stationsliste

Ein weiteres Beispiel, welche Zusatzinformationen Deviations bzw. Bias-Korrekturen enthalten können, wurde bereits in Kapitel 3 gegeben. Die gemeldeten Druckwerte werden mit den Höhenangaben in der benutzten Stationsliste auf Meeressniveau reduziert und anschließend der VERA-QC unterzogen. Für einige Stationen resultiert eine Bias-Korrektur des reduzierten Drucks, deren Ursache bereits auf falsche Höhenangaben in den verwendeten Stationsinformationen zurückgeführt werden konnte. Somit tragen die Deviations zur ständigen Verbesserung und Kontrolle der benutzten Stationsliste bei.

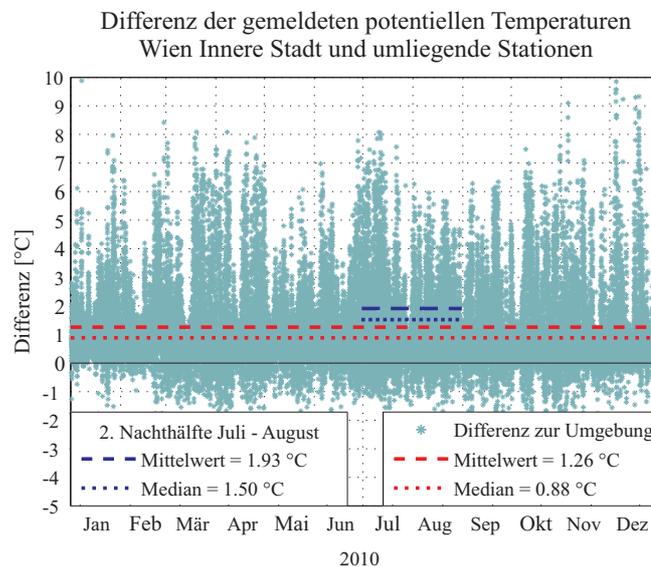
## 8.3 Erkennung mikrometeorologischer Phänomene

Wie bereits im Kapitel 1 erwähnt, kann ein synoptisches Messnetz nur jene meteorologischen Phänomene gut auflösen, deren räumliche (und zeitliche) Größenskala größer als der mittlere Stationsabstand (bzw. das mittlere Messintervall) ist. Kleinskaligere Vorgänge resultieren in einem Repräsentativitätsfehler. Ein Beispiel für ein, bezogen auf das vorliegende GTS-Messnetz kleinskaliges Phänomen, ist die städtische Überwärmung bzw. städtische Wärmeinsel (englisch: *urban heat island UHI*). Im Vergleich zur Umgebung erwärmen sich bebaute Gebiete stärker und können diese Energie auch länger speichern. Eine ausführliche Behandlung des Stadtklimas kann in Hupfer und Kuttler (2006) gefunden werden.

Im Abbildung 8.2 sind die Deviations der potentiellen Temperatur, berechnet für das Jahr 2010 für die Stationen Wien - Innere Stadt (WMO-Nr.: 11034) in Rot und die Deviations für die umliegenden Stationen Donauefeld (WMO-Nr.: 11090), Hohe Warte (WMO-Nr.: 11035), Mariabrunn (WMO-Nr.: 11080) und Wien - Schwechat (WMO-Nr.: 11036) in Blau dargestellt. Die Station Wien - Innere Stadt befindet sich im Zentrum der Stadt Wien, für die umliegenden Stationen hingegen lässt der urbane Einfluss bereits nach. Man erkennt, dass für die Station Wien - Innere Stadt durchwegs negative Korrekturen berechnet werden (rot), die Deviations für die umliegenden Stationen (blau) sind eher um Null verteilt. Dieser systematische Repräsentativitätsfehler der Station Innere Stadt wird von der Qualitätskontrolle richtig erkannt und korrigiert.



**Abbildung 8.2:** Darstellung der Deviations für die Station Wien - Innere Stadt als rote und für umliegende Stationen (Donauefeld, Hohe Warte, Mariabrunn und Schwechat) als blaue Sterne. Die jeweils roten und blauen Linien stellen die Regression durch die, für die einzelnen Stationen berechneten Korrekturen dar. Die Station Wien Innere Stadt erhält meist negative Korrekturen.



**Abbildung 8.3:** Durch grüne Sterne ist die Differenz zwischen den gemeldeten potentiellen Temperaturen der Station Wien - Innere Stadt und jenen der umliegenden Stationen (Donauefeld, Hohe Warte, Mariabrunn und Schwechat) dargestellt. In Rot ist zusätzlich der Mittelwert (strichliert) und der Median (punktiert) der Differenzen eingezeichnet. In Blau ist der Mittelwert (strichliert) und der Median (punktiert) der Differenzen in der jeweils zweiten Nachthälfte (00:00–06:00 UTC) der Monate Juli und August dargestellt.

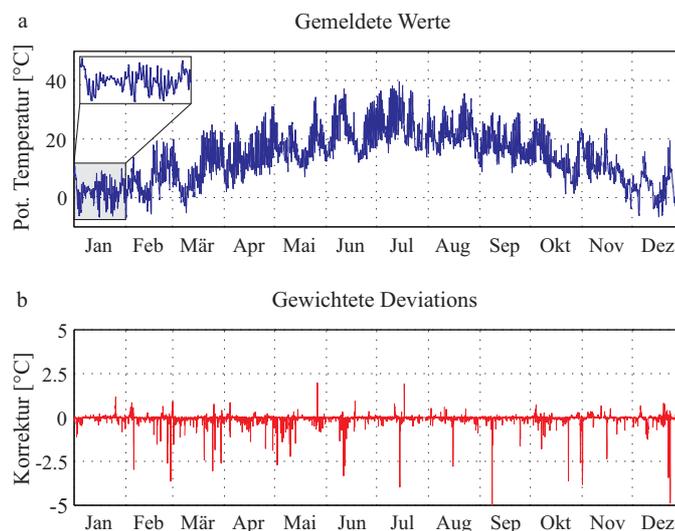
Um zu überprüfen, ob diese Deviations gerechtfertigt sind, werden in Abbildung 8.3 die Differenzen zwischen den gemeldeten potentiellen Temperaturen der Station Wien - Innere Stadt zu denen der umliegenden Stationen in Grün dargestellt. Positive Differenzen sagen aus, dass die Station in der Stadt höhere Werte gemeldet hat als die jeweilige Vergleichsstation in der Umgebung.

Mittelwert und Median der Differenzen geben an, dass die Station im Inneren der Stadt durchwegs um 1 °C höhere potentielle Temperaturen meldet. Diese Differenz wird auf den städtischen Wärmeinsel-Effekt zurückgeführt. In Hupfer und Kuttler (2006) ist beschrieben, dass das Maximum dieser Temperaturabweichung in den zweiten Nachthälften der Sommermonate Juli und August auftritt. Auch in den vorliegenden Daten aus dem Jahr 2010 können, bei ausschließlicher Betrachtung der Termine von 00:00–06:00 UTC im Juli und August, größere Differenzen zu den umliegenden Stationen gefunden werden. In Abbildung 8.3 ist dieses Ergebnis durch blaue Linien gekennzeichnet.

## 8.4 Analyse auf Periodizitäten

Betrachtet man die potentiellen Temperaturwerte aus dem Jahr 2010 der Station Innsbruck - Flughafen (Abbildung 8.4a), so ist der Jahrgang der Werte deutlich abgebildet. Ein Tagesgang ist ebenfalls vorhanden. Um diesen besser erkennbar zu

Station Innsbruck-Flughafen (Stationsnummer 11120), Jahr 2010



**Abbildung 8.4:** Zeitserien der Station Innsbruck - Flughafen (WMO-Stationsnummer 11120) des Jahres 2010. (a): gemeldete potentielle Temperatur. Der grau hinterlegte Bereich kennzeichnet die Daten vom Januar 2010 und ist zusätzlich gestreckt dargestellt, um den Tagesgang erkennbar darzustellen. (b): berechnete gewichtete Korrektur.

machen, sind die Werte vom Jänner 2010 (grau hinterlegt) auch gestreckt dargestellt, wodurch die Abfolge von höheren Temperaturen am Tag und niedrigeren in der Nacht erkennbar wird.

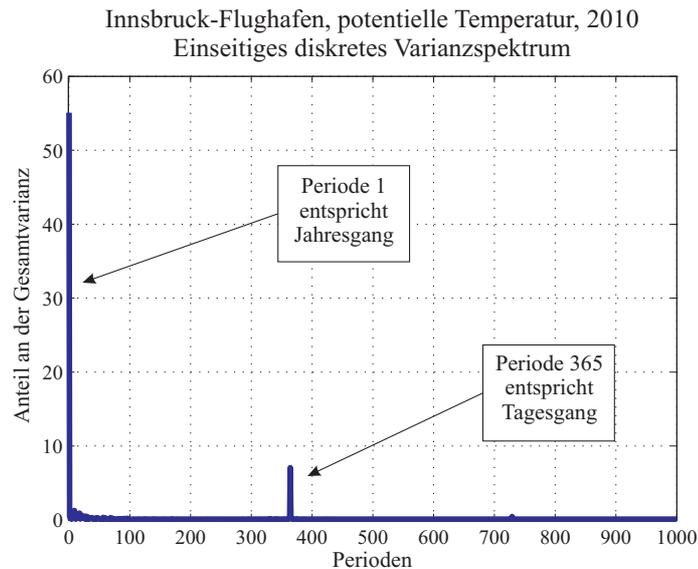
Um zu überprüfen, ob in der Zeitserie der gemeldeten Werte bzw. der Deviations auch noch weitere Prozesse mit unterschiedlichen Periodizitäten beteiligt sind, wird eine Fourier - Analyse durchgeführt. Der im Folgenden verwendete Formalismus ist im zweiten Kapitel in Broughton und Bryan (2009) beschrieben, für die diskrete Fouriertransformation wurde die in Matlab implementierte Funktion  $Y = \text{fft}(X)$  verwendet. Man beachte, dass bei dieser Implementierung der Vorwärtstransformation im Gegensatz zur üblichen Notation der Faktor  $1/N$  ( $N$  bezeichnet die Anzahl der Stützstellen) nicht vormultipliziert wird. Beim Erstellen eines Varianzspektrums (englisch: *power spectrum*) müssen dementsprechend die berechneten Fourierkoeffizienten  $c_k$  mit  $1/N^2$  multipliziert werden.

Im Falle unserer diskreten Zeitserie der potentiellen Temperatur  $f$ , die an  $N$  äquidistanten Stützstellen (stündliche Messungen im Jahr 2010,  $N = 24 * 365$ ) abgegriffen wurde, lassen sich die  $k = 0 \dots N - 1$  Fourierkoeffizienten  $c_k$  über

$$c_k = \frac{1}{N} \sum_{j=0}^{N-1} f_j e^{-ikj2\pi/N} \quad (8.1)$$

ermitteln. Der erste quadrierte Koeffizient  $c_0^2$  stellt den quadrierten Mittelwert der Zeitserie dar. Die weiteren Koeffizienten werden mit sich selbst bzw. mit den konjugiert komplexen Koeffizienten  $c_k^*$  multipliziert ( $c_k c_k^*$ ) und geben an, wie viel ein Prozess der Wellenzahl  $k$  zur Gesamtvarianz beiträgt.

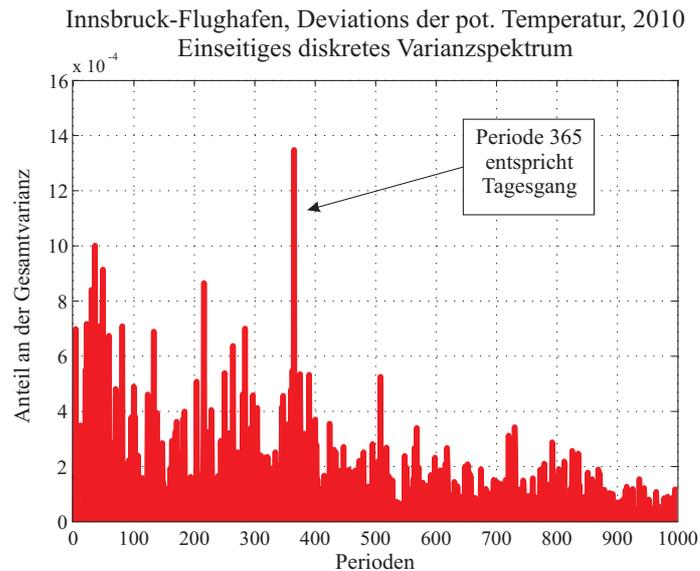
In Abbildung 8.5 ist die Energie  $c_k c_k^*$  gegen die Wellenzahl in einem sogenannten Varianzspektrum aufgetragen. Man beachte, dass es sich hierbei um ein mit zwei multipliziertes einseitiges Varianzspektrum handelt und auf die redundante zweite Hälfte verzichtet wurde. Zusätzlich ist nur ein relevanter Ausschnitt, der die ersten 1000 Perioden abdeckt, dargestellt. Auf der x-Achse ist die Anzahl der Perioden im betrachteten Zeitintervall angegeben. Zwei Perioden bzw. Prozesse tragen wesentlich dazu bei, die Gesamtvarianz der betrachteten Zeitserie zu erklären. Hohe Energien werden den Perioden 1 und 365 zugewiesen. Das Ereignis mit einer Periode im betrachteten Zeitintervall ist der Jahresgang. Das zweite Ereignis, das sich 365 mal im Jahr 2010 wiederholt hat, ist der Tagesgang. Beide Phänomene konnte man auch beim Betrachten der Zeitserie der potentiellen Temperatur (Abbildung 8.4a) erkennen.



