



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Eidgenössisches Departement des Innern EDI  
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz



# Werkzeuge für die Qualitätskontrolle und die Bearbeitung von Meteorologischen Daten

Datenarchitekturdokument 1f3

Madlaina Perl  
Christian Häberli  
Estelle Grüter  
Marc Musa  
Ursula Küng  
Markus Abbt

Version 2.0

Freigabedatum: 03. Juli 2009

# Dokumentenverwaltung

## Verteiler:

Name, Vorname	Firma/Prozess/Projekt	Funktion	Änderungsberechtigung
Ch. Häberli	MeteoSchweiz	PL DWH / Autor	ja
M. Perl	MeteoSchweiz	Autorin	ja
E. Grüter	MeteoSchweiz	Autorin	ja
U. Küng	MeteoSchweiz	Autorin	ja
M. Abbt	MeteoSchweiz	Autor	ja
M. Musa	MeteoSchweiz	Autor	ja
Kernteam DWH	MeteoSchweiz	Umsetzung	nein
Projektteam DWH	MeteoSchweiz	Umsetzung	nein
PAB DWH	MeteoSchweiz	Beratungsgremium	nein
OCMD Gremium	MeteoSchweiz	Beratungsgremium	nein

## Versionenkontrolle:

ErstellerIn	Freigabedatum	Version	Tätigkeit	Anmerkungen
diverse	14.03.01	0.1	Erstellen	
diverse	30.08.02	0.2	Überarbeiten	
M. Perl	22.07.05	0.3	Komplette Überarbeitung des Dokuments	Zusammenfassung verschiedener Dokumente
M. Perl	12.12.05	0.4	Überarbeitung aufgrund von Reviewkommentaren	Beseitigen von Inkonsistenzen
M. Perl	19.01.07	0.5	Erarbeitung der Kommentare aus dem DABES Workshop vom 15. Dezember 2005 mit KD und TM	
M. Perl	02.02.07	0.6	Überarbeitung aufgrund von Reviewkommentaren von C OCMD	an PAB DWH und OCMD Gremium zur Review
M. Perl	08.0.07	0.7	Einarbeitung der Reviewkommentare aus dem PAB DWH	
M. Perl§	28.03.07	1.0	Überarbeitung Abbildung 1	Genehmigt durch PA DWH
M. Musa	28.05.09	1.1	Einarbeitung Vorgehen historische Plausibilitätstests. Aufdatierung Kapitel 5 und 14. Erfassung neues Kapitel 9 Plausibilitätstests für Profildaten..	zum Review PAB DWH
M. Musa	19.06.09	1.2	Einarbeitung Kommentare PAB DWH	
M. Del Dangan	23.06.09	1.3	Uebertragen von Framemaker auf Word	
M. Musa	03.07.09	2.0	Erstellung Version 2.0	Genehmigt durch PA DWH

## *Offene Punkte*

<b>was</b>	<b>wer</b>	<b>bis wann</b>
In der nächsten Version des Dokuments werden die Kapitel 3.1.1 und 13 mit Beiträgen aus dem Prozess TM ergänzt	lab	12.2009
Bei einer Überarbeitung der Architektur sind die Skripts (Augenbeobachtungstest und Schneefüllskript) in die bestehenden Tools zu integrieren.	mam	offen
Abklärung der Plausibilitätstests, welche im Nudging Prozess des COSMO durchgeführt werden. -> sti 09.2009	sti	09.2009

# Inhaltsverzeichnis

<b>KAPITEL 1</b>	<i>Management Summary</i> .....	10
1.1	Datenbearbeitungsketten der MeteoSchweiz .....	10
1.2	Sollqualität .....	10
1.3	Qualitätsprüfung .....	10
1.4	Bearbeitungswerkzeuge .....	11
<b>KAPITEL 2</b>	<i>Begriffsdefinitionen</i> .....	12
<b>KAPITEL 3</b>	<i>Einleitung</i> .....	15
3.1	Arten der Qualitätsprüfung .....	15
3.1.1	<i>Technische Real-Time-Qualitätskontrolle</i> .....	15
3.1.2	<i>Fachliche Real-Time-Qualitätskontrolle</i> .....	16
3.1.3	<i>Längerfristige Qualitätskontrolle</i> .....	16
3.2	Prozesskette Datenbearbeitung .....	16
3.3	Inhalt des Dokuments .....	18
<b>KAPITEL 4</b>	<i>Zielvorgaben für Datenqualität und Verfügbarkeit</i> .....	21
4.1	Empfehlungen der WMO.....	21
4.2	Situation an der MeteoSchweiz .....	22
4.2.1	<i>Vorgaben aus dem Messkonzept 2010</i> .....	23
4.2.2	<i>Umsetzung in der Datenbearbeitung</i> .....	23
4.2.3	<i>Datenverfügbarkeit</i> .....	24

---

<b>KAPITEL 5</b>	<i>Plausibilitätstests und Qualitätsinformation</i> .....	25
5.1	Plausibilitätstests und Verwendung in den Applikationen .....	25
5.2	Qualitätsinformation .....	28
<b>KAPITEL 6</b>	<i>Qualitätskontrolle auf den Sammelsystemen</i> .....	29
6.1	Sammelsysteme der automatischen Messnetze .....	29
6.2	Erfassungssysteme mit integrierter Qualitätskontrolle .....	30
<b>KAPITEL 7</b>	<i>Calc&amp;Check Modul</i> .....	32
7.1	Aufgabenumschreibung .....	32
7.2	Positionierung im System .....	33
7.3	Ablauf .....	33
<b>KAPITEL 8</b>	<i>Assimilation COSMO</i> .....	35
<b>KAPITEL 9</b>	<i>Prüfung und Mutation - Automatische Bearbeitung (PuMAB)</i> .....	36
9.1	Aufgabenumschreibung und Positionierung im System .....	36
9.2	Ablauf .....	37
9.3	Testfehlertabelle.....	38
<b>KAPITEL 10</b>	<i>Prüfung und Mutation - Profildaten (CALRAS Plausibilitätskontrolle)</i> .....	39
10.1	Aufgabenumschreibung und Positionierung im System .....	39
10.2	Ablauf .....	39

---

<b>KAPITEL 11</b>	<i>Prüfung und Mutation - Interaktive Bearbeitung (PuMIB)</i> .....	40
11.1	Aufgabenbeschreibung Positionierung im System .....	40
11.2	Steuerung .....	41
11.3	Bearbeitung der Testfälle .....	42
11.4	Freie Datenwahl .....	42
11.5	Auswertung PuMIB .....	42
<b>KAPITEL 12</b>	<i>MANDEFpro</i> .....	43
12.1	Einleitung .....	43
12.2	Grobablauf .....	43
<b>KAPITEL 13</b>	<i>VERAQC</i> .....	44
13.1	Einleitung .....	44
13.2	Grobablauf .....	44
13.3	Das Analyseverfahren .....	45
<b>KAPITEL 14</b>	<i>Homogenisierung</i> .....	47
14.1	Einleitung .....	47
14.2	Grundsätze .....	47
14.3	Verlängerte und synthetische Reihen .....	48
14.4	THOMAS .....	48
14.4.1	<i>Bearbeitung</i> .....	48
14.4.2	<i>Resultate</i> .....	49
<b>KAPITEL 15</b>	<i>Lückentests</i> .....	50
15.1	Lücken bei Augenbeobachtungen .....	50

---

---

15.1.1 Aufgabenbeschreibung und Positionierung im System .....	50
15.1.2 Ablauf .....	50
15.1.3 Beispiel-Output .....	51
15.2 Lücken bei den Schneebeobachtungen .....	51
15.2.1 Aufgabenbeschreibung und Positionierung im System .....	51
15.2.2 Bearbeitung .....	51
15.2.3 Ermitteln von fehlenden Werten .....	52
15.2.4 Ablauf .....	52
<b>KAPITEL 16 Geberstörungsmeldungen .....</b>	<b>53</b>
16.1 Beschreibung .....	53
16.2 Ablauf .....	54
16.3 Beispiel einer Geberstörungsmeldung .....	55
<b>KAPITEL 17 Referenzen .....</b>	<b>56</b>
<b>ANHANG A Spezifikation Umsetzung Plausibilitätsinformation zu Testprotokoll</b>	<b>58</b>
A.1 Einführung .....	58
A.2 Details zur Struktur qual_byte .....	61
<b>ANHANG B Daten- und Limitenübergabe .....</b>	<b>62</b>
B.1 Allgemeines .....	62
B.2 Allgemeingültige Bestimmungen .....	62
<b>ANHANG C Bearbeitungsfunktionen .....</b>	<b>63</b>
C.1 Funktion I "rational_spline_method" .....	64
C.2 Funktion II "mean rational_spline_method" .....	65

---

C.3	Funktion III "to treat by PuMIB" .....	66
C.4	Funktion IVa "treatment of values outside Tagbogen" und Funktion IVb "treatment of values outside astr. sunset/rise" .....	67
C.5	Funktion V „Ersetzung der Werte durch Werte von redundanten Messungen“ .....	67
C.6	Anwendung von Funktionen I und II bei intermittierend auftretenden Messlücken (‘Stotterbetrieb’) .....	68
C.7	Detailspezifikation der Bearbeitungsfunktionen.....	68
C.8	Resultate aus dem Bericht AMUT - automatische Dateninterpolation im ANETZ.....	72

## KAPITEL 1 *Management Summary*

### *1.1 Datenbearbeitungsketten der MeteoSchweiz*

MeteoSchweiz betreibt eine Reihe von Werkzeugen, welche die Qualität der meteorologischen Messdaten überprüfen, automatische und interaktive Bearbeitungen zulassen und die Arbeitsschritte dokumentieren.

Jeder Datensatz auf den Datenbanken des MeteoSchweiz Data Warehouse ist gekennzeichnet mit Informationen zur Qualität und zum Bearbeitungszustand.

In der vorliegenden Dokumentation finden sich detaillierte Informationen zu den Qualitätskontrollen der Bodenstationsdaten und den Sondierungsdaten sowie den Bearbeitungswerkzeugen, mit denen die Qualität der meteorologischen Datenerhebung gesichert wird. Es werden die Konzepte der Plausibilitätsprüfungen, Bearbeitungsfunktionen und verschiedene Applikationen vorgestellt.

Das vorliegende Dokument ist ein Ausführungskonzept zum QMS Ablauf „Qualitätskontrolle Meteorologische Daten“.

### *1.2 Sollqualität*

Die WMO (World Meteorological Organisation) empfiehlt eine Reihe von Qualitätskriterien für meteorologische Messwerte und Datenreihen [11].

Das Ziel der Datenbearbeitung ist konsistente und vollständige Datenreihen in einer definierten Qualität zu erhalten.

### *1.3 Qualitätsprüfung*

Die Daten werden geprüft auf Über/Unterschreitung von Grenzwerten, Konsistenz innerhalb der Messreihe, auf die Variabilität und auf die räumliche Konsistenz. Für alle Parameter sind Grenzen und mögliche Variabilitäten vordefiniert. Diese können abhängig von der räumlichen Position sein.

Das Ziel dieser Prüfungen ist, eine vorgeschriebene Qualität der Daten aufrecht zu erhalten und die Rohdaten wo nötig zu korrigieren. Das vorliegende Dokument beschreibt, mit welchen Bearbeitungswerkzeugen diese Qualität der Daten erreicht wird.

## *1.4 Bearbeitungswerkzeuge*

Mit Werkzeugen zur Überprüfung und Bearbeitung der meteorologischen Daten werden verschiedenen Qualitätslevel erreicht, bis hin zu homogenen Klimareihen.

Es existieren automatische und interaktive Bearbeitungsprogramme. Für Bodenstationen ist im Messkonzept [2] vorgegeben, welche Messgrößen die interaktive Bearbeitung durchlaufen. Die automatische Bearbeitung umfasst grundsätzlich alle Daten, kann aber auch eingeschränkt werden.

## KAPITEL 2 *Begriffsdefinitionen*

AnalyticalDB	Datenbank, welche v.a. aggregierte Werte enthält und auf Datenabfragen optimiert ist. Sie dient auch als Basis für applikationsspezifische Datamarts.
AWS	Sammelsystem der ANETZ Stationen. Fragt die Daten des automatischen Netzes im 10 Minuten Intervall ab.
Calc&Check Modul	Berechnungs- und Überprüfungsmodul. Präprozessor zum Data Warehouse System.
CALRAS	The Comprehensive Alpine Radiosonde Dataset
CDAS/NIMDAS	Sammelzentrale der Daten des SwissMetNet
DWH-Dienst	Systemteil des Data Warehouse, der auf dem Datenbankserver implementiert ist und über eine Informatica Power Center Session gesteuert wird.
EZEN	Sammelzentrale der Daten des automatischen Ergänzungsnetzes (ENET)
Ferneingabe	Eingabesystems zur Erfassung der Daten von konventionellen Klimastationen oder AERO-Beobachtungsstationen. Vorgängersystem des OBS
GTS	Global Telecommunication System. Rechner auf dem das MHS und das Calc&Check Modul läuft. Der Name kommt von Global Telecommunication System, das für die weltweite Verteilung der meteorologischen Daten an der WMO zuständig ist.
KontextDB	Die Kontextdatenbank (DWH1 Schema) enthält alle Steuerinformationen des DWH zu den Stationen, Parametern, Berechnungen, zu den Testverfahren und weiteren Geschäftsabläufen.
Limiten	Physikalische Limiten (harte Limiten) beschreiben die Grenze zu physikalisch unmöglichen Messwerten. Klimatologische Limiten (weiche Limiten) beschreiben den Grenzwert, bei dem 95% der Messwerte innerhalb eines definierten Messintervalls liegen. Ausreisser werden kontrolliert.

Lückentest	Kornshellskript zum Test, ob Lücken bestehen in den Messreihen mit Terminwerten (Bodendaten, Augenbeobachtungen, Messdaten der konventionellen Klimastationen und Aero-Stationen)
Mandef	Manuelle Datenerfassung Programm zum Erfassen der Daten von phänologischen Stationen, Pollenstationen oder Niederschlagsstationen.
Messauflösung	Zeitliche Granularität der Messungen / Beobachtungen (Sekunden, Minuten, 10min, 30min, Stunde, Tag, Monat, Jahr)
Meteoblock	Gemessene Daten einer Station und eines gewissen Zeitfensters (für 10-min. Daten: 33h, von 16H Tag-2 bis 1H Tag), die zur Plausibilitätsprüfung bzw. Mutation aus der WorkDB ausgezogen werden.
MHS	Message Handling System. Läuft auf den GTS Rechnern und wird von WMAN gesteuert.
MI	Die Mutationsinformation beschreibt die Methode, mit der ein Ersatzwert für einen fehlenden oder fehlerhaften Wert bestimmt wurde.
NIME	Niederschlagsmessstationen
Norm90	Projekt zur Erstellung von klimatologischen Normwerten.
OBS	Teilprojekt des SwissMetNet zur Modernisierung des Eingabesystems der Augenbeobachtungen (Westa O = Wetterstation mit Augenbeobachtungen).
PI	Die Plausibilitätsinformation beschreibt die Plausibilität eines Wertes wie sie aufgrund der für jede Kombination von Station und Parameter festgelegten Kombination von Plausibilitätstests und Limiten festgestellt wird.
PuMAB	Modul zur Prüfung und Mutation mit automatischer Bearbeitung bei Lücken. Das Modul kennzeichnet alle Messwerte mit einer Qualitätsinformation und einem Reihenzustand.
PuMIB	Interaktive Applikation zur Prüfung und Mutation mit interaktiver Bearbeitung.
QI	Mit der Qualitätsinformation wird jeder Messwert gekennzeichnet. Sie setzt sich zusammen aus der Plausibilitätsinformation (PI), der Mutationsinformation (MI), dem Reihenzustand (RZ) und dem Homogenitätszustand (HZ).
Raw Data Repository	Datenbankinstanz mit unbearbeiteten Messdaten
RZ	Der Reihenzustand beschreibt die Bearbeitungsstufe, welche ein Messwert durchlaufen hat.

Schneefüllskript	Kornshellskript zum Füllen von fehlenden Schneemessungen in den Sommermonaten, während denen die Beobachter keinen Schnee eingeben müssen.
THOMAS	Tool for Homogenization of Monthly Data Series
VERAQC	Vienna Enhanced Resolution Analysis - Quality Control. Räumliche Qualitätskontrolle für meteorologische Daten basierend auf einem Thin-Plate-Spline-Algorithmus.
WMAN	Weather Man. Software, die das MHS betreibt.
WMO	World Meteorological Organization; Meteorologische Weltorganisation der UNO
WorkDB	Datenbank, welche für die Datenbearbeitung optimiert ist. Die Work DB "Surface" besitzt folgende Schnittstellen: Laden und Exportieren der Bulletins, Bereitstellung der Messwerte für die Anwendung PuMAB (Meteoblock Schnittstelle), Einlesen der von PuMAB geprüften und korrigierten bzw. interpolierten Messwerte, Bereitstellung einer Leseschnittstelle zu PuMIB über JDBC sowie Versorgen der Ladeprozesse für die Analytical DB.

