

Report No. 3

Überprüfung der Metadaten der Agromet-Datenbank der FAO als Vorbereitung zum Einlesen in Mirakel

Jürgen Grieser
(DWD/GPCC)

Inhalt

1. Einleitung	2
2. Ergebnisse der Prüfung der Metadaten	2
2.1. Übersicht und Prüfkriterien	2
2.2 Es existiert eine Vergleichsstation mit anderer Länderkennung	4
2.3 Die WMO-Nummer existiert in der WZN-Datenbank an anderem Ort	5
2.4 Es existiert eine Vergleichsstation mit einem anderen Namen	6
2.5 Die Vergleichsstation hat eine andere WMO-Nummer	6
2.6 Es existieren mehrere Vergleichsstationen mit teils passender Information	7
2.7 Die FAO-Station wird als neue Station erkannt	7
2.8 Abschliessende Überprüfung aller unklaren Fälle und neuen Stationen	7
3. Stand der FAO-Daten nach Kontrolle der Metainformationen	8

1. Einleitung

Erfahrungsgemäß ist keine Datensammlung oder Datenbank frei von Fehlern. In Datenbanken von Stationsdaten können zwei Arten von Fehlern auftreten:

1. Die Stationsmetainformationen wie Ortskoordinaten, Name, Länderkennung und WMO-Nummer können vollständig oder teilweise falsch sein und
2. Die Beobachtungsdaten selbst können falsch sein.

In dieser Untersuchung wird nur die erste Art von Fehlern betrachtet. Dabei ist es bei einem Stationsumfang von 5996 Temperaturstationen und 13350 Niederschlagsstationen aus über 150 Ländern praktisch unmöglich die Informationen über jede einzelne Station im Detail auf etwaige Fehler zu untersuchen.¹

Um dennoch die Fehler in der Datenbank des WZN möglichst gering zu halten, wird ein Abgleich der Stationsmetainformationen im Datensatz und der MIRAKEL-Datenbank entsprechend dem WZN-Lademodul (PDB 2000) durchgeführt. Dazu wird auf die bereits in der Datenbank des WZN verfügbaren Stationen (derzeit 59287) zurückgegriffen und nach Vergleichsstationen gesucht. So gelingt es, die Menge der FAO-Stationen in drei Gruppen einzuteilen. Einerseits sind dies die eindeutig erkannten und in Bezug auf die WZN-Datenbank als richtig anzusehenden Stationen. Andererseits werden Stationen, zu denen keine Vergleichsstationen gefunden werden, zunächst als ‚neu‘ angenommen und Stationen, die nicht eindeutig einer Vergleichsstation zugeordnet werden können, bilden die Gruppe der unklaren Fälle. Auf diese und die als neu eingestuft Stationen ist das Hauptaugenmerk dieser Untersuchung gerichtet.

2. Ergebnisse der Prüfung der Metadaten

2.1. Übersicht und Prüfkriterien

Verschiedene Gründe können dazu führen, dass eine Station in der FAO-Datenbank im Vergleich zur WZN-Datenbank eine unklare Station ist. Nach einem Plausibilitätscheck der Koordinaten wird in der WZN-Datenbank nach Vergleichsstationen gesucht, deren Horizontalkoordinaten um nicht mehr als .01° (pro Koordinate) von den vorgegebenen Koordinaten der Station in der FAO-Datenbank abweichen dürfen. Mit diesen Stationen wird dann der Vergleich durchgeführt. Als Vergleichsparameter dienen WMO-Nr., Länderkennung und Stationsname. Dabei muss berücksichtigt werden, dass für einige Stationen nicht alle Vergleichsparameter vorhanden sind. Tabelle 1 gibt einen Überblick darüber, bei wie vielen FAO-Stationen nur Teile der Vergleichsparameter zur Verfügung stehen.

Tabelle 1: Metainformationen zu Temperatur- und Niederschlagsstationen der FAO-Datenbank.

Stationen	Niederschlag	Temperatur
Gesamtzahl	13530	5996
Ohne WMO-Nr.	8173	2560
Ohne Namen	0	0
Ohne Länderkennung	66	42
Ohne Stationshöhe	303	39
Ohne Stationshöhe und WMO-Nr.	303	38
Ohne Länderkennung und WMO-Nr.	22	13

Aus den zur Verfügung stehenden Informationen ergibt sich die folgende Liste von Gründen, die im Lademodul des WZN zu unklaren Fällen führen.