**Abbildung 8.5:** Dargestellt ist ein einseitiges diskretes Varianzspektrum bis zur Periode 1000 für die potentielle Temperatur, die im Jahr 2010 von der Station Innsbruck - Flughafen gemeldet wurde. Der Jahres- und der Tagesgang der Messwerte tragen wesentlich dazu bei, die Gesamtvarianz der Zeitserie zu erklären.

Für die Berechnung der Deviations wird die Differenz der potentiellen Temperatur zu jener der Standardatmosphäre in Stationshöhe (D-Wert) betrachtet. Die Zeitserie des D-Wertes ist ebenso einem Jahres- und Tagesgang unterworfen wie die potentielle Temperatur selbst. Betrachtet man jedoch die Zeitserie der Deviations (Abbildung 8.4b), so kann man keine auffälligen Periodizitäten erkennen. Auch für diese Werte wurde eine Fourier-Analyse durchgeführt, welche in Abbildung 8.6 dargestellt wird.

Auffällig ist, dass die Energie auf mehrere Perioden verteilt wird und somit die Varianz der Zeitserie der Deviations durch viele verschiedene Prozesse entsteht. Dennoch wird dem Prozess mit der Periode 365, dem Tagesgang, die meiste Energie zugewiesen. Physikalisch kann diese Erkenntnis so interpretiert werden, dass räumlich betrachtet, die Unterschiede in den D-Werten der einzelnen Stationen einem Tagesgang unterliegen. Stationen in Tälern, die der täglichen lokalen Überhitzung stärker ausgesetzt sind, werden mit Stationen in der Ebene oder an Küsten verglichen. Daraus resultieren Deviations, die ebenfalls einen Tagesgang aufweisen. Dieser ist in den Deviations der Stationen, die durch die lokale Überhitzung stärker betroffen sind und deren Nachbarstationen in der Ebene liegen, auch stärker ausgeprägt.



**Abbildung 8.6:** Dargestellt ist ein einseitiges diskretes Varianzspektrum für die berechneten Deviations des D-Wertes der potentiellen Temperatur, bezogen auf die Werte, die im Jahr 2010 von der Station Innsbruck - Flughafen gemeldet wurden. Gezeigt werden nur die ersten 1000 Perioden.

Die Fourier-Analyse wurde für mehrere Stationen durchgeführt, die nach ihrer Lage in Tälern, auf Hochebenen, in der Ebene oder an der Küste unterteilt wurden. Natürlich sind die berechneten Deviations nicht nur von den lokalen orographischen Gegebenheiten abhängig, sondern auch davon, welche Nachbarn für die betroffene Station gefunden wurden. Dennoch konnten folgende einheitliche Eigenschaften bezüglich der Deviations für die potentielle Temperatur erkannt werden:

- Stationen in der Ebene weisen keine ausgeprägten Periodizitäten auf, lediglich der Jahresgang ist meistens erkennbar.
- Stationen an der Küste sind sowohl vom Jahres- als auch vom Tagesgang beeinflusst, wobei keine klare Aussage getroffen werden kann, welcher Prozess überwiegt. Des Weiteren ist noch eine 12-stündige Periodizität erkennbar.
- In den Deviations der Stationen der spanischen Hochebene sind sowohl Jahres- als auch Tagesgang bemerkbar, wobei ersterer deutlicher ausgeprägt ist. Neben einer 12-stündigen Periodizität tritt auch eine 8-stündige Schwingung hervor.
- An Stationen in Tälern ist der Tagesgang deutlicher erkennbar als der Jahresgang. Zusätzlich sind 12-, 8- und 6-stündige Periodizitäten vorhanden.

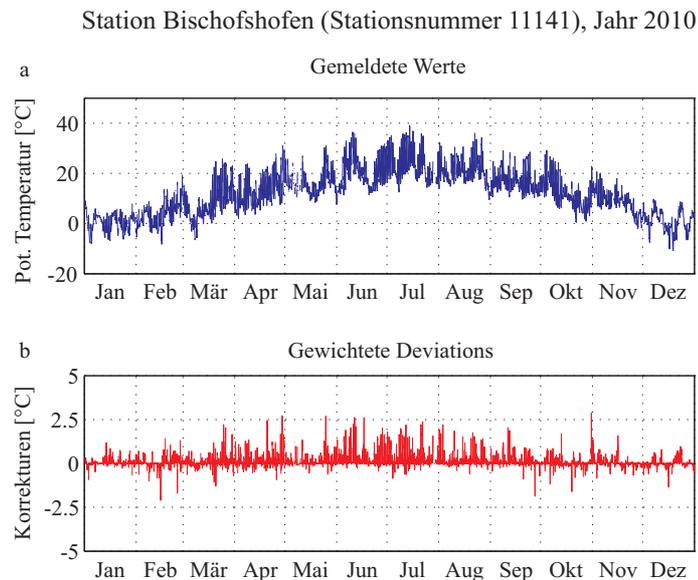
Diese Betrachtungen wurden mit vergleichbaren Ergebnissen auch für den Parameter reduzierter Druck durchgeführt. Der einzige nennenswerte Unterschied im

Vergleich zur potentiellen Temperatur ist, dass an den Stationen in den Tälern der Jahresgang des Drucks stärker ausgeprägt ist als der Tagesgang. Wie vermutet, hat die lokale Überhitzung am Tag bzw. die lokale Abkühlung in der Nacht auf den Druck einen geringeren Einfluss als die jahreszeitlichen Variationen der Druckregime.

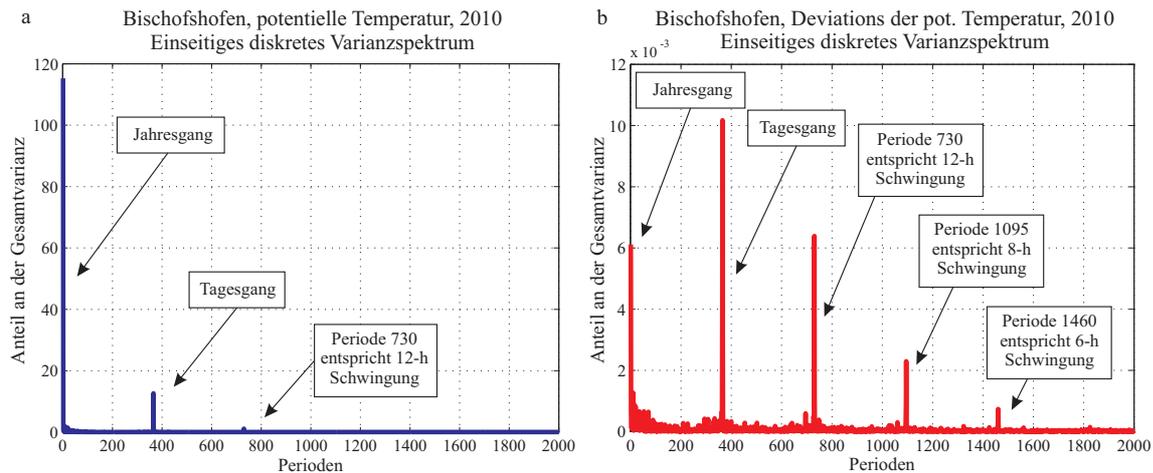
Die erwähnten 24-, 12-, 8- und 6-stündigen Oszillationen werden für Druckmessungen auch als „atmosphärische Gezeiten“ (englisch: *atmospheric tides*) bezeichnet, die bereits früh erkannt und von Haurwitz (1964) publiziert wurden. Diese werden zum größten Teil durch den täglichen Zyklus der Sonneneinstrahlung ausgelöst. Einen betragsmäßig geringeren Antrieb liefert auch die Gravitationskraft des Mondes.

Eine Station, die besonders stark ausgeprägte Periodizitäten in den berechneten Deviations für die potentielle Temperatur aufweist, ist die Station Bischofshofen (WMO-Nr.: 11141) auf 550 m. In Abbildung 8.7 sind die gemeldeten Werte und die gewichteten Deviations der potentiellen Temperatur für Bischofshofen dargestellt. Die einseitigen Varianzspektren für die gemeldeten Werte sowie für die daraus berechneten Deviations findet man in Abbildung 8.8.

Eine Erklärung für die erkannten 24-, 12-, 8-, und 6-h Schwingungen in den Deviations der Station Bischofshofen kann in der genauen Betrachtung der Lage der



**Abbildung 8.7:** (a): Zeitserie der potentiellen Temperatur der Station Bischofshofen (WMO-Stationsnummer 11141) aus dem Jahr 2010. (b): Zeitserie der gewichteten Deviations, berechnet für die potentielle Temperatur der Station Bischofshofen im Jahr 2010.



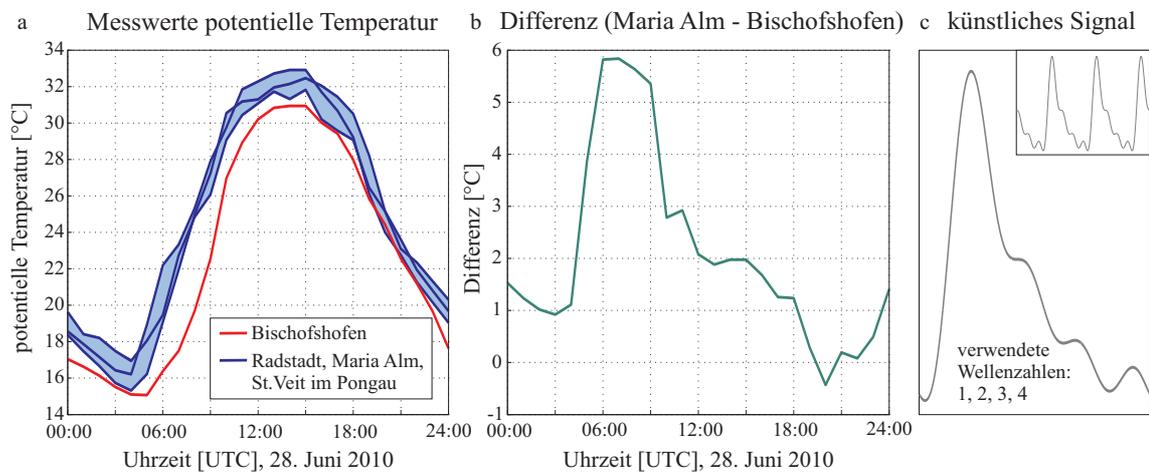
**Abbildung 8.8:** Einseitiges diskretes Varianzspektrum für die potentielle Temperatur (a) und für die daraus berechneten gewichteten Deviations (b) der Station Bischofshofen (WMO-Stationsnummer 11141) aus dem Jahr 2010.

Station 11141 selbst und der Lage ihrer umliegenden (primären) Nachbarstationen gefunden werden. Abbildung 8.9 soll helfen, diese Zusammenhänge anhand einer Fallstudie zu veranschaulichen. Für den Vergleich wurden die benachbarten Stationen Radstadt (WMO-Nr.: 11147), Maria Alm (WMO-Nr.: 11137) und St. Veit im Pongau (WMO-Nr.: 11345) herangezogen. Beachtenswert ist, dass diese drei Stationen in vergleichsweise breiten, West - Ost ausgerichteten Tälern aufgestellt sind, Bischofshofen hingegen im Nord - Süd gerichteten Salzachtal liegt. Nur ca. 200 m östlich der WMO-Messstation 11141, die nahe an der Salzach steht, beginnt der Anstieg eines Ausläufers des Hochgründecks bis auf über 1000 m. Die Station in Bischofshofen ist somit für die ersten möglichen Sonnenstrahlen abgeschattet und der morgendliche Anstieg der Temperaturen ist im Vergleich zu den umliegenden Stationen zeitlich nach hinten versetzt.