## KAPITEL 3 *Einleitung*

Dieses Dokument beschreibt das Konzept der Werkzeuge zur Qualitätskontrolle und zur Bearbeitung der meteorologischen Daten der MeteoSchweiz. Es ist kein technisches Handbuch der Datenbearbeitungswerkzeuge, sondern zeigt den Hintergrund der gesamten Datenbearbeitungskette auf und befasst sich mit der Qualitätskontrolle im allgemeinen, der angestrebten Sollqualität der Daten und den vorhandenen Bearbeitungswerkzeugen.

Als Teil der Datenarchitektur der MeteoSchweiz [13] dient das Dokument als übergeordnetes Konzept für folgende Dokumente:

- Qualitätsbezeichnung der Messwerte im MeteoSchweiz Data Warehouse System [5]
- Specifications of Inter-Parameter Constraints at MeteoSwiss (Internal Consistancies) [4]

sowie der Dokumentation der Applikationen ([6],[7],[8],[14],[15]).

### *3.1 Arten der Qualitätsprüfung*

Es gibt verschiedene Arten der Qualitätskontrolle. Einige werden in Echtzeit, d.h. laufend nach Eintreffen der Daten durchgeführt. Weitere Qualitätsprüfungen und Datenbearbeitungen erfolgen regelmässig, zum Beispiel täglich oder monatlich und werden nur für ausgewählte Messdaten, also für Daten von wichtigen Stationen und wichtigen Messgrössen, durchgeführt.

Während der Datenerfassung sollen in der Messkette zwei Typen von Qualitätskontrollen durchgeführt werden: die technische und die fachliche Real- Time Kontrolle.

#### *3.1.1 Technische Real-Time-Qualitätskontrolle*

Bei der technischen Real-Time Qualitätskontrolle werden primär technische Qualitätskontrollen der beteiligten Erfassungs-, Übermittlungs- und IT-Systeme durchgeführt, z.B. die Überprüfung von Statusanzeigen oder Spannungen. Sie umfasst eine Vielfalt von normalerweise automatischen Tests, die erlauben, die technische Funktionalität der beteiligten Systeme zu beurteilen und allenfalls Massnahmen einzuleiten. Bei Erfassungssystemen z.B. werden die unaufbereiteten Roh-Messungen (teilweise vor der Umwandlung der Messgrössen in ihre physikalischen Einheiten) überprüft und daraus Statusinformationen abgeleitet.

### 3.1.2 Fachliche Real-Time-Qualitätskontrolle

Nach der Umwandlung in meteorologische und klimatologische Grössen (Temperatur, Feuchtigkeit u.a.) werden die erfassten Messgrössen in der Messkette mit geeigneten Prozeduren einer fachlichen Real-Time Qualitätskontrolle unterzogen. Sie sollen eine automatische Beurteilung und Bezeichnung der Glaubwürdigkeit der Messwerte erlauben und allenfalls ermöglichen, geeignete Massnahmen einzuleiten (z.B. Unterdrückung oder Flaging bei der Auslieferung). Diese Qualitätskontrolle ist automatisiert und erstreckt sich über die gesamte Datenbearbeitungskette (Logger, Datensammelzentrale, Datenvermittlungssystem oder Ladesystem ins Data Warehouse, siehe dazu auch Kapitel 3.2). Sie umfasst die automatische Beurteilung und Bezeichnung der Plausibilität der Messwerte, sowie deren automatische und interaktive Bearbeitung. Dazu werden verschiedene Filterungskriterien benutzt wie Minima, Maxima, Variabilität oder Standardabweichung.

### 3.1.3 Längerfristige Qualitätskontrolle

Bei der längerfristigen Qualitätskontrolle werden die Testverfahren komplexer. Sie werden nur jährlich oder alle paar Jahre durchgeführt.

- Komplexe Tests mit engeren Limiten
- Vergleiche mit Referenzen
- Vergleiche mit anderen Messsystemen, die dieselben Grössen messen
- Zeitreihenanalysen
- Räumliche Konsistenztests
- Vergleich mit Vorhersagewerten aus numerischen Modellen

## 3.2 Prozesskette Datenbearbeitung

Die Bearbeitung der meteorologischen Messdaten folgt dem fünfstufigen generischen Transformationsprozess [12].

- *Level 1 Daten* werden an der Messstation produziert indem die Sensordaten (z.B. elektrische Spannung, Zählimpulse) in physikalische Einheiten umgewandelt werden, die für die weitere klimatologische und meteorologische Anwendung sinnvoll sind (z.B. Temperatur, Luftdruck, Niederschlagsmenge). Wenn möglich werden Fehlwerte eingefügt und gewisse Sensoren liefern zusätzlich eine Statusinformation.
- *Level 1 zu Level 2:* Die Level 1 Daten der verschiedenen Messstationen werden von Sammelsystemen zentralisiert. In einem weiteren Schritt werden auf bestimmte Daten Kalibrationspolynome angewendet. Mit einfachen Qualitätskontrollen (Limitentests, Variabilitätstests) werden die Daten auf offensichtliche Fehler geprüft. Ausreisser werden gekennzeichnet. Die Level 1 Daten sollen später im sogenannten ‚Raw Data Repository‘ abgelegt werden. Dieses dient als Datenquelle für den Fall, dass bestimmte Daten im Nachhinein mit verbesserten Kalibrationspolynomen noch einmal transformiert werden sollen (sogenannte Re-Evaluation). Ein Beispiel hierfür ist die Re-Evaluation einer Zeitreihe von Radiosondierungsdaten mit einer angepassten Funktion für die Strahlungskorrektur.
- *Level 2 zu Level 3:* Die Daten aus den verschiedenen Sammelsystemen werden integriert und in die Arbeitsdatenbanken des Data Warehouse Systems geladen. Gleichartige Daten aus verschiedenen Sammelsystemen (z.B. Daten aus verschiedenen Bodenstationsnetzen) werden einer einheitlichen Qualitätskontrolle unterworfen. Die

in diesem Transformationsschritt angewandten Plausibilitätstests (Limitentests, Konsistenztests) sind rigoroser und vielfältiger als jene in den Sammelsystemen. Aus plausiblen Werten werden in diesem Schritt abgeleitete Grössen (z.B. Dampfdruck, auf Meereshöhe reduzierter Luftdruck) berechnet. Level 3 Daten werden entsprechend gekennzeichnet, gegebenenfalls aggregiert und in die analytischen Datenbanken propagiert, wo sie den Anwendern und Kunden zur Verfügung stehen.

- *Level 3 zu Level 4:* Dieser Transformationsschritt ist der eigentliche Datenbereinigungsschritt von meteorologischen und klimatologischen Daten. In Erweiterung zur Qualitätskontrolle in den vorangehenden Transformationsschritten werden in diesem Schritt auch die zeitliche und räumliche Konsistenz der Daten kontrolliert (inklusive Variabilitätstests, Konsistenztests, Lückentests, Gebermeldungen, OBS). Korrekturen (automatisch oder interaktiv) werden angebracht, wo dies sinnvoll möglich ist. Der Transformationsschritt von Level 3 zu Level 4 kann mehrfach durchlaufen werden. Die verschiedenen Bearbeitungsschritte müssen nachvollziehbar sein (Historisierung der Korrekturläufe).
- *Level 4 zu Level 5:* Ausgewählte Daten aus Level 4 werden mit Hilfe von statistischen Verfahren und Kontextinformationen (insbesondere Stationsgeschichte) homogenisiert. Unter Homogenisierung wird in diesem Kontext das Entfernen aller nicht natürlichen Änderungen in den Daten (z.B. Stationsverschiebungen, Instrumentenaufstellung) verstanden, um die raum- zeitliche Konsistenz der Daten hinsichtlich veränderter Messbedingungen sicherzustellen. Resultat des Homogenisierungsvorgangs sind Zeitreihen von Homogenisierungswerten (Beträge oder Faktoren), die zu den zugehörigen Level 4 Daten addiert bzw. damit multipliziert die homogene Datenreihe, d.h. die Level 5 Daten, ergeben.

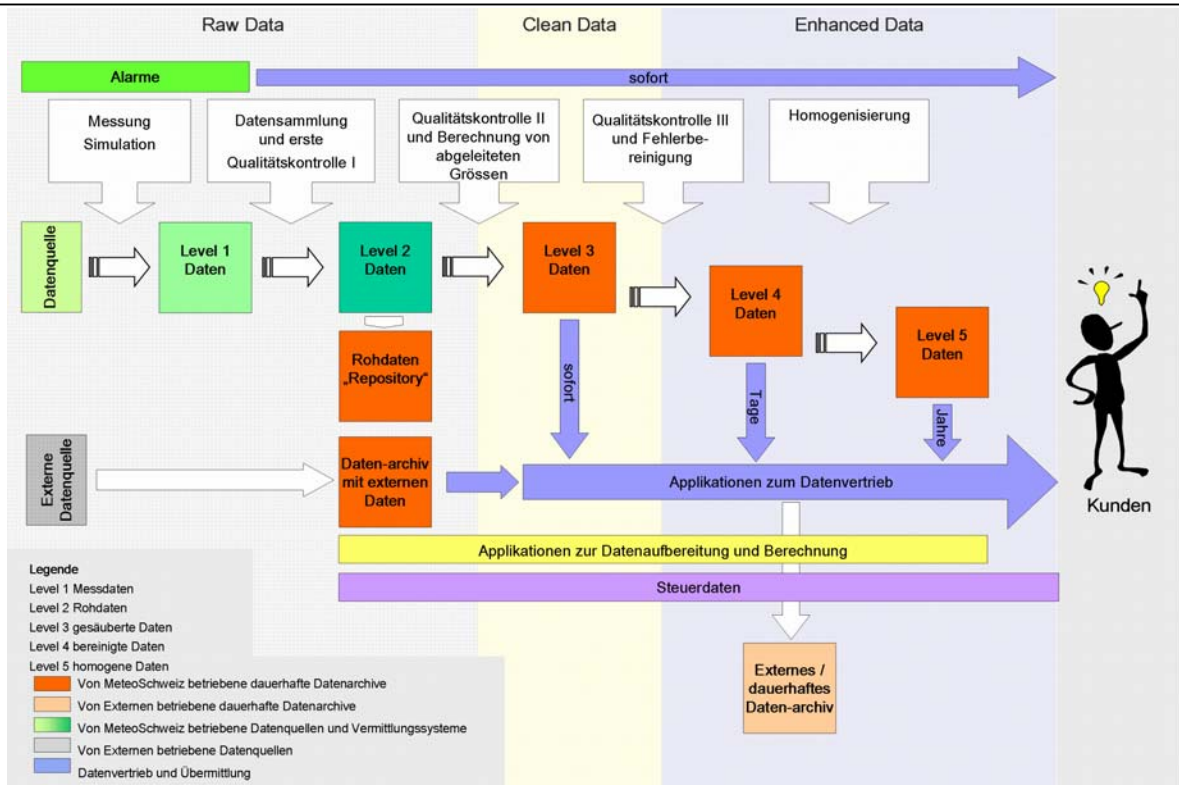


Abbildung 1. Prozesskette der Datenbearbeitung von der Messstation bis zur Homogenisierung. [1b stark überarbeitet]

### 3.3 Inhalt des Dokuments

Dieses Dokument behandelt die Qualitätsprüfung und die Bearbeitung der Daten von Level 2 bis Level 5. Es soll allen Anwendern und Betreuern der Applikationen 'Calc&Check', 'PuMAB' (Prüfung und Mutation automatische Bearbeitung), 'PuMIB' (Prüfung und Mutation, interaktive Bearbeitung), Lückentest, Schneeskript, Gebermeldungen, VERAQC (Tool zur räumlichen Konsistenzprüfung) und THOMAS (Tool zur Homogenisierung monatlicher Datenreihen) einen generellen Überblick verschaffen (siehe auch Abbildung 2).

Die oben erwähnten Applikationen dienen der Qualitätsprüfung von Bodenstationsdaten. Für die Qualitätsprüfung der Sondierungsdaten von Level 3 zu Level 4 steht die CALRAS Qualitätsprüfung zu Verfügung. Diese wird gleich nach dem Kapitel zum PuMAB beschrieben.

Für diese Applikationen gelten einige generelle Grundsätze:

- Die automatische Bearbeitung (Datenmutation) wird durch Konfiguration in der KontextDB unter Berücksichtigung der Qualitätsinformationen gesteuert.

- Die interaktive Mutation wird durch den Bearbeiter durchgeführt. Ihm stehen Hilfsinformationen wie Qualitätsinformation und Testprotokoll zur Verfügung. Das Testprotokoll wird nach erfolgter Bearbeitung gelöscht.
- Mutationen werden wenn möglich auf der Messauflösung durchgeführt. Mutationen an Aggregaten können mit einer speziellen Applikation vorgenommen werden. Dabei werden die mutierten mit der sogenannten ‚Aggregierungssperre‘ für erneutes Aggregieren aus nicht bearbeiteten Daten gesperrt.

Das Dokument klammert die organisatorischen Verantwortlichkeiten auf der Messkette aus.

Änderungen an den bestehenden Qualitätskontrollen und Bearbeitungsabläufen können über den QMS Ablauf „Änderungsmanagement für meteorologische Daten“ angemeldet werden.

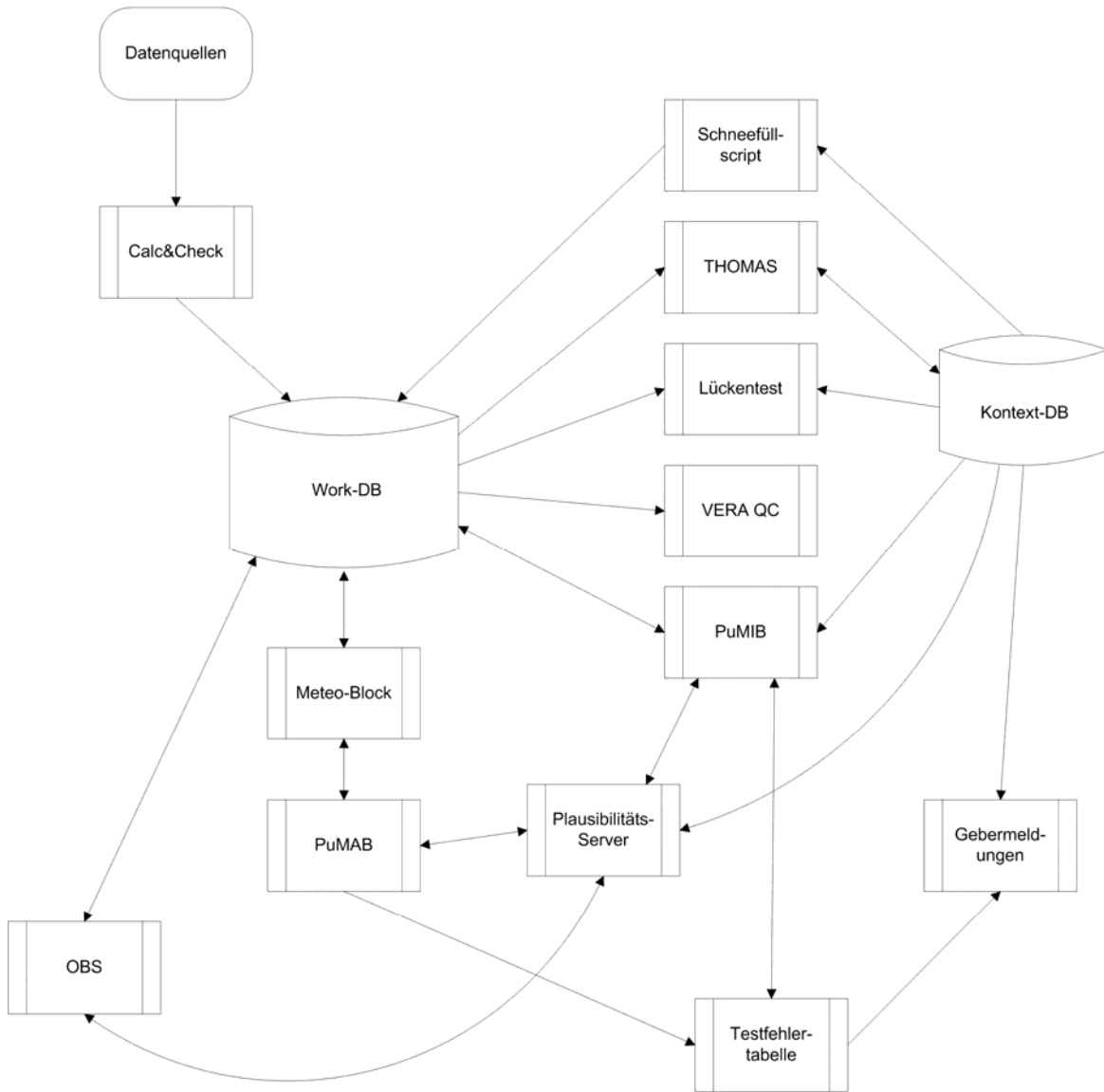


Abbildung 2. Prinzipschema der Datenbearbeitung für Bodenstationsdaten

## KAPITEL 4 *Zielvorgaben für Datenqualität und Verfügbarkeit*

### 4.1 *Empfehlungen der WMO*

Die Empfehlungen der WMO zur Qualität der Messdaten stammen aus „WMO 100: Leitfaden der klimatologischen Praxis“ [11]. Neben Vorgaben zur Datenerhebung und Datenspeicherung, werden auch Empfehlungen zur Kontrolle der Datenqualität gemacht.