¹ Dazu müsste mit den nationalen Einrichtungen, die die einzelnen Meßstationen betreuen bzw. betreut haben, Kontakt aufgenommen werden und eine genaue Liste erbeten werden, die selbst wiederum unkorrekt sein kann.

1. Die Ortskoordinaten Länge, Breite und Höhe sind nicht realistisch (Höhe über 8000m, Breite außerhalb $\pm 90^\circ$, Länge außerhalb $\pm 180^\circ$).
2. Es wird keine Vergleichsstation gefunden, die FAO-Station hat aber eine WMO-Nr., die abweichende Koordinaten in der WZN-Datenbank aufweist.
3. Es wird genau eine Vergleichsstation gefunden, die eine andere Länderkennung besitzt.
4. Es wird genau eine Vergleichsstation gefunden, die eine andere WMO-Nummer hat.
5. Es wird genau eine Vergleichsstation gefunden, die zwar die gleiche Länderkennung und/oder WMO-Nummer hat, deren Name aber nicht auf mindestens die ersten 10 Zeichen übereinstimmt.
6. Es wird genau eine Vergleichsstation gefunden, deren Länderkennung und WMO-Nummer nicht prüfbar sind und deren Name nicht auf voller Länge mit dem der FAO-Station übereinstimmt.
7. Es wird genau eine Vergleichsstation gefunden, jedoch sind weder Name, noch WMO-Nummer, noch Länderkennung prüfbar und die Stationskoordinaten werden nicht genau getroffen.
8. Es werden mehrere Vergleichsstationen gefunden, bei denen jeweils alle vorhandenen Vergleichsparameter übereinstimmen.
9. Es werden mehrere Vergleichsstationen gefunden, jedoch ist bei mindestens einer Vergleichsstation kein Vergleichsparameter vorhanden.
10. Es werden mehrere Vergleichsstationen gefunden, jedoch ist bei mindestens einer der Vergleichsstationen mindestens ein Vergleichsparameter (aber nicht alle) identisch.
11. Es werden mehrere Vergleichsstationen gefunden, die FAO-Station hat aber eine WMO-Nr., die bereits an einem anderen Ort in der WZN-Datenbank existiert.

Um im Vorfeld schon bekannte Fehlerquellen auszuschließen, wurden Stationsnamen nur in Großbuchstaben dargestellt.

Darüber hinaus ist von 144 WMO-Stationen bekannt, dass die Ortskoordinaten nachweislich teilweise falsch sind. Diese wurden zunächst entsprechend der im WZN vorhandenen Listen verändert (143 Stationen: U. Schneider; eine zusätzliche Station: H. Daan (WMO)). In der FAO-Datenbank waren 67 Niederschlagsstationen und 39 Temperaturstationen davon betroffen.

Von den oben angegebenen möglichen unklaren Fällen wurden im FAO-Datensatz die Nummern 1, 7, 8 und 9 sowohl im Fall der Temperaturstationen als auch der Niederschlagsstationen nicht realisiert.

Einen Überblick über die Anzahl der zu Beginn der Untersuchung auftretenden unklaren Fälle verschiedener Art gibt Tabelle 2.

Tabelle 2: Überblick über die zu Beginn auftretenden klaren und unklaren Fälle (unter Berücksichtigung von Groß- und Kleinschreibung).

Stationszahl	Niederschlag		Temperatur	
Im WZN-Lexikon eindeutig erkannt	13530	100%	5996	100%
Dem WZN-Lexikon neu	5675	41.9%	3128	52.5%
Unklare Fälle	4797	35.5%	975	16.3%
Anzahl der gefundenen Vergleichsstationen:				
0	3057	22.6%	1893	31.6%
1	5466	40.4%	1394	23.2%
2	7744	57.2%	4401	73.4%
3	281	2.1%	176	2.9%
4	32	.2%	21	.4%
4	7		4	
Details zu den unklaren Fällen:				
WMO-Nr. existiert an anderem Ort	669	4.9%	419	7.0%
Eine Vergleichsstation mit anderer WMO-Nr.	52	.4%	28	.5%
Eine Vergleichsst. mit anderer Länderkennung	173	1.3%	101	1.7%
Eine Vergl., gleiche WMO-Nr. anderer Name	1843	13.6%	1144	19.1%
Mehrere Vergl. mit teils gleicher Information	320	2.4%	201	3.4%

Im folgenden werden die verschiedenen Schritte im Umgang mit den unterschiedlichen vorkommenden unklaren Fällen besprochen. Dabei ist zu beachten, dass in einem Analyseschritt eine Station nur einer Art von unklarem Fall zugeordnet wird. Das bedeutet insbesondere, dass die Klärung eines unklaren Falls einen anderen unklaren Fall generieren kann.