In Abbildung 8.9a sind die Werte der potentiellen Temperatur für die drei umliegenden Stationen in Blau und jene der Station Bischofshofen in Rot für den 28. Juni 2010 (sowohl in der Höhe als auch am Boden gradientschwache Lage, kein Niederschlag) dargestellt. Man erkennt deutlich die zeitliche Verschiebung des Temperaturanstiegs für die Station in Bischofshofen, die mit der Abschattung (verursacht vom Gebirge im Osten) erklärt werden kann. Die am Nachmittag beginnende Abkühlung setzt für alle vier Stationen etwa zur selben Zeit ein. In (b) wird die Differenz der potentiellen Temperatur an den Stationen Maria Alm und Bischofshofen gezeigt. Vor allem in den Morgenstunden und am Vormittag ist ein Ausschlag in den Differenzen (zeitliche Verschiebung des Temperaturanstiegs) erkennbar. Dieses Signal ist von

den geographischen Gegebenheiten bedingt und somit von systematischer Natur. In (c) wurde ein analytisches Signal mit den in den gewichteten Deviations gefundenen 24-, 12-, 8-, und 6-h Schwingungen (entsprechen den Wellenzahlen 1, 2, 3, 4 bei der Betrachtung eines Tages; Amplitude mit steigender Wellenzahl abfallend gewählt) erzeugt, welches eine sehr gute Übereinstimmung mit (b) erkennen lässt.

Die beschriebenen Periodizitäten in den gewichteten Deviations der potentiellen Temperatur der Station Bischofshofen können somit durch die lokale Abschattung und den Vergleich mit Stationen, die einer derartigen Beeinflussung des Temperaturtagesgangs nicht unterliegen, erklärt werden.



**Abbildung 8.9:** Fallstudie vom 28. Juni 2010 zur Erklärung der gefundenen Periodizitäten in den gewichteten Deviations der potentiellen Temperatur der Station Bischofshofen (WMO-Nr.: 11141). In (a) ist der Tagesgang der potentiellen Temperatur der Station Bischofshofen in Rot und jener von drei primären Nachbarstationen (Radstadt, Maria Alm, St. Veit im Pongau) in Blau dargestellt. Man beachte die zeitliche Verschiebung des Temperaturanstiegs, der auch in den Differenzen der Messwerte in Maria Alm und jenen in Bischofshofen erkennbar ist (b). Dieses typische, orographisch bedingte systematische Signal kann alleine mit den Wellenzahlen 1, 2, 3 und 4 (entsprechen den gefundenen 24-, 12-, 8- und 6-h Schwingungen bei Betrachtung eines Tages) gut angenähert werden (c).

## 8.5 Meldungen der österreichischen internationalen Flughäfen

In diesem Abschnitt werden die dem Institut zur Verfügung stehenden Stationsmeldungen der österreichischen internationalen Flughäfen Linz-Hörsching, Wien-Schwechat, Innsbruck, Salzburg, Klagenfurt und Graz (ICAO-Codes: LOWL, LOWW, LOWI, LOWS, LOWK und LOWG) genauer betrachtet. Im zugrundeliegenden Zeitraum (Jahr 2010) kam es lediglich fünfmal vor (3 x LOWL, 1 x LOWS und 1 x LOWK), dass eine Stationsmeldung für den Parameter Druck oder Temperatur von einem Flughafen fehlte. Ebenso selten war es, dass unphysikalische Werte übermittelt wurden. Hierzu sei erwähnt, dass diese wahrscheinlich durch Fehler bei der Ver- oder Entschlüsselung entstanden sind, ein klassischer Messfehler wird bezüglich Meldungen von Flughäfen als eher unwahrscheinlich erachtet.

Bezüglich der Wind-Meldungen kam es jedoch überdurchschnittlich oft vor, dass der VERA-QC keine Beobachtung vorlag (vgl. Tabelle 8.1). Der Grund dafür kann bei der Dekodierung gefunden werden. Da die VERA-Analysen auf der Betrachtung von den beiden Windkomponenten  $u$  und  $v$  basiert, werden diese im Zuge der Dekodierung aus der gemeldeten Windrichtung und -geschwindigkeit berechnet. Die VERA-QC erhält ebenfalls nur die beiden Windkomponenten, jedoch keine decodierte Windgeschwindigkeit. Wird „umlaufender Wind“ gemeldet, ist die Berechnung von  $u$  und  $v$  nicht möglich und die VERA-QC interpretiert diese Fälle als „nicht gemeldet“.

**Tabelle 8.1:** Anzahl der stündlichen Analysetermine, an denen im Jahr 2010 von einem internationalen österreichischen Flughafen keine Meldung der Windkomponenten  $u$  und  $v$  vorliegt.

ICAO-Code	LOWL	LOWW	LOWI	LOWS	LOWK	LOWG
Anzahl nicht gemeldet:	1694	568	3341	1987	3791	3195

Handelt es sich bei den Betreibern einer Station um Flughäfen, so kann man davon ausgehen, dass auf die Qualität der Beobachtungen und auf die Wartung der Messgeräte großer Wert gelegt wird. Bei der Interpretation der folgenden Ergebnisse muss darauf geachtet werden, dass die VERA-QC Stationsmeldungen räumlich vergleicht und diese auf ihre Konsistenz überprüft. Wird eine Meldung verworfen, so kann das an einer fehlerhaften Messung liegen, aber auch daran, dass der Wert für die betrachtete Stationsverteilung nicht repräsentativ ist.



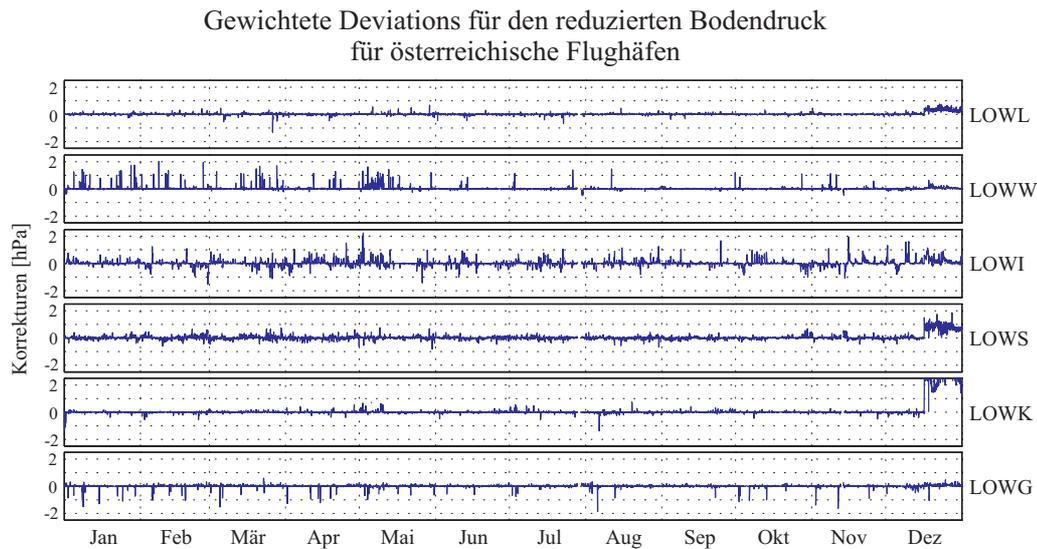
In Tabelle 8.2 ist prozentuell aufgeschlüsselt, welche Anteile der Meldungen von der VERA-QC akzeptiert, korrigiert oder verworfen wurden. Ebenso sind die Prozente der Stationsmeldungen, die als Gross Error erkannt wurden, und die Beteiligung an Stationscluster aufgelistet. Die Werte können mit jenen, die sich auf alle Stationen in Europa beziehen (siehe Tabelle 7.1), verglichen werden. Überdurchschnittlich hohe Prozentsätze an erkannten Gross Errors sind fett hervorgehoben.

Vor allem Meldungen der Station Salzburg - Flughafen (LOWS) werden überdurchschnittlich oft als Gross Error identifiziert. Sogar 44 % der gemeldeten Windgeschwindigkeiten werden von der VERA-QC als nicht repräsentativ beurteilt und verworfen. Vergleicht man die gemeldeten Windgeschwindigkeiten der Station Salzburg - Flughafen mit Meldungen umliegender Stationen (z.B.: Salzburg Freisaal (WMO-Nr.: 11350), Mattsee (WMO-Nr.: 11152), Mondsee (WMO-Nr.: 11153), Chiemsee (WMO-Nr.: 10982) und Lofer (WMO-Nr.: 11140)), so geht hervor, dass vor allem bei hohen Windgeschwindigkeiten von der exponierten Station LOWS immer die höchsten Werte gemeldet werden. Das resultiert in einer Bias-Korrektur für die Windgeschwindigkeit von -1.49 m/s (vgl. Tabelle 8.3).

**Tabelle 8.3:** Bias-Korrekturen (gültig für den 1. 1. 2011) für die betrachteten Parameter Niederschlag ( $NS$ ), Windkomponente ( $u$ ), Windkomponente ( $v$ ), potentielle Temperatur ( $\theta$ ), äquivalentpotentielle Temperatur ( $\theta_e$ ), reduzierter Druck ( $p_{red}$ ) und die Windgeschwindigkeit ( $|\vec{v}|$ ), berechnet für die österreichischen internationalen Flughäfen. Die betragsmäßig höchste Bias-Korrektur ist fett hervorgehoben.

ICAO-Codes	$NS$ [mm]	$u$ [m/s]	$v$ [m/s]	$\theta$ [°C]	$\theta_e$ [°C]	$p_{red}$ [hPa]	$ \vec{v} $ [m/s]
LOWL	0.00	-1.74	-0.28	-0.04	0.01	0.04	-0.88
LOWW	0.00	-0.34	0.60	0.16	0.22	-0.52	-1.42
LOWI	0.00	-1.23	0.30	-0.12	-0.05	0.38	-0.08
LOWS	0.00	-0.36	-0.42	-0.15	0.42	0.28	-0.91
LOWK	0.00	-1.36	0.42	-0.10	0.05	0.32	-0.18
LOWG	0.00	0.22	0.32	-0.26	-0.18	<b>2.09</b>	0.58

In der Tabelle 8.3 sind die für den 1. 1. 2011 geltenden Bias-Korrekturen für die Stationen an den Flughäfen bezüglich der einzelnen Parameter aufgelistet. Lediglich die Station LOWG erhält für den reduzierten Bodendruck eine Korrektur über 2 hPa. Da, wie bereits erwähnt, die gemeldeten Druckwerte im Rahmen der Dekodierung auch auf Meeressniveau reduziert werden, liegt die Vermutung nahe, dass die Höhenangabe der Station LOWG in der verwendeten Stationsliste falsch ist (Angaben in

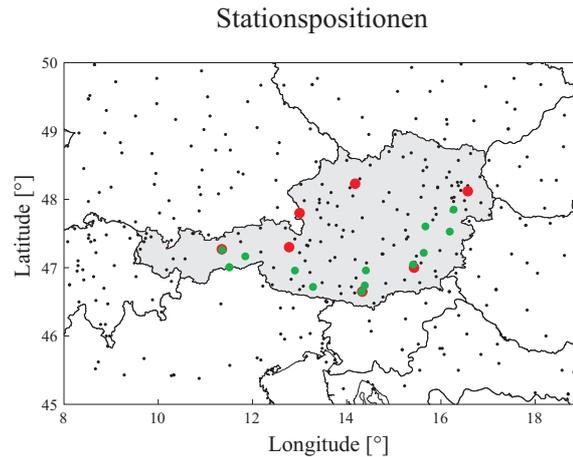


**Abbildung 8.10:** Blau: gewichtete (Bias-bereinigte) Deviations des reduzierten Bodendrucks, welcher aus den, von den österreichischen internationalen Flughäfen im Jahr 2010 gemeldeten (nicht reduzierten) Druckwerten berechnet wurde. Man beachte den Sprung in den Deviations, der Mitte Dezember vor allem für den Flughafen LOWK auftritt. Ab diesem Zeitpunkt stehen zusätzliche Meldungen von Stationen in der jeweiligen Umgebung zur Verfügung.