Zusammenfassend empfiehlt die WMO, dass bei der Datenerhebung alle Daten geprüft werden müssen. Die Qualitätskontrollen sollen bei jeder Umrechnung der Daten erfolgen. Das Ziel der Qualitätskontrollen ist, Fehler in den Datenquellen möglichst rasch zu entdecken. Es gibt keine quantitativen Vorgaben zur resultierenden Qualität der Daten, sondern mehr zu den Methoden der Qualitätsprüfung.

Eine erste Qualitätskontrolle soll bei der Datenquelle (z. B. vom Beobachter) vorgenommen werden. Es sollen die Exaktheit und Verlässlichkeit der erhobenen Daten, sowie die Kohärenz zwischen den ermittelten Grössen geprüft werden. Die Prüfung der Daten hat am zentralen Sammelort zu erfolgen und schliesst auch die Prüfung auf Lücken in den erhobenen Datenreihen mit ein.

- Fehler identifizieren  
Lücken erkennen und kennzeichnen
- Verifikation der inneren Kohärenz der Daten  
Es gibt verschiedenste Zusammenhänge z. B. zwischen den Messdaten und den Beobachtungen (beispielsweise Temperatur/Taupunkt, Sichtweite/ Luftfeuchtigkeit), die verschiedene Kohärenztests ermöglichen.
- Verifikation der physikalischen und klimatologischen Limiten  
Verschiedene Messgrössen können nur bestimmte Werte annehmen (Windrichtung), für anderen können klar definierte Grenzen festgelegt werden (Luftdruck, Temperatur).
- Verifikation der zeitlichen Kohärenz  
Einzelne Werte müssen auch immer in einer Zeitreihe betrachtet werden.
- Verifikation der räumlichen Kohärenz  
Für die räumliche Prüfung der Daten müssen klimatologisch sinnvolle Gruppen gebildet werden.

Idealerweise resultiert ein räumlich zeitlich kohärenter Datensatz mit genau definierten physikalischen und klimatologischen Grenzen, der keine Lücken beinhaltet. [11] und [18] Suppl. No. 4 (VII.1997)

---

Zur Datenverfügbarkeit an den RBSN Stationen unterscheidet die WMO zwei verschiedene Ziele

1. TRQ - Target requirements: Gewünschte Zielverfügbarkeit
2. MRQ - Minimum requirements: Grenzbereich für die Verfügbarkeit einer RBSN Station

	<b>TRQ Surface</b>	<b>MRQ Surface</b>
Parameter Bodenstationen	Druck Temperatur Wind Feuchtigkeit Niederschlag aktuelles Wetter Sicht Bewölkung Wolkenuntergrenze	Druck Temperatur Feuchtigkeit
Parameter Upper-Air Stationen	Druck  Temperatur  Feuchtigkeit  Wind	mindestens einen der angegebenen Parameter:  Druck  Temperatur  Feuchtigkeit  Wind
Verfügbarkeit der Daten	95-100%	50%

*Tabelle 1. Tabelle 1. Datenverfügbarkeit an Regional Basic Synoptic Network Stationen gemäss WMO - No. 544 [19] (Amendments to the Manual on GOS. Part II, 6.1.2)*

Wird eine Station bei der WMO als RBSN Station angemeldet, so muss die geforderte Zielverfügbarkeit (TRQ) angestrebt werden.

## 4.2 Situation an der MeteoSchweiz

In den Datenaufbereitungsketten der MeteoSchweiz durchlaufen alle Daten mehrere Qualitätskontrollen. Mit jeder Bearbeitungsstufe (Level 1 bis Level 5) werden die Daten getestet oder bearbeitet. Der Bearbeitungszustand eines Datensatzes kann jederzeit anhand des Reihenzustandes (series\_state\_nr), des Bearbeitungsinformation (mut\_info\_nr) und anhand der Plausibilitätsinformation (plaus\_info\_nr) abgelesen werden.

#### *4.2.1 Vorgaben aus dem Messkonzept 2010*

Bei der Ausarbeitung des Messkonzept 2010/07 wurden Überlegungen zum Qualitätsmanagement bei der Erfassung von meteorologischen und klimatologischen Daten in Boden- und Atmosphärensystemen gemacht. [2]

- Die initialen Massnahmen haben das Ziel, die Messsysteme in einem angemessenen technisch-wissenschaftlichen Umfeld zu platzieren und deren langfristige Qualität zu sichern. Zu den initialen Massnahmen werden das Messnetzkonzept 2010 selber und die daraus folgenden Umsetzungskonzepte sowie die Planung der weiteren Massnahmen des Qualitätsmanagements gerechnet.
- Die Stabilität einer Datenerfassungskette kann nur dann sichergestellt werden, wenn die notwendigen präventiven Routine-Massnahmen organisiert sind. Beispiele für solche präventive Routine-Massnahmen sind die periodischen technischen Unterhaltsarbeiten (z.B. an Messstationen oder beteiligten IT- Mitteln), die Sicherung des Ausbildungsstandes der beteiligten Personen, die Sicherstellung der notwendigen Mittel und die Sicherstellung der notwendigen Pikettorganisation.
- Das komplexe Zusammenspiel der vielen Einzelelemente einer Messkette muss mit geeigneten Massnahmen periodisch vom Anfang bis zum Ende ausgetestet werden. Beispiele für solche periodische Kontrollen sind die Überprüfung der Vermittlung von Alarmen als einer der kritischsten Punkte einer operationellen Messkette.

#### *4.2.2 Umsetzung in der Datenbearbeitung*

Bis zum Level 3 durchlaufen alle Messwerte der MeteoSchweiz eine oder mehrere Qualitätskontrollen. Die Sollqualität bezieht sich auf die Level 4 Daten und ist pro Station und Parameter im Messkonzept 2010/07 I festgelegt [2].

Die interaktive Bearbeitung wird für klimatologisch wichtige Stationen und Parameter der Level 3 durchgeführt und hat zum Ziel homogenisierte lange Klimareihen herzustellen..

Die geschätzte Datenqualität reicht von 99.3% korrekter und lückenloser Messwerte bei den interaktiv bearbeiteten Parametern und Stationen bis zu 98.2% bei den unbearbeiteten Stationen und Parameter.

<b>Bearbeitung/Art der Messstation</b>	<b>Automatisch Calc&amp;Check</b>	<b>Automatisch PuMAB</b>	<b>Interaktiv PuMIB</b>
automatische Stationen des SwissMetNet <sup>1</sup>	alle	alle	92
Beobachtungsstationen (OBS)	automatische Plausprüfung bei der Eingabe der Daten		nur Bewölkung und Schneehöhen
SMART	alle	-	-
IMIS	alle	-	-
Strassenwetter	-	-	-
NIME	Kontrolle der Eingaben direkt im Eingabetool		
Phänologie / Pollen	Kontrolle der Eingaben direkt im Eingabetool		

Tabelle 2. Bearbeitung automatisch und interaktiv bei den Stationsnetzen [2]

#### 4.2.3 Datenverfügbarkeit

An der MeteoSchweiz werden die Daten der wichtigsten automatischen Stationen (72 ANETZ und SMN und 3 ENET) alle 10 Minuten gesammelt, geprüft und auf dem zentralen Data Warehouse gespeichert. Die Daten von weiteren 41 automatischen Stationen des Ergänzungsnetzes (ENET) werden einmal in der Stunde gesammelt, geprüft und gespeichert. Dazu kommen die Beobachtungsdaten von 25 konventionellen Klimastationen, welche drei Mal am Tag direkt an die Zentrale geliefert werden, wo sie ebenfalls geprüft und dann gespeichert werden.

---

<sup>1</sup>. Vorläufig auch ANETZ, ENET und konventionelle Klimastationen

## KAPITEL 5 *Plausibilitätstests und Qualitätsinformation*

### 5.1 *Plausibilitätstests und Verwendung in den Applikationen*

Die Plausibilitätstests werden vornehmlich auf der Messauflösung (10-Min, Terminwerte) durchgeführt. Bei bestimmten Grössen sind auch Tests auf berechneten Grössen vorgesehen (z.B. NIME-Daten zusammen mit aggregierten Tagesniederschlägen).

Die detaillierten Testspezifikationen sind festgehalten im Dokument „Specifications of Inter-Parameter Constraints at MeteoSwiss“ [4].

Die Resultate der Tests werden innerhalb der Tools auf einem Testprotokoll gespeichert. Die Auswertung (Generalisierung) des Testprotokolls ergibt die im DWH abgelegte Plausibilitätsinformation (vgl. [5]).

<b>Gruppe</b>	<b>CC</b>	<b>PuMAB</b>	<b>PuMIB</b>	<b>Funktion von</b>
physikalische Limitentests	x	x	x	Station, Parameter
klimatologische Limitentests	x	x	x	Station, Parameter, Bezugszeitraum
maximale Variation		x	x	Station, Parameter
minimale Variation		x	x	Station, Parameter
Konsistenztests	x	x	x	Station, Test, Parameter

Tabelle 3. *Durchgeführte Tests in Calc&Check, PuMAB und PuMIB*

Die Durchführung der Plausibilitätstests beruht in allen Anwendungen auf derselben Methodenbibliothek und denselben Steuerinformationen.

Die Methodenbibliothek enthält die implementierten Tests (C-Funktionen) und die entsprechenden Schnittstellen. Alle Applikationen greifen auf dieselbe Methodenbibliothek zu.

Die Steuerinformationen werden der KontextDB entnommen. Darin wird festgelegt, für welche Stationen und Parameter welche Tests durchgeführt werden und welche Limiten gelten. Die KontextDB enthält den aktuell benötigten Testbestand. Ebenfalls sind die Limiten auf die heutige Messtechnik angepasst. Dies bedeutet, dass keine Informationen über historisch durchgeführte Tests oder Limiten der KontextDB entnommen werden können. Nicht mehr angewendete Tests sind aber im Dokument [4] aufgeführt.

Für die Prüfung historischer Daten bestehen daher folgende Einschränkungen:

- PuMAB läuft nur auf der Installation der aktuellen Station (nicht Masterinstallation). Bei der Umstellung ANETZ zu SMN laufen die Tests nur noch auf dem Datenbestand der SMN-Station.
- Verändertes Messprogramm hat verändertes Testset zur Folge. Z.B. Helligkeit wird nicht mehr, langwellige Strahlung wird neu gemessen.. Gleiches gilt für die Veränderung von Limiten.

Typ	Beschreibung	Beispiel/Kommentar
Vollständigkeitstest	Sind die gewünschten Mess-/Beobachtungswerte vorhanden?	Wird für Augenbeobachtungen vom Lückentest und für 10' Werte vom PuMAB gemacht.
Limitentest	Liegt der einzelne Mess-/Beobachtungswert zwischen oberen und unteren Grenzwerten?	T2m = 50 <sup>0</sup> > Limite → PI setzen! es wird auf physikalische und klimatologische Limiten (Grenzwerte) getestet
Test auf innere Konsistenz	Liegen zwei oder mehrere Mess-/Beobachtungswerte desselben Messintervalls innerhalb vernünftiger Beziehungen?	Sonnenschein und Blitz im gleichen 10-Min.-Intervall → PI setzen! es werden ca. 150 Tests angewendet
Variabilitätstest	Variieren die Mess-/Beobachtungswerte mehrerer aufeinanderfolgender Messintervalle innerhalb oder ausserhalb vorgegebener Limiten?	Messintervall der Windgeschw.: ...,2,2,1,0,0,0,0,0,0,0,0,0,4,3,6,... →Windmesser ev. eingefroren → PI setzen! es wird auf ‚zu grosse Sprünge‘ und auf ‚zu geringe Variabilität‘ getestet
Räumlicher Vergleichstest	Ist der Unterschied zw. gemessenem und mit einem räumlichen Interpolationsverfahren ermitteltem Wert grösser als ein bestimmter Grenzwert?	Luftdruck Luftfeuchtigkeit potentielle Temperatur  Applikation VERA (Vienna Enhanced Resolution Analysis)
Homogenitätstest	Weist eine Zeitreihe einen natürlichen Verlauf auf?	Störungen durch Instrumenten- wechsel, Standortverschiebungen u.a. (insgesamt 12 Tests)  Applikation THOMAS (Tool for Homogenization of Monthly Data Series)

Tabelle 4. Beschreibung und Beispiele für die Tests

## 5.2 *Qualitätsinformation*

Jedem Messwert auf dem MeteoSchweiz Data Warehouse wird aufgrund der Testresultate eine Qualitätsinformation zugeordnet<sup>2</sup>. Zudem hat jeder Wert eine Kennzeichnung über seinen Bearbeitungszustand<sup>3</sup>.

Die genaue Bedeutung der verwendeten Qualitätsinformation ist im Dokument „Qualitätsbezeichnung der Messwerte im MeteoSchweiz Data Warehouse“ [5] beschrieben.

Die Qualitätsinformation wird auch an die berechneten Werte weitergegeben. Dies erfolgt in Abhängigkeit der Plausibilitätsinformation der Inputwerte. Die genaue Beschreibung ist ebenfalls in [5] zu finden.

---

<sup>2</sup> Als Attribut *plaus\_info\_nr* in den Tabellen *t\_wrk\_surface* und *t\_ana\_surface*.

<sup>3</sup> Als Attribute *mutation\_info\_nr* (Mutationszustand) und *series\_state\_nr* (Reihenanzustand) in den Tabellen *t\_wrk\_surface* und *t\_ana\_surface*.

---

## KAPITEL 6 *Qualitätskontrolle auf den Sammelsystemen*

Die Qualitätskontrolle auf den Sammelsystemen ist eine Kombination von technischer und fachlicher Qualitätskontrolle. Sie bezieht sich auf Level 0 und Level 1 Daten.

### *6.1 Sammelsysteme der automatischen Messnetze*

Die Sammelsysteme rufen die Rohdaten (Level 1) der einzelnen Messstationen ab und versenden diese weiter an das MHS. Auf der Sammelzentrale werden die Messdaten der automatischen Stationen in physikalische Einheiten umgerechnet und einer ersten Qualitätsprüfung unterzogen. Zudem werden die technischen Kennzahlen der Geber geprüft und Defekte an den Gebern festgestellt. Bei Problemen mit einer Wetterstation oder speziellen meteorologischen Verhältnissen werden Extremwertmeldungen ausgesendet (z.B. bei hohen Windspitzen oder erhöhter Radioaktivität).

Die mit meteorologischen Daten durchgeführten Tests und dazu verwendete Limiten entsprechen nicht denjenigen aus der KontextDB des DWH System. Sie dienen primär dazu anhand der Datenwerte fehlerhafte Sensoren (Geber) zu erkennen, die Situation zu analysieren und zu intervenieren (→ Monitoring).

Von der Sammelzentrale werden die Messdaten in Bulletins an das MHS weiterverschickt.

Durchgeführte Test	AWS	EZEN	CDAS/NIMDAS
Netzzustand, Übermittlungs-probleme von Stationen	✓		✓
Ermitteln von defekten Gebern	✓		✓
Limitentests auf den meteorologischen Messdaten <sup>4</sup>	✓	✓	✓
Aussenden von Alarmen bei fehlerhaften Sensoren (z. B. bei fehlender Variabilität in der Messung)	✓		✓
Aussenden von Alarmen beim überschreiten von Extremwerten bei bestimmten Parametern	✓	<sup>5</sup>	✓
Zeitsynchronisation der Stationen		✓	✓
Austausch von Steuerdaten (Koeffizienten) mit dem DWH System		✓ <sup>6</sup>	✓

Tabelle 5. Qualitätskontrollen auf den Sammelsystemen

## 6.2 Erfassungssysteme mit integrierter Qualitätskontrolle

Für die Erfassung der Daten von Beobachtungsstationen bestehen an der MeteoSchweiz verschiedene Systeme. Zum einen können die Beobachter ihre Daten vor Ort selber erfassen und direkt übermitteln, es gibt aber auch Beobachtungsnetze, von denen die Daten im Nachhinein digitalisiert werden.

