

Deutscher Wetterdienst



Berichte des Deutschen Wetterdienstes

216

**Ablauf der
Datenprüfung und -vervollständigung
im Geschäftsfeld Landwirtschaft**

von
Wolfgang Jansen und Rainer Sedlatschek

Zur Herstellung dieses Buches wurde chlor- und säurefreies Papier verwendet.

ISSN 0072-4130
ISBN 3-88148-374-8

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten. Kein Teil dieses Werkes darf ohne schriftliche Einwilligung des Deutschen Wetterdienstes in irgendeiner Form (Fotokopie, Mikrofilm, oder ein anderes Verfahren), auch nicht für Zwecke der Unterrichtsgestaltung, reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden. Für den Inhalt ist der Autor verantwortlich.

Herausgeber und Verlag:
Deutscher Wetterdienst
Frankfurter Str. 135
63067 Offenbach am Main

Anschrift der Autoren:
Dipl.-Met. Wolfgang Janssen
Rainer Sedlatschek
Deutscher Wetterdienst
Geschäftsfeld Landwirtschaft
Frankfurter Str. 135
63067 Offenbach a. M.

Inhaltsverzeichnis

| | |
|---|-----------|
| Zusammenfassung/Abstract | 5 |
| 1 Einleitung | 7 |
| 2 Agrarmeteorologische Datenbank | 8 |
| 3 Allgemeine Vorbemerkungen | 11 |
| 4 Datenprüfung | 12 |
| 4.1 Formale Prüfung | 12 |
| 4.2 Klimatologische Prüfung | 12 |
| 4.2.1 Temperaturen | 12 |
| 4.2.2 Niederschläge | 14 |
| 4.2.3 Strahlung | 15 |
| 4.2.4 Andere Elemente | 16 |
| 4.3 Zeitliche Konsistenzprüfung | 17 |
| 4.3.1 Temperaturen | 17 |
| 4.3.2 Wind | 19 |
| 4.3.3 Andere Elemente | 19 |
| 4.4 Innere Konsistenzprüfung | 20 |
| 4.4.1 Temperaturen | 20 |
| 4.4.2 Niederschläge | 22 |
| 4.4.3 Strahlung | 24 |
| 4.4.4 Wind | 25 |
| 4.4.5 Andere Elemente | 26 |
| 4.5 Manuelle Bestätigung | 27 |
| 4.6 Fehlerstatistik | 29 |
| 5 Datenvervollständigung | 31 |
| 5.1 Manuell einfügen oder korrigieren | 31 |
| 5.2 Sonstige Verfahren | 31 |
| 5.2.1 Temperaturen | 32 |
| 5.2.2 Niederschläge | 33 |
| 5.2.3 Strahlung | 35 |
| 5.2.4 Andere Elemente | 37 |
| 5.3 Zeitliche Interpolation | 37 |
| 5.3.1 Temperaturen | 37 |
| 5.3.2 Niederschläge | 38 |
| 5.3.3 Strahlung | 38 |
| 5.3.4 Wind | 38 |
| 5.3.5 Andere Elemente | 39 |
| 5.4 Räumliche Interpolation | 39 |
| 5.4.1 Temperaturen | 39 |

| | | |
|----------|--|-----------|
| 5.4.4 | Wind | 41 |
| 5.4.5 | Andere Elemente | 41 |
| 5.5 | Erfolgsquote der Vervollständigung | 42 |
| 6 | Ausblick | 42 |
| 7 | Literatur | 43 |

Zusammenfassung/Abstract

Im Geschäftsfeld Landwirtschaft des Deutschen Wetterdienstes werden seit geraumer Zeit agrar-meteorologische Modelle betrieben, aus denen Beratungshinweise für die Landwirtschaft abgeleitet werden. Diese Modelle benötigen einen möglichst richtigen und unbedingt vollständigen Input-Datensatz in Form von gemessenen meteorologischen Größen. Aus diesem Grunde werden alle benutzten meteorologischen Größen intensiv geprüft und ggf. auch korrigiert und Lücken im Datengut vervollständigt. Diese Datenprüfung und -vervollständigung wird sehr zeitnah zum Meldetermin in vollautomatisierter Form durchgeführt. Gefundene Meßfehler werden in einem Fehlerprotokoll gesammelt und nachträglich manuell bearbeitet. Alle original gemeldeten Meßwerte werden zusammen mit dem "richtigen", vervollständigten Datensatz in eine eigene Datenbank geschrieben, um sie innerhalb des DWD oder auch an externe Kunden abgeben zu können.

Durch das tägliche Datenmonitoring und das Hinzunehmen weiterer Elemente in die Datenbank ist eine laufende Pflege und Verbesserung der hier in diesem Dokument beschriebenen Prüf- und Vervollständigungsverfahren notwendig und somit spiegeln die beschriebenen Module lediglich den jetzigen Stand wider.

For several years now agrometeorological models have been operated in Business Unit Agriculture at the German Meteorological Service with the aim of providing an advisory service for farmers. These models require an absolutely complete input data set that is as correct as possible. For this reason all measured data undergo an intensive quality check and, if necessary, are corrected and completed. This quality check and completion is carried out completely automatically and very close to real time. Measuring errors found are collected in a special log file and subsequently edited manually. At the same time the log file is a check for used algorithms. All originally measured values are filed together with the "correct" and completed data set in a separate database, from which all customers are provided.

Continuous modification of all algorithms described in this paper is necessary, due to the daily data monitoring and the inclusion of further elements in the database. Consequently, this paper describes the status quo only.

1 Einleitung

Die Ursprünge der aktuellen Datenprüfung gehen auf die Zeit von 1990 zurück. Im Rahmen eines Forschungsvorhabens, welches von dem Statistischen Amt der Europäischen Union finanziert wurde, wurden zahlreiche Prüfkventionen ausgearbeitet, die zum Teil in die jetzige Datenprüfung übernommen wurden. Die theoretische Ausarbeitung der Prüfkriterien wurde in einem Werkvertrag von G. Augter durchgeführt. Die Implementierung dieser Prüfkriterien in ein Programmsystem inklusive Datenbank wurde von den damaligen Mitarbeitern des Referates AM2 ausgearbeitet und von R. Vogelgesang (1993) beschrieben.

Zunächst soll erläutert werden, warum sich gerade das Geschäftsfeld Landwirtschaft seit geraumer Zeit intensiv mit der Problematik einer Datenprüfung und -vervollständigung beschäftigt. Alle im Geschäftsfeld betriebenen agrarmeteorologischen Modelle benötigen in irgendeiner Form meteorologischen Input in täglicher, aber auch stündlicher Auflösung. Die Modelle benötigen einen kompletten Datensatz, damit sie problemlos von der Vergangenheit in die Zukunft gerechnet werden können. Die Modellergebnisse werden täglich zu festen Zeiten zentral und dezentral in Produkten (Wetterfax für die Landwirtschaft, Agromet-Online ...) eingebunden und an Kunden abgegeben, so daß nicht darauf gewartet werden kann, bis alle Werte richtig und vollständig gemeldet werden bzw. über die verschlungenen Wege der nationalen und internationalen Datenkommunikation im meteorologischen Rechenzentrum angekommen sind. Die Datenprüfung und -vervollständigung werden zentral durchgeführt, um eine möglichst große Datenbasis nutzen zu können und um ein gemeinsames Datenkollektiv innerhalb des Geschäftsfeldes Landwirtschaft zu erzeugen. Eine dezentrale Vervollständigung der Daten könnte dieses nicht erreichen und würde in der Summe zu einer Mehrarbeit führen.

Kein Bereich innerhalb des Wetterdienstes sah bisher die Notwendigkeit, die gemeldeten Wetterdaten so zeitkritisch zu prüfen, geschweige denn zu vervollständigen. Selbst wenn im Geschäftsbereich Technische Infrastruktur bessere Prüfalgorithmen als die bisherigen mit entsprechenden Qualitätsbytes installiert werden würden, machen diese noch keinen kompletten Datensatz. Wenn als falsch erkannte oder fehlende Werte ersetzt werden, so müssen auch diese ersetzten Werte einer Prüfung unterzogen werden. Außerdem ist es von Vorteil, wenn die Ersetzalgorithmen eine "Ahnung" davon haben, warum ein Wert falsch gegeben wurde. Es kann festgehalten werden, daß eine Datenvervollständigung ohne gleichzeitige Datenprüfung nicht vorstellbar ist.

Die Datenprüfung und -ersetzung in dieser komplexen Form wird seit Anfang 1994 im Geschäftsfeld Landwirtschaft betrieben und fortlaufend aktualisiert, da das aktuelle Wettergeschehen weiteren Modifikationsbedarf auslöst. Somit kann nur eine Beschreibung des momentanen Zustandes erfolgen. Außerdem ist nicht ausgeschlossen, daß mit weiteren Elementen, die in die agrarmeteorologische Datenbank aufgenommen werden, zusätzliche Prüf- und Vervollständigungsverfahren installiert werden.

Zum besseren Verständnis der recht trockenen Materie ist es von Vorteil, auch das Umfeld in Form der zentralen agrarmeteorologischen Datenbank zu kennen, denn sie gibt den Rahmen für jegliches Handeln und Wandeln.

2 Agrarmeteorologische Datenbank

Mit der Migration von NOS/VE zu UNIX war zunächst eine kritische Bestandsaufnahme vom damaligen Stand durchzuführen. Hierbei fiel auf, daß es fast ausschließlich große Programmpakete gab, die vollkommen unabhängig voneinander lauffähig waren und identischen Input immer wieder neu vom Anfang der Verarbeitungskette aufbereiteten sowie die Ergebnisse in verschiedenen Dateien und Formaten ablegten. Zur weiteren Verarbeitung dieser Ergebnisse waren wieder eigenständige Programme notwendig, und kein Programmpaket hatte eine Ahnung von den anderen Programmen. Bei dieser Struktur ist ein "Auswuchern" des zu betreuenden Programmcodes und Rechenaufwandes mit Ausweitung der Aufgabestellung "vorprogrammiert". Somit verursachte die Migration nicht nur Arbeit, sondern bot auch die Chance, ein "ganz neues" Konzept einzuführen, welches langfristig den Programmieraufwand minimieren sollte.

Naturgemäß hat dieses natürlich am Anfang genau den gegenteiligen Effekt, da wenig von den alten Programmen übernommen werden kann. Bei diesem neuen Konzept hatte eine neu zu entwickelnde Datenbank die zentrale Schlüsselposition bei der gesamten Verarbeitung der Daten. Alle Daten, die in irgendeiner Form benötigt werden, sollten in dieser Bank abgelegt werden, damit bei der individuell zu gestaltenden Datenversorgung der Außenstellen und Kunden auf ein einziges Datenkollektiv zurückgegriffen werden kann.

Soviel zu der Theorie, nun zur Praxis. Welche Datenbank sollte genommen werden und was sollte sie leisten können. Zum Zeitpunkt dieser Überlegungen wurde im DWD noch nicht an eine kommerzielle Datenbank gedacht, geschweige denn betriebsbereit vorgehalten. Selbst wenn eine geeignete Software gefunden worden wäre, hätte diese in die Betriebslandschaft des Wetterdienstes passen müssen. Diese Randbedingungen versprachen keinen schnellen Erfolg und bestätigten sich im Nachhinein bei der Beschaffung von Oracle und wurden zugegebenermaßen vom Geschäftsfeld Landwirtschaft nicht überblickt. Aus diesem Grunde wurde eine eigene Datenbank unter Fortran realisiert, wie sie von R. Vogelgesang (1993) beschrieben wurde.

Anforderungen an diese Datenbank:

- Die Bank ermöglicht schnellen Zugriff auf die Daten.
- Die Bank beansprucht möglichst wenig Speicherplatz.

Es wurden nur die Elemente aufgenommen, die in irgendeiner Form für die agrarmeteorologische Beratung von Interesse oder für die innere Konsistenzprüfung benötigt werden. Die Daten werden komprimiert, und es wird nur soviel Platz in der Bank vorgehalten, wie es das Meldegutaukommen und der Wertebereich erforderlich machen. Der schnelle Zugriff wird über eine direkt adressierbare Datei gewährleistet. Da die agrarmeteorologische Beratung vorwiegend zeitkritisch arbeitet, wird nur ein Zeitraum von 55 Tagen (aktueller Tag -48 Tage bis aktueller Tag +7 Tage) vorgehalten, so daß zukünftige und rückwirkende agrarmeteorologische Betrachtungen durchgeführt werden können. Es wird in der Bank nur für die Stationen Platz vorgehalten, die auch benötigt werden. Bei den Stationen werden verschiedene Stationstypen unterschieden und für jeden Typ wird eine eigene Bank angelegt. Das Konzept der verschiedenen Bänke für jeden Stationstyp hat indirekt eine schnellere Datenverarbeitung zur Folge, weil z.B. die deutschen Synop-Stationen häufiger bearbeitet werden als die Länderstationen aus Hessen und Rheinland-Pfalz und somit nur diese Datei bzw. Bank der deutschen Synop-Stationen auf den Ausfallrechnern aktualisiert werden muß. Insgesamt gibt es fünf verschiedene Datentypen bzw. Bänke, wobei jede Station eindeutig einer Bank zugeordnet ist. Aktueller Stand Ende 1998:

- Deutsche Synop-Stationen (212 Stationen)
- Europäische Synop-Stationen (700 Stationen)
- Europäische Synop-Stationen im Grenzbereich zu Deutschland (95 Stationen)

- Stationen am Ort von Individualkunden (60 Stationen)
- Länderstationen aus Hessen, Rheinland-Pfalz und Bayern (35 Stationen)

Für jede Station und Tag stehen insgesamt maximal 10 Datensätze zur Verfügung, in denen die Stunden- bzw. Tageswerte der jeweiligen Datenbankelemente an vordefinierten Stellen abgelegt werden können.

- Die Bank soll zwischen verschiedenen Qualitätsstufen unterscheiden.
- Der Ursprung der gespeicherten Daten sollte nachvollziehbar sein.
Damit zwischen originalen und richtigen Werten (z.B. falsches Original wird durch einen richtigen Wert ersetzt) unterschieden werden kann, ist jeder Datensatz (Zeile in der Bank) in eine "rechte und linke Hälfte" geteilt, um ein Element doppelt vorzuhalten. Auf der linken Seite stehen die gemeldeten Originalwerte und auf der rechten Seite die "richtigen Werte", die natürlich identisch zu den Originalwerten sein können.
Neben dem eigentlichen Zahlenwert wird ebenfalls ein Qualitätsbyte für jedes Element auf beiden Seiten abgespeichert, aus dem der Ursprung und das Qualitätsniveau hervorgeht (siehe auch QB-Tabelle 2).

Bänke

| | | | | |
|-----------------------------|--------------------------------|---------------|--------------------------|----------------------|
| Deutsche SYNOP-Stationen | Europäische SYNOP-Stationen | Randstationen | Individual- Stationen | Länder- Stationen |
|-----------------------------|--------------------------------|---------------|--------------------------|----------------------|

Bankaufbau

| Datum | Stationen | Datensatz-Nr.: |
|----------|-----------|----------------|
| 19981126 | 10015 | 1 |
| ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | 9 |
| ⋮ | 10020 | 1 |
| ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | 9 |
| 19981127 | 10015 | 1 |
| ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | 9 |
| ⋮ | 10020 | 1 |

Datensatz

| | |
|---|---|
| Originale (Linke Seite) | Richtige Werte (Rechte Seite) |
| Wert1, QB1, Wert2, QB2, Wert3, QB3, ... | Wert1, QB1, Wert2, QB2, Wert3, QB3, ... |

Über die Qualitätsbytes und die zwei Seiten eines Datensatzes lassen sich über eine Einheitsausgabeschnittstelle gezielt Daten verschiedener Qualitätsstufen herausziehen. Es wird zur Zeit zwischen drei Qualitätsstufen, die nicht mit den Qualitätsbytes zu verwechseln sind, unterschieden:

- Richtige Originalwerte: Von den gemeldeten Werten (linke Seite des Datensatzes) werden nur die Werte ausgegeben, die die Datenprüfung bestanden haben.

- Richtige Werte: Alle Werte, die die Datenprüfung bestanden haben, wobei die Herkunft egal (rechte Seite vom Datensatz) ist.
 - Alle originalen Werte: Alle als falsch und richtigen gemeldeten Werte (gesamte linke Seite des Datensatzes).
- Die Bank soll alle benötigten Daten enthalten.
 - Die Einführung neuer Elemente und Stationen sollte möglichst einfach sein.
Um in dem abschließenden Prozess nach der Datenprüfung und -vervollständigung sowie den Modellrechnungen Daten möglichst bequem und in einem einheitlichen Format abgeben zu können, werden auch alle Ergebnisse der Modellrechnungen (Bodenfeuchte, -temperatur, Kornfeuchte ...) auf die rechte Seite des Datensatzes geschrieben. Auf diese Daten kann dann wieder über die Einheitsausgabeschnittstelle zugegriffen werden.
Im Laufe der Zeit hat sich die Anzahl der in den Bänken enthaltenen Stationen enorm ausgedehnt, so daß fast monatlich Stationsänderungen vorzunehmen sind. Diese sind über eine Stationsdatei –in der alle relevanten Stationsangaben enthalten sind– einfach zu realisieren. Über diese Stationsdatei und ein entsprechendes Programm werden die Bänke neu dimensioniert und gegebenenfalls Datensätze angelegt oder gelöscht. Fehleranfälliger und weniger komfortabel ist dagegen das Ändern von Datenbankelementen, weil hier der benötigte Platz und die Adresse im jeweiligen Datensatz definiert werden muß. Kleine Fehler können hier die gesamte Adressberechnung und Wertezuweisungen durcheinanderbringen. Daher werden in diesem Fall Parallelbänke mit den neuen Elementen angelegt und parallel zur Routine-Bank auf 55 Tage "angefüttert" und laufend kontrolliert, bis sie für die Routine-Datenverarbeitung genutzt werden.

Neben den Bänken und der Stationsdatei ist eine Definitionsdatei für den Betrieb der Bank notwendig, in der alle Informationen über die Bänke, Datensätze, Elemente deren Elementkennungen, Elementerklärung, Einheit, Meldetermine sowie die Position im jeweiligen Datensatz enthalten sind. Dies war nur eine grobe Beschreibung der agrarmeteorologischen Datenbank, die in ihrer jetzigen Form nach reichlichen Verbesserungen und Optimierungen die Erwartungen hinsichtlich Schnelligkeit und Brauchbarkeit als zentrales Datensammelsystem voll erfüllt und im Prinzip auf jedes Betriebssystem überführt werden kann. Falls zukünftig die Daten in die zentrale Oracle-Datenbank integriert werden sollen, ergibt sich hieraus hoffentlich nur ein geringer Programmieraufwand, weil "nur" die Ein- und Ausgabeschnittstelle geändert werden muß. Da der Ausgabeschnittstelle ein Programm (gibd) nachgeschaltet ist, welches neben der reinen Bankausgabe auch einfache Berechnungen durchführt und in einem einheitlichen Ausgabeformat ausgibt, wird sich bei der weiteren Verarbeitung auf Seiten der Abnehmer nichts ändern.

3 Allgemeine Vorbemerkungen

In den nachfolgenden Kapiteln wird permanent von verschiedenen Qualitätsniveaus gesprochen und akribisch monoton der jeweilige momentane Qualitätszustand erwähnt, denn dieses ist für die Aktionen der Datenprüfung sehr entscheidend. Als Orientierungshilfe für die nachfolgenden Kapitel werden zunächst die verschiedenen Prüfzustände bzw. Qualitätsniveaus (Zeilen) sowie als Vorgriff von Kapitel 5 die möglichen Herkünfte der Werte (Spalten) und hieraus resultierenden Qualitätsbytes dargestellt.