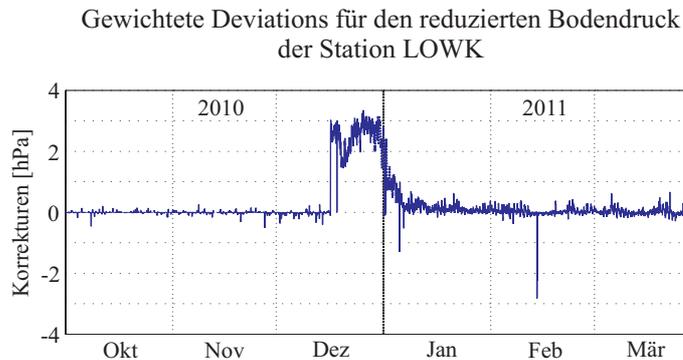
den zur Verfügung stehenden Stationslisten reichen von 337 m über 347 m bis zu 363 m).

Abbildung 8.10 stellt die Zeitserien der berechneten gewichteten Deviations des reduzierten Bodendrucks im Jahr 2010 für die Stationen an den Flughäfen dar. Die Korrekturen betragen für diese Stationen meist unter 1 hPa. Mitte Dezember erkennt man vor allem in den Deviations der Station LOWK einen Sprung. Von einem Termin an werden Korrekturen berechnet, die systematisch um 2 hPa höher sind als jene im restlichen Jahr. Dieses Signal ist ab dem 16. 12. 2010 um 16:00 UTC auch in den Deviations anderer österreichischer und italienischer Stationen erkennbar. Die Ursache dafür ist, dass ab diesem Zeitpunkt Meldungen von mehreren neuen Stationen für die Verarbeitung mit VERA zur Verfügung stehen. In Abbildung 8.11 sind diese Stationen als grüne Punkte eingezeichnet. Zusätzlich sind in Schwarz die regelmäßig meldenden Stationen und in Rot die Stationen der österreichischen internationalen Flughäfen eingetragen.

Werden neue Stationsmeldungen in die Analyse aufgenommen, wie es Mitte Dezember der Fall war, liegt noch keine statistische Information über die Vertrauenswürdigkeit der neuen Stationsmeldungen vor. Um eine Bias-Korrektur für den reduzierten Druck zu berechnen, benötigt die VERA-QC mit den aktuellen Einstellungen 30 Tage. Da die Bias-Korrektur über den Median der Deviations



**Abbildung 8.11:** Schwarz: Stationspositionen der im Jahr 2010 regelmäßig meldenden Stationen. Rot: internationale Flughäfen. Grün: Stationen, deren Meldungen seit 16.12.2010 (16:00 UTC) verwendet werden können.



**Abbildung 8.12:** Blau: gewichtete Deviations des reduzierten Bodendrucks der Station LOWK für den Zeitraum Oktober 2010–März 2011. Der Sprung in den Deviations wird durch neue, zusätzliche Informationen in der Umgebung der Station LOWK verursacht.

definiert ist, kann innerhalb der Hälfte dieses Zeitraums ein systematischer Fehler einer neuen Station die unmittelbare Umgebung beeinflussen.

Betrachtet man zum Beispiel die Zeitserie der gewichteten Deviations der Station LOWK (Abbildung 8.12), so erkennt man, dass der Einfluss der neuen Stationen nach einer Bias-Berechnungsperiode wieder verschwindet. Abbildung 8.12 hebt hervor, dass die VERA-QC mit der Bias-Korrektur im operationellen Betrieb lernfähig ist und sich auf neue Stationsverteilungen anpassen kann.

# Kapitel 9

## Conclusio

Die komplexe Qualitätskontrolle synoptischer Daten ist eine sehr junge wissenschaftliche Disziplin und erlebte mit dem Aufkommen der numerischen Wettervorhersage in der zweiten Hälfte des 20. Jahrhunderts ihren ersten Aufschwung. Im Dezember 1999 wurde durch das Orkantief „Lothar“ erneut die Bedeutung einer hochwertigen und auch anwendungsorientierten Qualitätskontrolle bzw. Datenassimilation hervorgehoben. Die Entwicklung dieses Sturmtiefs konnte von den Vorhersagemodellen nicht erfasst werden, da Messwerte im Rahmen der Datenassimilation falsch interpretiert bzw. von der QC verworfen wurden. Dies war ein weiterer Anstoß, um die Entwicklung besserer Datenassimilationsmethoden und der darin eingebetteten Qualitätskontrollen voranzutreiben. Mittlerweile ist ein vierdimensionaler variationeller Ansatz mit Bayes'scher Qualitätskontrolle bei führenden numerischen Wettervorhersagezentren Standard. Allerdings ist diese Methode für regionale, hochauflösende Modelle schwierig anzuwenden, besonders in komplex strukturierten Gebieten. Aus diesem Grund wurde am Institut für Meteorologie und Geophysik der Universität Wien die in dieser Diplomarbeit vorgestellte VERA-QC entwickelt und wird seit Jänner 2010 in der beschriebenen Konfiguration operationell ausgeführt. Sie hat sich als Preprocessing Tool für die operationelle zweidimensionale VERA-Analyse bewährt und hebt sich durch die folgenden Besonderheiten von herkömmlichen Qualitätskontrollmethoden ab:

- Der in sich konsistente variationelle Ansatz der VERA-QC kann auf alle Daten angewendet werden, die eine hohe Autokorrelation aufweisen. Dies trifft bezüglich des vorliegenden GTS-Messnetzes für die meisten meteorologisch interessanten Parameter zu (wie reduzierter Druck, potentielle oder äquivalentpotentielle Temperatur). Werden in Feldexperimenten hoch aufgelöste Niederschlagsdaten erhoben (vgl. COPS), bietet sich die VERA-QC, die ohne Vorwissen auskommt, auch für die Kontrolle dieser Daten an.

- Betrachtet man das zur Verfügung stehende GTS-Messnetz in Europa, so fällt auf, dass die Stationen sehr inhomogen verteilt sind. Mit dem Prinzip der „natürlichen Stationsvernetzung“ kann sich die VERA-QC automatisch an die unterschiedlichen Stationsdichten anpassen.
- Eng benachbarte Stationen stellen für die meisten Qualitätskontrollen eine Schwierigkeit dar, wenn zumindest eine dieser Stationen fehlerhafte Meldungen abgibt. Für die VERA-QC wurde eine spezielle Cluster-Behandlung entwickelt, die dieses Problem löst.
- Da vor allem große Fehler (Gross Errors) eine weiterführende Analyse der Daten stark beeinflussen können, werden diese so gut wie möglich im Rahmen der Gross Error Erkennung identifiziert und verworfen. Als Kriterium zur Erkennung der Gross Errors wird nicht nur die Glattheit bzw. Rauigkeit des betrachteten Feldes in der Umgebung der fehlerhaften Beobachtung herangezogen, es kommen auch variable, stationspezifische Grenzwerte zum Einsatz.
- Die Qualitätskontrolle von Daten in komplexem Gelände stellt eine besondere Herausforderung dar. Durch das reduzierte Talvolumen in Gebirgen oder die lokale Überhitzung auf Hochebenen können Temperaturwerte räumlich nicht ohne weitere Modifikationen miteinander verglichen werden. Aus diesem Grund wurde die Möglichkeit geschaffen, die Abweichungen zur Standardatmosphäre an den jeweiligen Stationspositionen zu kontrollieren.
- Mit der Sammlung der von der VERA-QC berechneten Korrekturvorschläge können auch systematische Fehler identifiziert werden. Operationell werden diese von einer Bias-Korrektur ausgebessert, die sich laufend an die vorherrschenden Gegebenheiten anpasst und somit „lernfähig“ ist.

Die im Rahmen dieser Diplomarbeit umgesetzten Erweiterungen der VERA-QC und die präsentierten Auswertungen der Deviations beinhalten noch viel Potential für weiterführende Bearbeitungen und zusätzliche Untersuchungen.

Im Kapitel 8 wurde gezeigt, wie sich das mikrometeorologische Phänomen der städtischen Wärmeinsel in den Deviations widerspiegelt. Um diesen interessanten Effekt auch quantitativ abzuschätzen, sollte die Station in der Wiener Innenstadt nur mit solchen Stationen in der Umgebung verglichen werden, die dem urbanen Einfluss nicht mehr ausgesetzt sind. Dazu müsste man in die Stationsvernetzung eingreifen und die lokale (sekundäre) Nachbarschaft der Station in der Wiener Innenstadt willkürlich gestalten. Es wird erwartet, dass der Effekt der Wärmeinsel, der in Zeiten zunehmender Urbanisierung immer öfter thematisiert wird, noch stärker hervortritt. Der Vorteil dieser Methode gegenüber einfachen Vergleichen

---

von Temperatur-Messzeitreihen wäre die zugleich durchgeführte Überprüfung auf klimatologisch mögliche Limits und die Gross Error Erkennung.

Die präsentierte Analyse auf Periodizitäten weist auf das Vorhandensein von Tagesgängen in den Deviations hin und weitere Auswertungen bieten sich an. Es wurde bereits gezeigt, dass die systematischen Signale vor allem bei Talstationen gut erkennbar sind. Man kann vermuten, dass diese umso größer ausfallen, umso eher Talstationen mit Flachlandstationen verbunden werden, da in Tälern durch das reduzierte Talvolumen der solare Antrieb verstärkt wird. Um dies zu überprüfen, könnte auch hier eine willkürliche Stationsvernetzung eingesetzt werden. Ebenso wäre es interessant, diskrete Varianzspektren der Deviations der potentiellen Temperatur jenen des reduzierten Bodendrucks auch betragsmäßig gegenüberzustellen. Am Beispiel der Station Bischofshofen wurde gezeigt, dass die Deviations einer einzelnen Station aufgrund der lokalen Bedingungen (z.B.: Orographie) von der Tageszeit abhängen können. Um diesen Zusammenhang auch bei der Korrektur systematischer Fehler zu berücksichtigen, könnte die Berechnung der Bias-Korrektur auf tageszeitlicher Basis erfolgen.

Die Kontrolle des D-Wertes der potentiellen Temperatur (vgl. Kapitel 5) könnte noch verbessert werden, indem man anstatt der Abweichung zur Standardatmosphäre die Abweichung zum lokalen klimatologischen vertikalen Temperaturgradienten betrachtet. Diese Erweiterung können dann vor allem in komplexem Gelände eine Verbesserung bewirken, da dort der feuchtisentrope vertikale Temperaturgradient keine optimale Näherung darstellt.

Phänomene, deren Größenskala kleiner ist als der mittlere Stationsabstand, werden von der VERA-QC nicht aufgelöst. Wird die VERA-QC nun auf ein sehr datenarmes Gebiet angewendet, kann es passieren, dass beispielsweise Messwerte innerhalb eines kleinen Tiefdrucksystems fälschlicherweise nach oben korrigiert werden. Dies kommt unter Umständen vor, wenn ein solches System über eine kleine Insel im Ozean zieht und in ihrer Umgebung keine weiteren Beobachtungen zur Verfügung stehen. Es wurde bereits der Versuch unternommen, für derartige Fälle die VERA-QC mit Modellinformationen zu unterstützen. Bevor eine hohe Korrektur angebracht wird, erfolgt ein Vergleich der fraglichen Messung mit dem auf die Stationsposition interpolierten Wert des zugehörigen Modellfeldes. Als Referenz wird das numerische Wettervorhersagemodell GFS („*Global Forecasting System*“ des amerikanischen Wetterdienstes) verwendet. Nur wenn auch dieser Vergleich eine zu hohe Diskrepanz ergibt, wird die berechnete Korrektur angebracht.

Diese Methode muss noch ausgetestet werden und passende benutzerdefinierte Schwellwerte sind festzulegen.

Abschließend ist zu erwähnen, dass die mathematische Formulierung, die Cluster-Behandlung, die Gross Error Erkennung sowie ein Vergleich der VERA-QC mit anderen Qualitätskontrollmethoden im Artikel „Data Quality Control Based on Self-Consistency“ zusammengefasst sind und in der meteorologischen Fachzeitschrift *Monthly Weather Review* publiziert wurden (Steinacker *et al.* 2011). Ein weiterer Artikel, der einen, der VERA-QC vorgeschalteten Stationsauswahlalgorithmus und Themen dieser Diplomarbeit wie die Bias-Korrektur, die Betrachtung des D-Wertes und die Erweiterung der Gross Error Erkennung sowie reale Fallstudien behandelt, ist bereits in Arbeit. Die Ergebnisse des operationellen Betriebes der VERA-QC, die damit verbundenen Probleme in komplexem Terrain und deren Lösungsvorschläge, sowie die Auswertungen der gesammelten Deviations wurden im September 2011 im Rahmen der meteorologischen Tagung ECAM (European Conference on Application of Meteorology) sowie im November 2011 am 4. Österreichischen MeteorologInnen-tag in Form eines Posters präsentiert.