Bei diesen Daten kann man nicht von Level 1 sprechen, da vor dem Abschicken einer Beobachtung Plausibilitäts-, Konsistenz- und Limitentests durchgeführt werden. Bei der Übermittlung an das System haben diese Daten bereits Level 3 erreicht. Dadurch erhalten sie eine spezielle Kennzeichnung (z.B. spezielle Mutationsinformation bei der manuellen Datenerfassung).

<sup>4</sup> Es werden keine Flags an die Folgesysteme weitergeleitet.

<sup>5</sup> Wird durch die ESTA (Wetterstation des Ergänzungsnetzes) direkt ausgeführt.

<sup>6</sup> Technisch möglich, wird aber nicht eingesetzt

<b>Durchgeführte Test</b>	<b>OBS<sup>7</sup></b>	<b>Manuelle Datenerfassung</b>
Plausibilitätstests bei der Eingabe von Wetterbeobachtungen	✓	
Konsistenztests bei der Eingabe von Wetterbeobachtungen	✓	
Limitentests bei der Eingabe von Messdaten	✓	
Integrierter Kontrollmodus mit Doppelerfassung bei den phänologischen und Pollendaten		✓
Summentests bei der Erfassung der Niederschlagsdaten		✓

*Tabelle 6. Qualitätskontrolle bei den Eingabesystemen der Beobachtungsstationen*

---

<sup>7</sup> *Eingabesystem für Beobachtungsstationen. Arbeitet mit den Steuerdaten aus der Kontextdatenbank des DWH Systems*

## KAPITEL 7 *Calc&Check Modul*

Die meteorologischen Daten gelangen von den Sammelsystemen als Level 2 Daten auf das MHS. Dort durchlaufen sie die Bearbeitung im Calc&Check Modul (Prüf- und Berechnungsmodul). Die von Calc&Check bearbeiteten Daten erreichen Level3, werden an das DWH weitergeleitet und auf der Datenbank gespeichert.

### *7.1 Aufgabenumschreibung*

Im Calc&Check Modul werden die Messdaten den ersten einheitlichen Tests unterzogen, bevor sie an die nachgelagerten Systeme weiterverteilt werden. Aus den angelieferten Werten werden abgeleitete Grössen in der zeitlichen Originalauflösung generiert. Alle Parameter werden im ersten Schritt auf physikalische Limiten getestet. Alle anderen Tests werden nur durchgeführt, wenn kein am Test beteiligter Parameter eine physikalische Limite verletzt hat.

Das Calc&Check Modul kann benutzt werden um ‘ergänzte’ Bulletins zu generieren, welche auch die berechneten Grössen und die detaillierte Plausibilitätsinformation enthalten.

Man kann unterscheiden, ob ein Bulletin nur berechnet (calc) oder berechnet und geprüft (calc&check) wird. Diese Unterscheidung ist wichtig, wenn die ersten Konsistenztests bereits auf der Station bei der Datenerfassung durchgeführt werden (z.B. bei konventionellen Klimastationen, OBS). [6]

## 7.2 Positionierung im System

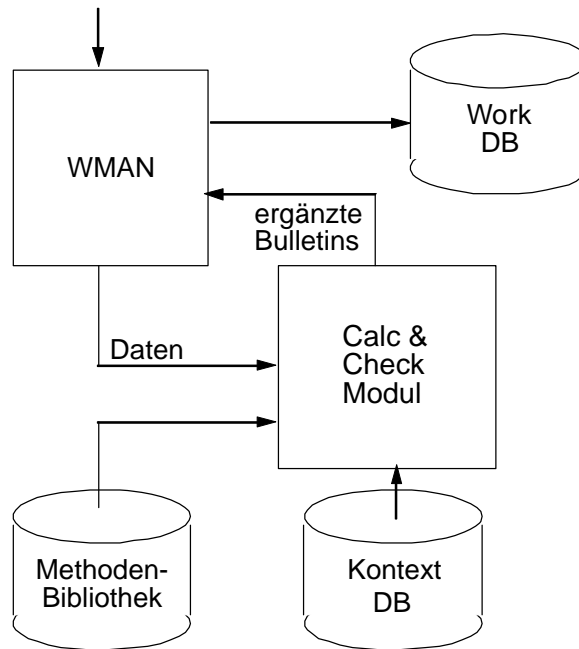


Abbildung 3. Skizze zum Berechnungs- und Plausibilitätsmodul ‚Calc&Check‘

Das Calc&Check Modul läuft auf dem GTS Server, verarbeitet Bulletins die von WMAN angeliefert werden und schickt die mit berechneten Grössen und Qualitätsinformationen ergänzten Bulletins an WMAN zurück. Von dort werden die Bulletins ans DWH geliefert.

Für Berechnungs- und Plausibilitätsprüfung werden dieselben Methoden wie im Berechnungs- und Aggregiermodul bzw. PuMAB und PuMIB verwendet. Die benötigten Steuerinformationen und Stationskenndaten sind in der KontextDB enthalten.

## 7.3 Ablauf

In der Initialisierungsphase werden alle Stationen eingelesen. Beim Aufbau der Stationsliste werden die Limiten und weitere, für die Berechnungen und Tests nötige Stationskoeffizienten zugeordnet.

In der Parameterliste werden alle Parameter eingetragen, die geprüft oder berechnet werden sollen.

Die Limitentests werden ebenfalls in einer Liste eingetragen und sortiert, damit sie beim Aufruf der Funktion `check_pum()` den Stationen und Parametern richtig zugeordnet werden können.

Sämtliche Testfunktionen werden in die Testfunktionsliste eingetragen. Hier sind die Testgruppe und die Methodennummern festgehalten. Mit der Funktion `check_pum()` werden die Funktionen zu einem bestimmten Parameter aufgerufen.

In den Steuerinformationen für die Berechnung werden alle Informationen festgehalten, welche für die Berechnungsfunktionen nötig sind.

Nach der Initialisierungsphase wird die Datenbankverbindung geschlossen. Das Calc&Check läuft danach ohne Verbindung zur Datenbank und schreibt die Testresultate und Berechnungen in Bulletins, die weitergeschickt werden. Werden Veränderungen an den Steuerungsinformationen vorgenommen, muss Calc&Check neu initialisiert werden. [6]

## **KAPITEL 8** *Assimilation COSMO*

Bei der Assimilation ins COSMO Modell werden diverse Plausibilitätsprüfungen der angelieferten Beobachtungsdaten durchgeführt. In der COSMODB werden die Qualitätstests des ECMWF durchgeführt gemäss [21] Kapitel 4.7.3. Im Nudging Prozess des COSMO werden auch Qualitätsprüfungen durchgeführt.

## KAPITEL 9 *Prüfung und Mutation - Automatische Bearbeitung (PuMAB)*

Eingangsdaten für diesen Bearbeitungsschritt sind die Level 3 Daten. Nach der Bearbeitung erreichen ein Teil der bearbeiteten Daten bereits den Level 4 Status, andere werden für die Bearbeitung zu Level 4 vorbereitet und mit PuMIB auf diesen Stand gebracht.

### *9.1 Aufgabenumschreibung und Positionierung im System*

Der Name PuMAB ist die Abkürzung für "**P**rüfung **u**nd **M**utation - **A**utomatische **B**earbeitung". Die automatische Bearbeitung bezieht sich darauf, dass PuMAB versucht, Messlücken und physikalisch unmögliche Werte mittels Bearbeitungsfunktionen (siehe ANHANG C) automatisch zu interpolieren, oder zu bearbeiten. Automatisch heisst nicht, dass PuMAB selber eine Bearbeitungsfunktion wählen kann, sondern diese ist abhängig vom Parameter und von der Länge der Messlücke. Die Informationen hierzu sind in den Steuertabellen in der Datenbank abgelegt.

Die automatische Bearbeitung findet auf der WorkDB statt. Die Daten (Meteoblöcke) werden von der WorkDB gelesen, anschliessend getestet, eventuell mutiert und Änderungen wieder auf die WorkDB zurück geschrieben. Tests und Mutationen werden auf der Stufe der Messauflösung durchgeführt (10-Minuten- und Terminwerte).

Neben dem Schreiben der neuen Meteoblockdatei trägt PuMAB Plausibilitätsfälle, welche vom Sachbearbeiter individuell nachbearbeitet werden müssen, in eine Tabelle in der Datenbank ein. Ebenso werden alle Testverletzungen von Gebermeldungen in diese Tabelle eingetragen.

PuMAB arbeitet sequentiell und führt folgende Aufgabe durch:

- Lesen von Messwerten aus sogenannten Meteoblockdateien
- Erstes Prüfen/Testen der gelesenen Messwerte
- Berechnen von fehlenden Werten und Werten, welche physikalische Limitentests verletzen, durch Interpolationsroutinen oder mittels redundanten Werten
- Erneutes Testen der so entstandenen, vollständigen Zeitreihen von Messwerten
- Schreiben einer neuen Meteoblockdatei mit den Änderungen zum Ursprungsmeteoblockfile für den Import in die WorkDB.

Die zu bearbeitenden Zeitperioden werden vom Reihenzustand definiert. Es werden alle Daten bearbeitet, deren Reihenzustand noch nicht dem Sollzustand entspricht oder deren Reihenzustand Auskunft darüber gibt, dass diese Daten noch nicht automatisch bearbeitet wurden. Der Reihenzustand wird nicht historisiert.

PuMAB wird jede Nacht um 3.30 UTC über einen PowerCenter-Batch gestartet und arbeitet in einer Schleife alle vorhandenen Meteoblockdateien einzeln ab.

PuMAB kann jederzeit manuell gestartet werden. Somit können auch historische Daten getestet und bearbeitet werden.

Die Meteoblockdatei enthält alle Parameter und definierten Tests von Messwerten aus der WorkDB einer bestimmten Station der letzten 52 oder 57 Stunden.

## 9.2 Ablauf

Nachfolgendes Schema gibt den Ablauf in \$DABES\_WORK/src/pumab/pumab.c wieder.

- PuMAB Start  
(Funktion main())
- Lesen der im Verzeichnis \$PUMAB\_OUTMEBL\_PATH vorhandenen Trigger- Dateien zur Steuerung der nachfolgenden Bearbeitung
- Schleife über alle vorhandenen Meteoblockdateien:
  - Lesen der Meteoblockdatei und der zugehörigen Messortdaten aus der Datenbank
  - Testen der gelesenen Meteoblock-Daten
  - Übertragen der Testresultate in die Meteoblockstruktur
  - Bearbeitung der Messlücken in den Meteoblockdaten
    - Schleife über Messfelder
    - Schleife über Parameter
    - Schleife über gefundene Messlücken
  - Testen der bearbeiteten Daten
  - Übertragen der Testresultate in die Meteoblockstruktur
  - Erzeugen der Datenbankeinträge für Messwerte mit Testverletzungen und Schreiben des Meteoblock-Importfiles in \$PUMAB\_INMEBL\_PATH
  - Verschieben des Triggerfiles von \$PUMAB\_OUTMEBL\_PATH nach \$PUMAB\_INMEBL\_PATH (im Testfall \$PUMAB\_TMP\_PATH)
  - ==> zurück zum Schleifenanfang und Bearbeitung der nächsten Meteoblock-Datei.
- PuMAB Schluss

### 9.3 Testfehlertabelle

Für die Steuerung der Datenbearbeitung auf dem DWH nimmt die Tabelle WORK1.T\_PUM\_TEST\_FAILURE eine Schlüssel Rolle ein. PuMAB ist die einzige Applikation, welche in diese Tabelle schreibt. Die Tabelle enthält alle Testverletzungen, die von PuMAB festgestellt werden, das heisst, dass PuMAB alle Resultate der Plausibilitätstest in diese Tabelle schreibt.

Attributname	Null?	Type
REFERENCE_TS	NOT NULL	DATE
INSTALLATION_ID	NOT NULL	NUMBER(8)
PARAMETER_ID	NOT NULL	NUMBER(8)
QC_METHOD_ID	NOT NULL	NUMBER(8)
PUM_OUTPUT_ID	NOT NULL	NUMBER(8)
CREATETIME_DT	NOT NULL	DATE

Tabelle 7. Aufbau der Tabelle work1.t\_pum\_test\_failure

Für jeden Messwert auf der Datenbank mit einer Testverletzung gibt es einen Eintrag in der Tabelle. Daraus ist ersichtlich, welcher der über 200 Tests angesprochen hat (QC\_METHOD\_ID) und für welche Applikation die Bearbeitung des vorliegenden Falles vorgesehen ist (PUM\_OUTPUT\_ID, z.B. PuMIB oder Gebermeldungen). Das System ist auf diese Weise einfach konfigurierbar und sehr flexibel.

## KAPITEL 10 *Prüfung und Mutation - Profildaten (CALRAS Plausibilitätskontrolle)*

Eingangsdaten für diesen Bearbeitungsschritt sind die hochaufgelösten Sondierungen von Payerne sowie die international versendeten Sondierungen TEMP und PILOT, welche bereits teilweise auf den Erfassungssystemen manuell korrigiert wurden, bevor die Daten weitervermittelt wurden.

### *10.1 Aufgabenumschreibung und Positionierung im System*

Als Teil des Projektes CALRAS (The Comprehensive Alpine Radiosonde Dataset) [20] wurde eine komplette Funktionenbibliothek mit Plausibilitätstests und Interpolationsmechanismen entwickelt. Dieser Funktionsumfang wird an der MeteoSchweiz für die automatische Plausibilitätsprüfung und Korrektur von Radiosondendaten verwendet. Die CALRAS Routinen entsprechen vom Umfang denjenigen von PuMAB für die Bodenstationsdaten. Deshalb kann man hier von einem Übergang von Level 3 zu Level 4 sprechen.

### *10.2 Ablauf*

Nachfolgendes Schema gibt den Ablauf der Anwendung der CALRAS Routinen im MeteoSchweiz Data Warehouse System wieder.

- Lesen der Daten aus dem DWH
- Überprüfung der Daten anhand der definierten Plausibilitätstests
- Schreiben der veränderten Messwerte, Plausibilitäts- und Mutationsinformationen ins DWH

Die Plausibilitätsprüfung erfolgt pro Sondierung und wird einmal täglich für definierte Stationen und Termine durchgeführt.

## KAPITEL 11 *Prüfung und Mutation - Interaktive Bearbeitung (PuMIB)*

### *11.1 Aufgabenbeschreibung Positionierung im System*

Die ‚Interaktive Bearbeitung‘ (PuMIB) ist eine Applikation mit grafischer Benutzeroberfläche, mit dem die von PuMAB gekennzeichneten Testverletzungen (Fälle in T\_PUM\_TEST\_FAILURE) bearbeitet werden können. Jeder Eintrag aus der Tabelle wird als ein Fall in PuMIB behandelt. Nach der interaktiven Bearbeitung werden die Änderungen direkt auf die WorkDB geschrieben. Zusätzlich können auch Daten von bestimmten Stationen oder Parametern in der sogenannten freien Datenwahl angeschaut und bearbeitet werden. Dies wird gebraucht, wenn zu einem fragwürdigen Datensatz keine Testverletzung vorliegt oder Lücken bei den Terminwerten gefüllt werden sollen (siehe Kapitel 15.1).

Die Bearbeiter bestimmen mit Hilfe graphischer und tabellarischer Darstellungen sowie mit berechneten Interpolationen die plausibelsten Ersatzwerte für Lücken und falsche Daten jedes Bearbeitungsfalles.

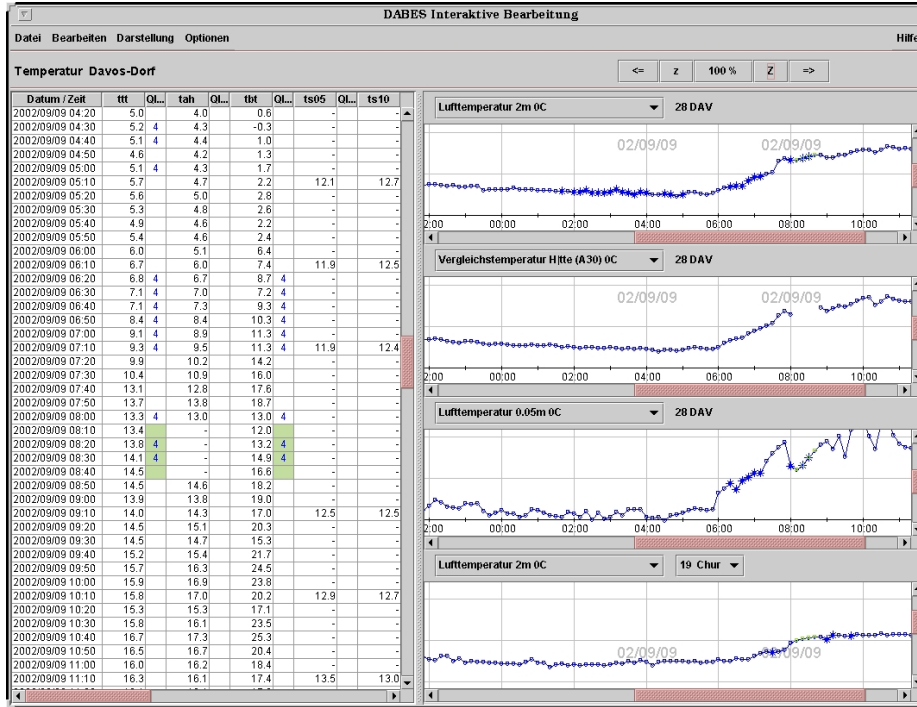


Abbildung 4. Graphical User Interface der interaktiven Bearbeitung

## 11.2 Steuerung

Die Steuerinformationen entnimmt PuMIB aus der KontextDB. In den Tabellen sind Informationen zu den

- physikalischen und klimatologischen Limiten der Parameter
- in PuMIB bearbeitenden Parameter
- verwendeten Parametergruppen
- Vergleichsstationen
- Vergleichsparametern
- bearbeitenden Personen
- und zu den Einstellungen und Darstellungen im GUI abgelegt.