Tabelle 2: Qualitätsbytes

| Prüfzustand | Herkunft | | | | | | |
|-------------------|----------|------------|-------------|----------|-------------|----------|---------|
| | Original | Vorhersage | 1. zeitlich | räumlich | 2. zeitlich | sonstige | manuell |
| ungeprüft | A | a | H | O | V | k | r |
| formal richtig | B | b | I | P | W | l | s |
| klimat. sinnvoll | C | c | J | Q | Y | m | t |
| zeitl. konsistent | D | d | K | R | Z | n | u |
| inn. konsistent | E | e | L | S | h | o | v |
| räuml. konsistent | F | f | M | T | i | p | w |
| manuell bestätigt | G | | N | U | j | q | y |
| fachlich falsch |) |] | - | ! | & | , | % |
| manuell falsch | ? | | | | | | |
| fehlend | X | | | | | | |

Mit Hilfe dieser Tabelle lassen sich eindeutig die Herkunft und das Qualitätsniveau eines Wertes nachvollziehen. Hat z.B. ein Wert das QB "o" so wurde er über *sonstige Verfahren* (in Abschnitt 5.2) ersetzt und hat die *innere Konsistenzprüfung* (Abschnitt 4.4) bestanden. Außerdem wird schon ersichtlich, daß die Daten nicht nur einmal geprüft werden –was für die Bestimmung der Güte von den original gemessenen Daten nötig wäre– sondern nach jedem in Kapitel 5 beschriebenen Ersetzungsschritt nochmals mit den gleichen Prüfalgorithmen geprüft werden.

Am Anfang der Kette steht die synoptische Datenbank des Rechenzentrums, aus der alle benötigten meteorologischen Input-Daten extrahiert werden und auf die linke Seite der in Kapitel 2 beschriebenen Bank mit Qualitätsstufe *ungeprüft* geschrieben werden. Auf diese ungeprüften originalen Daten setzt die Datenprüfung an. Nach dem gesamten in Kapitel 4 beschriebenen Prüfzyklus sind die meisten Werte als richtig erkannt worden und haben das entsprechende QB der zuletzt durchgeführten Prüfung. Wird ein Originalwert in der Prüfungshierarchie nur bis zur *klimatologischen Prüfung* geprüft, so bekommt es das QB "C". Die Elemente, die als falsch erkannt wurden oder nicht gemeldet bzw. *fehlend* sind und als vollständiger Datensatz vorliegen sollen, werden in den nachfolgenden Vervollständigungsschritten ersetzt und auf das Prüfniveau *ungeprüft* gesetzt und den gleichen Prüfschritten unterzogen. Am Ende aller Vervollständigungsschritte mit nachfolgender Prüfung sollten alle zu vervollständigende Elemente richtig sein, was in der Regel auch gelingt. Die Werte, welche während dieser Schritte einmal als richtig erkannt wurden, brauchen in den nachfolgenden Schritten nicht mehr geprüft werden, wodurch die Datenprüfung mit jedem Vervollständigungsschritt immer schneller wird. Am Ende des gesamten Verarbeitungszyklusses werden alle Werte mit entsprechendem QB auf die rechte Seite des Datensatzes geschrieben. Bei richtigen Originalwerten sind somit die linke und rechte Seite des Datensatzes identisch.

Um kleine Modifikationen in der Datenprüfung möglichst einfach durchführen zu können, werden alle freien Variablen oder Grenzwerte aus einer separaten Datei (Grenzwerte-Datei) in den verschiedenen Programmteilen eingelesen, so daß Änderungen sehr schnell und ohne Programmänderung umgesetzt werden können.

4 Datenprüfung

In der Datenprüfung werden jeweils die Werte eines Tages geprüft. Es können zwar mehrere Tage hintereinander geprüft werden, jedoch werden lediglich die Werte eines Tages innerhalb des gesamten Zeitraum einer Prüfung unterzogen. Je nach Element liegen unterschiedliche Bezugszeiträume vor. Um alle möglichen Bezugszeiträume abzudecken, werden für die Prüfung des gestrigen Tages, die Werte von vorgestern 21 UTC bis heute 6 UTC in der Prüfung berücksichtigt, wobei lediglich die Werte von gestern 7 bis heute 6 UTC einer Prüfung unterzogen werden. Die Qualitätsbytes der Werte von vorgestern 21 UTC bis gestern 6 UTC bleiben unverändert und dienen lediglich als Randinformation zur Bildung von Extrem- bzw. Summenwerten.

4.1 Formale Prüfung

Damit die Elemente der Datenbank über die Ausgabeschnittstelle ausgegeben werden können, dürfen sie nicht das Qualitätsbyte *ungeprüft* behalten, denn keine der drei in Kapitel 2 beschriebenen Qualitätsstufen holt ungeprüfte Werte. Dies hat zur Konsequenz, daß alle in der Bank liegenden Werte zumindest einer formalen Prüfung unterzogen werden müssen.

Die in Tabelle 3 aufgelisteten formalen Prüfkriterien der einzelnen Elemente geben somit gleichzeitig Aufschluß über den Bankinhalt an gemeldeten Elementen.

Die von den automatischen Stationen gemeldeten **W1** und **ww** werden vor dem Einbringen in die Bank in entsprechende **ww** von bemannten Stationen umgewandelt. Bei den Elementen **CL**, **TR00**, **TR06**, **TR12**, **TR18**, **TR24**, **N**, **VV** und **E** endet bereits an dieser Stelle die Datenprüfung, was nicht heißen soll, daß sie später nicht mehr benötigt oder vervollständigt (**N**) werden. So werden **CL**, **VV** und **N** bei der Überprüfung des **ww** benutzt und die Bezugszeiträume vom Niederschlag intern dazu verwendet, um im Bedarfsfall die Niederschläge von **RR06** und **RR18** –bei den Vorhersagen und im Ausland liegt der Niederschlag teilweise nur mit 6-stündigem Bezugszeitraum vor– auf 12-stündige Bezugszeiträume aufzusummieren, damit die Daten einheitlich vorliegen. Nach bestandener Prüfung erreichen die Werte die erste Qualitätsstufe *formal richtig*.

4.2 Klimatologische Prüfung

Die meisten der noch verbleibenden Elemente werden im nächsten Prüfschritt einer klimatologischen Prüfung unterzogen. Besteht der Wert diese Prüfung kommt er in die nächste Qualitätsstufe *klimatologisch sinnvoll*, im anderen Fall wird er *fachlich falsch*.

4.2.1 Temperaturen

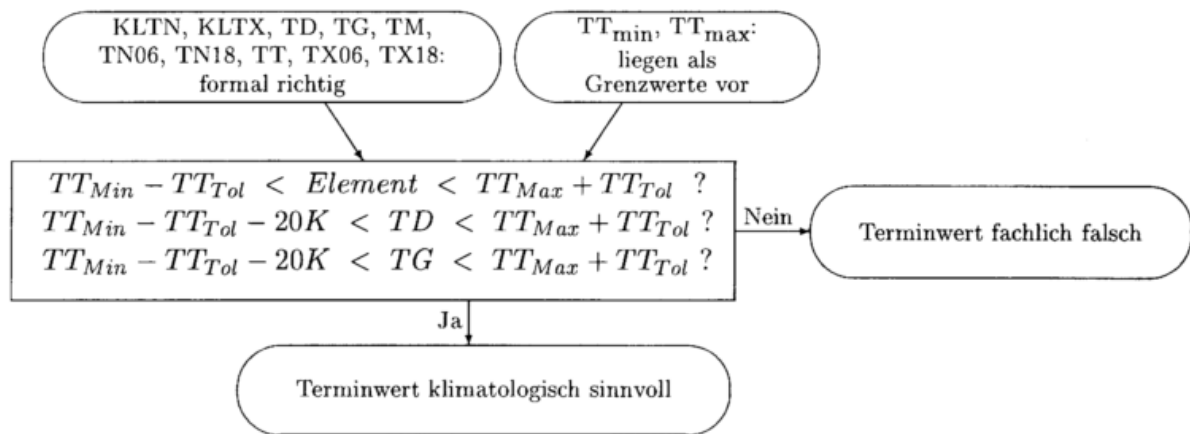
- **TT**, **TD**, **TM**, **KLTN**, **KLTX**, **TG**, **TN06**, **TN18**, **TX06**, **TX18**

In der ersten Version der Datenprüfung wurde Europa in 16 verschiedene Klimaregionen eingeteilt. Jede Station wurde einer Region zugeordnet, für die eine Minimum- TT_{Min} und eine Maximumtemperatur TT_{Max} monatsweise vorgegeben wurde. Mit zunehmender Ausweitung der europäischen Stationen wurde die regionale Zuordnung immer schwieriger und Spezialeffekte wie der Föhn im Alpenvorland bewirkten einen zu großzügigen Temperaturbereich in den jeweiligen Regionen. Aus diesem Grunde wurde in jüngster Zeit dazu übergegangen, monats- und stationsweise TT_{Min} und TT_{Max} vorzugeben. Es läßt sich sicherlich erahnen, welcher gigantische Aufwand hiermit verbunden war, für möglichst jede Station einen möglichst langen Zeitraum an Messungen aufzutreiben und diesen trotz vieler Falschmeldungen auf realistische Werte auszuwerten. Es ist klar, daß von einigen –z.B. neu eingerichteten– Stationen nur kurze Zeiträume oder bei anderen keine Werte zur Verfügung standen bzw. stehen, die das Durchhalten des stationsbezogen Konzeptes sehr erschweren. Aus diesem Grunde wurde für jeden Extremwert zusätzlich noch ein Wert für die Toleranz TT_{Tol} eingebaut, der je nach Länge des Auswertungszeitraumes zwischen 3 und 7 K variieren kann. Bei komplett fehlenden Extremwerten werden die der benachbarten Stationen verwendet. Diese klimatologischen

Tabelle 3: Gemessene Elemente in der AM-Datenbank und deren formalen Prüfkriterien.

| Element | Bezeichnung | Einheit | Wertebereich |
|---------|--------------------------------------|----------------|--------------|
| A | Art der Luftdruckänderung | Ziffer | 0 – 8 |
| CL | Art der tiefen Wolken | Ziffer | 0 – 10 |
| DD | Windrichtung | Grad | 0 – 360 |
| E | Erdbodenzustand | Ziffer | 0 – 9 |
| ES | Erdbodenzustand bei Schnee | Ziffer | 0 – 9 |
| FF | Windgeschwindigkeit | $1/10$ m/s | 0 – 990 |
| FX | Max. Böe | $1/10$ m/s | 108 – 990 |
| GS | Stündl. Globalstrahlung | J/cm^2 | 0 – 400 |
| GS24 | Tägliche Globalstrahlung | J/cm^2 | 0 – 4000 |
| KLRR24 | Klimat. 24h Niederschlag | $1/10$ mm | 0 – 3000 |
| KLTN | Klimat. Temperaturminimum | $1/10$ K | 2232 – 3232 |
| KLTX | Klimat. Temperaturmaximum | $1/10$ K | 2232 – 3232 |
| N | Bedeckungsgrad | Achtel | 0 – 9 |
| PPP | Druckänderung der letzten 3 Stunden | $1/10$ hPa | 0 – 200 |
| PPPP | Luftdruck auf NN | $1/10$ hPa | 8000 – 10800 |
| SD | Tägl. Sonnenscheindauer | $1/10$ Stunden | 0 – 240 |
| SN | Schneehöhe | cm | 0 – 998 |
| R1R1 | Stündl. Niederschlag | $1/10$ mm | 0 – 1200 |
| RR00 | 6-stünd. Niederschlag bis 00UTC | $1/10$ mm | 0 – 3000 |
| RR06 | 12-stünd. Niederschlag bis 06UTC | $1/10$ mm | 0 – 3000 |
| RR12 | 6-stünd. Niederschlag bis 12UTC | $1/10$ mm | 0 – 3000 |
| RR18 | 12-stünd. Niederschlag bis 18UTC | $1/10$ mm | 0 – 3000 |
| RR24 | 24-stünd. Niederschlag bis 06UTC | $1/10$ mm | 0 – 3000 |
| TB02 | Erdbodentemp. in 2cm 07/14/21UTC | $1/10$ K | 2232 – 3332 |
| TB05 | Erdbodentemp. in 5cm 07/14/21UTC | $1/10$ K | 2332 – 3232 |
| TB10 | Erdbodentemp. in 10cm 07/14/21UTC | $1/10$ K | 2432 – 3132 |
| TB20 | Erdbodentemp. in 20cm 07/14/21UTC | $1/10$ K | 2483 – 3082 |
| TB50 | Erdbodentemp. in 50cm 07/14/21UTC | $1/10$ K | 2532 – 3032 |
| TB100 | Erdbodentemp. in 100cm 14UTC | $1/10$ K | 2582 – 2982 |
| TD | Taupunkt | $1/10$ K | 2132 – 3132 |
| TG | Erdbodenminimum in 5cm Höhe 06UTC | $1/10$ K | 2132 – 3132 |
| TM | Tagesmittel der Lufttemperatur | $1/10$ K | 2232 – 3132 |
| TN06 | Temperaturminimum der Luft bis 06UTC | $1/10$ K | 2232 – 3132 |
| TN18 | Temperaturminimum der Luft bis 18UTC | $1/10$ K | 2232 – 3132 |
| TR00 | Bezugszeitraum RR00 | Stunden | 6 – 6 |
| TR06 | Bezugszeitraum RR06 | Stunden | 12 – 12 |
| TR12 | Bezugszeitraum RR12 | Stunden | 6 – 6 |
| TR18 | Bezugszeitraum RR18 | Stunden | 12 – 12 |
| TR24 | Bezugszeitraum RR24 | Stunden | 24 – 24 |
| TT | Lufttemperatur | $1/10$ K | 2232 – 3232 |
| TX06 | Temperaturmaximum der Luft bis 06UTC | $1/10$ K | 2232 – 3232 |
| TX18 | Temperaturmaximum der Luft bis 18UTC | $1/10$ K | 2232 – 3232 |
| VV | Sichtweite | Metern | 0 – 75000 |
| W1 | Wetter im Bezugszeitraum | Ziffer | 0 – 9 |
| ww | Wetter zum Termin | 0 | 0 – 99 |

Extremwerte werden fortlaufend aktualisiert, um so indirekt den Auswertungszeitraum zu erhöhen.



• **TB02, TB05, TB10, TB20, TB50, TB100**

Die Bodentemperaturen werden im wesentlichen in der gleichen Art und Weise wie die anderen Temperaturen klimatologisch geprüft. In den oberen Bodenschichten treten größere Temperaturschwankungen als in der Luft auf, so daß in Anlehnung an die Klimatologischen Grenzwerte der Lufttemperatur TT_{Min} und TT_{Max} die Grenzwerte tiefenabhängig vorgegeben werden müssen.

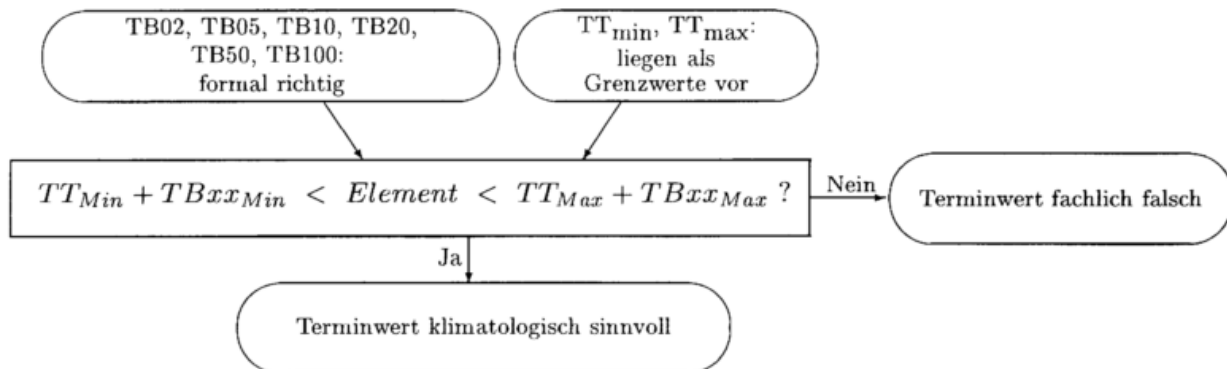


Tabelle 4: Tiefenabhängige Temperaturzuschläge TB_{xxMin} und TB_{xxMax} zu den klimatologischen Lufttemperaturgrenzwerten TT_{Min} und TT_{Max} .

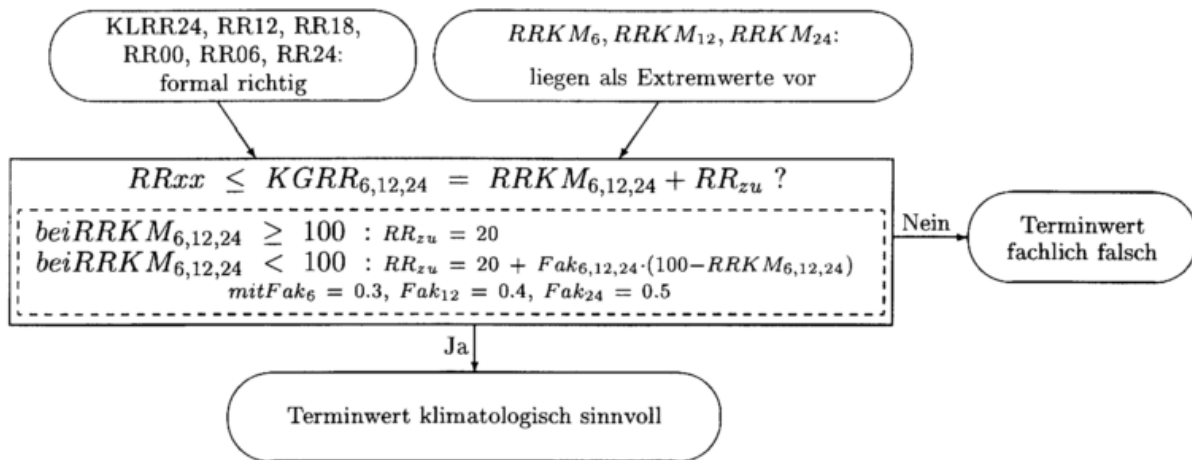
| | TB02 | TB05 | TB10 | TB20 | TB50 | TB100 |
|--------------------|------|------|------|------|------|-------|
| TB_{xxMin} in °C | -8 | -6 | -3 | 0 | 3 | 6 |
| TB_{xxMax} in °C | 15 | 13 | 7 | 2 | -1 | -4 |

4.2.2 Niederschläge

• **RR00, RR12, RR06, RR18, RR24, KLRR24**

Bei der klimatologischen Niederschlagsprüfung wurden früher die Grenzwerte für den 6-, 12- und 24-stündigen Niederschlag über Höhenintervalle monatsweise angegeben. Ähnlich wie bei der Temperatur werden die Toleranzen zu groß, wenn man z.B. in den Wintermonaten schwedische und mediterrane Stationen in einen Topf wirft. Deshalb wurde auch bei der klimatologischen Niederschlagsprüfung auf ein stationsbezogenes Konzept zurückgegriffen, wobei sich hier die gleichen Probleme wie bei der Temperatur ergaben und in der gleichen Art und Weise gelöst wurden. Jedoch wird die Toleranz nicht für jeden Extrem- bzw. Grenzwert angegeben, sondern für alle Stationen und Monate nach dem gleichen Verfahren berechnet. In Anbetracht der Tatsache, daß der Niederschlag kleinräumig sehr unterschiedliche Höchstwerte

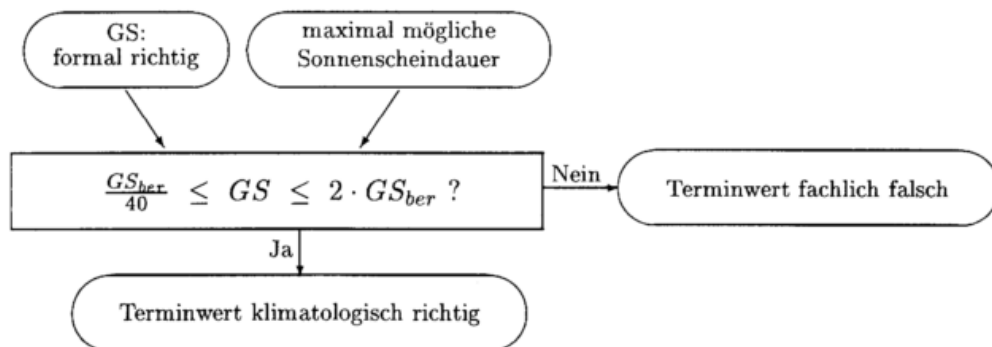
aufweisen kann, muß der klimatologische Grenzwert $KGRR_{6,12,24}$ recht großzügig angegeben werden. Grundsätzlich wird eine Toleranz von 20mm zum bisherigen gemeldeten maximalen Niederschlag $RRKM_{6,12,24}$ und im Bedarfsfall ein zusätzlicher Term RR_{zu} addiert.