# Literatur

- Andersson, E., und H. Järvinen, 1999: Variational quality control. *Quarterly Journal of the Royal Meteorological Society*, **125**, 697–722.
- Baker, N., 1992: Quality Control for the Navy Operational Atmospheric Database. *Weather and Forecasting*, **7**, 250–261.
- Barber, C. B., D. P. Dobkin, und H. Huhdanpaa, 1996: The Quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software*, **22**, 469–483.
- Barnes, S., 1964: A Technique for Maximizing Details in Numerical Weather Map Analysis. *Journal of Applied Meteorology*, **3**, 396–409.
- Bauer, S., R. Gutdeutsch, M. Hantel, H. Reuter, H. Rucker, G. Siedler, T. Spohn, R. Steinacker, und W. Zenk, 2001: *Bergmann-Schaefer: Lehrbuch der Experimentalphysik, Band 7, Erde und Planeten*. Walter de Gruyter & Co., 2nd edition, pp. 131 – 307.
- Bica, B., R. Steinacker, C. Lotteraner, und M. Suklitsch, 2007: A new concept for high resolution temperature analysis over complex terrain. *Theoretical and Applied Climatology*, **90**, 173–183.
- Broughton, S., und K. M. Bryan, 2009: *Discrete Fourier analysis and wavelets: Applications to Signal and Image Processing*. John Wiley & Sons, Inc.
- Emeis, S., 2010: *Measurement Methods in Atmospheric Sciences - In situ and remote*. Gebrüder Borntraeger, Stuttgart, pp. 7–28.
- Feng, S., Q. Hu, und W. Qian, 2004: Quality control of daily meteorological data in china, 1951-2000: a new dataset. *International Journal of Climatology*, **24**, 853–870.
- Fiebrich, C., und K. Crawford, 2001: The Impact of Unique Meteorological Phenomena Detected by the Oklahoma Mesonet and ARS Micronet on Automated Quality Control. *Bulletin of the American Meteorological Society*, **82**, 2173–2187.
- Gandin, L., 1988: Complex Quality Control of Meteorological Observations. *Monthly Weather Review*, **116**, 1137–1156.

- Google Inc., 2011: Google Earth 2010 (Version 5.1) [Software]. Erhältlich bei <http://www.google.de/intl/de/earth/> [Zugriff am 28.1.2011].
- Goovaerts, P., 1997: *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, pp. 266–271.
- Haurwitz, B., 1964: Atmospheric Tides. *Science*, **144**, 1415–1422.
- Hubbard, K., und J. You, 2005: Sensitivity Analysis of Quality Assurance Using the Spatial Regression Approach - A Case Study of the Maximum/Minimum Air Temperature. *Journal of Atmospheric and Oceanic Technology*, **22**, 1520–1530.
- Hupfer, P., und W. Kuttler, 2006: *Witterung und Klima - Eine Einführung in die Meteorologie und Klimatologie*. B.G. Teubner Verlag, Wiesbaden, 12th edition.
- Ingleby, N. B., und A. Lorenc, 1993: Bayesian quality control using multivariate normal distributions. *Quarterly Journal of the Royal Meteorological Society*, **119**, 1195–1225.
- Kalnay, E., 2003: *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, Cambridge, pp. 136–204.
- Klausen, J., H.-E. Scheel, und M. Steinbacher, 2010: WMO/GAW Glossary of QA/QC - Related Terminology. <http://gaw.empa.ch/glossary/glossary.html> [Zugriff am 14.12.2010].
- Lorenc, A., und O. Hammon, 1988: Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. *Quarterly Journal of the Royal Meteorological Society*, **114**, 515–543.
- Lorenc, A. C., 1981: A Global Three-Dimensional Multivariate Statistical Interpolation Scheme. *Monthly Weather Review*, **109**, 701–721.
- Pötschacher, W., R. Steinacker, und M. Dorninger, 1996: VERA - a high resolution analysis scheme for the atmosphere over complex terrain. *MAP Newsletter*, **5**, 64–65.
- Reek, T., S. R. Doty, und T. W. Owen, 1992: A Deterministic Approach to the Validation of Historical Daily Temperature and Precipitation Data from the Cooperative Network. *Bulletin of the American Meteorological Society*, **73**, 753–762.
- Salvati, M., und E. Brambilla, 2008: *Data quality control procedures in Alpine meteorological services*. Regional Agency for Environmental Protection of Lombardia.

- Shafer, M., C. Fiebrich, D. Arndt, S. Fredrickson, und T. Hughes, 2000: Quality Assurance Procedures in the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology*, **17**, 474–494.
- Steinacker, R., C. Häberli, und W. Pöttschacher, 2000: A Transparent Method for the Analysis and Quality Evaluation of Irregularly Distributed and Noisy Observational Data. *Monthly Weather Review*, **128**, 2303–2316.
- Steinacker, R., M. Mayer, und A. Steiner, 2011: Data Quality Control Based on Self Consistency. *Monthly Weather Review*, doi: 10.1175/MWR-D-10-05024.1.
- Steinacker, R., M. Ratheiser, B. Bica, B. Chimani, M. Dorninger, W. Gepp, C. Lotteraner, S. Schneider, und S. Tschannett, 2006: A Mesoscale Data Analysis and Downscaling Method over Complex Terrain. *Monthly Weather Review*, **134**, 2758–2771.
- Strangeways, I., 2003: *Measuring the Natural Environment*. Cambridge University Press, Cambridge, 2nd edition.
- Wade, C. G., 1987: A Quality Control Program for surface mesometeorological data. *Journal of Atmospheric and Oceanic Technology*, **4**, 435–453.
- Wahba, G., 1990: *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, 5th edition, pp. 1–20.
- WMO, 1993: *Guide on the Global Data-processing System*. World Meteorological Organization, WMO (Series); no 305.
- WMO, 2003: *Manual on the Global Observing System*. World Meteorological Organization, WMO (Series); no 544.
- WMO, 2007: *Guide to the Global Observing System*. World Meteorological Organization, WMO (Series); no 488.
- WMO, 2008: *Guide to Meteorological Instruments and Methods of Observation*. World Meteorological Organization, WMO Nr. 8, 7th edition.
- Working Group 2 of the Joint Committee for Guides in Metrology, 2008: *International vocabulary of metrology - Basic and general concepts and associated terms (VIM)*. Bureau International des Poids et Mesures (BIPM) (Sèvres, France) and International Organization for Standardization (ISO) (Geneva, Switzerland), 3rd edition.
- Wulfmeyer, V., A. Behrendt, C. Kottmeier, U. Corsmeier, C. Barthlott, G. C. Craig, M. Hagen, D. Althausen, F. Aoshima, M. Arpagaus, H. S. Bauer, L. Bennett, A. Blyth, C. Brandau, C. Champollion, S. Crewell, G. Dick, P. D.

Girolamo, M. Dorninger, Y. Dufournet, R. Eigenmann, R. Engelmann, C. Flamant, T. Foken, T. Gorgas, M. Grzeschik, J. Handwerker, C. Hauck, H. Höller, W. Junkermann, N. Kalthoff, C. Kiemle, S. Klink, M. König, L. Krauss, C. N. Long, F. Madonna, S. Mobbs, B. Neining, S. Pal, G. Peters, G. Pigeon, E. Richard, M. W. Rotach, H. Russchenberg, T. Schwitalla, V. Smith, R. Steinacker, J. Trentmann, D. D. Turner, J. van Baelen, S. Vogt, H. Volkert, T. Weckwerth, H. Wernli, A. Wieser, und M. Wirth, 2011: The Convective and Orographically-induced Precipitation Study (COPS): the scientific strategy, the field phase, and research highlights. *Quarterly Journal of the Royal Meteorological Society*, **137**, 3 – 30.

Zahumenský, I., 2004: *Guidelines on Quality Control Procedures for Data from Automatic Weather Stations*. World Meteorological Organization.

# Anhang A

## A.1 Konfigurationseinstellungen der VERA-QC

```

$ Zeitstempel          # 2011040900          % Zeitinformation entweder als Zeitstempel yyyyymmddHH oder durch Angabe von aktuell,
                        #                               Zeitinkremente erlaubt
$ Domaenen_Name       # Europa          % Domaenenennamen

$ Stammverzeichnis    # E:\Veraxx\Vera_Operationell\      % Oberstes Verzeichnis in dem sich alle weiteren Ordner befinden
$ Hauptverzeichnis    # ~Stammverzeichnis\Daten\        % Pfad in dem sich die Datenfiles befinden

$ Ordner_Eingang      # ~Hauptverzeichnis~QC_Messwerte~\Domaenen_Name~\yyyy\mm\dd\      % Pfad mit den decodierten und zu überprüfenden Messwerten
$ Ordner_Ausgang      # ~Hauptverzeichnis~QC_Resultate~\Domaenen_Name~\                    % Pfad für die qualitätskontrollierten Werte
                                                                (die nicht für Vera bestimmt sind, wenn leer, nicht verwendet)
$ Ordner_Ausgang_Vera # ~Hauptverzeichnis~Veraxx_Eingang~\Domaenen_Name~\yyyy\mm\dd\      % Pfad für die qualitätskontrollierten Werte für die Verwendung
                                                                von Vera
$ Ordner_Deviations   # ~Hauptverzeichnis~QC_Deviations~\Domaenen_Name~\yyyy\mm\        % Pfad für die zu sammelnden Korrekturvorschläge
$ Ordner_Bias         # ~Hauptverzeichnis~QC_Bias~\Domaenen_Name~\yyyy\mm\              % Pfad für die täglichen Bias-Daten
$ Ordner_IQR          # ~Hauptverzeichnis~QC_IQR~\Domaenen_Name~\yyyy\mm\              % Pfad für die monatlichen IQR-Daten
$ Ordner_Klima_Limits # ~Hauptverzeichnis~QC_Klima_Limits~\Domaenen_Name~\              % Pfad für die Matrix mit den klimatologischen Limits
$ Ordner_Stationskennungen # ~Hauptverzeichnis~QC_Stationskennungen\                          % Pfad für das File mit den Stationskennungen
$ Ordner_Stationsnummern # ~Hauptverzeichnis~QC_Messwerte~\Domaenen_Name~\yyyy\mm\dd\      % Pfad für das File mit den Stationsnummern
$ Ordner_QC_Vera_Sensitiv # ~Hauptverzeichnis~QC_Vera_Sensitiv~\Domaenen_Name~\yyyy\mm\dd\ % Pfad für die Files mit den Vera-Sensitiv Daten
$ Ordner_Bogus        # ~Hauptverzeichnis~GFS_Bogus~\Domaenen_Name~\yyyy\mm\dd\        % Pfad für die Files mit den Bogus Werten (GFS-Gitterpunktwerte)

$ Name_Eingang        # Input_yyyymmddHH.mat      % Bezeichnung des Dateinamens mit den zu überprüfenden Messwerten
$ Name_Ausgang        # Ergebnis.mat              % Bezeichnung des Dateinamens der qualitätskontrollierten Werte (die nicht für Vera
                                                                bestimmt sind, wenn leer, nicht verwendet)
$ Name_Ausgang_Vera   # Input_yyyymmddHH.mat      % Bezeichnung des Dateinamens mit den qualitätskontrollierten Messwerten für Vera
$ Name_Deviations_gewichtet # Deviations_gewichtet_yyyy_mm_par % Dateiname der monatlichen Sammlung mit den gewichteten Deviations (ohne Unterstrich,
                                                                Parameternummer und Endung)
$ Name_Deviations_ungewichtet # Deviations_ungewichtet_yyyy_mm_par % Dateiname der monatlichen Sammlung mit den ungewichteten Deviations (ohne Unterstrich,
                                                                Parameternummer und Endung)
$ Name_Bias           # Koord_und_Bias_yyyymmdd.mat % Dateinamen für die täglichen Bias-Daten
$ Name_IQR_gewichtet  # IQR_gewichtet_yyyy_mm_par % Dateinamen für die monatlichen gewichteten IQR-Daten (ohne Unterstrich,
                                                                Parameternummer und Endung)
$ Name_IQR_ungewichtet # IQR_ungewichtet_yyyy_mm_par % Dateinamen für die monatlichen ungewichteten IQR-Daten (ohne Unterstrich,
                                                                Parameternummer und Endung)
$ Name_Klima_Limits   # Klima_Limits.mat         % Dateiname der Matrix mit den klimatologischen Limits