Es fließen die Daten von rund 92 Stationen in die interaktive Bearbeitung ein (siehe dazu [2]).

### *11.3 Bearbeitung der Testfälle*

Alle Tätigkeiten der Bearbeitungspersonen erfolgen im GUI.

- Aus einer Liste können die Testfälle ausgewählt werden. Die Sortierung ist nach Parametergruppe, Station oder Zeit möglich.
- Die Bearbeitung eines Falles erfolgt in der tabellarischen oder graphischen Oberfläche. Die Datenreihen von Vergleichsinstrumenten oder Vergleichsstationen können als Hilfen eingeblendet werden.
- Werte können gelöscht, ersetzt, interpoliert, als zweifelhaft gekennzeichnet oder als richtig bestätigt werden.
- Nach dem Abschluss der Bearbeitung werden wieder Tests durchgeführt und allenfalls wird ein neuer Testfall geschrieben.
- Ein Fall gilt als definitiv abgeschlossen, wenn keine Tests mehr verletzt sind.

### *11.4 Freie Datenwahl*

Über die freie Datenwahl können jederzeit Datensätze selber zusammengestellt und angeschaut werden. Es sind alle Stationen und Parameter aufrufbar, für die eine Bearbeitung mit PuMIB vorgesehen ist, das heisst es muss einen entsprechenden Eintrag geben in den Kontextdaten.

### *11.5 Auswertung PuMIB*

Nach dem Abschluss der Bearbeitung wird der Testfall aus der „Pendenzenliste“ gestrichen.

Aus dem Testprotokoll werden neue Qualitätsinformationen generalisiert:

- die Plausibilitätsinformation wird angepasst
- die Mutationsinformation wird angepasst
- der Reihenzustand wird auf ‘IB absolviert’ gesetzt (Ausnahme erzwungener un abgeschlossenes Verlassen eines Falles).

Die korrigierten Werte inklusive der Qualitätsinformationen werden auf die WorkDB zurückgeschrieben.

## KAPITEL 12 *MANDEFpro*

Die Applikation PUMIB wurde entwickelt für die Bearbeitung von 10 Minuten Daten. Gemäss den Vorgaben des Messkonzept 2010/07 I werden längere Datenlücken nicht mehr auf 10 Minutenbasis sondern je nach Datenlücke auf Tages- oder Monatsbasis interpoliert. Dazu und auch für die Bearbeitung der historischen Stundenwerte zwischen 1981 und 2004 ermöglicht die Applikation MANDEFpro das Einfüllen neuer Daten und das Korrigieren bestehender Daten.

### 12.1 Einleitung

Die Applikation MANDEFpro beinhaltet neben der Mutation von Daten auf Stunden-, Tages- oder Monatsstufe auch die Funktionalitäten für die Eingabe der NIME-, Phäno- und Pollendaten. Die Applikation ist als WEB-Applikation implementiert wo sich der Benutzer einloggen muss.

### 12.2 Grobablauf

Das Verfahren für die Interpolation eines Datensatzes erfolgt über X Schritte:

1. Der Datensatz wird vorändig interpoliert. MANDEFpro bietet keine Interpolationshilfe an und wird nur für die Eingabe verwendet.
2. Der Benutzer wählt die Station, Parameter und Station aus.
3. MANDEFpro bestimmt anhand der gemachten Auswahl die korrekten InstallationsID.s
4. Der Benutzer gibt den Datensatz mit MANDEFpro ein. Der Datensatz wird auf T\_WRK\_SURFACE abgespeichert.
5. Zusätzlich wird automatisch ein entsprechender Eintrag für die Aggregierungssperre gesetzt.

## KAPITEL 13 *VERAQC*

Für die räumliche Qualitätskontrolle von Luftdruck, Luftfeuchtigkeit und potentieller Temperatur wird an der MeteoSchweiz das System VERAQC eingesetzt. Die Analyse erfolgt mit Level 3 Daten ist jedoch im Gegensatz zu den Prüfungs- und Mutationssystemen (PuMAB und PuMIB) in der Lage, längere Zeitreihen zu analysieren.

Es gibt noch keinen Warnmechanismus, mit dem auf die von VERA erkannten Problemfälle aufmerksam gemacht werden kann.

### *13.1 Einleitung*

VERAQC (Vienna Enhanced Resolution Analysis Quality Control) ist ein mesoskaliges Echtzeit Verfahren zur Prüfung von meteorologischen Daten und wurde am Institut für Meteorologie und Geophysik der Universität Wien entwickelt. Es wird eingesetzt, um für meteorologische Messnetze die räumliche Verteilung von Luftdruck, Temperatur, Wind oder Niederschlag zu berechnen. Somit ist VERA ein objektives, automatisches Analyseverfahren meteorologischer Daten über komplexer Orographie. Das Projekt wurde 1995 begonnen und wird seither laufend verbessert und weiterentwickelt. Es können mesoskalige Strukturen aufgelöst werden, die kleiner als die Skala des mittleren Stationsabstandes sind, indem Vorwissen über das Verhalten der Felder in die Analyse eingebracht wird. VERA kann alle skalaren und vektoriellen Felder, die an unregelmäßig verteilten Stationen vorliegen, auf ein regelmäßiges Gitter bringen. [15]

### *13.2 Grobablauf*

Das dreistufige Verfahren besteht aus [15]:

- Erstens aus einer Fehlerkorrektur, die die einzelnen Messwerte auf ihre räumliche physikalische Konsistenz prüft. Die Daten werden von Fehlern befreit. Sowohl systematische Fehler durch falsch kalibrierte Messgeräte sowie zufällige Fehler in der Folge von Übertragungsproblemen werden erkannt und eliminiert. Dazu wird ein eigens entwickeltes Verfahren verwendet, das die räumliche Konsistenz der Daten überprüft.
- Zweitens aus der Interpolation, für die ein Variationsansatz verwendet wird, der auf die thin plate spline Methode zurückzuführen ist. Die Einflüsse der Topographie auf das zu analysierende skalare oder vektorielle Feld gehen in Form der so genannten "Fingerprints" ein. Dabei werden bekannte orographische Einflüsse auf das betrachtete Feld erkannt und zum "downscaling" genutzt. Die Methode braucht kein numerisches Modell oder first guess Feld als Anfangszustand.

- Der letzte Schritt ist die graphische Aufbereitung der Ergebnisse. Die berechneten Felder werden mit Isolinienkarten, Farbflächendarstellungen oder Vektoren visualisiert. Die Abbildung zeigt ein Beispiel für eine Analysekarte.

### *13.3 Das Analyseverfahren*

Das Herzstück von VERA ist ein Variationsalgorithmus, der die Quadrate der räumlichen ersten und/oder zweiten Ableitungen des Messfeldes minimiert. Dabei wird davon ausgegangen, dass sich viele Feldverteilungen meteorologischer Parameter (z.B. Druck, Temperatur) so verhalten, dass ihre Krümmung minimal ist. Das entspricht dem Prinzip des minimalen Aufwandes, das in Natur oft zu beobachten ist (vgl. Oberfläche einer Seifenblase).

Das Gitterpunktfeld mit konstanten Abständen ist teilweise mit bekannten Werten gefüllt, die von Messungen an Stationen stammen. An die unbekanntes Gitterpunkte soll das Messfeld interpoliert werden. Für diesen Zweck gibt es zwei Forderungen an die Interpolationsfunktion: Erstens soll sie durch alle gegebenen Punkte exakt durchführen. Zweitens sind wir nur an den Werten der Funktion an den Gitterpunkten interessiert und nicht an der stetigen Funktion zwischen den Gitterpunkten.

Die Minimierung der Quadrate der zweiten Ableitungen des 1D-Feldes nach der räumlichen Koordinate führt zu einer Kurve mit minimaler Krümmung. Beim Variationsansatz kommt ein weiterer Term hinzu, der die Differenz zwischen tatsächlichem und erwartetem Wert an den Gitterpunkten enthält. Dieser ist gleich null. Um den Ansatz zu lösen wird diskretisiert und ein Gleichungssystem aufgestellt.

Für die praktische Anwendung in VERA wird dieses simple Prinzip auf 3 Dimensionen erweitert. Weiter wird auf die Tatsache eingegangen, dass die Stationen nicht auf den Gitterpunkten liegen. [15]

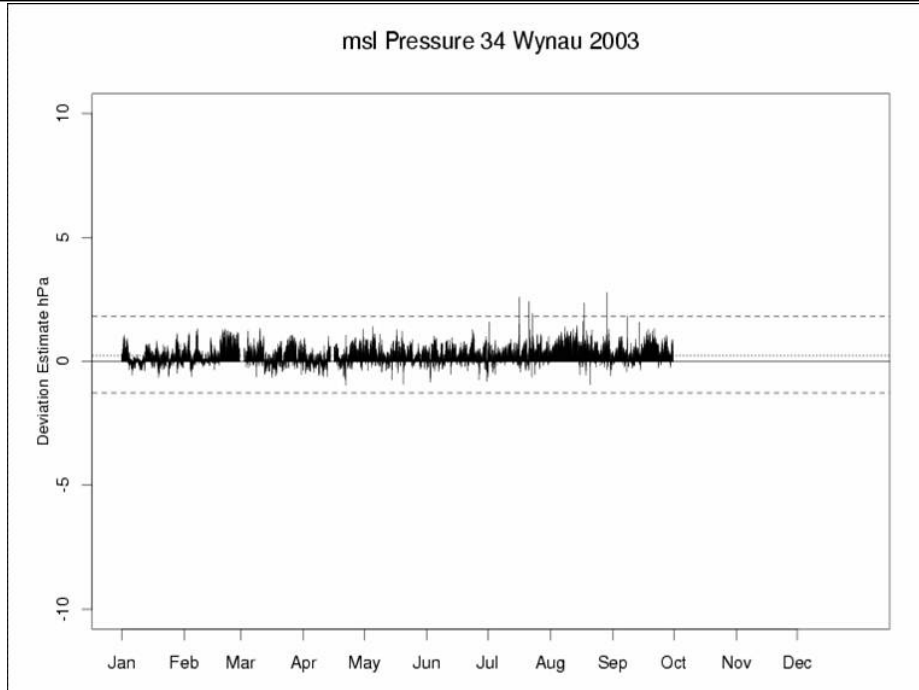


Abbildung 5. Abweichung für den reduzierten Druck der Station Wynau 2003

## KAPITEL 14 *Homogenisierung*

### 14.1 *Einleitung*

Mit der Homogenisierung erreichen die Daten den Level 5. Mit dem Projekt NORM90 [3] wurde die WMO-Standardperiode 1961-1990 aufgearbeitet. Bei der Homogenisierungsarbeit wird versucht, Veränderungen in einer Messreihe zu finden, die auf nicht klimatologische Einflüsse zurückgeführt werden können. Dabei ist es wichtig, die Geschichte der Messungen an einer Station nachzuvollziehen.

Für die Parameter Luftdruck, Temperaturmittel, mittlere Maximum- und Minimumtemperatur, Niederschlag, Sonnenscheindauer, Dampfdruck, Wind (Richtung und Geschwindigkeit), Globalstrahlung und Bewölkung stehen nach Abschluss des Projekts Norm90 von allen operationellen Klima- und Niederschlagsstationen bearbeitete Reihen und die entsprechenden Norm- bzw. Mittelwerte zur Verfügung. Zudem wurden 12 lange Temperatur- und Niederschlagsreihen ab Messbeginn homogenisiert.

Die Methodik der Homogenisierung wurde im Rahmen des Projekts weitgehend standardisiert. Das erstellte Softwarepaket THOMAS (Tool for Homogenization of Monthly Data Series) stellt 12 verschiedene Homogenitätstests zur Verfügung und vereinigt alle Schritte, die zur Homogenisierung von Messreihen auf Monatsstufe nötig sind. Das Tool besteht aus einer grafischen Benutzeroberfläche, einem Datenbankschema und einer Software-Bibliothek, die alle im Homogenisierungsprozess verwendeten mathematischen Funktionen enthält.

### 14.2 *Grundsätze*

- Der Homogenisierungsbetrag bzw. -faktor wird abgelegt (keine Korrektur an Daten).
- Die Homogenisierungsbeträge/-faktoren werden auf Monatsbasis berechnet und wo möglich auf Tage umgelegt.
- Auswahl der Gründe für Inhomogenitäten:
  - Beobachterwechsel
  - Einführung der automatischen Messung
  - Stationsverlegungen
  - Instrumentenwechsel
  - Wechsel des Instrumententyps
  - ungenaue Eichungen

- Stationswartungen
- Beobachtungszeitenwechsel

### *14.3 Verlängerte und synthetische Reihen*

Als homogene Reihe wird die mit den Homogenisierungsbeträgen / -faktoren verrechnete originale Datenreihe bezeichnet, wobei sie ein bestimmtes Anfangsdatum (Anfangsdatum der Homogenisierung) und Enddatum (Enddatum der Homogenisierung = Referenzdatum) hat. Um die Reihe bis zum aktuellen Datum verwenden zu können, werden ab dem Enddatum der Homogenisierung Originaldaten angehängt. Es wird dabei davon ausgegangen, dass sich ab Zeitpunkt Referenzdatum keine Veränderungen in den Messbedingungen mehr ergeben haben. Diese zusammengesetzte Reihe wird als verlängerte homogene Reihe bezeichnet. Originaldaten dürfen jedoch nur am Ende, nicht aber am Anfang der Reihe angehängt werden, da sich die Homogenisierung auf den Referenzdatum bezieht. [14]

Synthetische Reihen entstehen beim Zusammenhängen von klimatologisch ähnlichen Stationen oder Standorten. Dies wird zum Beispiel nötig, wenn Stationen verschoben werden müssen oder eine Erneuerung nicht am selben Standort möglich ist. Verschiedene Stationen der MeteoSchweiz werden so zu einer Messreihe zusammengeführt (Bsp. Samedan-Bever-synth).

### *14.4 THOMAS*

THOMAS (Tool for Homogenization of Monthly Data Series) benutzt als Datenquelle die originalen Monats- und Tageswerte, die im DWH abgelegt sind. Stationsinformationen entnimmt THOMAS aus den `t_series`-Tabellen (`t_series`, `t_series_definition`, `t_series_element`), die zum Umgang mit synthetischen Reihen (zusammengesetzte Reihen aus mehreren Stationen, z.B. Blatten\_Ried-Lötschen) eingeführt wurden. [14]

#### *14.4.1 Bearbeitung*

Als kurze Reihen werden alle diejenigen Datenreihen bezeichnet, für die bei der Homogenisierung die Bestimmung der Normwerte und nicht die homogene Reihe im Vordergrund steht. Für diese Reihen steht eine möglichst rasche Aktualisierung der Normwerte nach einer Inhomogenität im Vordergrund. Für diese Reihen erfolgt eine provisorische Bestimmung der Homogenisierungsbeträge, damit die Abweichungen von der bzw. Verhältnisse zur Norm konsistent bleiben.

Als lange Reihen werden alle diejenigen Datenreihen bezeichnet, für die als Resultat der Homogenisierung die homogene Reihe im Vordergrund steht, wobei jedoch aktualisierte Normwerte ebenfalls möglichst rasch nach einer Inhomogenität erwünscht sind. In der Regel handelt es sich um Reihen, die im 19. Jahrhundert beginnen und deren Erhalt besondere Aufmerksamkeit erfährt. Bei der Homogenisierung der langen Reihen besteht der Unterschied zu den kurzen Reihen darin, dass nach der provisorischen Bearbeitung einer Inhomogenität zwingend eine definitive folgt. Die dazwischen liegende Periode kann nicht pauschal angegeben werden, beträgt aber sicher mehrere Jahre. [14]

#### *14.4.2 Resultate*

Nach einer Homogenisierung werden von THOMAS für alle homogenen Parameter folgende Resultate ausgegeben:

- monatliche Homogenisierungsbeträge/-faktoren der homogenisierten Reihen
- entsprechende monatliche Normwerte
- ein Update des Homogenitätszustandes (nicht homogen, homogen)

Für einzelne definierte Parameter werden zudem folgende Resultate ausgegeben:

- Tägliche Homogenisierungsbeträge/-faktoren für lange Reihen (momentan Tmittel und Niederschlagssumme).
- Tägliche Homogenisierungsbeträge/-faktoren für Parameter, für die Normwerte bestimmt werden, die auf der Auszählung von Tageswerten beruhen (Tmin, Tmax, Niederschlagssumme).
- Abgeleitete monatliche Normwerte (Eis-, Frost-, Sommer-, Hitze- und Niederschlagstage; Relative Sonnenscheindauer noch zu implementieren).
- Tägliche Normwerte, abgeleitet aus den monatlichen Werten (Tmittel, Tmin, Tmax).