4.2.3 Strahlung

- **GS**

Bei GS werden alle Stundenwerte überprüft. Über Sonnenaufgang und -untergang werden die Nachtstunden bestimmt, in der GS immer 0 sein muß, was allerdings aufgrund einer ungenauen Nulljustierung einiger Strahlungsmeßgeräte nicht immer der Fall ist. Am Tage darf sich GS nur innerhalb bestimmten Grenzen bewegen, die über das im Kapitel 5.2.3 beschriebene Verfahren zur Berechnung von GS_{ber} ermittelt werden.



- **GS24**

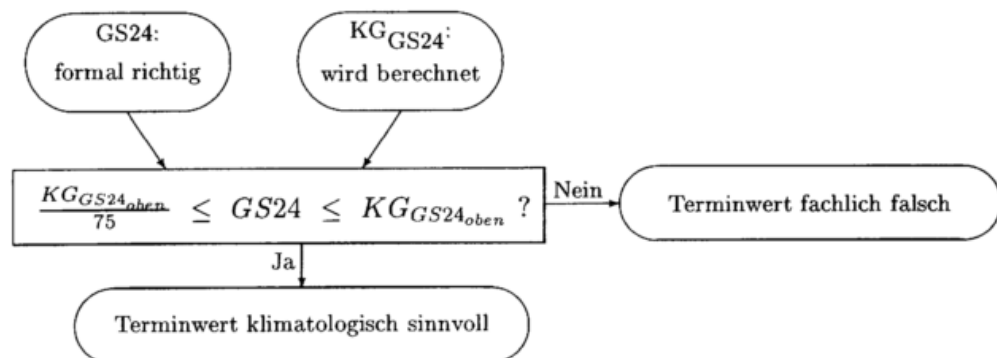
Die klimatologischen Grenzwerte für die Globalstrahlung wurden in der Vergangenheit häufig abgeändert und scheinen in ihrem jetzigen Zustand endlich ausgereift zu sein. Die Bestimmung der oberen $KG_{GS24_{oberen}}$ und unteren $KG_{GS24_{unteren}}$ klimatologischen Grenze für GS24 ist recht aufwendig und basiert im wesentlichen auf physikalischer Grundlage. Über das Datum wird die nicht konstante Solarkonstante SK_{sol} und über die geographische Breite und Datum der mittlere Sonnenstand SW_{mit} während des Tages und die astronomisch mögliche Tageslänge tl wird die maximale kurzwellige Strahlung kS_{max} berechnet, die die Erdoberfläche erreichen kann. Hierbei werden die Polarnacht bzw. -tag mitberücksichtigt werden.

$$kS_{max} = SK_{sol} \cdot tl \cdot 0.75 \sin(SW_{mit}) \quad (1)$$

Der Faktor 0.75 ist auf die Absorption der solaren Strahlung innerhalb der Atmosphäre zurückzuführen. Zusätzlich zu kS_{max} wird eine Toleranz von

$$GS24_{Tol} = \max(0, \frac{2500 - kS_{max}}{10000}) \cdot kS_{max} + 50 \quad (2)$$

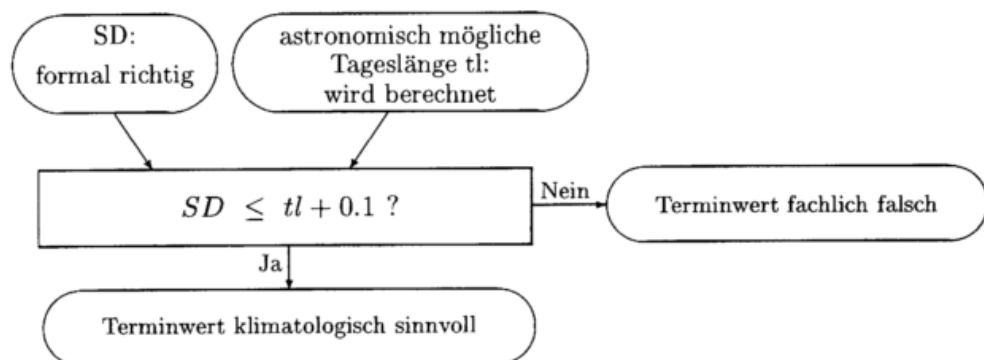
bei der Festlegung des oberen Grenzwertes aufaddiert.



Der untere Grenzwert wird hier direkt vom oberen Grenzwert abgeleitet, könnte jedoch auch über Gl. 6 mit einer Sonnenscheindauer von 0 und einer zu definierenden Toleranz bestimmt werden.

- **SD**

Die Sonnenscheindauer wird gegen die astronomisch mögliche ($SD_{astr.}$) geprüft. Eine Toleranz von 0,1 Stunde wird zugelassen, da eine Ablesung am Registrierstreifen leichte Ungenauigkeiten mit sich bringt.



4.2.4 Andere Elemente

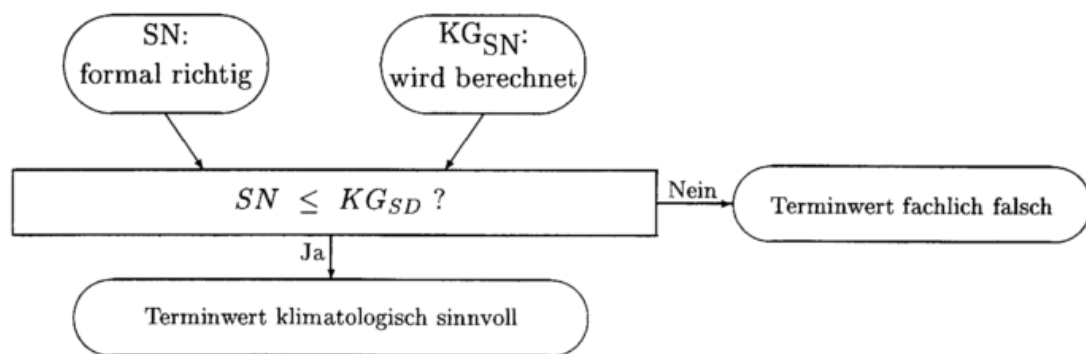
- **SN**

Der obere klimatologische Grenzwert der Schneehöhe KG_{SN} wird z.Zt. vollempirisch über die Stationshöhe sth , geogr. Breite gb , geogr. Länge gl und jeweiligen Monat berechnet. Da es zur Zeit noch keine zeitliche Prüfung der Schneehöhe gibt wurde die klimatologische Prüfung recht scharf an vorliegende Extremwerte der Schneehöhe **SN** angepaßt, damit die Prüfung nicht zu großzügig wird. In nachfolgender Tabelle 5 sind exemplarisch die beobachteten und berechneten KG_{SN} -Werte für einige Stationen gegenübergestellt.

Nur in wenigen Einzelfällen liegen die Werte von KG_{SN} unter den bisher an den Stationen beobachteten Schneehöhen, die allerdings nur in Zeitspannen beobachtet wurden, die die Lebensdauer dieser Prüfung sicherlich überschreitet.

Tabelle 5: Gegenüberstellung der berechneten klimatologischen Grenzwerte der Schneehöhe KG_{SN} zu den gemessenen Extremwerten SN

| Station | Monat | | | | | | | | | | | |
|-----------------|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 10035 SN | 62 | 70 | 51 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 50 |
| 10035 KG_{SN} | 65 | 81 | 53 | 19 | 0 | 0 | 0 | 0 | 0 | 5 | 28 | 47 |
| 10384 SN | 35 | 55 | 35 | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 24 | 23 |
| 10384 KG_{SN} | 73 | 73 | 47 | 16 | 0 | 0 | 0 | 0 | 0 | 16 | 47 | 73 |
| 10410 SN | 32 | 30 | 15 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 18 |
| 10410 KG_{SN} | 65 | 81 | 53 | 20 | 0 | 0 | 0 | 0 | 0 | 5 | 28 | 48 |
| 10453 SN | 235 | 255 | 315 | 380 | 290 | 40 | 0 | 0 | 2 | 80 | 115 | 200 |
| 10453 KG_{SN} | 324 | 375 | 421 | 351 | 252 | 141 | 46 | 0 | 27 | 121 | 199 | 266 |
| 10544 SN | 105 | 105 | 147 | 70 | 13 | 0 | 0 | 0 | 0 | 30 | 39 | 111 |
| 10544 KG_{SN} | 170 | 200 | 227 | 171 | 110 | 45 | 0 | 0 | 0 | 52 | 98 | 136 |
| 10635 SN | 77 | 90 | 103 | 56 | 22 | 0 | 0 | 0 | 0 | 13 | 44 | 70 |
| 10635 KG_{SN} | 139 | 165 | 171 | 132 | 77 | 21 | 0 | 0 | 0 | 37 | 76 | 109 |
| 10637 SN | 20 | 34 | 10 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 15 | 19 |
| 10637 KG_{SN} | 65 | 65 | 42 | 14 | 0 | 0 | 0 | 0 | 0 | 14 | 42 | 65 |
| 10836 SN | 72 | 106 | 115 | 53 | 9 | 0 | 0 | 0 | 0 | 14 | 42 | 68 |
| 10836 KG_{SN} | 117 | 140 | 102 | 53 | 3 | 0 | 0 | 0 | 0 | 26 | 61 | 91 |
| 10908 SN | 220 | 335 | 350 | 310 | 210 | 27 | 0 | 0 | 18 | 140 | 160 | 245 |
| 10908 KG_{SN} | 361 | 420 | 472 | 411 | 301 | 176 | 65 | 0 | 23 | 131 | 220 | 296 |
| 10961 SN | 560 | 650 | 600 | 780 | 680 | 670 | 480 | 260 | 200 | 395 | 300 | 480 |
| 10961 KG_{SN} | 646 | 717 | 780 | 836 | 778 | 627 | 442 | 291 | 235 | 367 | 475 | 567 |
| 10963 SN | 110 | 110 | 93 | 60 | 17 | 0 | 0 | 0 | 0 | 25 | 50 | 90 |
| 10963 KG_{SN} | 109 | 131 | 128 | 91 | 43 | 0 | 0 | 0 | 0 | 22 | 55 | 84 |
| 10980 SN | 408 | 340 | 300 | 320 | 280 | 120 | 12 | 30 | 40 | 140 | 150 | 245 |
| 10980 KG_{SN} | 430 | 494 | 552 | 511 | 397 | 259 | 132 | 48 | 55 | 175 | 273 | 357 |



4.3 Zeitliche Konsistenzprüfung

Für eine zeitliche Konsistenzprüfung müssen mindestens zwei Werte pro Prüfzeitraum vorliegen. Da sich unser Prüfzeitraum in der Regel auf die Periode 07 bis 06UTC erstreckt, können nur wenige Elemente dieser Prüfung unterzogen werden.

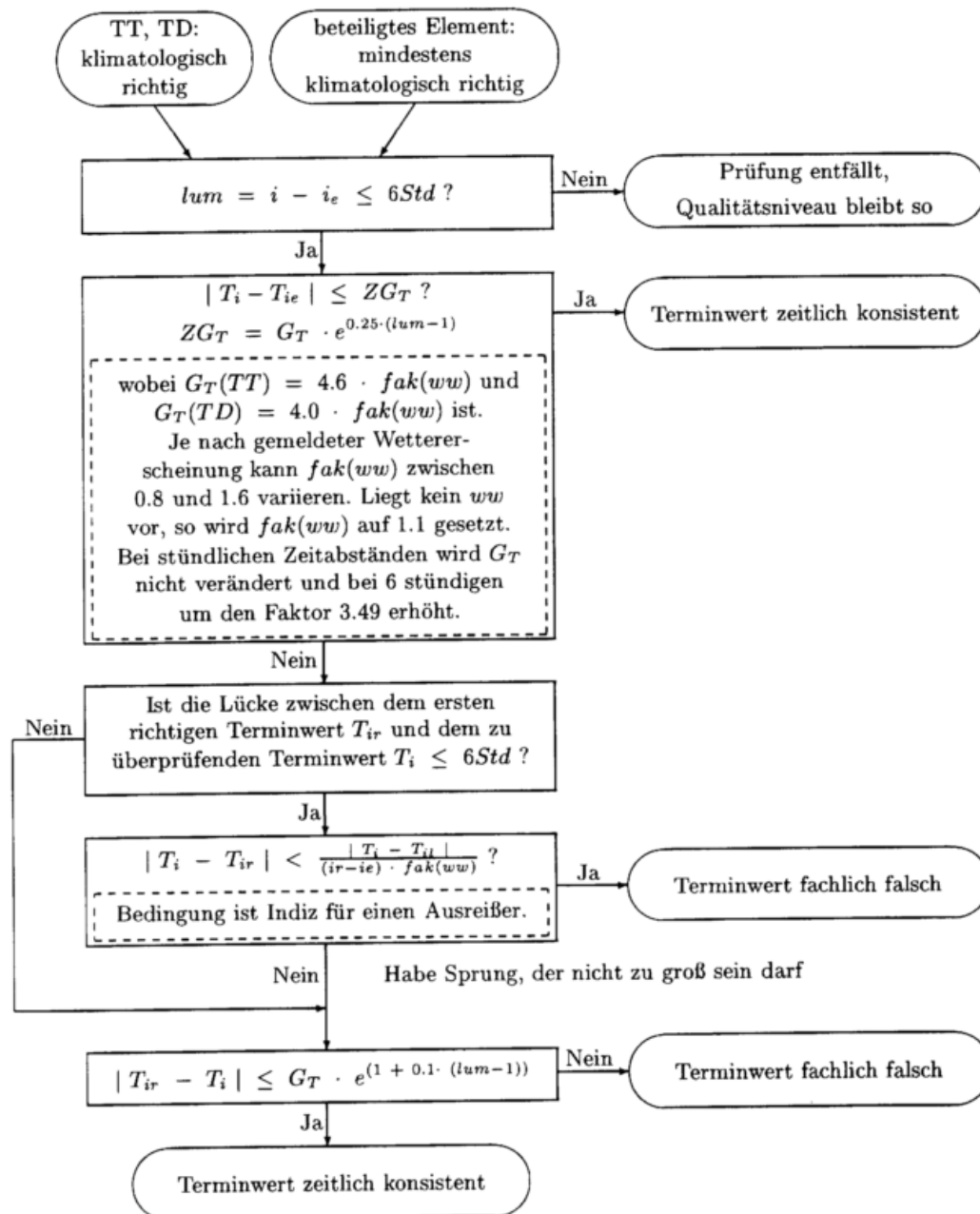
Bei bestandener Prüfung wird das geprüfte Element auf das Qualitätsniveau *zeitlich konsistent* angehoben.

4.3.1 Temperaturen

- **TT, TD**

Häufig treten Vorzeichenfehler oder 5 K Fehler bei den Meldungen auf, die nachfolgend als

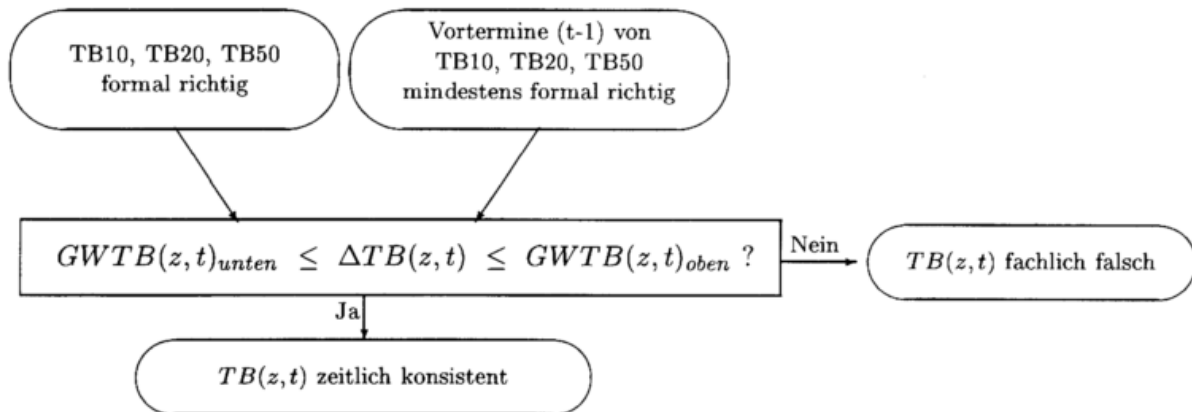
Ausreißer bezeichnet werden sollen. Diese Ausreißer dürfen jedoch nicht mit markanten Temperaturänderungen an Fronten verwechselt werden, die durchaus über 10 K in der Stunde betragen können und nachfolgend als Sprünge bezeichnet werden sollen. Um zwischen Ausreißer und Sprung unterscheiden zu können, werden mindestens drei Terminwerte benötigt, die nicht zu weit auseinander liegen dürfen. Die maximal zugelassene Lücke lum zwischen dem zu prüfenden Termin Index i und dem letzten davorliegenden Termin Index il sowie dem ersten nachfolgenden Termin ir mit richtigen Werten beträgt 6 Stunden.



• **TB10, TB20, TB50**

Aufgrund der Vielzahl an Böden mit unterschiedlichen Temperaturleitfähigkeiten ist das Spektrum der möglichen Temperaturschwankungen von Termin zu Termin sehr groß. Eigentlich müsste für jede Station besondere Prüfparameter herangezogen werden, um die Fehlertoleranz niedrig zu halten. Da die Bodentemperaturen jedoch nicht direkt in Modellberechnungen einfließen, wäre der nötige Aufwand nicht gerechtfertigt. Somit wurden für alle Stationen der gleiche Ansatz verwendet, wobei allerdings darauf verzichtet wurde die obersten Tiefen von 2

und 5 cm einer zeitlichen Konsistenzprüfung zu unterziehen, weil die Fehlertoleranz einfach zu groß wäre und hierdurch zu wenig Werte angemeckert werden würden. Selbst in den geprüften tieferen Schichten von 10, 20 und 50 cm sind die unterschiedlichen Charakteristika der Böden noch stark erkennbar. Da bei der Prüfung eines Tages auch die Werte ab 21 UTC Vortag zur Verfügung stehen, lassen sich alle zeitlichen Änderungen $\Delta TB(z, t) = TB(z, t) - TB(z, t - 1)$ zwischen den Terminen t und Vortermin $t - 1$ untersuchen. Für jeden Termin und jede Tiefe werden unterschiedliche obere $GWTB(z, t)_{oberen}$ und untere $GWTB(z, t)_{unten}$ Grenzwerte bestimmt, die nicht überschritten werden dürfen.



Bei der Bestimmung der Grenzwerte $TB(z, t)_{untereGr.}$ und $TB(z, t)_{obereGr.}$ geht neben einem festen Wert auch noch anteilig der Wert des Vortermine ein, um indirekt den Einfluß der Jahreszeit mitzuberechnen, denn bei geringerem Energieangebot können die zeitlichen Änderungen nicht so groß sein.

4.3.2 Wind

- **FF**

Die zeitlichen Konsistenzprüfung der Windgeschwindigkeit **FF** wird in der gleichen Art und Weise, wie bei Temperatur durchgeführt. Der korrespondierende Grenzwert für den Wind G_F wird folgendermaßen berechnet:

$$G_F = fak(w) \cdot (4.0 + 0.3 \cdot FF_i) \quad (3)$$

Neben dem bereits erwähnten Faktor fak , der von dem gemeldeten Wetter w abhängt, werden 30% des zu Überprüfenden Terminwertes von FF und eine Konstante von 4 m/s berücksichtigt.