```

```

$ Name_Stationskennungen      # Stationsliste_Erde.dat          % Dateiname des Files mit den Stationskennungen
$ Name_Stationsnummern        # Stat_nr_yyyyymmddHH.mat          % Name des Files mit den Stationsnummern
$ Name_QC_Vera_Sensitiv        # Ergebnis_QC_yyyyymmddHH_par_nr.mat % Name des Files mit den für Vera Sensitiv relevanten Namen (ohne Unterstrich,
                                nr wird durch Parameternummer ersetzt)

$ Name_Bogus                    # Bogus_yyyyymmddHH.mat            % Name des Files mit den Bogus Werten (GFS-Gitterpunktwerte)

$ Flag_Parameter_testen        # [2,1,1,1,1,0,0,1,0,0,0,0,1,0,0] % Flag, ob Parameter kontrolliert werden soll [N,u,v,Theta,Theta_e,Tpot_min,Tpot_max,p_red,
                                dp_dt,p,T,Td,(+3)](1 immer, 2 Haupttermin, 0 nie)
$ Flag_D_Wert                  # [0,0,0,1,0,0,0,0,0,0,0,0,0,0,0] % Flag, ob der ursprüngliche Wert (0) oder der D-Wert (1) analysiert werden soll
$ Flag_Schwelle_individuell    # [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0] % Flag, ob für die Stationen konstante (0) oder individuelle (1) Schwellwerte verwendet werden
$ Flag_Gradient_Bereinigung    # [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0] % Flag, ob der Gradient bei den Messwerten abgezogen werden soll (1) Gradient, (0) keiner
$ Flag_lokal_global            # [1,1,1,1,1,1,1,1,1,1,1,1,1,1,1] % Flag, ob die Qualitätskontrolle mit Subdomänen (0) oder global (1) durchgeführt werden soll
$ Flag_Kost_fun_red            # [1,1,1,1,1,1,1,1,1,1,1,1,1,1,1] % Flag, ob Kostenfunktionsreduktion durchgeführt werden soll. (0) nein, (1) ja
$ Flag_Cluster_Bereinigung     # [1,1,1,1,1,1,1,1,1,1,1,1,1,1,1] % Flag, ob Clusterbereinigung erfolgen soll. (0) nein, (1) ja
$ Flag_GE_Bereinigung          # [1,1,1,1,1,1,1,1,1,1,1,1,1,1,1] % Flag, ob Stationen mit Gross Errors gelöscht werden sollen (0) nein, (1) ja
$ Flag_Sammlung_Deviations     # [3,3,3,3,3,3,3,3,3,3,3,3,3,3,3] % Flag, ob Korrekturvorschläge in die Sammlung aufgenommen werden sollen (0: nein, 1: un-,
                                2: gewichtete, 3: beide)

$ Flag_IQR_Parameter           # [0,1,1,1,1,0,0,1,0,0,0,0,1,0,0] % Flag, ob für Parameter die IQR berechnet werden soll
$ Flag_Bias_anbringen          # [0,1,1,1,1,0,0,1,0,0,0,0,1,0,0] % Flag, ob Bias-Korrekturen angebracht werden sollen (0) nein, (1) ja
$ Flag_Bias_Korrektur          # [0,1,1,1,1,0,0,1,0,0,0,0,1,0,0] % Flag, ob für Parameter eine Bias-Korrektur berechnet werden soll (0) nein, (1) ja
$ Flag_zu_wenig_Meldungen      # [0,1,1,1,1,0,0,1,0,0,0,0,1,0,0] % Flag, ob Meldungen nur berücksichtigt werden sollen, wenn in Vergh. genug vorhanden
$ Flag_Flaechenmittel_abziehen # [0,0,0,1,1,0,0,0,0,0,0,0,0,0,0] % Flag, ob das Flächenmittel über primäre Nachbarn den Deviations abgezogen werden soll
$ Flag_Klimalimit_ueberpruefen # [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0] % Flag, ob Meldungen auf Einhaltung klimatologischer Limits geprüft werden sollen

$ Flag_Dimensionen            # [1,1,0,0]                          % Zu verwendende Dimensionen [x,y,z,t] (1 ja, 0 nein)
$ Flag_Ableitungen_gemischt    # 1                                    % Verwendung der gemischten zweiten Ableitungen (1 ja, 0 nein)
$ Flag_Windgeschwindigkeit     # 1                                    % Flag, ob die Windgeschwindigkeit (1) oder u und v getrennt (0) kontrolliert werden sollen
$ Flag_Koordinaten             # 0                                    % Koordinatensystem der Eingangsdaten (0: lambda, phi, 1: x,y)

$ Format_Eingang                # 3                                    % Format der Eingangsdaten (1: Ascii, 2: Fortran Binär, 3: Matlab Binär)
$ Format_Ausgang                # 3                                    % Format der Ausgabedaten (1: Ascii, 2: Fortran Binär, 3: Matlab Bin
$ Intervall_Fehlwerte          # [9000,10000]                        % Intervall für Stationswerte bei Fehlzeiten
$ Fehlwert                     # 9999.9                              % Wert für nicht gemeldete Stationen und Gross Errors in den Ausgabefiles
$ Multiplikationsfaktor_z       # 1                                    % Multiplikationsfaktor Höhe (wenn Höhe in km)
$ Multiplikationsfaktor_t       # 1                                    % Multiplikationsfaktor Zeit (wenn Zeit in Stunden)

```

```

$ ID_Glaettung          # 0.000001          % Glättungsfaktor für den Nenner bei der Inverse Distance Interpolation
$ ID_Exponent           # 2                  % Exponent für den Abstand bei der Inverse Distance Interpolation
$ ID_Gitterabstand     # 100               % Abstand zwischen Gitterpunkten um eine Station als Bruchteil der mittleren Abstände in der Domäne

$ Parameter_Untergrenze_phys # [0,-50,-50,-50,-50,-60,-60,900,-30,0,-60,-60,0,0,0] % Absolute untere Grenze (wenn darunter, wird Station nicht miteinbezogen)
$ Parameter_Obergrenze_phys # [200,50,50,80,100,60,60,1100,30,1100,60,60,50,0,0] % Absolute obere Grenze (wenn darüber, wird Station nicht miteinbezogen)
$ Parameter_Schwelle    # [50,1,1,1,1,5,5,0,5,1,5,5,1,5,1] % Maximaler Korrekturvorschlag, bei dem noch nicht korrigiert wird
$ Abstand_min          # 0.1               % Kleinster zulässiger Abstand (Clusterbereinigung) als Vielfaches des Median der Entfernungen
$ Abstand_max          # 5                 % Maximaler Abstand für Verbindung der Stationen als Vielfaches des Medians der Entfernungen

$ GE_Kost_fun_red_min  # [85,85,85,85,85,85,85,85,85,85,85,85,85,85] % Kostenfunktionsreduktion in Prozent, ab welcher Gross Error vorliegt (0 oder 1)
$ GE_Vielfaches_Median # [30,30,30,30,30,30,30,30,30,30,30,30,30,30,30] % Korrekturvorschlag als Vielfaches des Medians der Vorschläge, ab welcher Gross Error vorliegt
$ GE_Min_IQR_Schwelle # [0 4 4 4 4 0 0 4 0 0 0 0 4 0 0] % Minimale, variable IQR-Schwelle, ab der ein Gross Error erkannt werden soll
$ GE_Vielf_IQR        # 3                 % Vielfaches der IQR (als variabler maximaler Korrekturvorschlag für das Finden von Gross Errors)

$ Min_Anzahl_Deviations_Bias # 30                % Mindestanzahl vorhandener Deviations ab der eine Bias-Berechnung erfolgt
$ Bias_Zeitspanne      # [30,30,30,-210,-210,30,30,30,30,30,30,30,30,30,30] % Anzahl der Tage, die für die Bias-Berechnung herangezogen werden
$ Bias_Intervallbreite_Tage # [30,30,30,30,30,30,30,30,30,30,30,30,30,30,30] % Anzahl der Tage der beiden Intervalle am Anfang und Ende des Bias-Berechnungszeitraums, die verwendet werden, wenn die Bias-Zeitspanne negativ ist
$ Min_Anzahl_Deviations_IQR # [50]              % Mindestanzahl vorhandener Deviations ab der eine IQR-Berechnung erfolgt

$ Max_Korrektur       # [1000,20,20,20,20,20,1000,1000,20,1000,1000,1000,1000,20,1000,1000] % Überschreitet eine Korrektur diese maximale Korrektur, wird der Stationswert mit dem interpolierten Wert der 4 umliegenden GFS-Gitterpunkte verglichen
$ Delta_Bogus        # [1000,20,20,15,10,1000,1000,15,1000,1000,1000,1000,10,1000,1000] % Überschreitet die Differenz zwischen Stationswert und interpoliertem GFS-Gitterpunktwert, so wird der Stationswert verworfen

```

# Anhang B

## B.1 Statistische Auswertungen der Stationsmeldungen

**Tabelle B.1:** Durchschnittliche Anzahl (gerundet) der verfügbaren stündlichen Stationsmeldungen im Jahr 2010 nach Ländern aufgeschlüsselt. Die Zählung erfolgte, sobald für einen Parameter eine Meldung zur Verfügung steht, ungeachtet der Qualität der Meldung. Zusätzlich wurde noch zwischen allen Terminen, Haupt- (00:00,6:00,12:00 und 18:00 UTC) und Nebenterminen unterschieden. Die Angabe der Länderkürzel entspricht der ISO-3166-1 ALPHA-3 Kodierung, die von der Internationalen Organisation für Normung (englisch: *International Organization for Standardization*, ISO) geführt wird.

Länderkürzel	alle Termine	Haupttermine	Nebentermine	Länderkürzel	alle Termine	Haupttermine	Nebentermine
ALB	1	1	1	LBY	4	9	3
AUT	123	149	118	LIE	1	1	1
BEL	9	16	7	LTU	6	10	5
BGR	12	31	8	LUX	1	1	1
BIH	5	9	4	LVA	4	7	4
BLR	13	13	13	MAR	12	16	11
CHE	51	67	48	MDA	1	1	1
CZE	32	34	32	MKD	3	2	3
DEU	208	221	206	MLT	1	1	1
DNK	18	17	20	MNE	2	3	2
DZA	27	38	24	NLD	50	50	50
EGY	2	2	2	NOR	100	113	97
ESP	25	37	23	POL	20	49	15
EST	10	23	7	ROU	18	28	16
FIN	44	84	36	RUS	3	4	2
FRA	103	154	93	SRB	10	26	7
GBR	165	166	164	SVK	19	21	18
GRC	21	28	20	SVN	8	14	6
HRV	21	19	21	SWE	69	151	52
HUN	32	33	31	TUN	14	20	13
IRL	7	13	6	TUR	11	19	9
ITA	72	87	69	UKR	5	12	4
				<b>total</b>	<b>1363</b>	<b>1800</b>	<b>1274</b>

**Tabelle B.2:** In den Spalten „Anzahl Meld.“ ist die durchschnittliche Anzahl (gerundet) der verfügbaren stündlichen Stationsmeldungen nach Ländern aufgeschlüsselt angegeben. Die Spalten „Fläche in km<sup>2</sup>“ beinhalten die, auf 100 km<sup>2</sup> gerundeten Flächenangaben der beteiligten Länder ungeachtet der Fläche, die in der betrachteten Europadomäne liegt (vgl. Tabelle B.3). Die Spalten „Anz. pro 100 000 km<sup>2</sup>“ enthalten die durchschnittliche Anzahl der stündlichen Meldungen des jeweiligen Landes bezogen auf 100 000 km<sup>2</sup>. Die Zählung erfolgte, sobald für einen Parameter eine Meldung zur Verfügung steht, ungeachtet der Qualität der Meldung. Die Angabe der Länderkürzel entspricht der ISO-3166-1 ALPHA-3 Kodierung, die von der Internationalen Organisation für Normung (englisch: *International Organization for Standardization*, ISO) geführt wird.