Auf der Datenbank werden folgende Daten abgelegt:

- Monatliche und tägliche homogene Reihen: Entstehen erst durch Verrechnung der abgelegten Betrags- oder Faktorenreihen mit den Originaldaten.
- Monatliche und tägliche Abweichungen bzw. Verhältnisse von bzw. zu der Norm: Entstehen durch Verrechnung der Normwerte mit den aktuellen Messwerten.

## KAPITEL 15 *Lückentests*

Bei den Lückenskripten handelt es sich um Werkzeuge zum Erkennen und gegebenenfalls Füllen von Lücken in den Messreihen für spezielle Grössen, die in der automatischen Prüfung- und Mutationsbearbeitung nicht berücksichtigt werden.

Die Lückentests beziehen sich auf Level 3 Daten, wobei die Behandlung der Lücken nicht in Realtime erfolgt.

### *15.1 Lücken bei Augenbeobachtungen*

#### *15.1.1 Aufgabenbeschreibung und Positionierung im System*

Mit der Applikation PuMAB werden Datenlücken nur auf Zehnminutenstufe detektiert und interpoliert. Deshalb werden Datenlücken bei den nicht in 10 Minuten-Takt angelieferten Parametern (im ANETZ z.B. Augenbeobachtungsparameter und Bodentemperaturen) nicht gefunden. Um diese Datenlücken trotzdem herauszufinden und der interaktiven Bearbeitung mit PuMIB zuzuführen, wurde der DWH-Dienst Lückentest implementiert.

Der Lückentest gibt für die Beobachtungsstationen die fehlenden Augenbeobachtungen der 3 Haupttermine (06Z, 12Z und 18Z) in Listenform aus. Bei den ANETZ Stationen werden zusätzlich alle drei Bodentemperaturen (5cm, 10cm und 20cm Tiefe) auf Lücken überprüft, wobei jeweils auf die HH:10 Werte abgesehen wird.

#### *15.1.2 Ablauf*

Täglich um 7:00 Uhr Lokalzeit läuft das Lückenskript für den Vortag.

1. für alle ANETZ- und konv. Klimastationen werden fehlende Augenbeobachtungen zu den Terminen 06:00 Uhr, 12:00 Uhr und 18:00 Uhr gesucht. Dies betrifft die Parameter 1760, 1763, 1764, 1765, 1766, 1767, 1769, 1770, 1771, 1773 und 1774. Ausgegeben wird pro Zeitstempel und pro Station die Anzahl fehlender Beobachtungen.
2. für die ANETZ-Stationen werden zusätzlich fehlende Bodentemperaturdaten gesucht: Parameter 104, 105 und 106.

Auf dem Server DWHPROD werden die erzeugten Listen direkt auf einen definierten Drucker geschickt. Das Skript kann für die Prüfung von beliebigen Zeiträumen auf Lücken manuell gestartet werden.

### 15.1.3 Beispiel-Output

Fehlende Augenbeobachtungen ANETZ 19.08.2004 10:32 KD/US

\*\*\*\*\*

Station-Id	Name	Zeit	Fehlende Werte
27	Chur	19.08.2004 18:0013	
28	Robbia	19.08.2004 12:0013	
28	Robbia	19.08.2004 18:0013	
29	Samedan-Flugplatz	19.08.2004 06:0011	
32	Payerne	19.08.2004 06:002	

---

## 15.2 Lücken bei den Schneebeobachtungen

### 15.2.1 Aufgabenbeschreibung und Positionierung im System

Vom 1. Mai bis zum 31. September müssen die Beobachter der ANETZ-Stationen keine Schneebeobachtungen eingeben, sofern kein Schnee liegt und es keinen Neuschnee gibt. Während dieser Zeit sollen deshalb die fehlenden Beobachtungen automatisch mit 0 gefüllt werden. Allenfalls eingegebene Werte dürfen auf keinen Fall überschrieben werden.

Wenn alle Stationen auf OBS umgestellt sind, wird der Test überflüssig.

### 15.2.2 Bearbeitung

Momentan werden nur ANETZ-Stationen bearbeitet (Liste mit den entsprechenden Installation-Ids wird dynamisch über T\_MEAS\_SITE\_TYPE\_CURR/ T\_MEAS\_SITE\_TYPE\_REF ermittelt, welche zum gewünschten Termin auch beobachten (Information aus T\_OBS\_TIME, WMO-Termin ist relevant).

Es werden die Parameter hns000s0 (155) und hto000s0 (156) gefüllt (mit reference\_ts aus T\_OBS\_TIME.real\_time\_dt).

Das Skript kann für einen bestimmten Termin aufgerufen werden. Wird es ohne Argument aufgerufen und findet der Aufruf nach 20:00 Uhr UTC statt, bearbeitet der Job den 06:00 und 18:00 Uhr Termin des aktuellen Tages. Falls der Aufruf vor 20:00 Uhr UTC stattfindet, wird der 06:00 Uhr Termin des aktuellen und der 18:00 Uhr Termin des Vortages bearbeitet.

---

### 15.2.3 Ermitteln von fehlenden Werten

Es wird geprüft, ob der zugehörige WMO-Parameter (hns000sw (1758), hto000sw (1759)) zum WMO Zeitpunkt vorhanden ist. Wenn dieser fehlt, dann wird ein neuer Datensatz zum Basisparameter erzeugt.

Die Daten werden auf den Beobachtungszeitpunkt (aus T\_OBS\_TIME.real\_time\_minutes\_utc\_hhmm) abgelegt, der bezüglich Season-Id zum Beobachtungszeitpunkt gültig ist.

Spalte	Befüllung
REFERENCE_TS	Beobachtungstag (in der Regel sysdate) und Zeitpunkt aus T_OBS_TIME.real_time_minutes_utc_hhmm
INSTALLATION_ID	aus T_OBS_TIME.installation_id
PARAMETER_ID	155 und 156 (hns000s0 und hto000s0)
MEAS_CAT_NR	1
LOAD_TS	Sysdate
VALUE_NU	0
SERIES_STATE_NR	35
MUTATION_INFO_NR	33
PLAUS_INFO_NR	0
TIME_GRANULARITY_TX	'T'

Tabelle 8. Beschreibung des eingefügten Datensatzes in der WorkDB

### 15.2.4 Ablauf

Das Skript wird vor dem Meteoblock ausgeführt. Es werden folgende Schritte durchlaufen:

1. Bestimmen der Beobachtungstermine, falls nicht ein spezieller Termin vorgegeben wurde.
2. Ermitteln einer Liste der Installation-Ids, welche bearbeitet werden sollen.
3. Ermitteln fehlender Werte und Ausgeben dieser Werte in Form eines Reports an stdout.
4. Durchführen des Inserts für die fehlenden Werte.
5. Aufräumen, d.h. das Skript selber erzeugt keine Logdatei, sondern gibt alle Infos an stdout aus.

## KAPITEL 16 *Geberstörungsmeldungen*

Als Geberstörungsmeldungen werden Warnmeldungen bezeichnet, die aufgrund von Plausibilitätsverletzungen direkt verschickt werden.

Diese Auswertung bezieht sich auf Level 3 Daten.

Ein Spezialfall ist das Versenden von Geberstörungsmeldungen für die Stationen der Regionalflugplätze. Bei diesen Geberstörungsmeldungen werden stündlich die vom Calc&Check Modul gesetzten Qualitätsinformationen ausgewertet und bei Bedarf eine entsprechende Geberstörungsmeldung generiert und an Skyguide gesendet.

### *16.1 Beschreibung*

Mit der Geberstörungsmeldung werden die Testverletzungen ausgewertet, die von PuMAB festgestellt wurden. Eine Störungsmeldung wird erzeugt, wenn in einem definierten Zeitraum eine gewisse Anzahl von gleichen Testverletzungen überschritten wird. Wenn ein solcher Fall vorliegt, wird eine Datei generiert und über das MHS versandt.

Die Störungsmeldungen können über längere Zeit ausgewertet werden und fliessen zum Beispiel in den Bericht zuhänden des Eidgenössischen Nuklearsicherheitsinspektorat (ENSI) ein oder werden zur Überwachung der Regionalflughäfen an Skyguide gesendet.

Für die Überwachung beim Prozess TM werden die Fälle der nicht mit PUMIB bearbeiteten Geber in eine weitere Datenbanktabelle geschrieben von wo aus, die Meldungen mit einer WEB-Applikation quittiert werden können.

NetworkControl: Aktuelle Gebermeldungsstörungen

	Datum	Sta	Station	Meas site	Par	Parameter	Methode
<input type="checkbox"/>	17.06.2009	PAY	Payeme	1708	fk010z0	Windgeschwindigkeit skalar; Zehnminutenmittel	7
<input type="checkbox"/>	17.06.2009	STG	St. Gallen	29	tso010s0	Bodentemperatur 10 cm Tiefe; Momentanwert	112
<input type="checkbox"/>	17.06.2009	STG	St. Gallen	29	tso005s0	Bodentemperatur 5 cm Tiefe; Momentanwert	112
<input type="checkbox"/>	17.06.2009	GLA	Glarus	30	ua1200s0	Vergleichsfeuchtigkeit 1; z.T. Wetterhütte	7
<input type="checkbox"/>	17.06.2009	VIS	Visp	37	dk010z0	Windrichtung; Zehnminutenmittel	0
<input type="checkbox"/>	17.06.2009	VIS	Visp	37	fve010z0	Windgeschwindigkeit vektoriell; Zehnminutenmittel	0
<input type="checkbox"/>	17.06.2009	DEM	Delémont	96	fk010z1	Böenspitze (Sekundenböe); Maximum	166
<input type="checkbox"/>	17.06.2009	DEM	Delémont	96	dk010z1	Windrichtung bei Böenspitze (Sekundenböe); Zehnminuten-Intervall	166
<input type="checkbox"/>	18.06.2009	PAY	Payeme	1708	fk010z0	Windgeschwindigkeit skalar; Zehnminutenmittel	7
<input type="checkbox"/>	18.06.2009	STG	St. Gallen	29	tso020s0	Bodentemperatur 20 cm Tiefe; Momentanwert	113
<input type="checkbox"/>	18.06.2009	STG	St. Gallen	29	tso010s0	Bodentemperatur 10 cm Tiefe; Momentanwert	113
<input type="checkbox"/>	18.06.2009	TIT	Tittlis	392	tre200s0	Lufttemperatur 2 m über Boden; Momentanwert	15
<input type="checkbox"/>	18.06.2009	TIT	Tittlis	1729	tsa000s0	Temperatur Schneehöhenmesser	15

Logout

Löschen

Abbildung 6. Screenshot NetworkControl

## 16.2 Ablauf

Die Benachrichtigung über die Störung ist als DWH-Dienst implementiert, läuft auf dem Datenbankserver und wird über PowerCenter gesteuert. Es wird mit den Testverletzungen aus WORK1.T\_PUM\_TEST\_FAILURE gearbeitet.

Es werden zwei Schleifen ausgeführt, die erste über alle Messfelder und die zweite über alle Geber eines Messfeldes.

Die Störungsmeldungen werden einmal pro Tag um 06:00 UTC erzeugt. Sie können für einen beliebigen Zeitpunkt über ein Skript manuell gestartet werden.

### *16.3 Beispiel einer Geberstörungsmeldung*

298

VQFA02 LSSW 230500

Meteorologische Messstation Kernanlage Leibstadt

Meldung Testverletzungen vom 23.11.2008 05:00 UTC.

Zwischen 22.11.2008 00:00 UTC und 22.11.2008 23:50 UTC traten bei folgenden Gebern 17 oder mehr identische Testverletzungen auf:

Towerlevel oben:

- Lufttemperatur Instrument 1
- Lufttemperatur Instrument 2
- Windgeschwindigkeit vektoriell; Zehnminutenmittel; Instrument 1
- Windgeschwindigkeit vektoriell; Zehnminutenmittel; Instrument 2

## KAPITEL 17 *Referenzen*

- [1] Häberli, Ch.: Zuständigkeit für Daten und Kundenprodukte Version 2.0
- [2] Ch. Häberli, C. Stocker, Th. Konzelmann: Das automatische Stationsnetz. Messkonzept 2010/07 I (automatische Bodenstationen) V3.0 11.09.2008
- [3] Begert, M., Seiz, G., Schlegel, T., Musa, M., Baudraz, G., und Moesch M.: Homogenisierung von Klimamessreihen der Schweiz und Bestimmung der Normwerte 1961-1990 - Schlussbericht des Projekts NORM90. Veröffentlichungen Nr. 67, MeteoSchweiz, Zürich.
- [4] Häberli, Ch.: Specifications of Inter-Parameter Constraints at MeteoSwiss (Internal Consistancies) - Specification Document DABES.2. Version 4.6
- [5] Häberli, Ch., Konzelmann, T., Steinegger, U., Mettler J., Mühlhäuser, C., Musa, M., Dössegger, R., Kiene, M.: Qualitätsbezeichnung der Messwerte im MeteoSchweiz Data Warehouse System - Ergebnisdokumentation DD02 Teilprojekt Datenmodelle und Detailspezifikation. Version 4.0
- [6] Küng, U.: Calc & Check: SW-Design Dokumentation.
- [7] Abbt, M. und Musa, M.: AP17.5 Beschreibung des DWH-Dienstes “Lückentest”. Version 1.0
- [8] Abbt, M.: Beschreibung des DWH-Dienstes “Füllen von fehlenden Schneebeobachtungen” - Spezifikation. Version 1.0
- [9] Hohenegger, C. und Musa, M.: Plausibilitätstests und Interpolation - Bericht. Version 2.1
- [10] Häberli, C.: Automatische Dateninterpolation im ANETZ - AMUT 1. Dokumentation. Dat Dok 1 (331)
- [11] WMO: Guide to climatological practices. WMO-No 100, Geneva 1983
- [12] Häberli, Ch.: Hintergrundinformationen und Situationsbeurteilung für die Erarbeitung der Datenarchitektur. Version 1.0
- [13] Häberli, Ch., von der Crone, M.: Hintergrundinformationen und Situationsbeurteilung für die Erarbeitung der Datenarchitektur. Version 1.1

- [14] Begert, M., Schlegel Th.: Konzept zur Implementation von THOMAS und seiner Resultate ins MeteoSchweiz DWH System. Version 1.2
  
- [15] <http://www.univie.ac.at/IMG-Wien/vera>. Besucht am 18.1.2007
  
- [16] WMO (1992a): Manual on the Global Telecommunication System, Global Aspects, No. 386 Vol. Ic
  
- [17] Stocker, C., Häberli, C., Perl, M.: Stationen und Bulletins für das Global Telecommunication System: Vorgaben der WMO, eingegangene Verpflichtungen von MeteoSchweiz und Vorgehen bei Änderungen. Standardisierungsdokument. Internes Dokument, Version 1.6.
  
- [18] WMO (1992b): Manual on the Global Data-Processing System, No. 485 Vol. I
  
- [19] WMO (1995): Manual on the Global Observing System. No-544 Vol. II. Regional Aspects. Amendments PART II.
  
- [20] Ch. Häberli: The Comprehensive Alpine Radiosonde Dataset (CALRAS). Wiener Meteorologische Schriften Heft 4 (2006).
  
- [21] M. Bassi et al.: MeteoSchweiz Data Warehouse. Bericht Konzept. Version 1.0 20.12.2000

## ANHANG A *Spezifikation Umsetzung Plausibilitätsinformation zu Testprotokoll*

Dieser Anhang beschreibt die Umsetzung der Plausibilitätsinformation (kondensierte Plausibilitätsbezeichnungen als Integerzahl, codiert in dualer Weise).

### A.1 *Einführung*

Zu jedem Wert im DWH System ist eine sogenannte Qualitätsinformation vorhanden [5]. Diese Qualitätsinformation lässt sich wiederum aufteilen in Plausibilitätsinformation, in eine Mutationsinformation und einen Reihenzustand.