4.3.3 Andere Elemente

- **PPPP**

Die zeitlichen Konsistenzprüfung des Bodendruckes über NN **PPPP** wird in der gleichen Art und Weise, wie bei Temperatur durchgeführt. Der korrespondierende Grenzwert für den Wind G_P wird folgendermaßen berechnet:

$$G_P = fak(w) \cdot 4.5 \quad (4)$$

Neben dem bereits erwähnten Faktor fak wird ein Grenzwert von 4.5 hPa verwendet.

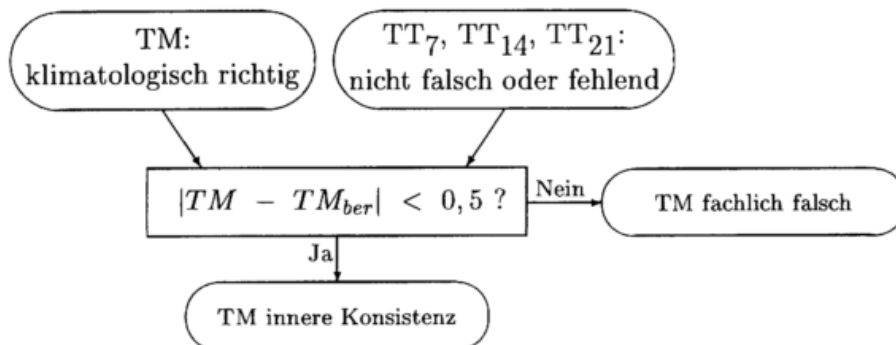
4.4 Innere Konsistenzprüfung

Die innere Konsistenzprüfung ist die "erfolgreichste" Prüfungsstufe, weil hier die Datenprüfung am häufigsten "zuschlägt" und man sich wundern muß, wie schlecht einige Elemente in der meteorologischen Datenbank vorliegen, die durch Verschlüsselungsfehler, Meldefehler oder auch Automaten-defekte hervorgerufen sein mögen.

4.4.1 Temperaturen

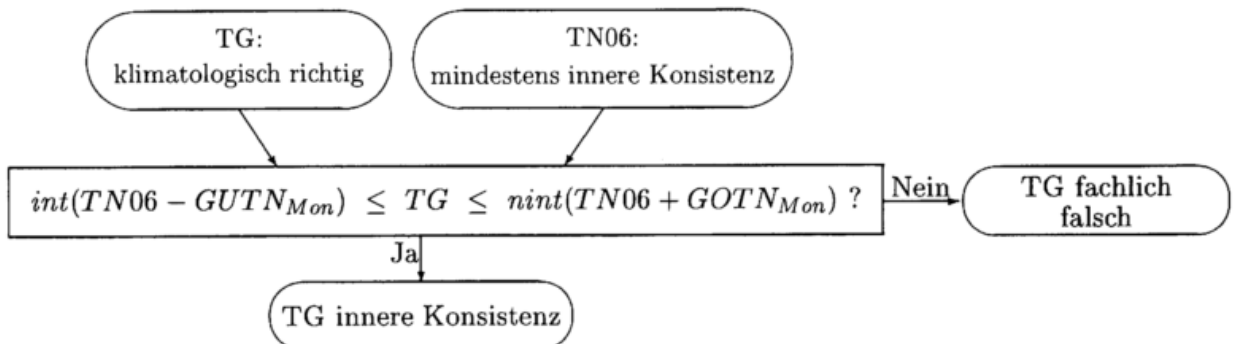
- **TM**

TM_{ber} wird nach Klimakonventionen ($\frac{I+II+2\cdot III}{4}$) berechnet und gegen **TM** geprüft.



- **TG**

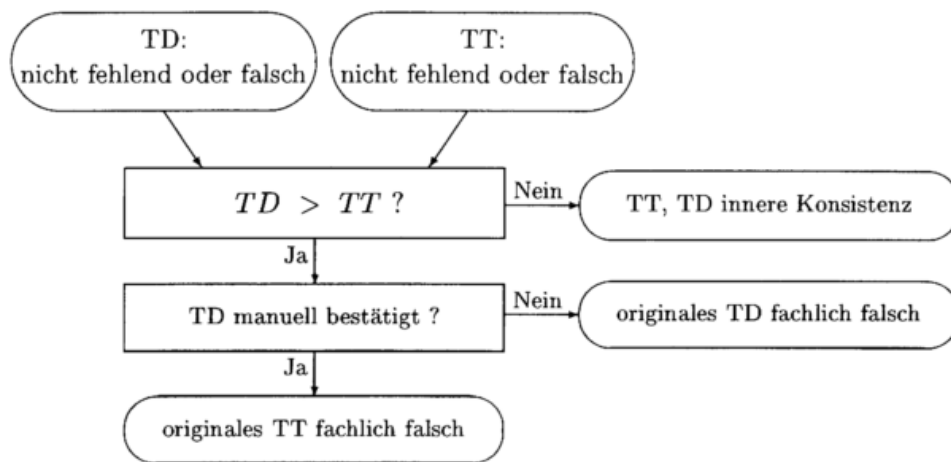
Da das Erdbodenminimum **TG** nur in ganzen °C angegeben ist, muß auch **TN06** für die Prüfung gerundet werden.



Der untere Grenzwert $GUTN_{Mon}$ liegt im Winter bei $9^{\circ}C$ und im Sommer bei $6^{\circ}C$ und der obere Grenzwert $GOTN_{Mon}$ im Winter bei $2^{\circ}C$ und im Sommer bei $0.5^{\circ}C$.

- **TD**

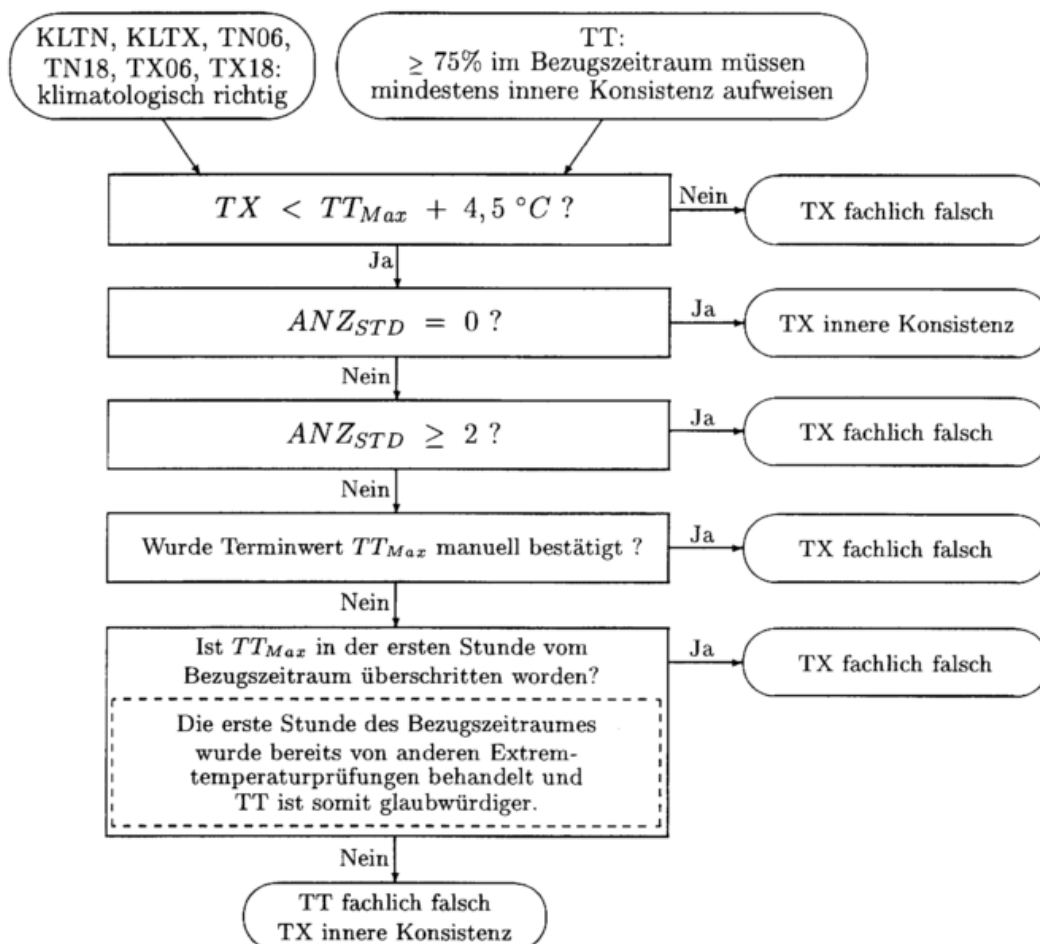
Unsaubererweise werden in diesem Modul falsch gegebene Werte des Taupunktes sofort ersetzt. Dabei wird der als falsch erkannte Wert durch den des beteiligten Elementes ersetzt und das QB auf sonstige Verfahren ungeprüft gesetzt.



• **TN06, TN18, TX06, TX18, KLTN, KLTX**

Die Temperaturextrema während der jeweiligen Bezugszeiträume werden nach dem gleichen Algorithmus geprüft, welches anhand der Maximalwerte verdeutlicht werden soll:

Es werden die Stunden im Bezugszeitraum ANZ_{Std} gezählt, in denen $TT > TX + 0.2$ K ist, zusätzlich wird der höchste Stundenwert TT_{Max} bestimmt.



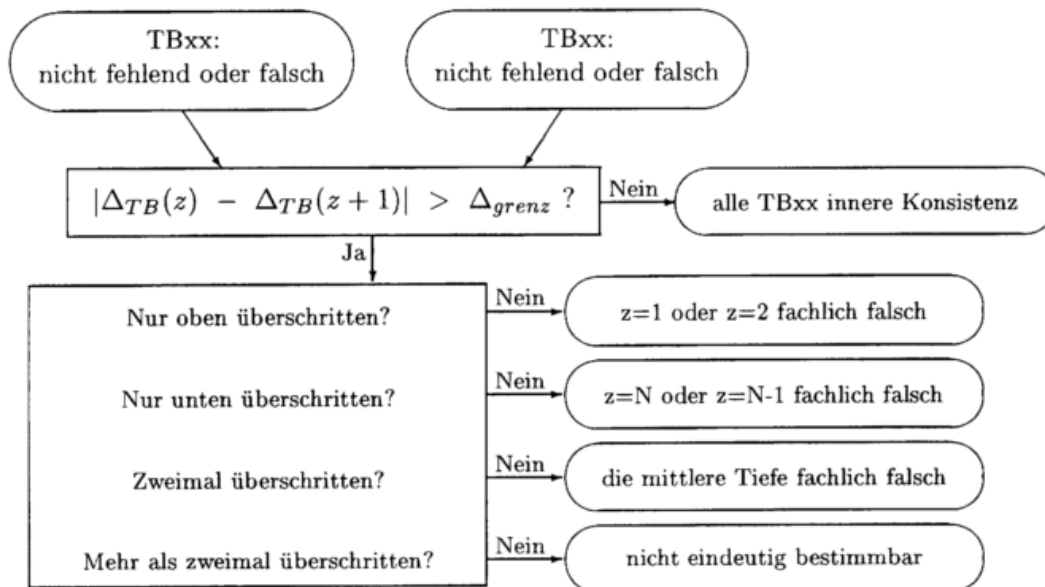
Die Extremwertprüfung ist überraschend ergiebig, was auf unterschiedliche Thermometer bzw. auf die Meldevorschriften der Bundeswehr¹ zurückzuführen ist.

¹Bei der Bundeswehr ist es Vorschrift, den abgelesenen Extremwert zu melden, auch wenn stündliche Werte diesem widersprechen.

- **TB02, TB05, TB10, TB20, TB50, TB100**

Die innere Konsistenzprüfung der Erdbodentemperaturen ist sehr schwierig, weil durch die Vielfältigkeit der Böden sehr unterschiedliche Temperaturleitfähigkeiten vorliegen. Beim Testen des Verfahrens wurden sehr viele Fehler gefunden, die bei anderen Stationen allerdings zu unrecht angemekert werden. Somit heißt die Devise bei diesem Modul, nicht unbedingt jeden Fehler entdecken, aber die gefundenen Fehler sollten auch solche sein. Die Qualität der Bodenmessungen ist erschreckend schlecht und einige Phänomene lassen sich nicht so ohne weiteres erklären. So überschreiten bei einigen Stationen z.B. die Bodentemperaturen im Erdboden bei Strahlungstagen niemals die Lufttemperatur, was wohl auf eine bewachsene Meßfläche schließen läßt. Bei bemannten Stationen werden gelegentlich 5 bzw. 10 °C Fehler verschlüsselt, da die Temperaturspanne in den verschiedenen Schichten bis zu 40 °C betragen kann. Um nicht auch die Stationen angemekert zu bekommen, die weit unter dem Lufttemperaturniveau liegen, wurden einfach die Änderung des Temperaturgradienten $\Delta_{TB}(z)$ zwischen den Tiefen z und $z+1$ ($1 \leq z \leq N$) im Erdboden untersucht, die je nach Strahlungsgegebenheiten nur einen "ausprobierten" Grenzwert Δ_{grenz} erreichen kann.

$$\Delta_{TB}(z) = \frac{TB(z) - TB(z+1)}{\ln(z) - \ln(z+1)} \quad (5)$$



Dieses Verfahren zeigt ziemlich zuverlässig, daß ein Fehler vorliegt, jedoch ist die genaue Lokalisation der fehlerhaften Tiefe sehr schwierig und es werden weitere Abfragen (im Diagramm nicht weiter aufgeschlüsselt) zur Bestimmung der falschen Tiefe durchgeführt. In lediglich 10% der Fälle wird die falsche Tiefe lokalisiert, so daß eine nachträgliche manuelle Kontrolle bei diesem Verfahren unumgänglich ist. Da dieses Verfahren lediglich eine Tiefe falsch geben kann, werden bei mehreren falsch gemeldet oder verschlüsselten Tiefen gelegentlich gerade eine richtige Tiefe als falsch deklariert.

Allein mit diesem Verfahren werden zusätzlich zu den bereits als falsch erkannten Bodenmeldungen pro Jahr noch ca. 500 falsche Meldungen (0,25%) zusätzlich bei allen deutschen Synop-Stationen gefunden.

4.4.2 Niederschläge

- **RR12, RR18, RR00, RR06, RR24**

Die innere Konsistenzprüfung des Niederschlages wurde zuletzt im November 1998 komplett auf ein neues Verfahren umgestellt. Durch die unterschiedlichen Bezugszeiträume und

Verschlüsselungsvorschriften ist gerade der Niederschlag sehr fehleranfällig. Die eigentliche Prüfung ist im Prinzip einfach, jedoch kann es mit einem unzureichenden Verfahren recht schwierig werden, den bzw. die falschen Werte zu lokalisieren. Natürlich könnten alle an der Prüfung beteiligten Niederschläge falsch gegeben werden, doch würden dann zuviele Informationen verloren gehen.

Bei der Prüfung werden alle synoptischen Niederschlagselemente **RR12**, **RR18**, **RR00**, **RR06** und **RR24** gegeneinander unter Berücksichtigung von **ww** und **W1** sowie **R1R1** geprüft. Das Verfahren beruht im Prinzip darauf, daß alle 6-, 12- und 24-stündigen Niederschläge in mehreren Schritten miteinander verglichen werden und bei Unstimmigkeiten "Fehlerzähler" jedes betroffenen Elementes erhöht werden. Am Ende wird nur der Wert falsch gegeben, dessen Zähler am höchsten ist. In Tabelle 6 werden die Prüfschritte aufgelistet, wobei natürlich

Tabelle 6: Erste RR-Prüfung

| Prüfung | Im Fehlerfall Zählererhöhung | | | | |
|-------------------------|------------------------------|-------|-------|-------|-------|
| | ZRR12 | ZRR18 | ZRR00 | ZRR06 | ZRR24 |
| $RR18 + RR06 = RR24$ | 0 | 2 | 0 | 2 | 2 |
| $RR18 \leq RR24$ | 0 | 2 | 0 | 0 | 2 |
| $RR06 \leq RR24$ | 0 | 0 | 0 | 2 | 2 |
| $RR18 + RR00 \leq RR24$ | 0 | 2 | 2 | 0 | 2 |
| $RR12 + RR06 \leq RR24$ | 2 | 0 | 0 | 2 | 2 |
| $RR12 \leq RR18$ | 2 | 3 | 0 | 0 | 0 |
| $RR00 \leq RR06$ | 0 | 0 | 2 | 3 | 0 |

Tabelle 7: Zusätzliche RR-Prüfung

| Prüfung | Im Fehlerfall Zählererhöhung | | | | |
|---------------------------------------|------------------------------|-------|-------|-------|-------|
| | ZRR12 | ZRR18 | ZRR00 | ZRR06 | ZRR24 |
| $\sum R1R1 \leq RR12$ | 1 | 0 | 0 | 0 | 0 |
| $\sum R1R1 \leq RR18$ | 0 | 1 | 0 | 0 | 0 |
| $\sum R1R1 \leq RR00$ | 0 | 0 | 1 | 0 | 0 |
| $\sum R1R1 \leq RR06$ | 0 | 0 | 0 | 1 | 0 |
| $\sum R1R1 \leq RR24$ | 0 | 0 | 0 | 0 | 1 |
| $RR12 = 0$ ww mit RR | 4 | 0 | 0 | 0 | 0 |
| $RR12 > 0$ alle ww ohne RR | 2 | 0 | 0 | 0 | 0 |
| $RR18 = 0$ ww mit RR | 0 | 4 | 0 | 0 | 0 |
| $RR18 > 0$ alle ww ohne RR | 0 | 2 | 0 | 0 | 0 |
| $RR00 = 0$ ww mit RR | 0 | 0 | 4 | 0 | 0 |
| $RR00 > 0$ alle ww ohne RR | 0 | 0 | 2 | 0 | 0 |
| $RR06 = 0$ ww mit RR | 0 | 0 | 0 | 4 | 0 |
| $RR06 > 0$ alle ww ohne RR | 0 | 0 | 0 | 2 | 0 |
| $RR24 = 0$ ww mit RR | 0 | 0 | 0 | 0 | 4 |
| $RR24 > 0$ alle ww ohne RR | 0 | 0 | 0 | 0 | 2 |
| $RR12 = 0$ W1 mit RR | 2 | 0 | 0 | 0 | 0 |
| $RR12 > 0$ zumind. Hauptt. W1 ohne RR | 2 | 0 | 0 | 0 | 0 |
| $RR18 = 0$ W1 mit RR | 0 | 2 | 0 | 0 | 0 |
| $RR18 > 0$ zumind. Hauptt. W1 ohne RR | 0 | 2 | 0 | 0 | 0 |
| $RR00 = 0$ W1 mit RR | 0 | 0 | 2 | 0 | 0 |
| $RR00 > 0$ zumind. Hauptt. W1 ohne RR | 0 | 0 | 2 | 0 | 0 |
| $RR06 = 0$ W1 mit RR | 0 | 0 | 0 | 2 | 0 |
| $RR06 > 0$ zumind. Hauptt. W1 ohne RR | 0 | 0 | 0 | 2 | 0 |
| $RR24 = 0$ W1 mit RR | 0 | 0 | 0 | 0 | 2 |
| $RR24 > 0$ zumind. Hauptt. W1 ohne RR | 0 | 0 | 0 | 0 | 2 |

die Rundungsvorschriften bei den 6- und 12-stündigen Niederschlägen sowie Niederschlag

gefallen, aber nicht meßbar (Trace) mitberücksichtigt werden. Liegt eine Inkonsistenz vor, werden zusätzlich die in Tabelle 7 angegebenen Prüfungen durchgeführt. Die Prüfung gegen **W1** findet nur statt, wenn noch keine eindeutige Aussage über das **ww** getroffen werden konnte, weil z.B. nicht alle Stunden komplett als Original vorlagen. Wurde über Tabelle 6 kein Fehler gefunden, so erhöht sich für alle beteiligten Niederschläge das Qualitätsniveau auf *innere Konsistenz*.