Länderkürzel	Anz. Meld.	Anz. pro 100 000 [km <sup>2</sup> ]	Fläche in [km <sup>2</sup> ]	Länderkürzel	Anz. Meld.	Anz. pro 100 000 [km <sup>2</sup> ]	Fläche in [km <sup>2</sup> ]
ALB	1	3.5	28 700	LBY	4	0.2	1 775 500
AUT	123	146.6	83 900	LIE	1	500	200
BEL	9	29.5	30 500	LTU	6	9.2	65 300
BGR	12	10.8	111 000	LUX	1	38.5	2 600
BIH	5	9.8	51 200	LVA	4	6.2	64 600
BLR	13	6.3	207 600	MAR	12	2.7	446 600
CHE	51	123.5	41 300	MDA	1	3.0	33 800
CZE	32	40.6	78 900	MKD	3	11.7	25 700
DEU	208	58.3	357 100	MLT	1	333.3	300
DNK	18	41.9	43 000	MNE	2	14.5	13 800
DZA	27	1.1	2 381 700	NLD	50	120.5	41 500
EGY	2	0.2	1 001 400	NOR	100	26.0	385 200
ESP	25	5.0	504 600	POL	20	6.4	312 700
EST	10	22.1	45 200	ROU	18	7.6	238 400
FIN	44	13.0	338 400	RUS	3	0.0	17 075 400
FRA	103	15.3	674 800	SRB	10	12.9	77 500
GBR	165	67.4	244 800	SVK	19	38.8	49 000
GRC	21	15.9	132 000	SVN	8	39.4	20 300
HRV	21	37.2	56 500	SWE	69	15.3	450 300
HUN	32	34.4	93 000	TUN	14	8.6	163 600
IRL	7	10.0	70 200	TUR	11	1.4	814 600
ITA	72	23.9	301 300	UKR	5	0.8	603 700

**Tabelle B.3:** In den Spalten „Anzahl Meld.“ ist die durchschnittliche Anzahl (gerundet) der verfügbaren stündlichen Stationsmeldungen nach Ländern aufgeschlüsselt angegeben. Die Spalten „Fläche in km<sup>2</sup>“ beinhalten die auf 100 km<sup>2</sup> gerundeten Flächenangaben innerhalb der betrachteten Europadomäne. Für Länder, die nicht vollständig in dieser Domäne liegen, wurde der beteiligte Flächenanteil geschätzt und kursiv markiert. Die Spalten „Anz. pro 100 000 km<sup>2</sup>“ enthalten die durchschnittliche Anzahl der stündlichen Meldungen, bezogen auf 100 000 km<sup>2</sup>. Die Zählung erfolgte, sobald für einen Parameter eine Meldung zur Verfügung stand, ungeachtet ihrer Qualität. Die Angabe der Länderkürzel entspricht der ISO-3166-1 ALPHA-3 Kodierung, die von der Internationalen Organisation für Normung (englisch: *International Organization for Standardization*, ISO) geführt wird.

Länderkürzel	Anz. Meld.	Anz. pro 100 000 [km <sup>2</sup> ]	Fläche in [km <sup>2</sup> ]	Länderkürzel	Anz. Meld.	Anz. pro 100 000 [km <sup>2</sup> ]	Fläche in [km <sup>2</sup> ]
ALB	1	3.5	28 700	LBY	4	1.1	<i>360 000</i>
AUT	123	146.6	83 900	LIE	1	500	200
BEL	9	29.5	30 500	LTU	6	9.2	65 300
BGR	12	10.8	111 000	LUX	1	38.5	2 600
BIH	5	9.8	51 200	LVA	4	6.2	64 600
BLR	13	7.2	<i>180 000</i>	MAR	12	7.1	<i>170 000</i>
CHE	51	123.5	41 300	MDA	1	3.0	33 800
CZE	32	40.6	78 900	MKD	3	11.7	25 700
DEU	208	58.2	357 100	MLT	1	333.3	300
DNK	18	41.9	43 000	MNE	2	14.5	13 800
DZA	27	4.5	<i>600 000</i>	NLD	50	120.5	41 500
EGY	2	1.3	<i>150 000</i>	NOR	100	50	<i>200 000</i>
ESP	25	5.0	504 600	POL	20	6.4	312 700
EST	10	22.1	45 200	ROU	18	7.6	238 400
FIN	44	38.3	<i>115 000</i>	RUS	3	4.6	<i>65 000</i>
FRA	103	15.3	674 800	SRB	10	12.9	77 500
GBR	165	67.4	244 800	SVK	19	38.8	49 000
GRC	21	15.9	132 000	SVN	8	39.4	20 300
HRV	21	37.2	56 500	SWE	69	23	<i>300 000</i>
HUN	32	34.4	93 000	TUN	14	8.6	163 600
IRL	7	10.0	70 200	TUR	11	7.3	<i>150 000</i>
ITA	72	23.9	301 300	UKR	5	1.7	<i>290 000</i>
				<b>total</b>	<b>1363</b>	<b>20.8496</b>	<b>6 537 300</b>

## B.2 Anteil von akzeptierten, korrigierten und verworfenen Meldungen

**Tabelle B.4:** Die Stationsmeldungen für den Parameter reduzierter Druck im Jahr 2010 wurden von der VERA-QC beurteilt. Die gerundeten Anteile der akzeptierten, korrigierten und verworfenen Stationsmeldungen werden, nach Ländern aufgeschlüsselt, gegenübergestellt. Die Prozentangaben beziehen sich auf die Summe der, für diesen Parameter und diesen Zeitraum zur Verfügung stehenden Meldungen des jeweiligen Landes. Anteile größer 1.5 [%] sind fett hervorgehoben. Die Angabe der Länderkürzel entspricht der ISO-3166-1 ALPHA-3 Kodierung, die von der Internationalen Organisation für Normung (englisch: *International Organization for Standardization*, ISO) geführt wird.

Länderkürzel	akzept. [%]	korrig. [%]	verw. [%]	Länderkürzel	akzept. [%]	korrig. [%]	verw. [%]
ALB	15.06	84.85	0.09	LBY	15.03	80.97	<b>3.99</b>
AUT	15.37	83.38	1.25	LIE	43.52	56.48	0.00
BEL	17.93	81.83	0.24	LTU	22.15	77.27	0.58
BGR	16.00	83.41	0.59	LUX	4.72	95.28	0.00
BIH	10.96	88.60	0.45	LVA	21.49	78.51	0.00
BLR	24.67	75.24	0.09	MAR	15.99	82.49	<b>1.52</b>
CHE	13.41	85.90	0.70	MDA	5.65	94.28	0.07
CZE	15.75	83.46	0.79	MKD	15.81	82.91	1.28
DEU	17.25	82.11	0.64	MLT	18.28	81.72	0.00
DNK	18.84	77.56	<b>3.61</b>	MNE	22.30	77.50	0.20
DZA	18.25	79.76	<b>1.99</b>	NLD	23.43	76.52	0.05
EGY	10.41	89.50	0.08	NOR	22.01	77.57	0.43
ESP	17.80	82.14	0.07	POL	20.89	78.11	1.00
EST	26.28	73.59	0.13	ROU	20.82	78.91	0.26
FIN	17.25	81.82	0.93	RUS	21.98	77.06	0.96
FRA	17.22	82.42	0.36	SRB	18.08	80.89	1.03
GBR	14.49	84.70	0.81	SVK	17.85	82.07	0.07
GRC	16.78	83.05	0.17	SVN	17.32	79.10	<b>3.58</b>
HRV	13.88	85.64	0.48	SWE	18.83	80.45	0.72
HUN	19.39	80.44	0.17	TUN	17.75	81.34	0.91
IRL	15.25	84.70	0.05	TUR	16.78	82.78	0.44
ITA	12.18	86.88	0.95	UKR	19.20	80.75	0.05
				<b>total</b>	<b>17.07</b>	<b>82.24</b>	<b>0.69</b>

**Tabelle B.5:** Die Stationsmeldungen für den Parameter potentielle Temperatur im Jahr 2010 wurden von der VERA-QC beurteilt. Die gerundeten Anteile der akzeptierten, korrigierten und verworfenen Stationsmeldungen werden, nach Ländern aufgeschlüsselt, gegenübergestellt. Die Prozentangaben beziehen sich auf die Summe der, für diesen Parameter und diesen Zeitraum zur Verfügung stehenden Meldungen des jeweiligen Landes. Anteile größer 1.5 [%] sind fett hervorgehoben. Die Angabe der Länderkürzel entspricht der ISO-3166-1 ALPHA-3 Kodierung, die von der Internationalen Organisation für Normung (englisch: *International Organization for Standardization*, ISO) geführt wird.

Länderkürzel	akzept. [%]	korrig. [%]	verw. [%]	Länderkürzel	akzept. [%]	korrig. [%]	verw. [%]
ALB	99.14	0.76	0.10	LBY	96.01	3.03	0.96
AUT	97.66	2.05	0.29	LIE	97.34	2.21	0.44
BEL	98.72	0.98	0.30	LTU	97.59	1.84	0.58
BGR	96.62	2.96	0.41	LUX	99.81	0.19	0.00
BIH	94.48	5.33	0.19	LVA	97.50	2.36	0.14
BLR	98.25	1.70	0.05	MAR	88.93	8.93	<b>2.15</b>
CHE	96.33	3.25	0.43	MDA	99.51	0.42	0.07
CZE	97.75	1.93	0.32	MKD	96.93	2.81	0.26
DEU	98.49	1.19	0.33	MLT	99.75	0.25	0.00
DNK	95.86	1.91	<b>2.23</b>	MNE	88.60	11.22	0.18
DZA	93.93	5.45	0.62	NLD	98.88	1.1	0.02
EGY	88.14	11.78	0.08	NOR	94.01	5.57	0.42
ESP	94.49	5.17	0.34	POL	97.58	1.45	0.98
EST	97.18	2.67	0.15	ROU	96.17	3.24	0.59
FIN	95.13	4.19	0.68	RUS	97.96	1.87	0.16
FRA	97.05	2.68	0.27	SRB	97.37	2.36	0.26
GBR	97.39	1.91	0.71	SVK	96.15	3.70	0.15
GRC	96.92	2.88	0.20	SVN	96.93	2.83	0.24
HRV	95.07	4.35	0.58	SWE	96.59	3.11	0.30
HUN	97.92	2.04	0.04	TUN	94.72	5.04	0.24
IRL	98.63	1.29	0.08	TUR	92.29	7.38	0.33
ITA	93.41	5.94	0.65	UKR	98.72	1.22	0.06
				<b>total</b>	<b>96.71</b>	<b>2.87</b>	<b>0.42</b>

**Tabelle B.6:** Die Stationsmeldungen für den Parameter Windgeschwindigkeit im Jahr 2010 wurden von der VERA-QC beurteilt. Die gerundeten Anteile der akzeptierten, korrigierten und verworfenen Stationsmeldungen werden, nach Ländern aufgeschlüsselt, gegenübergestellt. Die Prozentangaben beziehen sich auf die Summe der, für diesen Parameter und diesen Zeitraum zur Verfügung stehenden Meldungen des jeweiligen Landes. Anteile größer 1.5 [%] sind fett hervorgehoben. Die Angabe der Länderkürzel entspricht der ISO-3166-1 ALPHA-3 Kodierung, die von der Internationalen Organisation für Normung (englisch: *International Organization for Standardization*, ISO) geführt wird.

Länderkürzel	akzept. [%]	korrig. [%]	verw. [%]	Länderkürzel	akzept. [%]	korrig. [%]	verw. [%]
ALB	99.39	0.60	0.01	LBY	95.16	3.59	1.25
AUT	95.86	3.45	0.69	LIE	97.23	2.62	0.15
BEL	98.07	1.31	0.62	LTU	96.66	1.73	<b>1.61</b>
BGR	93.38	5.66	0.96	LUX	99.72	0.28	0.00
BIH	96.97	2.63	0.40	LVA	97.84	2.14	0.02
BLR	98.10	1.86	0.04	MAR	88.18	9.11	<b>2.71</b>
CHE	95.86	3.68	0.46	MDA	99.51	0.28	0.21
CZE	95.93	3.57	0.49	MKD	93.05	4.59	<b>2.36</b>
DEU	96.51	2.82	0.68	MLT	98.42	1.56	0.03
DNK	91.46	5.03	<b>3.52</b>	MNE	93.46	6.46	0.08
DZA	90.45	7.97	<b>1.57</b>	NLD	96.12	3.61	0.27
EGY	84.56	14.18	1.26	NOR	91.34	7.87	0.79
ESP	95.05	4.58	0.36	POL	96.45	2.24	1.30
EST	97.39	2.47	0.13	ROU	96.25	3.28	0.47
FIN	93.68	5.75	0.56	RUS	98.31	0.76	0.93
FRA	93.84	5.59	0.57	SRB	95.72	3.68	0.60
GBR	94.15	4.96	0.90	SVK	92.22	7.50	0.28
GRC	92.38	7.07	0.55	SVN	96.33	3.52	0.16
HRV	91.08	7.63	1.29	SWE	94.24	5.15	0.61
HUN	96.92	3.03	0.05	TUN	92.61	7.10	0.29
IRL	96.31	3.55	0.14	TUR	93.21	6.38	0.41
ITA	90.61	8.36	1.03	UKR	97.61	2.33	0.07
				<b>total</b>	<b>94.49</b>	<b>4.80</b>	<b>0.71</b>

## B.3 Anteil der Gross Errors

**Tabelle B.7:** Für den Parameter reduzierter Bodendruck stehen im Jahr 2010 in der betrachteten Europa-Domäne 10 210 113 Meldungen zur Verfügung. Davon wurden 33 528 Meldungen als Gross Error identifiziert. In dieser Tabelle ist in den jeweils zweiten Spalten die Gesamtanzahl der Meldungen für den betrachteten Parameter (zu allen Terminen) angegeben, in den jeweils dritten Spalten ist die absolute Anzahl der erkannten Gross Errors je Land angeführt. Die Prozentangaben in den jeweils vierten Spalten beziehen sich auf die Anzahl der Meldungen, die von den einzelnen Ländern für den Parameter reduzierter Bodendruck zur Verfügung stehen (vgl. zweite Spalte). Gross-Error Anteile größer 0.5 [%] sind fett gedruckt. Die Angabe der Länderkürzel entspricht der ISO-3166-1 ALPHA-3 Kodierung, die von der Internationalen Organisation für Normung (englisch: *International Organization for Standardization*, ISO) geführt wird.