Zurzeit sind knapp 150 Tests definiert, die einen Mess- oder Augenbeobachtungswert auf seine Plausibilität prüfen. Es können vier verschiedene Arten von Tests unterschieden werden, wobei die vierte Art noch nicht integriert ist:

1. Limitentest. Wo es sinnvoll ist, wird jeder Wert auf obere und untere Limite getestet. Es werden harte und weiche Limiten unterschieden. Werte ausserhalb der harten Limiten sind (physikalisch) unmöglich, Werte ausserhalb der weichen Limiten sind zweifelhaft (2 Tests).
2. Variabilitätstest. Hier wird unterschieden zwischen einer maximalen und einer minimalen Variabilität. Bei der maximalen Variabilität müssen zeitlich benachbarte Werte innerhalb eines Variationsbereichs bleiben (auch hier wird zwischen einem harten und einem weichen Bereich unterschieden). Die minimale Variabilität verlangt, dass ein Intervall von mehreren hintereinander folgenden Werten sich ausserhalb eines Variationsbereiches befinden muss (dead band range). Bei der minimalen Variation werden ebenfalls weiche und harte Limiten unterschieden (4 Tests).
3. Test auf innere Konsistenz. Werte gleicher Zeit und gleichen Ortes aber verschiedener Parameter müssen in einem stimmigen Verhältnis sein (ca. 150 Tests).
4. Räumlicher Vergleichstest. Werte gleicher Zeit und gleichen Parameters benachbarter Stationen dürfen nicht zu stark voneinander abweichen (1 Test).

Für jeden Test wird festgehalten, ob er positiv oder negativ<sup>8</sup> ist.

---

<sup>8</sup> Ein Test ist positiv, wenn dieser angesprochen hat, der getestete Wert also verdächtig ist. Der Test ist negativ, wenn der Wert die Plausibilitätsregeln nicht verletzt.

---

Diese ausführliche Plausibilitätsinformation speichert nicht nur die Testresultate im Testprotokoll, sondern hält auch die Grenzwerte fest, mit denen getestet wurde. Die folgenden Werte werden festgehalten:

- physikalisch obere Limite
- physikalisch untere Limite
- klimatologisch obere Limite
- klimatologisch untere Limite
- klimatologisch maximale Variabilität
- klimatologisch minimale Varianz

Bit <sup>9</sup>	Bedeutung
0	testgroup 1, in grounddata missing/hard
1	testgroup 2, in grounddata soft limit violation
2	testgroup 3, in grounddata inner_inconsistency
3	testgroup 4, in grounddata spatial_inconsistency
4	testgroup 5, in grounddata inhomogenous value
5	testgroup 6, in grounddata reserve, in 2nd aggregation formed by aggreg_gappy data
6	testgroup 7, in grounddata reserve, in 2nd aggregation formed by aggreg_doubt data
7	aggregated value which is formed from partly missing values
8	aggregated value which is formed from partly doubtful values
9	hard limit violation of aggregation
10	soft limit violat. of aggregation
11	inner_inconsistency of aggregation
12	spatial_inconsistency of aggregation
13	inhomogeneity of aggregation
14	reserve
15	reserve

Tabelle 9. Plausibilitätsinformation im Rahmen der Qualitätsinformation

Die Plausibilitätsinformation PI ist eine Zahl, die aus der Struktur qual\_byte berechnet worden ist; d.h. die Struktur qual\_byte ist die binäre Darstellung der PI. Ebenso lässt die PI sich auch wieder in die einzelnen Werte von qual\_byte zerlegen (siehe Anhang). Ein Testprotokoll (ein Array von Charactern) besteht aus verschiedenen Abschnitten, die je einer bestimmten Gruppe von Testfunktionen zugeordnet sind.

---

<sup>9</sup> 0 bis 6: PI aus PuMAB

7 bis 15: PI aus Aggregation

## A.2 Details zur Struktur qual\_byte

Im File agg.h ist die Struktur qual\_byte von Mario Rohrer folgendermassen festgelegt worden:

```
struct qual_byte {
    unsigned int v_hard_limit : 1;
    unsigned int v_soft_limit : 1;
    unsigned int inner_incons : 1;
    unsigned int space_incons : 1;
    unsigned int inhomogenous : 1;
    unsigned int reserve_1 : 1;
    unsigned int reserve_2 : 1;
    unsigned int aggreg_gappy : 1;
    unsigned int aggreg_doubt : 1;
    unsigned int hard_testagg : 1;
    unsigned int soft_testagg : 1;
    unsigned int inne_testagg : 1;
    unsigned int spac_testagg : 1;
    unsigned int inho_testagg : 1;
    unsigned int rese_testagg : 1;
};
```

Die Plausibilitätsinformation wird aus dieser Struktur berechnet mit der Funktion convert\_dual\_decimal() aus dem File functions\_agg\_qual\_bit.c .

Mit convert\_decimal\_dual() kann aus der Plausibilitätsinformation die Struktur qual\_byte gesetzt werden.

Für die Kennzeichnung der einzelnen Testgruppen wäre der folgende Aufzählungstyp sinnvoll:

```
typedef enum {
    HARD_LIMIT = 0,
    SOFT_LIMIT,
    INNER_CONS,
    SPACE_INCONS,
    INHOMOGENOUS,
    RESERVE_1,
    RESERVE_2,
    AGGREG_GAPPY,
    AGGREG_DOUBT,
    HARD_TESTAGG,
    SOFT_TESTAGG,
    INNE_TESTAGG,
    SPAC_TESTAGG,
    INHO_TESTAGG,
    RESE_TESTAGG,
    END_TESTGRP
} E_TESTGRP;
#define MAX_TESTGRP (int) END_TESTGRP /* Anzahl Testgruppen */
```

## ANHANG B *Daten- und Limitenübergabe*

### *B.1 Allgemeines*

Es wird jeweils die zum Test gehörende Information bezüglich Testparameter, Vergleichsparameter, Installation und Limiten aus der KontextDB gelesen (Steuerinformation).

Für einige Tests werden Messreihen eines oder mehrerer Parameter benötigt. Diese Reihen beziehen sich immer auf ein Zeitintervall vor der Messzeit des zu testenden Parameters. Die Messfrequenzen sind jeweils abhängig vom Parameter und der Station (10 Minuten bis 24 Stunden), der Beobachtungszeit (obs\_time) und den Informationen aus T\_INPUT\_AUX und ARRAY\_LENGTH.

### *B.2 Allgemeingültige Bestimmungen*

Sämtliche Werte in den Arrays werden zeitlich aufsteigend geordnet.

Die Messfrequenz eines Parameters einer Station wird als qc\_freq\_id bzw. obs\_time abgelegt.

## ANHANG C *Bearbeitungsfunktionen*

Für die Bearbeitung der verschiedenen Testfälle stehen den Applikationen PuMAB und PuMIB 7 Funktionen zur Verfügung (siehe Tabelle 11). Die Zuordnung eines bestimmten Parameters zu einer Bearbeitungsfunktion wird über die KontextDB gesteuert (T\_PUM\_AB\_PARAM\_WORK\_FCT).

Nr.	ID <sup>10</sup>	Beschreibung	max. Fehlwerte	PuM AB	PuM IB
I	1	rational spline method	6	x	
II	2	mean rational spline method	36	x	x
III	3	treated by PuMIB		x	
IVa	4	treatment of values outside ‘Tagbogen’		x	
IVb	5	treatment of values outside astr. sunset/sunrise		x	
V	6	sunshine-param.: filling gaps with default value	200	x <sup>11</sup>	
VI	7	global-radiation: filling gaps with default value	200	x	

Tabelle 10. Zusammenstellung der 7 Bearbeitungsfunktionen

---

<sup>10</sup> Entspricht der WORK\_FCT\_ID aus DWH1.T\_PUM\_AB\_WORK\_FCT.

<sup>11</sup> Wird momentan nicht verwendet.

<b>Funktion</b>	<b>I</b>	<b>II</b>	<b>III</b>
Luftdruck	6	36	33
Lufttemperaturen, Taupunkt	6	36	33
Niederschlag 10'-Summe	6	0	33
Sonnenscheindauer 10'-Summe	6	0	33
Dosisleistung Radioaktivität 10'-Mittel	0	0	33
Globalstrahlung	6	0	33
Helligkeit	6	0	33
Luftfeuchte 2m	6	36	33
Nah-, Fernblitze 10'-Summe <sup>12</sup>	-	-	-
Windspitze 1 Sekunde	6	0	33
Windrichtung bei Sekundenspitze	6	0	33
Windgeschwindigkeit 10'-Mittel	6	36	33
Windrichtung 10'-Mittel <sup>13</sup>	6	36	33
Bodentemperaturen	6	36	33
Verdunstungssumme	6	36	33
RASTA Daten	6	0	33
<i>Tabelle 11. Maximal zulässige Länge der Messlücken bei den Funktionen I, II und III in PuMAB</i>			

Bei der Messlückenbearbeitung in PuMIB werden die Stützstellen interaktiv gesetzt und es können beliebig grosse Lücken gefüllt werden.

### C.1 Funktion I "rational\_spline\_method"

Die Funktion wird benutzt, um kleine Messlücken (in der Regel bis 6 Messwerte) in einer Messreihe von 10-Minuten-Werten aufzufüllen. Wie der Name schon sagt, arbeitet sie mit splines. Um die Funktion anwenden zu können, braucht

---

<sup>12</sup> Seit Einführung des Blitzortungssystems von MeteorAge werden die Blitzdaten nicht mehr interpoliert.

<sup>13</sup> interpoliert werden die Windkomponenten (kartesische Koordinaten)

es mindestens 3 gültige Werte vor und nach der Messlücke, welche als Stützwerte benützt werden können. Einzelne gültige Werte ausserhalb einer Messlücke werden als Stützwerte verwendet und werden nicht überschrieben. Ist diese Bedingung nicht erfüllt (z.B. am Anfang einer Messreihe oder weniger als 3 gültige Werte zwischen zwei Messlücken), so muss diese Datenlücke mit PuMIB bearbeitet werden.

So bearbeitete Werte werden mit Mutationsinfo 33 gekennzeichnet.

## *C.2 Funktion II "mean\_rational\_spline\_method"*

Diese Funktion wird für grössere Messlücken (in der Regel bis Messlücken mit 36 fehlenden Werten) eingesetzt. Sie arbeitet analog zur Funktion "rational\_spline\_method", nur dass hier als Stützwerte nicht 3 Werte vor und nach der Messlücke benützt werden, sondern aus den letzten 18 Werten vor und nach der Messlücke jeweils 3 arithmetische Durchschnittswerte von 6 Messwerten gebildet werden, und diese Durchschnittswerte als Stützstellen für die Berechnung der Splines benützt werden. Es werden auch arithmetische Mittel gebildet, wenn innerhalb der 18 Datensätze Lücken bestehen, jedoch muss mindestens innerhalb von 6 Werten mindestens ein gültiger Wert vorhanden sein. Ansonsten wird die Lücke gelassen und muss interaktiv gearbeitet werden.

Anschliessend wird die Messlücke analog zur Funktion I mit einer rationalen Splinefunktion interpoliert. Das wird in vielen Fällen dazu führen, dass zwischen dem letzten gemessenen und dem ersten interpolierten bzw. zwischen dem letzten interpolierten und ersten gemessenen Wert ein grosser Sprung auftritt. Um dies zu verhindern wird beim letzten und beim ersten gemessenen Wert die Differenz zur interpolierten Kurve bestimmt (in der Skizze rote, mit a gekennzeichnete Strecken). Anschliessend wird dieser Wert zu den Stützstellen addiert (neue Stützstellen in der Grafik mit x2 bezeichnet). Danach wird mit Hilfe dieser neuen Stützstellen die Splineinterpolation erneut gerechnet.

So bearbeitete Werte werden auch mit Mutationsinfo 33 gekennzeichnet.

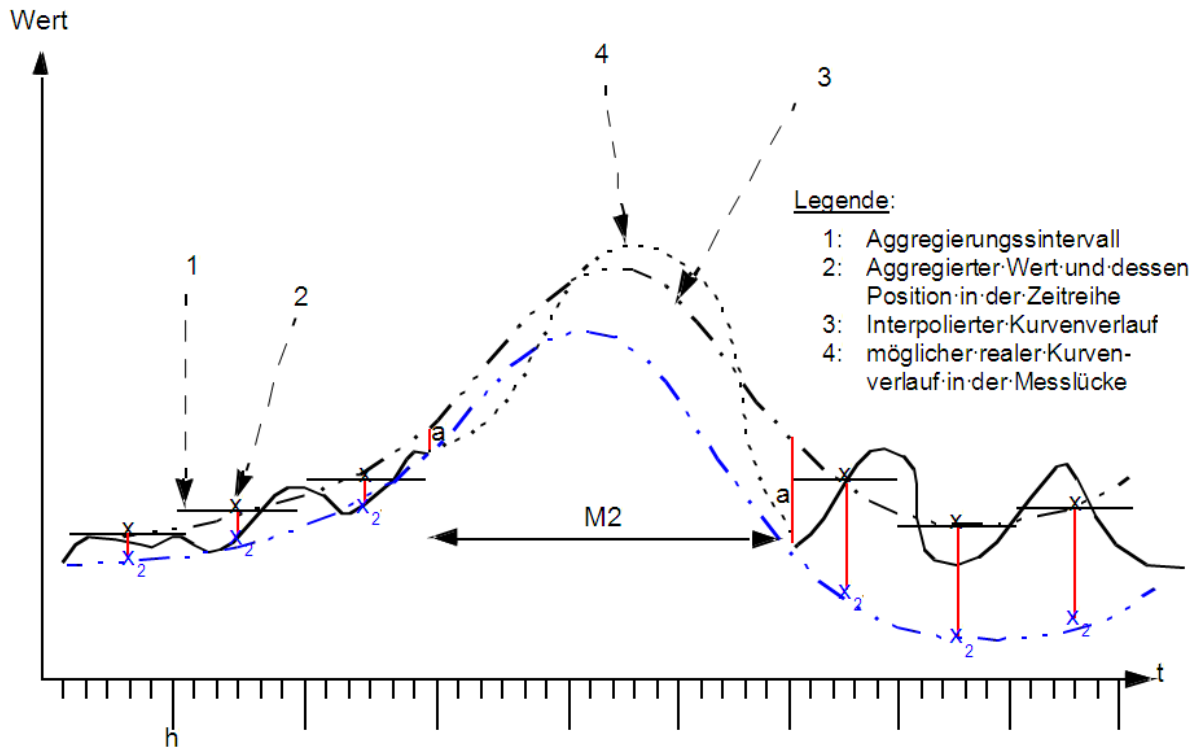


Abbildung 7. Veranschaulichung der Funktion II (kommt zur Anwendung bei der maximal zulässigen Lückenlänge M2; vgl. Tabelle 12)

### C.3 Funktion III "to treat by PuMIB"

Mit dieser Funktion werden all jene Datenlücken bearbeitet, welche aus irgendeinem Grund nicht mit den anderen Bearbeitungsfunktionen bearbeitet werden konnten. (Z.B. zu grosse Messlücke oder „Stotterbetrieb“). Die Funktion führt keine eigentliche Interpolation durch, sondern setzt die fehlenden Werte einfach auf -10000000.0 (minus PACK\_NULL), welche dann im nachfolgenden Testen zu einer Verletzung der harten, unteren Limite führten. Die so erzeugten Testverletzungen werden dann wie alle anderen Testverletzungen in die Tabelle T\_PUM\_TEST\_FAILURE eingetragen und durch die Sachbearbeiter individuell bearbeitet.

Momentan werden so bearbeitete Werte mit Mutationsinfo 35 gekennzeichnet.

#### *C.4 Funktion IVa "treatment of values outside Tagbogen" und Funktion IVb "treatment of values outside astr. sunset/-rise"*

Mit der Funktion IVa werden all jene Werte der Globalstrahlung auf die in T\_PUM\_AB\_PARAM\_WORK\_FCT spezifizierte Schranke gesetzt, welche ausserhalb des Tagbogens sind. Der Tagbogen wird für jedes Messfeld für jeden Tag neu nach folgender Formel berechnet:

Sonnenaufgang = 11:40 Uhr - Anzahl Minuten aus T\_FACTOR.value\_nu (Typ s\_vut)

Sonnenuntergang = Sonnenaufgang + Anzahl Minuten aus T\_FACTOR.value\_nu (Typ s\_max).

Mit der Funktion IVb werden all jene Werte der Sonnenscheindauer auf die in T\_PUM\_AB\_PARAM\_WORK\_FCT spezifizierte Schranke gesetzt, welche ausserhalb des astronomischen Tagbogens sind. Der Tagbogen wird für jedes Messfeld für jeden Tag neu aus der geographischen Länge und Breite des Messfeldes berechnet.

In der Tabelle T\_FACTOR kann pro Messfeld Beginn und Ende des Tagbogens für bestimmte Bezugsperioden definiert werden. Über die Umgebungsvariable TOLERANCE\_HORIZON wird der Interpolation eine Verschiebungskonstante mitgegeben, welche den Sonnenaufgang früher und den Sonnenuntergang später macht. Dies führt dazu, dass weniger Werte an der Grenze versehentlich auf Null gesetzt werden. Diese Funktion wird momentan nicht verwendet.