Nach der Prüfung genießt der **RR24** eine etwas tolerantere Behandlung, weil dieses Element in der Regel zuverlässiger gemeldet wird. Gelegentlich kommt es vor, daß einige Zähler den gleichen Wert haben. Ist dies der Fall und der **RR24** ist unter diesen Elementen, so wird **ZRR24** um 1 reduziert. Wurde ein Niederschlag *manuell bestätigt*, so wird dessen Zähler auf 0 gesetzt. Nun wird der oder die Niederschläge mit dem höchsten Fehlerzähler auf *fachlich falsch* gesetzt.

- **R1R1**

Alle Einzelstunden des Niederschlages **R1R1** werden über den jeweiligen Bezugszeitraum aufsummiert und mit den entsprechenden **RRxx** Werten unter Berücksichtigung der Rundungsgenauigkeit und Trace verglichen.

Eine besondere Behandlung von Vorhersagen ist darauf zurückzuführen, daß der "heutige" Tag aus *gemeldeten* und *vorhergesagten* **R1R1** Werten zusammengesetzt sein kann, die mehrstündigen Niederschläge jedoch ausschließlich aus der Vorhersage stammen. Weichen die vorhergesagten mehrstündigen Niederschläge von der Summe der stündlichen ab, so wird entgegen der sonstigen Konvention die vorhergesagte Niederschlagssumme *fachlich falsch* gegeben.

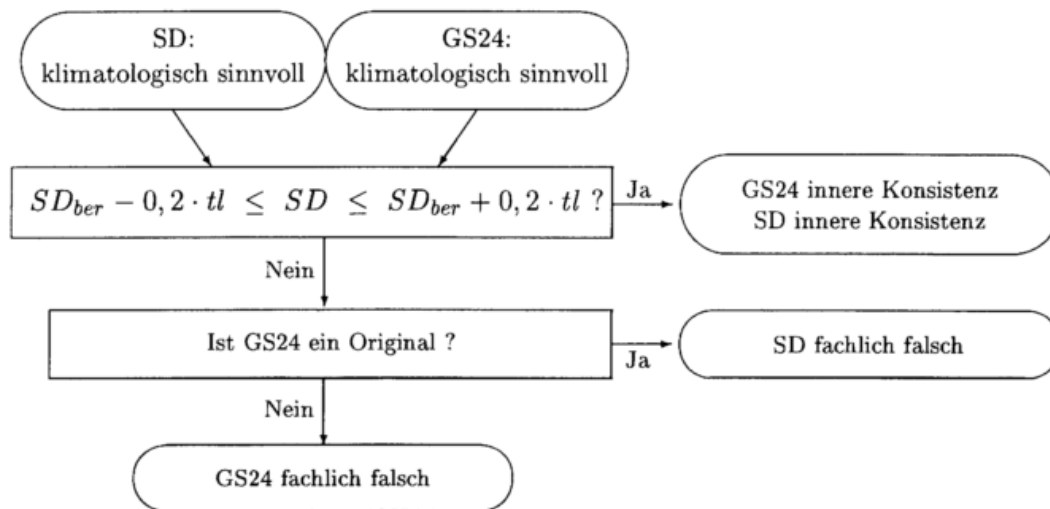
4.4.3 Strahlung

- **SD, GS24**

Mit Hilfe von **GS24**, des Datums sowie der geogr. Breite und Länge eine Sonnenscheindauer SD_{ber} berechnet. Dieses geschieht in etwas abgewandelter Form nach dem gleichen Verfahren wie bei der klimatologischen Prüfung von **GS24**, denn es wird die Strahlung kS_{max} berechnet, die im Mittel für diesen Ort und Tag die Erdoberfläche erreichen kann. Über nachfolgende Beziehung nach Wendling, welche im DVWK erschienen ist, läßt sich eine Sonnenscheindauer SD_{ber} ableiten, die mit der tatsächlichen Sonnenscheindauer **SD** verglichen wird.

$$SD_{ber} = tl \cdot \left(1.43 \cdot \frac{GS24}{kS_{max}} - 0.43 \right) \quad (6)$$

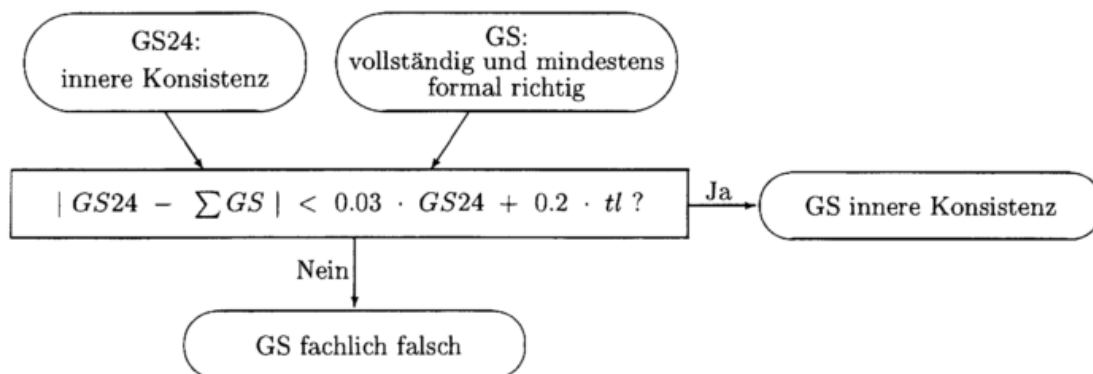
Da es durchaus Fälle gibt, wo **GS24** größer als kS_{max} ist, wird SD_{ber} auf die astronomisch mögliche Tageslänge tl beschränkt. Da diese Verfahren im Mittel sehr gute Ergebnisse liefern, jedoch in Einzelfällen größere Abweichungen aufweisen können, muß die Fehlertoleranz recht groß sein.



Die letzte Entscheidung geht davon aus, daß die Globalstrahlungsmessung nicht so fehleranfällig ist wie die Messung der Sonnenscheindauer. Diese Erfahrung haben wir bisher bestätigt gefunden, weil lediglich ein Grenzwert darüber entscheidet, ob die Sonne scheint oder nicht. Bei verdreckter Haube oder flachen Sonnenständen ist dieses Meßsystem eher fehleranfällig, weil die Globalstrahlungsmeßgeber bei gleichen Bedingungen zumindest noch einen –wenn auch zu kleinen– Wert liefern.

- **GS**

In dieser Prüfung wird die Tagessumme **GS24** gegen die Summe der stündlichen Werte **GS** geprüft. Hierbei wird eine geringe Differenz zugebilligt, weil **GS** ein Element ist, welches in Kapitel 5 vervollständigt und anschließend wieder geprüft wird. Außerdem können nächtliche Falschmeldungen von $GS > 0 \text{ J/cm}^2$ im gemeldeten **GS24** enthalten sein. Neben einem Prozentualen Anteil der Tagessumme wird noch 20% der astronomisch mögliche Sonnenscheindauer $tl(Std)$ berücksichtigt.



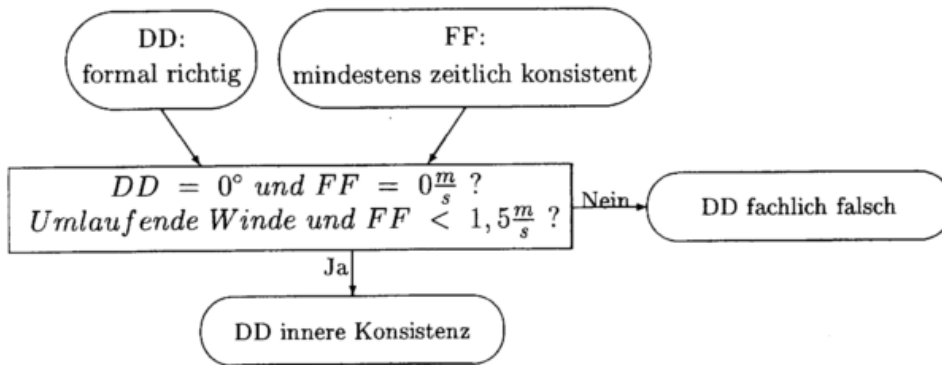
- **SD1**

Der Tageswert der Sonnenscheindauer wird in der gleichen Art und Weise wie die Globalstrahlung gegen die Summe der stündlichen Werte geprüft. Hierbei werden 5% der Tagessumme und 2% der astronomisch möglichen Sonnenscheindauer $tl(Std)$ als Fehlertoleranz zugelassen.

4.4.4 Wind

- **DD**

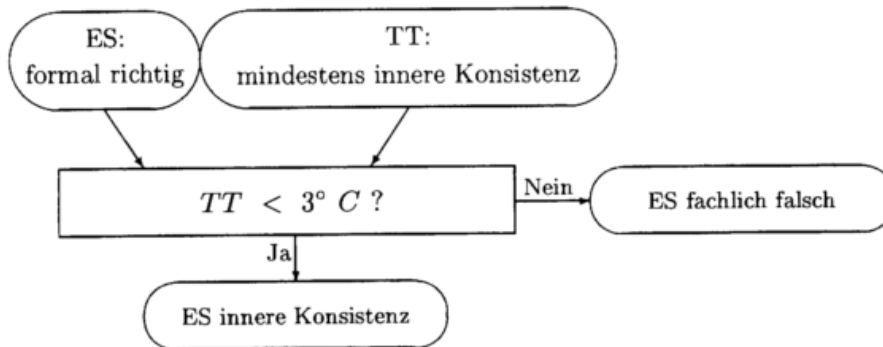
Umlaufende Winde dürfen nur bei $FF < 1.5 \text{ m/s}$ und $DD = 0^\circ$ darf nur bei $FF = 0 \text{ m/s}$ gegeben werden. Sind beide Bedingungen erfüllt, kommt der Wert in das Qualitätsniveau *genügt innerer Konsistenz*.



4.4.5 Andere Elemente

- **ES**

Es wird lediglich der lockere und trockene Schnee gegen die Temperatur **TT** zum Meldetermin geprüft. Eine Prüfung des festen oder nassen Schnees über **TT** macht keinen Sinn.

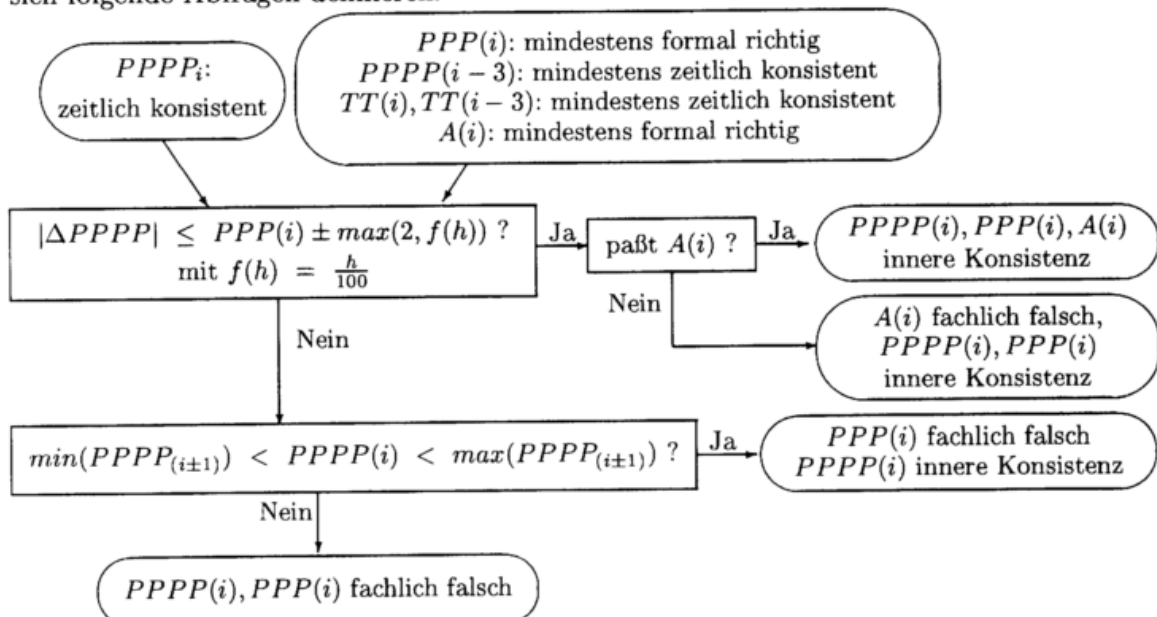


- **PPPP, PPP, A**

Da der gemessene Druck in Stationshöhe $PPPP_h$ nicht in der Datenbank enthalten ist und die 3-stündige Druckänderung **PPP** und -tendenz **A** sich auf dieses Druckniveau beziehen, muß diese zunächst mit Hilfe der Lufttemperatur **TT** über

$$PPPP_h = PPPP \cdot e^{-\frac{q}{R \cdot TT} \cdot h} \quad (7)$$

reduziert werden. Mit Hilfe der Druckdifferenz $\Delta PPPP = PPPP_h(i) - PPPP_h(i-3)$ lassen sich folgende Abfragen definieren.



Die Überprüfung der Drucktendenz $A(i)$ wurde im Flußdiagramm nicht näher erklärt, weil sie recht einfach ist. Es muß nur beachtet werden, daß bei $PPP(i) = 0$ nicht nur $A = 4$, sondern auch 0, 3, 5 und 8 gemeldet werden kann.

Die kompliziertere Behandlung der Bergstationen ist noch nicht in der Prüfung implementiert, weil diese Stationen lediglich ihr Geopotential bezogen auf das jeweilige Druckniveau melden, welches nicht in die Bank geschrieben wird. Beim Einbringen der ungeprüften Originalwerte müßte zunächst eine Umrechnung von Geopotential und Druckniveau auf Bodendruck erfolgen.

- **W1**

Dieses Modul wurde erst Ende 2000 neu in die Datenprüfung aufgenommen und bedurfte einiger Korrekturen, weil die Beobachteranleitung bzgl. dieses Elementes recht unklar formuliert ist. In dieser Prüfung werden nur original gemeldete **ww** und **W1** zur Prüfung herangezogen, die mindestens formal richtig sind. Aufgrund der unterschiedlichen Bezugszeiträume werden die Nebentermine anders geprüft als die Zwischen- und Haupttermine. Wird zum Nebentermin ein **W1** ≥ 3 gemeldet, so muß das aus dem zugehörigen **ww** abgeleitete **W1** > 0 sein.

Zu den Zwischen- und Hauptterminen müssen die gemeldeten **W1** zumindest gleich oder größer als die während des Bezugsraumes in den Vorstunden gemeldeten **W1** > 2 sein. Da zu den Zwischenterminen beim **ww** auch die Nachwetter (**ww** 20-29) in Kombination mit einem Bewölkungs-**W1** gemeldet werden können, müssen die in den Vorstunden aus **ww** abgeleiteten **W1** zumindest gleich oder kleiner als das zum Termin gemeldete **W1** sein.

- **ww**

Voraussetzung für eine Prüfung nach Tabelle 8 ist:

1. Original **ww** *formal richtig*
Beteiligtes Element auch Original
2. **ww** nicht original, aber *formal richtig*
Bei beteiligten Elementen Herkunft egal

Besteht das **ww** eine Prüfung nicht, wird es auf *fachlich falsch* gesetzt, ansonsten auf das Qualitätsniveau *genügt innerer Konsistenz* angehoben. Bei fehlerhaftem **ww** 10 (feuchter Dunst) wird dieses durch **ww** 05 (trockener Dunst) ersetzt.

4.5 Manuelle Bestätigung

Aus der Qualitätsbytetabelle (2) geht hervor, daß theoretisch eine räumliche Prüfung eingebaut werden kann. Letztlich wurde diese bis zu diesem Zeitpunkt noch nicht realisiert. Maßgebend für diese Entscheidung war, daß bisher immer nur über einen Tag geprüft wird. An einem Tag können auf engstem Gebiet sehr unterschiedliche Wettererscheinungen mit sehr unterschiedlichen Auswirkungen auf die meteorologischen Elemente auftreten, wie z.B. an Fronten, Schauern, Gewittern oder spezielle Vertikalschichtungen, die eine große Toleranzspanne bei den Elementen erforderlich machen würde und im wesentlichen schon durch andere Prüfschritte wie der zeitlichen Konsistenzprüfung abgefangen werden. Eine räumliche Prüfung über einen langen Zeitraum, in der Summen, Mittel, Minima oder Maxima einzelner Elemente miteinander verglichen werden, ist sinnvoll. So könnten hierdurch z.B. defekte Niederschlagsgeber erkannt werden, die maximal Trace melden und diesen intern richtig aufsummieren und somit *innerer Konsistenz* genügen, jedoch bei der Betrachtung des Monatsniederschlags sofort ins Auge springen.

Damit im Detail nachvollzogen werden kann, welche Fehler durch die Datenprüfung erkannt und falsch gegeben wurden, wird standardmäßig eine "Meckerliste" dieser Fehler erzeugt. Aus dieser Liste geht hervor, bei welchem Prüfverfahren, gegebenenfalls gegen welche Grenzwerte oder gegen

Tabelle 8: Prüfbedingungen für ww's

| ww | Prüfung | ww | Prüfung | ww | Prüfung |
|---------|---|---------|---|---------|---|
| 01 - 03 | $N \neq 9$ | 10 | $Rel.Feuchte \geq 77\%$ $VV < 8000m$ | 14 - 16 | $1 \leq N \leq 8$ |
| 17 | $N \neq 9$ | 18 | $FF > 0.5$ | 21 | $-3.0 \leq TT \leq 37.0$ |
| 22 - 23 | $-33.0 \leq TT \leq 7.0$ | 24 | $-7.0 \leq TT \leq 7.0$ | 25 | $-3.0 \leq TT - 10.0$ |
| 26 | $-33.0 \leq TT \leq 7.0$ | 27 | $-33.0 \leq TT \leq -40.0$ | 30 - 35 | $FF > 0.5$ |
| 36 - 39 | $FF > 0.5$ $-40.0 \leq TT \leq 0.0$ | 41 | $Rel.Feuchte > 79\%$ | 42 | $Rel.Feuchte > 84\%$ $-3.0 \leq TT \leq 27.0$ |
| 43 - 47 | $Rel.Feuchte > 85\%$ $-3.0 \leq TT \leq 27.0$ | 48 - 49 | $Rel.Feuchte > 79\%$ $-25.0 \leq TT \leq 0.0$ | 50 - 51 | $-1.0 \leq TT \leq 32.0$ |
| 52 - 53 | $N > 2$ $-1.0 \leq TT \leq 32.0$ | 54 - 55 | $N > 4$ $-1.0 \leq TT \leq 32.0$ | 56 | $-7.0 \leq TT \leq 3.0$ |
| 57 | $N > 2$ $-7.0 \leq TT \leq 3.0$ | 58 | $-4.0 \leq TT \leq 27.0$ | 59 | $N > 2$ $-4.0 \leq TT \leq 27.0$ |
| 60 - 61 | $-4.0 \leq TT \leq 36.0$ | 62 - 63 | $N > 2$ $-4.0 \leq TT \leq 36.0$ | 64 - 65 | $N > 4$ $-4.0 \leq TT \leq 36.0$ |
| 66 | $-7.0 \leq TT \leq 6.0$ | 67 | $N > 2$ $-7.0 \leq TT \leq 6.0$ | 68 | $-1.0 \leq TT \leq 7.0$ |
| 69 | $N > 2$ $-1.0 \leq TT \leq 7.0$ | 70 - 71 | $-25.0 \leq TT \leq 6.0$ | 72 - 73 | $N > 2$ $-25.0 \leq TT \leq 6.0$ |
| 74 - 75 | $N > 4$ $-25.0 \leq TT \leq 6.0$ | 76 | $-40.0 \leq TT \leq 0.0$ | 77 | $-30.0 \leq TT \leq 4.0$ |
| 78 | $-30.0 \leq TT \leq 0.0$ | 79 | $-30.0 \leq TT \leq -3.0$ | 80 | $-1.0 \leq TT \leq 37.0$ |
| 81 | $N > 2$ $-1.0 \leq TT \leq 37.0$ $CL > 1$ und $CL \neq 6$ | 82 | $N > 4$ $-1.0 \leq TT \leq 37.0$ $CL > 1$ und $CL \neq 6$ | 83 | $0.0 \leq TT \leq 6.0$ $CL > 1$ und $CL \neq 6$ |
| 84 | $N > 2$ $0.0 \leq TT \leq 6.0$ $CL > 1$ und $CL \neq 6$ | 85 | $-20.0 \leq TT \leq 4.0$ $CL > 1$ und $CL \neq 6$ | 86 | $N > 2$ $-20.0 \leq TT \leq 4.0$ $CL > 1$ und $CL \neq 6$ |
| 87 | $-20.0 \leq TT \leq 7.0$ $CL > 1$ und $CL \neq 6$ | 88 | $N > 2$ $-20.0 \leq TT \leq 7.0$ $CL > 1$ und $CL \neq 6$ | 89 | $CL = 3, 5, 7, 8, 9$ |
| 90 | $N > 2$ $CL = 3, 5, 7, 8, 9$ | 91 | $CL = 3, 5, 7, 8, 9$ | 92 | $N > 2$ |
| 94 | $N > 2$ | 95 | $CL = 3, 5, 7, 8, 9$ | 96 | $N > 2$ $CL = 3, 5, 7, 8, 9$ |
| 97 | $N > 4$ $CL = 3, 5, 7, 8, 9$ | 98 | $N > 2$ | 99 | $N > 4$ $CL = 3, 5, 7, 8, 9$ |

welche beteiligten Elemente der Wert falsch gegeben wurde. Über ein Programm und eine Prozedur lassen sich diese als Fehler erkannten Werte anschaulich sichtbar machen, und es kann hier noch einmal entschieden werden, ob als *falsch* erkannte Originalwerte *manuell bestätigt* werden (linke Seite des Datensatzes) oder *manuell korrigiert* (rechte Seite des Datensatzes) werden. Wurde ein Wert *manuell bestätigt*, so hat dieses die höchste Priorität und bewirkt, daß dieser Wert bei einer nachfolgenden Prüfung auf jeden Fall akzeptiert wird. Wurde ein Wert *manuell korrigiert*, so wird er in einer nachfolgenden Prüfung nochmals geprüft und kann hier wieder als *falsch* erkannt werden. Bei der manuellen Korrektur ist es von Vorteil, die beteiligten Elemente genau zu kennen, damit der Wert nicht noch einmal falsch gegeben wird.