Länderkürzel	Anzahl Meldungen	Anzahl GE	GE in %	Länderkürzel	Anzahl Meldungen	Anzahl GE	GE in %
ALB	7777	1	0.01	LBY	28 376	61	0.22
AUT	841 101	1836	0.22	LIE	8129	0	0.00
BEL	75 105	182	0.24	LTU	45 884	265	<b>0.58</b>
BGR	97 022	297	0.31	LUX	8648	0	0.00
BIH	38 548	102	0.27	LVA	26 500	0	0.00
BLR	36 413	30	0.08	MAR	71 241	1030	<b>1.45</b>
CHE	330 826	494	0.15	MDA	2834	2	0.07
CZE	255 575	379	0.15	MKD	21 706	62	0.29
DEU	1 750 577	7812	0.45	MLT	8658	0	0.00
DNK	35 884	23	0.06	MNE	19 928	10	0.05
DZA	170 064	1254	<b>0.74</b>	NLD	301 412	119	0.04
EGY	12 195	0	0	NOR	612 652	262	0.04
ESP	216 340	100	0.05	POL	173 664	519	0.30
EST	84 966	6	0.01	ROU	153 858	402	0.26
FIN	339 857	1584	0.47	RUS	21 608	17	0.08
FRA	861 992	2517	0.29	SRB	84 200	197	0.23
GBR	1 338 070	8233	<b>0.62</b>	SVK	157 072	93	0.06
GRC	180 816	274	0.16	SVN	23 417	28	0.12
HRV	180 025	678	0.38	SWE	410 574	717	0.18
HUN	268 295	79	0.03	TUN	124 273	1026	<b>0.83</b>
IRL	60 449	1	0.00	TUR	93 917	339	0.36
ITA	584 503	2478	0.42	UKR	45 164	19	0.04
				<b>total</b>	<b>10 210 115</b>	<b>33 528</b>	<b>0.33</b>

**Tabelle B.8:** Für den Parameter potentielle Temperatur stehen im Jahr 2010 in der betrachteten Europa-Domäne 10 512 621 Meldungen zur Verfügung. Davon wurden 38 628 Meldungen als Gross Error identifiziert. In dieser Tabelle ist in den jeweils zweiten Spalten die Gesamtanzahl der Meldungen für den betrachteten Parameter (zu allen Terminen) angegeben, in den jeweils dritten Spalten ist die absolute Anzahl der erkannten Gross Errors je Land angeführt. Die Prozentangaben in den jeweils vierten Spalten beziehen sich auf die Anzahl der Meldungen, die von den einzelnen Ländern für den Parameter potentielle Temperatur zur Verfügung stehen (vgl. zweite Spalte). Gross-Error Anteile größer 0.5 [%] sind fett gedruckt. Die Angabe der Länderkürzel entspricht der ISO-3166-1 ALPHA-3 Kodierung, die von der Internationalen Organisation für Normung (englisch: *International Organization for Standardization*, ISO) geführt wird.

Länderkürzel	Anzahl Meldungen	Anzahl GE	GE in %	Länderkürzel	Anzahl Meldungen	Anzahl GE	GE in %
ALB	7777	3	0.04	LBY	24 582	55	0.22
AUT	990 608	2698	0.27	LIE	8129	36	0.44
BEL	75 915	228	0.30	LTU	45 884	265	<b>0.58</b>
BGR	99 282	409	0.41	LUX	8648	0	0.00
BIH	38 548	70	0.18	LVA	26 500	38	0.14
BLR	36 413	18	0.05	MAR	72 241	1428	<b>1.98</b>
CHE	410 398	1725	0.42	MDA	2835	2	0.07
CZE	258 315	817	0.32	MKD	21 737	23	0.11
DEU	1 772 071	5766	0.33	MLT	8658	0	0.00
DNK	35 847	35	0.10	MNE	19 964	35	0.18
DZA	172 725	739	0.43	NLD	301 412	46	0.02
EGY	12 195	5	0.04	NOR	628 202	2148	0.34
ESP	216 340	739	0.34	POL	173 664	482	0.28
EST	84 966	71	0.08	ROU	153 859	902	<b>0.59</b>
FIN	339 554	1097	0.32	RUS	21 608	29	0.13
FRA	868 473	2232	0.26	SRB	83 538	188	0.23
GBR	1 348 230	9351	0.69	SVK	157 072	234	0.15
GRC	180 816	346	0.19	SVN	23 672	20	0.08
HRV	179 898	1040	<b>0.58</b>	SWE	409 073	861	0.21
HUN	272 223	113	0.04	TUN	124 286	253	0.20
IRL	60 449	18	0.03	TUR	93 917	306	0.33
ITA	594 957	3730	<b>0.63</b>	UKR	47 150	27	0.06
				<b>total</b>	<b>10 512 621</b>	<b>38 628</b>	<b>0.37</b>

**Tabelle B.9:** Für den Parameter äquivalentpotentielle Temperatur stehen im Jahr 2010 in der betrachteten Europa-Domäne 10 484 136 Meldungen zur Verfügung. Davon wurden 67 859 Meldungen als Gross Error identifiziert. In dieser Tabelle ist in den jeweils zweiten Spalten die Gesamtanzahl der Meldungen für den betrachteten Parameter (zu allen Terminen) angegeben, in den jeweils dritten Spalten ist die absolute Anzahl der erkannten Gross Errors je Land angeführt. Die Prozentangaben in den jeweils vierten Spalten bezieht sich auf die Anzahl der Meldungen, die von den einzelnen Ländern für den Parameter äquivalentpotentielle Temperatur zur Verfügung stehen (vgl. zweite Spalte). Gross-Error Anteile größer 0.5 [%] sind fett gedruckt. Die Angabe der Länderkürzel entspricht der ISO-3166-1 ALPHA-3 Kodierung, die von der Internationalen Organisation für Normung (englisch: *International Organization for Standardization*, ISO) geführt wird.

Länderkürzel	Anzahl Meldungen	Anzahl GE	GE in %	Länderkürzel	Anzahl Meldungen	Anzahl GE	GE in %
ALB	7774	202	<b>2.60</b>	LBY	24 565	231	<b>0.94</b>
AUT	989 723	4649	0.47	LIE	8129	20	0.25
BEL	75 883	309	0.41	LTU	45 874	463	<b>1.01</b>
BGR	99 282	779	<b>0.78</b>	LUX	8648	0	0.00
BIH	38 534	174	0.45	LVA	26 499	97	0.37
BLR	36 398	81	0.22	MAR	72 200	2143	<b>2.97</b>
CHE	410 263	1821	0.44	MDA	2833	1	0.04
CZE	258 286	1581	<b>0.61</b>	MKD	21 719	108	<b>0.50</b>
DEU	1 771 022	7742	0.44	MLT	8658	58	<b>0.67</b>
DNK	35 679	108	0.30	MNE	19 964	117	<b>0.59</b>
DZA	172 047	2986	<b>1.74</b>	NLD	301 397	219	0.07
EGY	12 195	2	0.02	NOR	623 043	3070	<b>0.49</b>
ESP	216 328	2112	<b>0.98</b>	POL	173 517	645	0.37
EST	84 947	305	0.36	ROU	153 859	1049	<b>0.68</b>
FIN	339 283	1703	<b>0.50</b>	RUS	21 604	106	0.49
FRA	868 006	4495	<b>0.52</b>	SRB	83 524	407	0.49
GBR	1 344 358	10687	<b>0.80</b>	SVK	157 064	532	0.34
GRC	180 809	1844	<b>1.02</b>	SVN	23 636	113	0.48
HRV	179 881	1710	<b>0.95</b>	SWE	398 477	1957	0.49
HUN	271 855	442	0.16	TUN	124 249	1556	<b>1.25</b>
IRL	60 346	86	0.14	TUR	93 916	955	<b>1.02</b>
ITA	590 723	10109	<b>1.71</b>	UKR	47 139	85	0.18
				<b>total</b>	<b>10 484 136</b>	<b>67 859</b>	<b>0.65</b>

# Anhang C



# Andrea Steiner

---

## Persönliche Daten

Name Andrea Steiner  
Geburtsdatum 07.06.1985  
Geburtsort Villach  
Nationalität Österreich

## Ausbildung

2005 – heute **Universität Wien**, *Diplomstudium der Meteorologie und Geophysik*, erster Studienabschnitt abgeschlossen: Dezember 2006, gewählter Studiengang: Meteorologie; Sommersemester 2008: ERASMUS - Auslandssemester Madrid  
1999 – 2004 **Handelsakademie**, *Spittal/Drau*, Ablegung der Reifeprüfung: Juni 2004  
1995 – 1999 **Hauptschule**, *Möllbrücke*  
1991 – 1995 **Volksschule**, *Kolbnitz*

## Berufstätigkeit

10/2009 – heute **Universität Wien**, *Wissenschaftliche Mitarbeiterin im Projekt: 'MANNO'*  
03/2010 – 07/2010 **Universität Wien**, *Tutorin für die Lehrveranstaltung Wetterbesprechung*  
07/2009 – 08/2009 **UBIMET GmbH**, *Praktikum - Abteilung Forschung & Entwicklung und Abteilung Meteorologie*  
03/2009 – 07/2009 **Universität Wien**, *Tutorin für die Lehrveranstaltung Mikrometeorologie*  
09/2008 – 09/2008 **MeteoServe Wetterdienst GmbH**, *Praktikum - Projektmitarbeit*  
08/2008 – 08/2008 **Austro Control Abteilung Meteorologie**, *Praktikum - Verifikation eines Radarextrapolationsalgorithmus*  
2006 – 2008 **LernQuadrat, privat**, *Nachhilfeunterricht*  
09/2004 – 10/2005 **Au-Pair Aufenthalt in Amerika**

Sommermonate 2000 – 2007 **diverse Ferialtätigkeiten**, 07/2000: *Praktikum Hypo Alpe-Adria-Bank Villach*; 07/08 2001 & 07/08 2002: *Betreuung der Bergstation Kreuzeckbahn, Kolbnitz*; 08/2003: *Betreuung von Jugendlichen innerhalb eines zweiwöchigen Intensivsprachkurses in Lignano - Italien*; 2001 - 2007: *Tennisunterricht (Kinder - Erwachsene)*

---

## Sprachkenntnisse

Deutsch Muttersprache  
Englisch fließend  
Italienisch Maturaniveau  
Spanisch Grundkenntnisse

---

## EDV-Kenntnisse

Programmiersprachen Matlab, Fortran  
Betriebssysteme MS Windows, MacOSx, Linux  
Anwendungen MS Office, L<sup>A</sup>T<sub>E</sub>X, CoreIDRAW

---

## Publikationen

**Steinacker, R., D. Mayer, und A. Steiner, 2011**, *Data Quality Control Based on Self Consistency*, Monthly Weather Review, doi: 10.1175/MWR-D-10-05024.1

**Steiner, A., D. Mayer, und R. Steinacker, 2011**, *VERA-QC\*: operational application, challenges, and how to cope with them*, Posterpräsentation 11th EMS / 10th ECAM 2011, Berlin

**Steiner, A., D. Mayer, und R. Steinacker, 2011**, *Die Qualitätskontrolle von VERA: operationelle Anwendung, Herausforderungen im komplexen Gelände und Erweiterungen der VERA-QC*, Posterpräsentation 4. Österreichischer MeteorologInnenntag 2011, Klagenfurt, ausgezeichnet mit dem dritten Posterpreis der Österreichischen Meteorologischen Gesellschaft