Bei der **Globalstrahlung** werden jene Ersatzwerte für die Messlücke, die nach dem astronomischen Sonnenuntergang bzw. vor dem astronomischen Sonnenaufgang liegen auf 0 gesetzt (Funktion IVa).

Bei der **Sonnenscheindauer** werden jene Ersatzwerte für die Messlücke, die nach dem Sonnenuntergang bzw. vor dem Sonnenaufgang liegen auf 0 gesetzt (Funktion IVb).

Als **Steuerinformation** benötigt diese Methode stationsabhängige Angaben über den Tagbogen bzw. über die Länge und Breite der Station in geographischen Koordinaten.

So bearbeitete Werte werden mit Mutationsinfo 34 gekennzeichnet.

#### *C.5 Funktion V „Ersetzung der Werte durch Werte von redundanten Messungen“*

Dies wurde nicht als eigentliche Bearbeitungsfunktion modelliert und implementiert, sondern wird folgendermassen gehandhabt: Falls eine Messlücke vorhanden ist, ein redundanter Parameter vorgesehen und die Werte im Meteoblockfile ohne eigene Messlücke in diesem Bereich vorhanden sind, werden die Werte vom redundanten Parameter in den Basisparameter linear geglättet übertragen (geglättet wird über das Intervall: letzter gültiger Messwert vor und nach der Messlücke).

Als **Steuerinformation** benötigt diese Methode eine stationsabhängige Zuweisung von Ersatzparametern zu jedem zu bearbeitenden Parameter.

So bearbeitete Werte werden mit Mutationsinfo 32 gekennzeichnet.

---

## C.6 Anwendung von Funktionen I und II bei intermittierend auftretenden Messlücken (‘Stotterbetrieb’)

Intermittierend auftretende Messlücken werden automatisch interpoliert, sofern der Abstand zwischen 2 gemessenen Punkten nicht grösser als M1 bzw. M2 ist und die Anzahl der vorhandenen randlichen Stützwerte die Anwendung der entsprechenden Methode zulässt, siehe Abbildung 7.

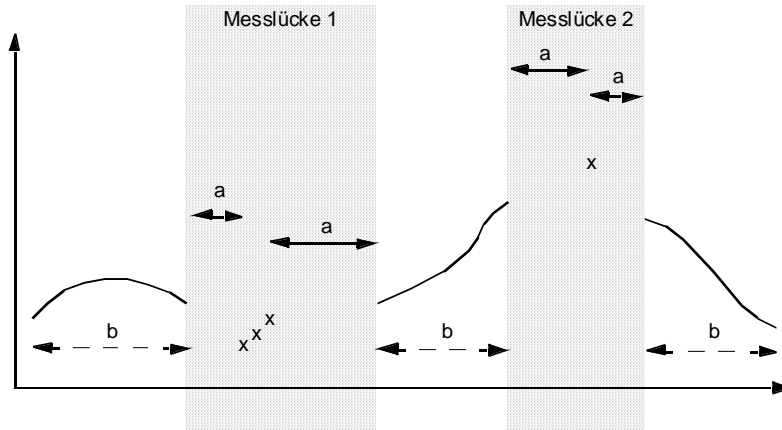


Abbildung 8. Anwendung von Methoden I und II bei intermittierend auftretenden Messlücken ( $x$  = vereinzelt, innerhalb von Messlücken auftretende Messwerte)

Im dargestellten Fall kann Methode I sowohl für Messlücke 1 als auch Messlücke 2 angewendet werden, wenn jeweils beide Strecken  $a \leq M1$  sind und die Strecken  $b$  jeweils  $\geq 3$  sind. Die Messwerte (mit  $x$  dargestellt) innerhalb der Messlücken werden als zusätzliche Stützwerte verwendet.

Methode II kann im dargestellten Fall (Abb. 7) dann angewendet werden, wenn jeweils beide Strecken  $a \leq M2$  sind und die Strecken  $b \geq 18$  10'-Intervalle sind. Auch hier werden die  $x$  innerhalb der Messlücken als zusätzliche Stützwerte verwendet.

## C.7 Detailspezifikation der Bearbeitungsfunktionen

Auf den folgenden Seiten werden als Beispiele die Ablaufdiagramme einiger Bearbeitungsfunktionen oder Module dargestellt.

### Bedeutung der Grafik und Abkürzungen

Entscheidungen: Ausgang rechts: nein, nach unten: ja  
ML: Messlücke  
IB: Interaktive Bearbeitung

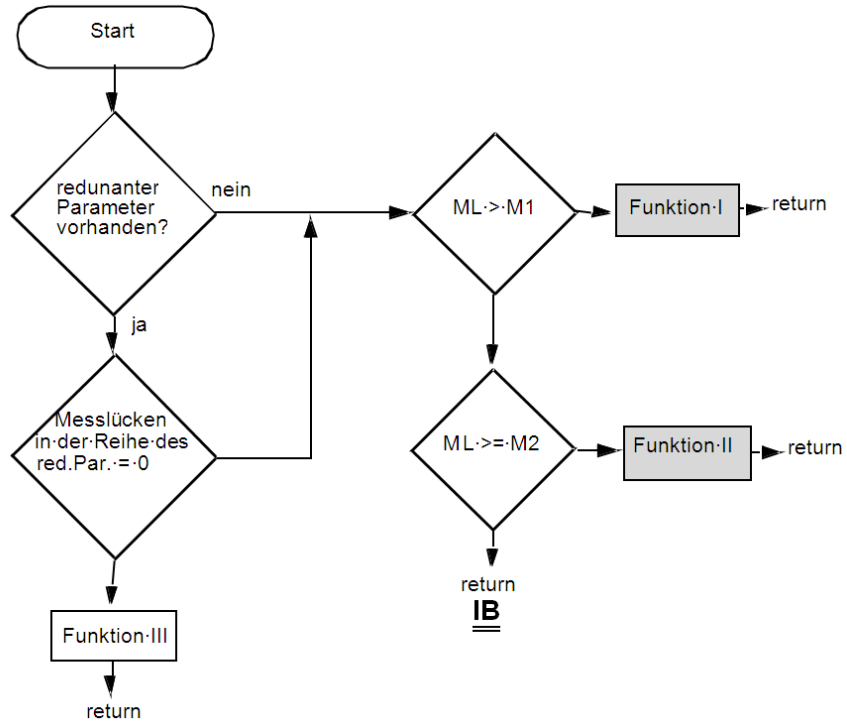


Abbildung 9. Ablauf des übergeordneten Moduls ‚Messlücke interpolieren‘

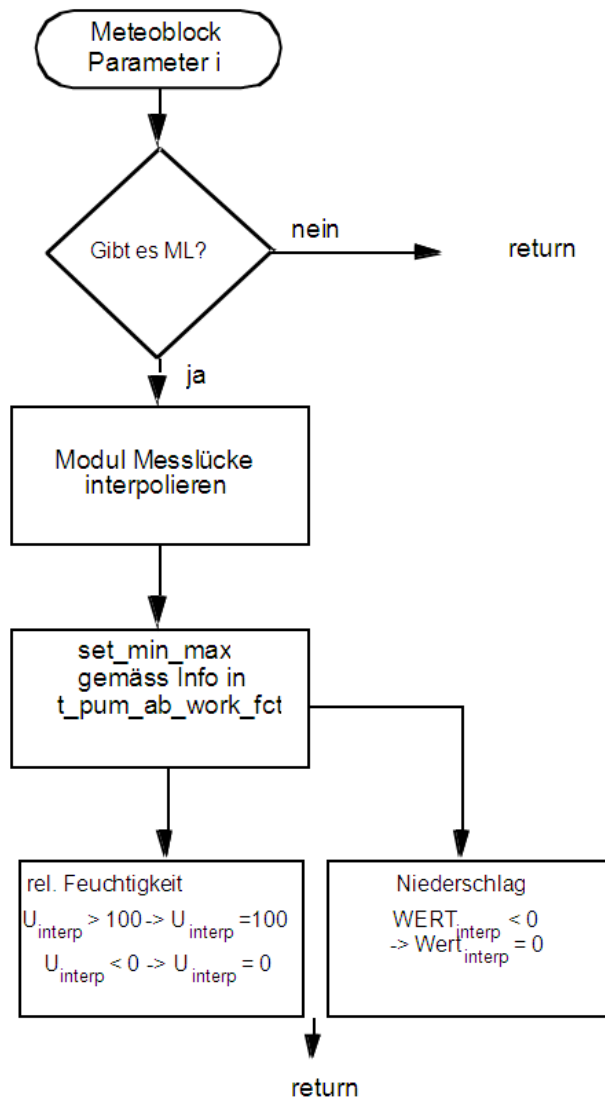


Abbildung 10. Ablauf der Bearbeitungsfunktionen I und II

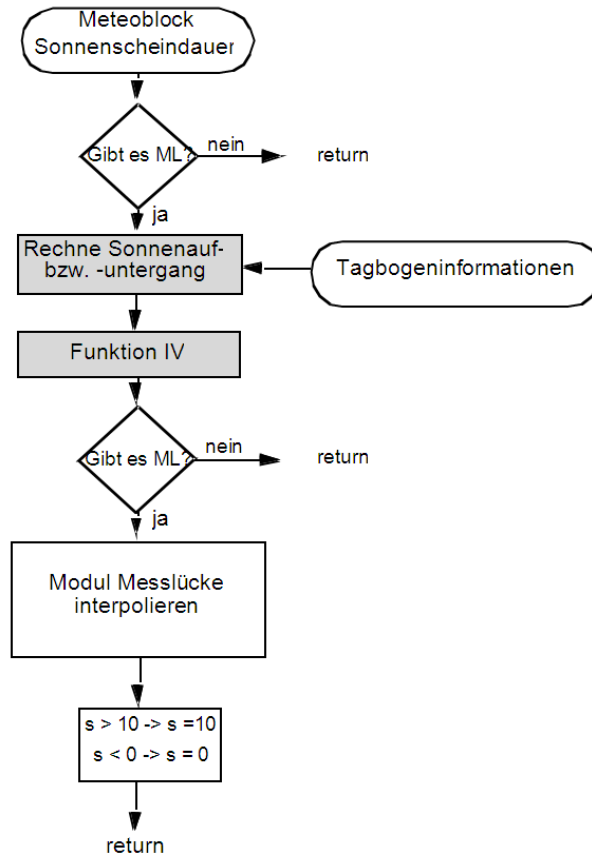


Abbildung 11. Ablaufdiagramm der Bearbeitungsfunktion IVa

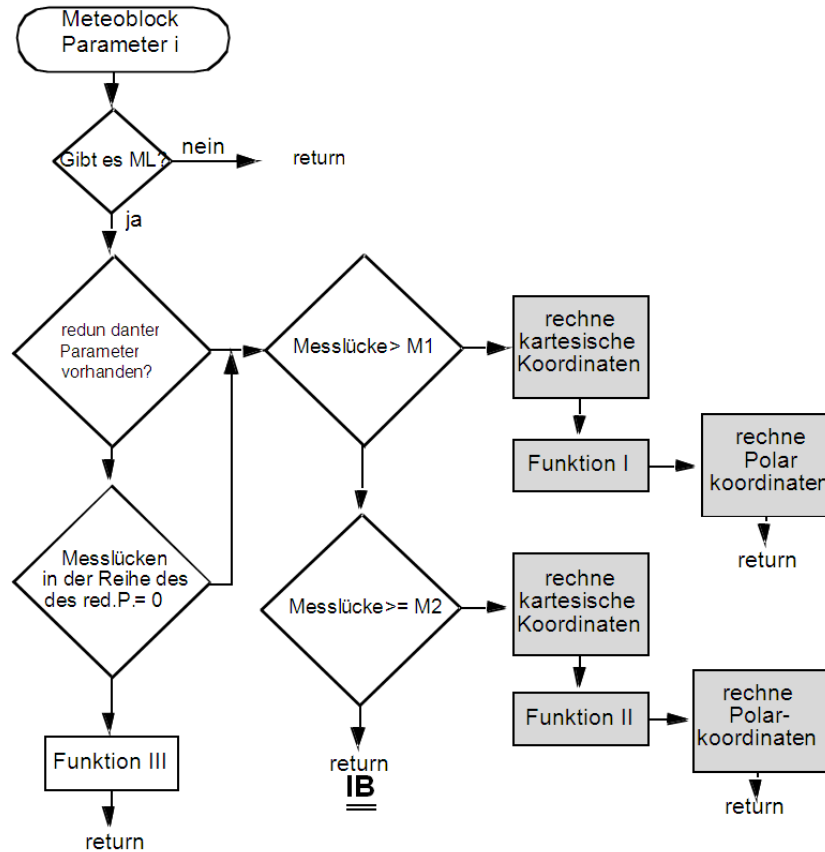


Abbildung 12. Ablauf am Beispiel der Interpolation von Winddaten

### C.8 Resultate aus dem Bericht AMUT - automatische Dateninterpolation im ANETZ

Der Bericht über die automatische Dateninterpolation im ANETZ [10] kommt zum Schluss, dass die Differenzen zwischen geschätzten und gemessenen Werten für die untersuchten 5 Stundenlücken klein sind. Bei Verwendung der rationalen Splinefunktion sind sie in der Vielzahl der Fälle kleiner als bei der kubischen. Ebenso korrelieren die mit den rationalen Splinefunktionen interpolierten Schätzwerte häufig besser mit den Messwerten als die mit den kubischen Splines interpolierten (siehe Abbildungen in Anhang 2). Dazu ist zu bemerken, dass die Stützpunkte ohne manuellen Eingriff gesetzt wurden und dass unmögliche Werte (z.B. Niederschlag < 0 mm) nicht korrigiert wurden.

Die Resultate sind wesentlich besser bei den kontinuierlich ändernden Parametern (z.B. Temperatur) als bei den "erratisch" ändernden Parametern (z.B. Niederschlag). Für einige Parameter können etwas detailliertere Aussagen gemacht werden:

1. Die Windkomponenten werden bei gewissen Wetterlagen sowohl mit kubischen als auch mit rationalen Splines zu tief geschätzt und eine gewisse Abhängigkeit des Schätzfehlers von der Stärke kann nicht ausgeschlossen werden.
2. Die Temperatur wird vor allem von den rationalen Splines dann unterschätzt, wenn die zu interpolierende Periode im Bereich des Maximums auftritt. Dies ist auf den dämpfenden Einfluss der rationalen Splines zurückzuführen. Bei den kubischen Splines wird dieser Effekt nicht oder nur schwach beobachtet.
3. Die Globalstrahlung wird häufig überschätzt, wenn die zu interpolierende Periode im Anstieg bzw. Abfallen liegt (Teilperioden 2 und 6. Am Morgen haben die Anfangsstützpunkte noch tiefe, die Endstützpunkte bereits hohe Werte, ab Abend ist die Situation umgekehrt). Deutlich unterschätzt wird sie, wenn die Lücke im Bereich des Maximums liegt. Ähnliches kann bei der Helligkeit und bei der Sonnenscheindauer beobachtet werden. Auch dieses ungünstige Verhalten tritt bei den rationalen Splines häufiger auf und ist wohl auf deren dämpfende Eigenschaft zurückzuführen. Die Sonnenscheindauer wird zudem generell unterschätzt. Dabei gibt es ausgeprägte Unterschiede zwischen Stationstypen und Wetterlagen.
4. Da Niederschlagsereignisse nur wenige Stunden dauern können, war abzusehen, dass die Interpolation eher zu tiefe Werte liefern wird: Ein Niederschlagsereignis kann genau in die Lücke fallen, während an den Stützpunkten 0 mm Niederschlag gemessen wird.

Fazit:

- Die Qualität der mit diesem Verfahren erzielten Resultate hängt wesentlich von der Wahl der Stützpunkte ab.
- Schätzung von Ersatzwerten für Niederschlag und Sonnenscheindauer sind mit diesem Verfahren problematisch. Dies zeigt sich auch in den sehr tiefen Korrelationen zwischen Schätz- und Messwerten bei diesen Grössen. Rationale Splines bringen nur hier geringfügig bessere Resultate. Für die Sonnenscheindauer kann dieses Problem mit Hilfe von Informationen über Sonnenauf- bzw. untergang und Setzen eines Stützpunktes beim Maximum etwas verbessert gelöst werden, während beim Niederschlag wohl auch weiterhin ein beträchtlicher Teil der Ersatzwerte von Hand eingegeben werden muss.
- Bei den übrigen Grössen können die negativen Eigenschaften der Splinefunktionen (Unter- bzw. Überschätzen bei Maximum bzw. Minimum) weitgehend eliminiert werden, indem die Stützpunkte innerhalb der zu interpolierenden Periode mit Hilfe von Information von Nachbarstationen beim zu erwartenden Maximum bzw. Minimum gesetzt werden (vgl. Anhang 2).