Die manuelle Überarbeitung der Meckerlisten ist sehr zeitaufwendig, wird jedoch weitestgehend nach jedem Prüfungslauf durchgeführt. Die Kontrolle der Meckerlisten und die hierdurch ausgelöste Kontrolle und gegebenenfalls erforderlichen manuellen Eingriffe kommen zum einen der Qualität der Daten, zum anderen auch der Güte der Datenprüfung zugute. Bei sehr widersprüchlichen Meldungen fällt die "manuelle" Entscheidung häufig sehr schwer.

4.6 Fehlerstatistik

In diesem Abschnitt soll einmal anhand der deutschen Synop-Stationen (insgesamt 212 Stationen) beispielhaft aufgelistet werden, wie häufig Originalwerte durch die Datenprüfung in welchem Prüfschritt falsch gegeben werden. Es ist klar, daß sich die Fehlerart mit jeder Wettersituation stark ändern kann. Daher werden zwei Fallbeispiele über einen Zeitraum von jeweils fünf Tagen ausgewählt, wobei sich die Auswertung auf ausgesuchte Elemente beschränken soll. In der Aufstellung sind natürlich nur die entdeckten Fehler von den Originaldaten und nicht die der ersetzten aufgelistet. In beiden Fallbeispielen können theoretisch 1060 Tageswerte und 25440 Stundenwerte (jeweils eines Elementes) gemeldet werden.

Als erster Fall (Tabelle 9) wurde ein Zeitraum gewählt, in dem sehr hohe Niederschläge beobachtet wurden.

Tabelle 9: Fehlerstatistik für den Zeitraum 26.10. bis 31.10.1998

| Element | Originale | Im Prüfschritt | | | | |
|---------|-----------|----------------|---------|--------|--------|---------|
| | | formal | klimat. | zeitl. | innere | manuell |
| RR12 | 822 | | | | | |
| RR18 | 1023 | | 1 | | 12 | |
| RR00 | 819 | | | | | 1 |
| RR06 | 1026 | 1 | | | 10 | 1 |
| RR24 | 1020 | | | | 1 | 1 |
| R1R1 | 17738 | | | | 648 | |
| TG | 1009 | | | | 2 | |
| TX18 | 1026 | | | | 1 | |
| TX06 | 1018 | | | | 1 | |
| TN18 | 1013 | | | | 5 | |
| TN06 | 1029 | | | | 7 | |
| TM | 1029 | | | | 5 | |
| TT | 24571 | | | 2 | 17 | |
| TD | 24569 | | 1 | 2 | 2 | |
| SD | 864 | | | | 1 | |
| GS24 | 206 | | | | 5 | |
| GS | 2860 | | 488 | | | |
| FF | 24562 | | | 9 | | |
| ww | 17397 | | | | 32 | |

Im zweiten Fallbeispiel (Tabelle 10) wurde der Zeitraum vom 9.8. bis 14.8.1998 untersucht, der durch sehr hohe Temperaturen geprägt war.

Die Elemente, die erst nach 1998 in die Bank aufgenommen wurden **PPPP**, **PPP** sowie **A** oder deren Prüfmodule **TBxx** und **W1** später eingeführt wurden, sind nicht in den Tabellen enthalten. Im großen und ganzen sind die Ergebnisse recht erfreulich, denn das Verhältnis Anzahl der gemeldeten Originale zur Anzahl der entdeckten Fehler ist für fast alle Elemente sehr gering. Die relativ hoch erscheinende Fehlerquote bei **GS** ist ausschließlich auf die "nächtliche Globalstrahlung" zurückzuführen. Die Fehlerrate bei **R1R1** erscheint recht hoch, relativiert sich jedoch vor dem Hintergrund, daß bei Beanstandungen mindestens 6 Einzelstunden falsch gesetzt werden. Es überrascht jedoch, da **R1R1** ausschließlich von Automaten gemeldet wird, die die Addition beherrschen sollten und somit immer innerer Konsistenz genügen müßten. Die Erklärung hierfür ist, daß diese Stationen auch teilweise bemannt sind und bei defekten Niederschlagsgebern Korrekturmeldungen abschicken, von denen der Automat nichts weiß und im Hintergrund seine Niederschläge fleißig weiter aufsummiert. Werden nun nicht alle Korrekturmeldungen abgeschickt, drängt sich der Automat mit seinen Meldungen immer wieder in den Vordergrund, so daß eine automatisierte innere Konsistenzprüfung nur bedingt richtige Ergebnisse liefern kann. Sind die automatischen Stationen nicht bemannt, so scheint alles in Ordnung zu sein, weil kein Fehler gefunden werden kann, jedoch weiß man auch nicht, ob zuverlässig gemessen wird. Es ist in der Vergangenheit schon zwei-

Tabelle 10: Fehlerstatistik für den Zeitraum 09.08. bis 14.08.1998

| Element | Originale | Im Prüfschritt | | | | |
|---------|-----------|----------------|---------|--------|--------|---------|
| | | formal | klimat. | zeitl. | innere | manuell |
| RR12 | 831 | | | | 2 | |
| RR18 | 1031 | | | | 9 | |
| RR00 | 822 | | | | 3 | |
| RR06 | 1040 | | | | 5 | |
| RR24 | 1019 | | | | 5 | |
| R1R1 | 16692 | | | | 210 | |
| TG | 1021 | | | | 7 | |
| TX18 | 1038 | | | | 16 | |
| TX06 | 1026 | | | | 28 | |
| TN18 | 1020 | | | | 1 | |
| TN06 | 1041 | | | | 14 | |
| TM | 982 | | | | 1 | |
| TT | 24679 | | 1 | 3 | 35 | |
| TD | 24681 | 1 | | 28 | | |
| SD | 857 | | | | | |
| GS24 | 210 | | | | | |
| GS | 2803 | | 122 | | | |
| FF | 24678 | | | 8 | | |
| ww | 17446 | | | | 3 | |

mal vorgekommen, daß solche Stationen über ein Jahr lang ohne Bezug zur Wirklichkeit vor sich hingemessen haben, ohne daß dieses von irgendeiner Stelle bemerkt wurde. Relativ gesehen ist die Fehlerquote bei den 12-stündigen Niederschlagssummen am höchsten, weil hier "häufig" vergessen wird, den 6-stündigen Niederschlag mitzuberücksichtigen oder bei der Addition inklusive Rundung Fehler gemacht werden. Sicherlich werden nicht alle Fehler durch die Datenprüfung erkannt, jedoch zeigt diese Aufstellung sehr anschaulich, daß es allein mit einer formalen und klimatologischen Prüfung nicht getan ist, denn der Großteil der Fehler wird erst in den nachfolgenden Schritten entdeckt.

Es ist nicht auszuschließen, daß bei der zeitlichen Prüfung gelegentlich mal ein Wert unberechtigterweise *fachlich falsch* gegeben wird oder bei der inneren Konsistenzprüfung der Niederschlagssummen nicht die eigentliche Falschmeldung richtig erkannt wird, jedoch liegen in diesen Fällen entweder unbestreitbare Widersprüche im Datengut oder überprüfenswerte Wetterphänomene vor.

5 Datenvervollständigung

Die im Geschäftsfeld Landwirtschaft betriebenen agrarmeteorologischen Modelle brauchen zum Teil Daten in stündlicher Auflösung. Fehlen diese Werte oder wurden Werte in der Datenprüfung als falsch erkannt, so müssen diese möglichst durch richtige Werte ersetzt werden. Es liegt auf der Hand, daß die "richtigen" Werte bei kompletten Stationsausfällen über Tage lediglich plausible, jedoch keinesfalls mit den tatsächlichen Verhältnissen vor Ort identisch sein können. Oberstes Streben bei der Datenvervollständigung ist es jedoch, aus dem vorhandenen Datenmaterial die richtigen Werte zu bestimmen. Insgesamt wird zwischen vier verschiedenen Vervollständigungsverfahren unterschieden, die mehrmals hintereinander aufgerufen werden können. Wurden Werte durch ein Verfahren ersetzt oder vervollständigt, so werden sie intern zunächst als ungeprüfte Werte abgespeichert und nach Beendigung des entsprechenden Vervollständigungsverfahrens einer Prüfung unterzogen. Die Reihenfolge und Häufigkeit der Verfahrensanwendungen wurde in der Vergangenheit häufig geändert und ist im jetzigen Zustand nicht mehr identisch zu der in Tab. 2 angegebenen Reihenfolge. Das räumliche Ersetzungsverfahren hat den geringsten Bezug zum Ort und wird daher in einem Ersetzungszyklus als letztes Verfahren benutzt. Da nicht alle Elemente räumlich ersetzt werden, muß mindestens ein weiterer Zyklus von Ersetzungsverfahren nachfolgen, um alle ungewünschten Datenlücken zu eliminieren.

Die in Kapitel 4.6 angegebenen Fallbeispiele enthalten Elemente, die auch ersetzt bzw. vervollständigt werden. Aus dem Verhältnis der tatsächlichen zu den möglichen Meldungen geht schon hervor, daß die Datenvervollständigung nicht unbeträchtliche Arbeit zu vollbringen hat.

5.1 Manuell einfügen oder korrigieren

Nicht nur ein zu vervollständigendes, sondern jedes beliebige Element der Datenbank kann manuell eingefügt werden. Dieses Verfahren ist bei fehlenden Werten sinnvoll, wenn keine Synop-Meldungen in der Rechenzentrumsdatenbank vorliegen, die Werte jedoch über telefonische Rücksprache bekannt sind. Das manuelle Korrigieren darf nicht mit dem manuell bestätigen verwechselt werden, weil in dem letzteren Fall der Wert als Original unverändert bleibt und automatisch die höchste Priorität zugewiesen wird und in dem anderen Falle lediglich der Wert auf der rechten Seite des Datensatzes verändert wird, wobei das evtl. vorliegende Original auf der linken Seite unverändert bleibt. Alle manuell eingefügten oder korrigierten Werte werden auf die linke Seite des Datensatzes auf das Qualitätsniveau ungeprüft geschrieben und in der nachfolgenden Datenprüfung als richtig oder falsch erkannt. Normalerweise werden im ersten Prüfschritt nur die original vorliegenden Werte (linke Seite) einer Prüfung unterzogen, wurde jedoch manuell korrigiert, so werden diese Werte (rechte Seite) als "Originalwerte" verwendet. Somit "greift" dieses Ersetzungsverfahren bereits vor der ersten Datenprüfung.

Eine manuelle Korrektur findet nur nach Ablauf eines kompletten Prüf- und Ersetzungslaufes und Ausgabe der als falsch erkannten Werte in Form der Meckerliste statt (siehe Kapitel 4.5). Bei der nachträglichen Bearbeitung dieser Meckerliste besteht die Möglichkeit, einen evtl. als falsch erkannten Originalwert manuell zu bestätigen oder auch manuell zu korrigieren. Aber auch vollautomatisch ersetzte Werte lassen sich nachträglich noch manuell korrigieren, wenn sie unplausibel erscheinen, was in der Regel nur geschieht, falls die ersetzten Werte nach einem Ersetzungsverfahren und anschließender Prüfung ebenfalls als falsch erkannt wurden und somit in der Meckerliste erscheinen. Wurden ersetzte Werte jedoch gleich nach der anschließenden Prüfung als "richtig" erkannt, so taucht kein Hinweis in der Meckerliste auf und eine weitere Überprüfung dieser Werte wird in der Regel nicht mehr stattfinden.

5.2 Sonstige Verfahren

Bei den sonstigen Verfahren wird in der Regel versucht anhand anderer Elemente einen fehlenden oder falschen Wert möglichst realistisch aus dem eigenen Datengut abzuleiten, was zum Teil auch eindeutig möglich ist.

Tabelle 11: Auflistung der zu vervollständigenden Elemente und deren Behandlung in den einzelnen Vervollständigungsverfahren

| Elemente | Vervollständigungsverfahren | | | |
|----------|-----------------------------|----------|-----------|-----------|
| | manuell | sonstige | zeitliche | räumliche |
| DD | x | x | | |
| DD | x | | x | x |
| FF | x | | x | x |
| GS | x | x | | |
| GS24 | x | x | x | |
| KLRR24 | x | | x | |
| KLTN | x | | x | |
| KLTX | x | | x | |
| N | x | | x | x |
| PPP | x | x | | |
| PPPP | x | x | x | x |
| R1R1 | x | x | | |
| RR00 | x | x | x | |
| RR06 | x | x | x | x |
| RR12 | x | x | x | |
| RR18 | x | x | x | x |
| RR24 | x | x | x | |
| SD | x | x | | x |
| SD1 | x | x | x | |
| TD | x | | x | x |
| TG | x | x | | |
| TM | x | x | | |
| TN06 | x | | x | |
| TN18 | x | | x | |
| TT | x | x | x | x |
| TX06 | x | | x | |
| TX18 | x | | x | |
| W1 | x | x | | |
| ww | x | x | | x |

5.2.1 Temperaturen

- **TG**

Die Ersetzung der Erdbodenminimumtemperatur ist empirisch hergeleitet worden, wobei die wichtigsten Abhängigkeiten in Form von Temperaturminimum **TN06**, mittlerer Bedeckungsgrad und mittlere Windgeschwindigkeit während der Nacht, Schneedecke sowie die Lage im Gelände berücksichtigt werden. Lediglich in 3,3% aller Fälle haben die so generierten **TG** Werte eine Abweichung von mehr als 2K zu den gemessenen Werten. In Tabelle 12 sind die mittlere Abweichung und Streuung der so berechneten Werte zu den original gemessenen Werten für **TG** für das Jahr 1998 monatsweise dargestellt. Zur Einordnung der Ergebnisse wurde ein altes Verfahren gegenübergestellt, bei dem lediglich 1,5K von **TN06** abgezogen wurde.

Besonders in den Fällen mit Schneedecke werden die Ergebnisse mit dem neuen empirischen Ansatz besser als bei dem alten Verfahren, wo über das Jahr gesehen ca. 6 % der Fälle eine Abweichung von mehr als 2K aufweisen. Die mittlere Streuung bei dem neuen Verfahren ist über das Jahr gesehen konstant niedrig, jedoch wurden die Fälle, bei denen **TG** größer als **TN06** ist –kann allein durch Rundung begründet sein– nicht berücksichtigt, was die relativ hohen mittleren Abweichungen in den Sommermonaten erklären mag.

Bei einer Fehlerschwelle von 5K liefert das neue Verfahren lediglich in 0,06% aller Fälle, das

Tabelle 12: Güte des Ersetzungsverfahrens für **TG** bei den deutschen Synop-Stationen im Jahre 1998.

| | Monat | | | | | | | | | | | | Jahr |
|--------------------------------------|-------|------|------|-----|------|-----|------|-----|-----|-----|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | |
| TG über empirisches Verfahren | | | | | | | | | | | | | |
| Mittel | -1,3 | -0,4 | -0,3 | 3,2 | 6,3 | 9,9 | 10,4 | 9,7 | 8,3 | 5,0 | -2,1 | -2,7 | 3,9 |
| Abw. | 0,1 | -0,1 | 0,0 | 0,1 | 0,2 | 0,5 | 0,5 | 0,3 | 0,4 | 0,3 | 0,3 | 0,2 | 0,2 |
| Streu. | 1,2 | 1,3 | 1,1 | 1,1 | 1,3 | 1,2 | 1,1 | 1,2 | 1,2 | 1,0 | 1,3 | 1,3 | 1,2 |
| TG über TN06 - 1.5 K | | | | | | | | | | | | | |
| Mittel | -1,1 | -0,1 | 0,0 | 3,2 | 6,6 | 9,9 | 10,4 | 9,7 | 8,2 | 4,6 | -1,7 | -2,3 | 4,0 |
| Abw. | -0,1 | -0,4 | -0,3 | 0,1 | -0,1 | 0,5 | 0,6 | 0,3 | 0,5 | 0,7 | -0,1 | 0,1 | 0,1 |
| Streu. | 1,5 | 1,8 | 1,6 | 1,1 | 1,5 | 1,3 | 1,3 | 1,2 | 1,3 | 1,4 | 1,4 | 1,7 | 1,5 |

alte Verfahren 0,3% aller Fälle größere Abweichungen.

• **TM, TT_{7UTC}, TT_{14UTC}, TT_{21UTC}**

Über die Konvention zur Berechnung der klimatologischen Tagesmitteltemperatur läßt sich eindeutig ein fehlender Wert berechnen, wenn die anderen drei Werte vorliegen. Bei der Ersetzung von **TM** muß außerdem die Rundungsvorschrift beachtet werden, daß bei 5/100 K immer zum nächsten geraden Zehntel gerundet werden muß. Dieses Modul wurde bis auf weiters nicht mehr zur Ersetzung der Stundenwerte von **TT** benutzt, weil durch falsche oder unterschiedliche ausländische Bestimmungsrichtlinien von **TM** falsche Stundenwerte generiert wurden.

• **TT**

Fehlen bei einer Station Werte von **TT** bei gleichzeitiger Kenntnis der Extremtemperaturen, so wird versucht diese Extremtemperaturen einzubauen.

- Liegt bei einer der Ersatzstationen eine vollständige Zeitreihe von **TT** vor, so werden die Termine mit Höchst- bzw. Tiefstwerten an der Ersatzstation im jeweiligen Bezugszeitraum bestimmt und diese Termine zur ersetzenden Station übertragen. Liegt zu diesen Zeitpunkt der aufgetretenen Extremwerte eine Fehlstunde an der Station vor, so werden die Extremwerte für **TT** übernommen.
- Liegt keine vollständige Zeitreihe von **TT** an der Station vor, so wird zusätzlich versucht die evtl. vorliegende Extremwerte **TN06** und **TX18** mitzuberechnen. Die Extremwerte werden allerdings nur eingebaut, wenn ein "normaler" Temperaturverlauf mit den Minimumwerten um Sonnenaufgang und den Maximumwerten um den Sonnenhöchststand vorliegen. Damit **TX18** zwei Stunden nach Sonnenhöchststand eingebaut wird, muß **TT_{12UTC} > TT_{6UTC}** und **TT_{18UTC} < TX18** sein. **TN06** wird zur Sonnenaufgangsstunde eingebaut, wenn **TT_{0UTC} < TT_{18UTC}** ist.

Durch die Berücksichtigung der Maximal- und Minimaltemperaturen wird der Verlauf von **TT** realistischer. Bei dem Taupunkt gibt es keine zusätzlichen Informationen im Vorhersagezeitraum vom GM, so daß der Taupunkt am Nachmittag in der Regel zu hoch liegt, was indirekt bei den agrarmeteorologischen Berechnungen über das Sättigungsdefizit zu etwas geringeren Verdunstungswerten führt.

5.2.2 Niederschläge

Die Verschlüsselungsvorschrift –ob nun auf ganze oder 1/10 mm gemeldet werden muß– erschwert beträchtlich die Behandlung des Niederschlages, weil permanent der Bezugszeitraum und die Höhe des gemeldeten Niederschlages berücksichtigt werden muß. Aufgrund dieser erzeugten Ungenauigkeit

kann z.B. aus zwei gerundeten 12-stündigen Niederschlagshöhen kein exakter Tagesniederschlag abgeleitet werden. Die maximal mögliche Ungenauigkeit würde sich in diesem Beispiel zwischen -1,0 mm bis 0,8 mm bewegen. Ohne im Detail auf die Rundungsvorschriften einzugehen, sei an dieser Stelle lediglich darauf hingewiesen, daß diese und die besondere Behandlung des Trace mit der maximal möglichen Genauigkeit behandelt werden.

• **RR12, RR18, RR00, RR06, RR24**

Tabelle 13: Ergänzung fehlender Niederschlagshöhen

| Nr. | RR | | | | | Aktion |
|-----|----|----|----|----|----|---|
| | 12 | 18 | 00 | 06 | 24 | |
| 0 | | | | | | alles komplett |
| 1 | | | | | f | RR24=RR18+RR06 |
| 2 | | | | | | RR06=RR24-RR18 |
| 3 | | | | f | f | siehe Spezialbehandlung |
| 4 | | | f | | | wenn RR06=0 ⇒ RR00=0 |
| 5 | | | f | | f | RR24=RR18+RR06; weiter bei Nr. 4 |
| 6 | | | f | f | | RR06=RR24-RR06; weiter bei Nr. 4 |
| 7 | | | f | f | f | keine Ersetzung möglich |
| 8 | | f | | | | RR18=RR24-RR06 |
| 9 | | f | | | f | siehe Spezialbehandlung |
| 10 | | f | | f | | wenn RR24=RR12+RR00 ⇒ RR18=RR12 RR06=RR00 |
| 11 | | f | | f | f | keine Ersetzung möglich |
| 12 | | f | f | | | RR18=RR24-RR06; weiter bei Nr. 4 |
| 13 | | f | f | | f | wenn RR06=0 ⇒ RR00=0 + weiter mit Nr. 9 |
| 14 | | f | f | f | | wenn RR24=RR12 ⇒ RR18=RR12 + weiter mit Nr. 6 |
| 15 | | f | f | f | f | keine Ersetzung möglich |
| 16 | f | | | | | wenn RR18=0 ⇒ RR12=0 |
| 17 | f | | | | f | RR24=RR18+RR06; weiter mit Nr. 16 |
| 18 | f | | | f | | RR06=RR24-RR06; weiter mit Nr. 16 |
| 19 | f | | | f | f | wenn RR18=0 ⇒ RR12=0 + weiter mit Nr. 3 |
| 20 | f | | f | | | wenn RR18=0 ⇒ RR12=0 + weiter mit Nr. 4 wenn RR06=0 ⇒ RR00=0 + weiter mit Nr. 16 |
| 21 | f | | f | | f | RR24=RR18+RR06; weiter mit Nr. 20 |
| 22 | f | | f | f | | RR06=RR24-RR18; weiter mit Nr. 20 |
| 23 | f | | f | f | f | wenn RR18=0 ⇒ RR12=0 + weiter mit Nr. 7 |
| 24 | f | f | | | | RR18=RR24-RR06; weiter mit Nr. 16 |
| 25 | f | f | | | f | keine Ersetzung möglich |
| 26 | f | f | | f | | wenn RR24=RR00 ⇒ RR06=RR00 + weiter mit Nr. 24 |
| 27 | f | f | | f | f | keine Ersetzung möglich |
| 28 | f | f | f | | | RR18=RR24-RR06; weiter mit Nr. 20 |
| 29 | f | f | f | | f | wenn RR06=0 ⇒ RR00=0 + weiter mit Nr. 25 |
| 30 | f | f | f | f | | wenn RR24=0 ⇒ RR12=RR18=RR00=RR06=0 |
| 31 | f | f | f | f | f | keine Ersetzung möglich |

Die Fälle 3 und 9 in Tabelle 13 haben bei ausländischen Stationen, die kein **RR24** melden, eine besonders wichtige Bedeutung und werden daher noch speziell behandelt. Liegen von dem fehlenden 12 -stündigen Niederschlagsintervall die 6 -stündige Niederschlagsmeldung und alle original gemeldeten **ww** oder zumindest ersatzweise das **W1** vom entsprechenden Haupttermin (bezieht sich auf die letzten 6 Stunden) vor, so kann die 6 -stündige Niederschlagssumme direkt übernommen werden, falls die **ww** oder das **W1** keinen Niederschlag hergeben.

Aus dem Zwang heraus, das Datengut auf jeden Fall zu vervollständigen, wird zusätzlich noch

eine spezielle Ersetzungsvorschrift durchgeführt, die bewußt Ungenauigkeiten in Kauf nimmt. Konnten die zwei 12-stündigen Niederschlagshöhen im ersten Ersetzungszyklus bestehend aus allen vier verschiedenen Verfahren nicht ersetzt werden, so werden sie mit den evtl. vorliegenden 6-stündigen Niederschlagshöhen gleichgesetzt. Da dieses Verfahren erst ab dem zweiten Ersetzungszyklus eingesetzt wird, kommt es nur sehr selten zum tragen.

- **R1R1**

Selbstverständlich wird **R1R1** nur dann unter Berücksichtigung von **ww**, **R1R1**, **RR12**, **RR18**, **RR00**, **RR06** und **RR24** ersetzt bzw. vervollständigt, wenn ein Wert fehlt. Dieses gilt auch für Stationen, die keine stündlichen Niederschläge melden. Dieses Modul ist recht aufwendig und soll daher nur prinzipiell erläutert werden. Die vorliegenden Niederschlagssummen werden über die **ww**'s der Einzelstunden verteilt. Grundlage hierfür ist eine langjährige durchschnittliche Niederschlagshöhe jedes **ww**'s, aus der die Gewichtung bzw. Normierung der verschiedenen **ww**'s untereinander hervorgeht. Liegen keine Wettermeldungen vor, wie es z.B. im Vorhersagezeitraum des globalen Modells immer der Fall ist, so wird die Niederschlagshöhe der mittleren Stunde des Bezugszeitraumes zugeordnet. Schwierig bei der Verteilung ist die Tatsache, daß die Niederschlagshöhen mit 6- und 12-stündigen Bezugszeiträumen im Gegensatz zum Tagesniederschlag bei Bedarf gerundet werden. Die Gesamtsumme der zum Teil vorliegenden und zu ersetzenden **R1R1** muß exakt mit der Tagessumme übereinstimmen, jedoch gleichzeitig unter Berücksichtigung der Rundungsvorschriften mit den evtl. vorliegenden Niederschlagshöhen der kürzeren Bezugszeiträume übereinstimmen. Hierbei kommt erschwerend hinzu, daß die im Original vorliegenden **R1R1** unverändert bleiben müssen. Liegt ein Trace für den jeweiligen Bezugszeitraum vor, so werden alle Stunden, in denen das **ww** Niederschlag zuläßt, mit einem Trace belegt. Die zu verteilenden Niederschlagshöhen der Bezugszeiträume haben gegenüber dem **ww** in dem Fall Priorität, wenn alle **ww** innerhalb eines Bezugszeitraumes Niederschlag eindeutig ausschließen, jedoch Niederschläge zu verteilen sind. In diesem Fall werden sie der mittleren Stunde des Bezugszeitraumes zugeordnet. Das **ww** hat auf jeden Fall Priorität, wenn es im Gegensatz zu den Niederschlagshöhen des Bezugszeitraumes als Original vorliegt.

5.2.3 Strahlung

- **GS24, SD**

Zwischen der Sonnenscheindauer und der Globalstrahlung besteht ein enger Zusammenhang, den man sich bei der Ersetzung zunutze macht. Diese Beziehung wurde bereits bei der inneren Konsistenzprüfung von **SD** in Form der Gl. 6 genutzt. Diese Beziehung kann selbstverständlich nach **GS24** umgeformt werden, so daß man bei vorliegendem **SD** die Globalstrahlung berechnen kann.

Um einen Eindruck von der Güte dieses Verfahrens zu bekommen, sollen einmal beispielhaft alle Fälle aus dem Jahre 1997 untersucht werden, an denen die deutschen Synop-Stationen beide Elemente original gemeldet haben. Die über Gl. 6 berechneten Werte werden mit den jeweiligen Originalwerten verglichen und die mittlere Abweichung und Streuung bestimmt.

Die Ergebnisse sind recht erfreulich, denn die mittleren Abweichungen sind in Bezug zum Mittelwert zu vernachlässigen. Dagegen ist die Streuung in den Wintermonaten aufgrund der geringen Mittelwerte in einer nicht zu vernachlässigenden Größenordnung. Bei der Betrachtung wurde allerdings vorausgesetzt, daß die gemessenen Werte keine Meßungenauigkeit aufweisen, was bei einigen Stationen definitiv nicht der Fall ist. Somit ist in der Streuung nicht nur die des Verfahrens sondern auch die der Meßungenauigkeit enthalten. Bei der mittleren Abweichung werden diese Meßungenauigkeiten sozusagen "weggemittelt".

Da in der Regel jede Synop-Station die Sonnenscheindauer, aber nur sehr wenige die Globalstrahlung melden, wird das Verfahren fast ausschließlich zur Ersetzung von **GS24** genutzt. Nur bei den Vorhersagezeiträumen liegen ausschließlich vorhergesagte Globalstrahlungswerte

Tabelle 14: Güte des Ersetzungsverfahrens für **SD** und **GS24** bei den deutschen Synop-Stationen im Jahre 1997.

| | Monat | | | | | | | | | | | | Jahr |
|-------------|-------|-------|-------|-------|-------|-------|------|------|------|-------|-------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | |
| SD | | | | | | | | | | | | | |
| Mittel | 2,2 | 2,7 | 4,0 | 6,2 | 6,8 | 7,5 | 6,4 | 7,9 | 7,0 | 3,8 | 2,0 | 0,9 | 4,8 |
| Abw. | 0,03 | -0,05 | -0,04 | -0,01 | -0,06 | -0,03 | 0,03 | 0,42 | 0,01 | -0,14 | -0,04 | 0,01 | 0,01 |
| Streu. | 1,2 | 1,0 | 1,1 | 1,3 | 1,4 | 1,6 | 1,5 | 1,5 | 1,2 | 1,1 | 0,8 | 0,7 | 1,2 |
| GS24 | | | | | | | | | | | | | |
| Mittel | 285 | 490 | 880 | 1474 | 1812 | 2009 | 1774 | 1744 | 1343 | 684 | 326 | 175 | 1097 |
| Abw. | 4 | -1 | -9 | -3 | 9 | 22 | -8 | -37 | 23 | 18 | 4 | -51 | 1 |
| Streu. | 82 | 110 | 153 | 201 | 228 | 239 | 245 | 203 | 146 | 127 | 72 | 57 | 169 |

vor und kein **SD**, so daß hier immer Gl. 6 benutzt wird. Zum Glück treten bei den Vorhersagen keine Meßungenauigkeiten auf, so daß das Ersetzungsverfahren in Anbetracht der Vorhersagegenauigkeit hinreichend exakte Werte berechnet.

• **GS, SD1**

Bis zum Jahre 1998 wurde die stündliche Globalstrahlung ausschließlich mit einem Verfahren nach Kasten et al. (1984) ersetzt, welches über den Sonnenstand und den Gesamtbedeckungsgrad eine Globalstrahlung berechnen kann. Dieses Verfahren beinhaltet einige Tücken, weil z.B. beim Bedeckungsgrad nicht zwischen den hohen und tiefen Wolken unterschieden wird sowie mittlere monatliche optische Durchlässigkeiten der Atmosphäre angenommen werden. Bei Kaltlufteinbrüchen im Sommer mit hohen optischen Durchlässigkeiten oder bei bedecktem Himmel mit Cirrusbewölkung bestehen große Abweichungen zwischen den berechneten und gemessenen Werten von **GS**.

Seit 1999 steht zusätzlich die stündliche Sonnenscheindauer **SD1** zur Berechnung von **GS** zur Verfügung. Über eine vergleichbare Beziehung wie Gl. 6, die auf den Bezugszeitraum von einer Stunde angepaßt wurde, läßt sich **SD1** aus vorhandener **GS** oder umgekehrt berechnen. Entsprechend der Tabelle 14 sind die Ergebnisse dieser Ersetzungsverfahren in Tabelle 15 für das Jahr 1999 dargestellt. Es wurden lediglich die Tagesstunden ausgewertet, an denen **SD1**, **GS** und **N** (für das Verfahren nach Kasten) als richtige Originale vorlagen.

Tabelle 15: Güte des Ersetzungsverfahrens für **SD1** und **GS** bei den deutschen Synop-Stationen im Jahre 1999.

| | Monat | | | | | | | | | | | | Jahr |
|---|-------|-------|------|------|------|-------|------|------|------|------|------|-------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | |
| SD1 in 1/10 Std. | | | | | | | | | | | | | |
| Mittel | 2,19 | 2,20 | 2,75 | 3,56 | 4,73 | 3,81 | 4,55 | 4,06 | 5,10 | 3,02 | 2,04 | 1,31 | 3,51 |
| Abw. | 0,06 | -0,36 | 0,05 | 0,10 | 0,21 | -0,08 | 0,18 | 0,17 | 0,69 | 0,10 | 0,03 | -0,02 | 0,11 |
| Streu. | 1,57 | 1,51 | 1,42 | 1,50 | 1,51 | 1,34 | 1,43 | 1,47 | 1,82 | 1,43 | 1,46 | 1,23 | 1,48 |
| GS in J/cm² | | | | | | | | | | | | | |
| Mittel | 31 | 48 | 65 | 91 | 116 | 108 | 114 | 99 | 93 | 55 | 32 | 21 | 81 |
| Abw. | -3 | 0 | -6 | -4 | -4 | -1 | -4 | -3 | -8 | -4 | -5 | -4 | -4 |
| Streu. | 11 | 15 | 20 | 21 | 22 | 22 | 20 | 19 | 18 | 14 | 13 | 9 | 18 |
| GS in J/cm² nach Kasten | | | | | | | | | | | | | |
| Mittel | 31 | 49 | 67 | 93 | 123 | 115 | 120 | 104 | 97 | 57 | 33 | 22 | 83 |
| Abw. | 0 | 1 | -3 | -1 | 3 | 0 | 2 | 1 | -1 | -3 | 0 | 0 | 0 |
| Streu. | 14 | 21 | 24 | 31 | 37 | 38 | 35 | 34 | 26 | 21 | 15 | 12 | 29 |

Die Berechnung von **SD1** über **GS** ist zufriedenstellend, da die mittlere Abweichung verschwindend gering und die Streuung nahezu konstant über die Monate bei 15/100 Stunden liegt. Wie nicht anders zu erwarten war, ist die Berechnung von **GS** über **SD1** exakter als die über den Gesamtbedeckungsgrad. Die mittleren Abweichungen sind bei beiden Verfahren sehr gut, jedoch ist die Streuung bei dem Verfahren von Kasten signifikant schlechter. Das Verfahren nach Kasten wird nur benutzt, wenn **GS** und **SD1** fehlen.

Damit die ersetzten Stundenwerte von **GS** und **SD1** mit den Tagessummen übereinstimmen, werden die ersetzten $GS(SD1)_{ers}$ Werte noch mit dem Faktor f_{ak}

$$f_{ak} = \frac{GS24\{SD\} - \sum GS\{SD1\}_{gem}}{\sum GS\{SD1\}_{ers}} \quad (8)$$

normiert. Damit bei der Normierung von **SD1** bei einem Faktor > 1.0 die Einzelwerte nicht über 10 kommen können wird hier noch einmal eine Spezialbehandlung durchgeführt. Somit werden die Verfahren lediglich dazu verwandt, die relative Verteilung von **GS** und **SD1** über den Tag zu bestimmen.

5.2.4 Andere Elemente

- **ww**

Fehlt ein beobachtetes Wetter zum Termin, so wird anhand der davorliegenden Stunde versucht, ein "20-er Wetter" zu erzeugen. Die Vorstunde muß ein richtiges $ww \geq 40$ oder $ww = 17$ haben, damit eine Vervollständigung erfolgt. Die so erzeugten **ww**'s spiegeln somit das minimal mögliche Wetter zum Termin wider, weil ja nicht ausgeschlossen werden kann, daß z.B. der Niederschlag angehalten oder sogar ein höherwertiges **ww** aufgetreten ist. Außerdem wird Nebel mit Hilfe der original gemeldeten Sichtweite und der rel. Feuchte abgeleitet.

- **W1**

Das vergangene Wetter hat je nach Meldetermin unterschiedliche Bezugszeiträume. Es werden nur die Zwischentermine mit Bezugszeitraum von 3 Stunden und die Haupttermine mit 6 Stunden ersetzt. Aus dem Bezugszeitraum wird das höchste vorliegende oder das über **ww** abgeleitete **W1** ersetzt, welches aus einem korrespondierendem Zuordnungsfeld entnommen wird. In diesem Zuordnungsfeld werden alle $ww < 20$ auf $W1 = 0$ gesetzt und die Spezialbehandlungen für **ww** 17, 20 bis 29, 40, 58, 59, 68, 69, 76 sowie 90 berücksichtigt.

- **PPPP, PPP A**

Bei Kenntnis von **PPPP** und **TT** lassen sich unter Verwendung der Gl. 7 **PPP** und **A** eindeutig ableiten, wobei bei **A** der stündliche Verlauf von **PPPP** berücksichtigt werden muß. **PPPP** läßt sich dagegen nur eindeutig ersetzen, wenn **PPPP(i-3)**, **PPP**, **A** und **TT** bekannt sind.

5.3 Zeitliche Interpolation

Bei der zeitlichen Interpolation muß prinzipiell zwischen zwei verschiedenen Vervollständigungsverfahren unterschieden werden. Einmal lassen sich Summen oder Extremwerte von stündlich vorliegenden Elementen aus diesen herleiten, zum anderen können stündliche Elemente, die über einen "kurzen" Zeitraum fehlen, linear interpoliert werden. In diesen Interpolationsverfahren werden nur die Werte von der zu behandelnden Station herangezogen.

5.3.1 Temperaturen

- **TT, TD**

Fehlende Stunden in einem Zeitraum werden mit Hilfe der richtigen Werte vor und nach

der Lücke linear interpoliert. Es wird bis zu einer Lücke von 5 Stunden linear interpoliert, andernfalls wird auf die räumliche Ersetzung vertraut. Die maximale Lücke von 5 Stunden ist zugegebenerweise zu groß und kann zu unrealistischen Temperaturverläufen führen. Die maximal zulässige Lücke von 5 Stunden muß genommen werden, damit aus den Vorhersagedaten des GM, die nur in 6-stündiger Auflösung vorliegen, stündliche Werte erzeugt werden können, denn auch räumlich stehen keine Zwischentermine zur Verfügung. Intern werden die linear ersetzten Taupunkte mit den Werten von **TT** verglichen. Ist $TD \geq TT$, so bekommt **TD** den Wert von **TT**. Bei den gemessenen Daten wird in der Regel nicht über einen Zeitraum von 5 Stunden linear interpoliert, weil die Fehlzeiträume typischerweise kürzer oder bei Stationsausfällen oder bei Stationen der Meldesollgruppe II größer sind.

- **TX06, TX18, KLTX, TN06, TN18, KLTN**

Die Minimumtemperaturen während des jeweiligen Bezugszeitraumes werden mit Hilfe der stündlich vorliegenden **TT** -Werte ersetzt. Im Bezugszeitraum werden TT_{min} und TT_{max} bestimmt. Liegt TT_{min} am Ende oder Anfang des Bezugszeitraumes, so wird TT_{min} als Minimumwert übernommen, andernfalls wird die Minimumtemperatur **TN** folgendermaßen

$$TN = TT_{min} - \frac{TT_{max} - TT_{min}}{24} \quad (9)$$

ersetzt. In der gleichen Art und Weise wird die Maximumtemperatur des jeweiligen Bezugszeitraumes ersetzt. Der zusätzliche Auf- bzw. Abschlag bei der Temperatur zu TT_{max} bzw. TT_{min} ist mit der Temperaturschwankung innerhalb des Bezugszeitraumes gekoppelt.

5.3.2 Niederschläge

- **RR12, RR18, RR00, RR06, RR24, KLRR24**

Fehlende Niederschlagshöhen während eines Bezugszeitraumes lassen sich mit Hilfe der komplett vorliegenden stündlichen Niederschlagshöhen berechnen. Bei der Summation werden die Rundungsvorschriften berücksichtigt, und bei mehreren Stunden mit Trace, wird als Summe ebenfalls Trace gegeben.

5.3.3 Strahlung

- **GS24**

Bei allen vorliegenden stündlichen Globalstrahlungswerten, kann die Tagessumme aufsummiert und somit ersetzt werden. Fehlen am Tage weniger als drei original gemeldete **GS** Werte und liegen in diesen Stunden originale **SD1** Werte vor, so werden die fehlenden **GS** Werte über **SD1** ersetzt und für die Tagessummation verwendet.

- **SD**

Bei allen am Tage vorliegenden originalen stündlichen Sonnenscheindauerwerten, kann die Tagessumme aufsummiert und somit ersetzt werden. Bei bis zu drei fehlenden originalen Werten von **SD1** am Tage werden diese indirekt aus der Globalstrahlung abgeleitet und für die Tagessummation verwendet, weil so der Fehler im Vergleich zu den anderen Ersetzungsverfahren zu vernachlässigen ist.

5.3.4 Wind

- **FF**

Bei fehlenden Zeiträumen, die kleiner als 6 Stunden sind, werden die Fehlstunden einfach linear zwischen dem zuletzt vorliegenden und dem Terminwert nach der Lücke interpoliert.

- **DD**

Es wird nur durch zeitliche Interpolation ersetzt, wenn der fehlende Zeitraum kleiner als 6

Stunden ist. Über die zuletzt gemeldete Stunde mit definierter Windrichtung (keine umlaufenden Winde oder $DD = 0^\circ$) und dem ersten Terminwert nach der Lücke mit definierter Windrichtung wird linear interpoliert. Hierbei werden Spezialfälle, wie Winddrehungen über Nord mitberücksichtigt. Es wird immer in die Richtung gedreht, bei der die Differenz zwischen den noch vorliegenden DD Werten kleiner als 180° ist.

Die so berechneten Windrichtungen werden nun noch mit der dazugehörigen Windgeschwindigkeit verglichen. Ist $FF = 0m/s$ dann wird dd ebenfalls auf 0 gesetzt, ist $FF < 1,5m/s$ so wird DD auf umlaufenden Wind gesetzt.

5.3.5 Andere Elemente

- **N**

Auch in diesem Modul beträgt die maximal zulässige Datenlücke 5 Stunden, bei der eine lineare Interpolation über die die Lücke definierenden Werte erfolgt. Im Gegensatz zur Temperatur oder dem Taupunkt schwankt der Gesamtbedeckungsgrad in der Regel nicht so gravierend innerhalb von 5 Stunden, so daß die so erzeugten Ungenauigkeiten vertretbar sind.

- **PPPP**

Korrespondierend zu den anderen Elementen wird bei einer maximal zulässigen Datenlücke von zur Zeit 5 Stunden linear interpoliert. Dieses Modul ist noch sehr neu und muß noch genauer verifiziert werden. Die sicherlich zu große zulässige Lücke von 5 Stunden ist bei gemessenen Werten auf jeden Fall zu hoch, wurde jedoch aufgrund der zeitlich nur in 6 Stunden vorliegenden Vorhersagen des globalen Modells gewählt, weil hier keine Zwischenwerte räumlich –fehlen bei den Ersatzstationen auch– ersetzt werden können. Dieses ist ein schönes Beispiel, daß die Reihenfolge der Ersetzungsverfahren einen großen Einfluß auf die Qualität der erzeugten Werte hat. Es sollte auf jeden Fall zuerst das in diesem Fall bessere räumliche Ersetzungsverfahren benutzt werden, bevor zeitlich interpoliert wird.

5.4 Räumliche Interpolation

Jede in der agrarmeteorologischen Datenbank enthaltene Station besitzt genau zwei Ersatzstationen, die sich in unmittelbarer Nachbarschaft befinden und in etwa die gleiche Stationshöhe haben. Anhand eines kompletteren Datensatzes dieser Ersatzstationen werden Lücken, die durch die anderen Ersetz- bzw. Vervollständigungsverfahren noch nicht geschlossen werden konnten, an der Station aufgefüllt. Dieses Vervollständigungsverfahren ist z.B. bei den fiktiven Stationen (siehe Kap. 2) die einzige Möglichkeit, einen kompletten Datensatz zu generieren.

5.4.1 Temperaturen

- **TT, TD**

Sind zu diesem Zeitpunkt der Datenersetzung noch Datenlücken vorhanden, so wird die Zeitspanne –in der Regel über 6 Stunden– bestimmt, in der keine Werte vorliegen.

- Fehlen während des gesamten Prüf- bzw. Ersetzzeitraums Werte, so werden von den Fehlstunden, die Werte der beiden Ersatzstationen herangezogen und zur Vervollständigung genutzt. Liegen in einer Fehlstunde Werte beider Ersatzstationen vor, so wird unter Berücksichtigung einer Höhenkorrektur von $0,6\text{ K}/100\text{ m}$ von beiden Werten das arithmetische Mittel gebildet und dieser Wert ersetzt. Liegt nur ein Werte vor, so wird dieser unter Berücksichtigung der Höhenkorrektur ersetzt. Bei **TD** und **TT** wird die gleiche Höhenkorrektur verwendet, damit der an den Ersatzstationen beobachtete Spread erhalten bleibt und Werte von **TD** nicht “künstlich“ über die korrespondierenden Werte von **TT** gerechnet werden.

- Fehlt nicht der gesamte Ersatzzeitraum, so wird die Ersatzstation zum Auffüllen herangezogen, die die meisten Einzelstunden in dem Fehlzeitraum vorliegen hat. Zunächst werden die jeweiligen positiven $\Delta_{Ers.Sta.}^+$ und negativen $\Delta_{Ers.Sta.}^-$ stündlichen Temperaturänderungen der Ersatzstation innerhalb des Ersatzzeitraumes aufsummiert. Außerdem werden die Wertedifferenzen $DIF_{Sta.}$ der zu ersetzenden und $DIF_{Ers.Sta.}$ der Ersatzstation zwischen den Zeitpunkten ie (Zeitpunkt, an dem wieder ein Wert an der zu ersetzenden Station vorliegt) und ia (letzter vorliegender Zeitpunkt von der zu ersetzenden Station) gebildet. Von dem Temperaturverlauf der Ersatzstation ausgehend muß die Differenz $DIF = DIF_{Sta.} - DIF_{Ers.Sta.}$ noch in irgendeiner Form auf die zu ersetzende Station verteilt werden. Dieses wird dadurch realisiert, daß jeweils die $\Delta_{Ers.Sta.}^+$ und $\Delta_{Ers.Sta.}^-$ so auf die zu ersetzende Station übertragen werden, daß deren Summe genau $DIF_{Sta.}$ entspricht.

$$\Delta_{Sta.}^+ = \frac{\Delta_{Ers.Sta.}^+}{|\Delta_{Ers.Sta.}^-| + \Delta_{Ers.Sta.}^+} DIF + \Delta_{Ers.Sta.}^+ \quad \text{und} \quad \Delta_{Sta.}^- = DIF_{Sta.} - \Delta_{Sta.}^+ \quad (10)$$

Von dem an der zu ersetzenden Station vorliegenden Wert $T_{Sta.}(ia)$ ausgehend, wird nun mit Hilfe der stündlichen Temperaturänderungen der Ersatzstation $T_{Ers.Sta.}(ia + 1) - T_{Ers.Sta.}(ia)$ und den jeweiligen Faktoren $\frac{\Delta_{Sta.}^+}{\Delta_{Ers.Sta.}^+}$ (bei positiven Temperaturänderungen) bzw. $\frac{\Delta_{Sta.}^-}{\Delta_{Ers.Sta.}^-}$ (bei negativen Temperaturänderungen) die Folgestunde $T_{Sta.}(ia + 1)$ berechnet.

$$T_{Sta.}(i + 1) = T_{Sta.}(i) + (T_{Ers.Sta.}(i) - T_{Ers.Sta.}(i - 1)) \frac{\Delta_{Sta.}^+}{\Delta_{Ers.Sta.}^+} \quad \text{bzw.} \quad \frac{\Delta_{Sta.}^-}{\Delta_{Ers.Sta.}^-} \quad (11)$$

Da **TT** und **TD** voneinander keine Kenntnis haben, kann es zur Zeit passieren, daß die Werte von **TD** über denen von **TT** liegen, bei der nächsten Prüfung falsch gegeben und im nächsten Ersetzungsschritt auf den Wert von **TT** gesetzt werden. Evtl. muß Gl. 10 dahingehend modifiziert werden, daß die Information von **TT** und **TD** gemeinsam berücksichtigt werden.

5.4.2 Niederschläge

- **RR18, RR06**

Wenn an der Station die Summe des Tagesniederschlages **RR24** fehlt, so muß zu diesem Zeitpunkt der Vervollständigung mindestens ein 12 -stündiger Niederschlag fehlen. Ist dieses der Fall, so wird der fehlende Wert mit Hilfe der Ersatzstationen räumlich ersetzt. Liegen von beiden Ersatzstationen Werte vor, so wird das arithmetische Mittel gebildet, wobei das Trace als niederschlagsfrei interpretiert wird. Um nicht gleich bei der nachfolgenden Datenprüfung mit dem so ersetzten Wert zu scheitern, wird die so ersetzte 12 -stündige Niederschlagshöhe mit der evtl. vorliegenden korrespondierenden 6 -stündigen Niederschlagshöhe verglichen. Ist die 6 -stündige Niederschlagshöhe bereits größer als die ersetzte 12 -stündige Niederschlagshöhe, so wird die vorliegende 6 -stündige Niederschlagshöhe ersetzt und es findet keine räumliche Ersetzung statt.

Durch das Ersetzen der 12 -stündigen Niederschlagshöhen ist der Grundstein für das endgültige Vervollständigen aller Niederschlagshöhen in 5.2.2 gelegt.

5.4.3 Strahlung

- **SD**

Liegt an der Station bis zu diesem Zeitpunkt der Vervollständigung keine Sonnenscheindauer vor, so wird sie über die Ersatzstationen räumlich interpoliert. Liegen beide Werte von den

Ersatzstationen vor, so wird das arithmetische Mittel gebildet, andernfalls wird lediglich der vorliegende Wert einer Ersatzstation übernommen. Eine Ersetzung muß entfallen, wenn beide Ersatzstationen keine Sonnenscheindauer haben.

5.4.4 Wind

- **FF**

Wie bei der räumlichen Ersetzung der Temperaturen werden die Fehlstunden bestimmt, in denen keine Werte vorliegen.

- Fehlen während des gesamten Prüf- bzw. Ersetzzeitraums Werte, so werden von den Fehlstunden wie bei der Temperatur über die Werte der beiden Ersatzstationen aufgefüllt.
- Fehlt nicht der gesamte Ersetzzeitraum, so wird die Ersatzstation zum Auffüllen herangezogen, die die meisten original gemeldeten Einzelstunden in dem Fehlzeitraum vorliegen hat. Nun werden die Wertedifferenzen $D(ia)$ und $D(ie)$ zwischen der zu ersetzenden Station und der Ersatzstation zu den Zeitpunkten ia (letzter vorliegender Zeitpunkt von der zu ersetzenden Station) und ie (Zeitpunkt, an dem wieder ein Wert an der zu ersetzenden Station vorliegt) gebildet. Über die Beziehung

$$FF_{Sta.(i)} = FF_{Ers.Sta.(i)} + D(ia) + \frac{D(ie) - D(ia)}{ie - ia}(i - ia) \quad (12)$$

werden innerhalb der Fehlstunden i ($ia < i < ie$), die Werteänderungen der Ersatzstation unter Berücksichtigung der verschiedenen Ausgangsniveaus zu Beginn und am Ende der Fehlstunden berücksichtigt. Werte unter $0m/s$ werden auf $0m/s$ gesetzt.

- **DD**

Vergleichbar zu Abschnitt 5.4.1 wird die Ersatzstation ermittelt, die während der Datenlücke an der zu ersetzenden Station am meisten Meldungen vorliegen hat. Nun werden die Werte der Ersatzstation einfach räumlich interpoliert, wobei allerdings noch eine Prüfung gegen die Windgeschwindigkeit am Ort erfolgt. Ist $FF = 0m/s$, so wird **DD** ebenfalls auf 0° gesetzt. Die Windrichtungsunterschiede zwischen der Ersatz- und der zu ersetzenden Station vor und nach der Datenlücke werden in diesem Verfahren nicht berücksichtigt.

5.4.5 Andere Elemente

- **N**

Fehlende Zeiträume des Gesamtbedeckungsgrades werden nach dem gleichen Verfahren wie die Windgeschwindigkeit interpoliert.

- **ww**

Anhand der bei den Ersatzstationen vorliegenden Meldungen von **ww** werden fehlende oder falsche Werte räumlich ersetzt. Es wird zunächst versucht, einen original gemeldeten richtigen Wert von einer Ersatzstation zu ersetzen. Liegt kein originales **ww** vor, so spielt die Herkunft keine Rolle mehr und es wird der zuerst gefundene richtige Wert einer Ersatzstation räumlich übernommen.

- **PPPP**

Fehlende Zeiträume des auf NN reduzierten Bodendruckes **PPPP** werden über das evtl. zu bildende arithmetische Mittel der Werte beider Ersatzstationen räumlich interpoliert. Liegt zu einer Fehlstunde nur an einer Ersatzstation ein Wert vor, so wird dieser direkt übernommen. Aufgrund möglicher Höhenunterschiede der beteiligten Stationen darf die 3-stündige Luftdruckänderung nicht räumlich interpoliert werden.

5.5 Erfolgsquote der Vervollständigung

Anhand des Jahres 1997 und aller deutschen Synop-Stationen soll abschließend der Vervollständigungsgrad aller zu vervollständigenden Elemente ausgewertet werden. Zur besseren Bewertung ist ebenfalls das prozentuale Verhältnis von Anzahl der original gemeldeten zu theoretisch möglichen Werte in Tabelle 16 angegeben.

Tabelle 16: Vervollständigungsgrad aller zu ersetzenden Elemente bei den deutschen Synop-Stationen im Jahre 1997.

| Elemente | Originale in % | | Anzahl der Werte | | |
|----------|----------------|----------|-------------------|---------|---------|
| | gemeldet | richtige | falsche Originale | ersetzt | fehlend |
| DD | 94,64 | 94,63 | 161 | 92845 | 1 |
| FF | 94,62 | 94,59 | 436 | 93578 | 1 |
| GS | 11,02 | 9,41 | 27830 | 1571525 | 0 |
| GS24 | 20,07 | 19,47 | 435 | 58197 | 0 |
| KLRR24 | 77,65 | 77,32 | 239 | 16374 | 20 |
| KLTN | 77,81 | 77,67 | 100 | 16128 | 7 |
| KLTX | 77,86 | 77,77 | 68 | 16068 | 1 |
| N | 69,81 | 69,81 | 0 | 523669 | 0 |
| R1R1 | 66,33 | 65,90 | 7445 | 591555 | 36 |
| RR00 | 76,64 | 76,56 | 58 | 16395 | 3 |
| RR06 | 96,05 | 95,13 | 665 | 3516 | 2 |
| RR12 | 78,72 | 78,65 | 48 | 15426 | 2 |
| RR18 | 95,43 | 94,28 | 828 | 4129 | 2 |
| RR24 | 94,62 | 93,68 | 682 | 4568 | 1 |
| SD | 81,56 | 81,21 | 247 | 13426 | 0 |
| TD | 94,82 | 94,73 | 1477 | 91401 | 4 |
| TG | 94,62 | 93,95 | 482 | 4370 | 0 |
| TM | 94,02 | 93,64 | 273 | 4593 | 0 |
| TN06 | 96,59 | 94,77 | 1317 | 3781 | 1 |
| TN18 | 94,66 | 93,50 | 842 | 4697 | 2 |
| TT | 94,85 | 94,62 | 4008 | 93327 | 32 |
| TX06 | 95,29 | 94,31 | 707 | 4112 | 0 |
| TX18 | 95,93 | 94,53 | 1012 | 3953 | 1 |
| ww | 69,58 | 69,47 | 1899 | 528379 | 1217 |

Allein die Anzahl der *falsch erkannten originalen Werte* aller deutschen Synop-Stationen in Tabelle 16, die –zur nachträglichen Überprüfung– in die Meckerliste eingetragen werden, läßt den täglichen Überprüfungs- und Korrekturaufwand für alle ca. 1100 Stationen erahnen. Die relativ hohe Anzahl der Fehlwerte bei **ww** ist nicht so tragisch, weil dieses Element lediglich zur Verteilung des stündlichen Niederschlags benutzt wird. Fehlen alle **ww** innerhalb des Bezugszeitraums, so wird der Niederschlag der mittleren Stunde zugeordnet.

6 Ausblick

Mit Einführung der neuen Klimatermine im Laufe des Frühjahres 2001 werden sich einige Probleme mit der selbstgestrickten Datenbank ergeben, weil die alten drei Klimatermine aus Gründen der Speicherersparnis fest vorgegeben sind. Die Datenbank müßte daher die neuen und alten Klimatermine aufnehmen können und die Datenprüfung und -ersetzung müßte automatisch erkennen, welches Verfahren in Abhängigkeit der Klima-Meldetermine ausgewählt werden muß.

Diese Änderungen sind machbar, jedoch ist die Möglichkeit zur Aufnahme neuer Elemente in der jetzigen Bank nahezu ausgeschöpft, was zukünftige schnelle Veränderungen behindern wird. Aus

diesem Grunde soll nicht unnötiger Aufwand mit der jetzigen Datenbank betrieben werden. Vielmehr bietet es sich nun an, ein komplett neues Bankkonzept zu entwickeln, welches die nun vorhandenen DWD internen Ressourcen in Form der Oracle-Datenbank nutzt. Dieses hätte den Vorteil, daß auch andere interne Nutzer leichter auf die Daten des Geschäftsfeldes Landwirtschaft zugreifen könnten. Eine gemeinsame Plattform bietet außerdem die Möglichkeit Datenprüfmodule oder zumindest im ersten Schritt Fehlerprotokolle innerhalb des Wetterdienstes zu vereinheitlichen, um so schneller und koordinierter auf Fehler in Meßwerten aufmerksam zu werden und letztendlich ein besseres Qualitäts-Management-System einzuführen.

7 Literatur

- **DVWK:** Ermittlung der Verdunstung von Land- und Wasserflächen. DVWK-Merkblätter zur Wasserwirtschaft, Heft 238 (1996), Bonn
- **Vogelgesang, Dr. R.:** Plausibilitätskontrollen meteorologischer Daten. Abschlußbericht zum FE-Vorhaben AM1 -Nr. 90/08, Deutscher Wetterdienst, 1993, Offenbach