2.2 Es existiert eine Vergleichsstation mit anderer Länderkennung

Die Länderkennung orientiert sich sowohl in der FAO-Datenbank als auch im WZN am Dreibuchstabencode der WMO. Probleme treten auf, da einerseits durch Grenzverlegungen und Kriege im Laufe der Zeit auch der Ländercode einzelner Stationen einer Veränderung unterworfen sein kann. Andererseits können innerpolitische Spannungen zu einer Änderung von Ländernamen und somit auch Länderkennungen führen. Nicht zuletzt sei der Fall von (Wieder-)Vereinigungen genannt, dem Ländernamen und somit auch Kennungen zum Opfer fallen. Viele unklare Fälle dieser Art sind zumindest semiautomatisch handhab- und lösbar.

Darüber hinaus wird in der WZN-Datenbank feiner unterschieden als es der Dreibuchstabencode der WMO vorgibt. So werden einzelne Gebiete, die vom Hauptgebiet eines Landes geographisch getrennt sind, zum Teil mit einem eigenen Ländercode versehen. Für die USA betrifft dies z.B. Alaska, Hawaii, Puertorico und die Midway Islands, für Portugal sind Madeira und die Azoren betroffen, für Spanien die Kanarischen Inseln.

So wurden die folgenden Länderkennungen ersetzt, neu eingeführt bzw. gestrichen:

- CSK (Czechoslovakia existiert nicht mehr) → CZE, SVK,
- BUR → MMR (Burundi → Myanmar),
- BYS → BLR (Belarus),
- AAA → ATA (Antarktis),
- DDR → DEU (Wiedervereinigtes Deutschland),
- DMK → DNK (Dänemark),
- YMD → YEM (Yemen),
- ZAR → COD (Democratic Republic of Congo),
- SUN → RUS (Russian Federation),
- REU → IOT (Reunion ist in der WZN-Datenbank nicht geführt),
- TMP → IDN (East Timor → Indonesien),
- FLK → SAI (Falkland → South American Islands),
- BIH = Bosnien Herzegowina,
- AZU = Azoren (Portugal),
- CAI = Canarian Islands (Spanien),
- ERI = Eritrea,
- HAW = Hawaii (USA),
- HRV = Croatia (Hrvatska),
- JON = Johnston Islands,
- MAD = Madeira (Portugal),
- MID = Midway Island (USA),
- MKD = Makedonia (former Yugoslav Republic),
- SAI = South American Islands,
- SVN = Slovenia.

Bei einigen Stationen der FAO-Datenbank ist keine Länderkennung gegeben. Falls diese Stationen jedoch durch die restliche vorhandene Information eindeutig zugeordnet werden können, wird die Länderkennung nachgetragen. Auf die gleiche Weise werden auch falsche Länderkennungen identifiziert und durch richtige ersetzt.

In zwei Fällen liegt die FAO-Station unmittelbar auf der einen Seite einer Landesgrenze und die Vergleichsstation der WZN-Datenbank auf der anderen Seite. Hier unterscheiden sich zwar die Länderkennungen, ohne das jedoch eine der Stationen eine falsche Länderkennung hat. Die entsprechenden Stationen der FAO-Datenbank sind

- Pafuri: Die FAO-Station Pafuri liegt in Mozambique während die Vergleichsstation in der WZN-Datenbank richtigerweise in Südafrika liegt.
- Moyale: Die FAO-Station Moyale liegt tatsächlich in Äthiopien, während die Vergleichsstation gleichen Namens in Kenia liegt.

Zwei weitere unklare Fälle treten auf, weil die Kennung der Stationen in Alaska innerhalb der WZN-Datenbank nicht einheitlich ALA statt USA ist. Aus Konsistenzgründen müssen hier die Änderungen in der FAO-Datenbank und der WZN-Datenbank vorgenommen werden. Betroffen sind die Stationen

- Cape Saint Elias und
- Holy Cross.

2.3 Die WMO-Nummer existiert in der WZN-Datenbank an anderem Ort

Die Ortskoordinaten werden in der FAO-Datenbank in Grad und Minuten angegeben. In der WZN-Datenbank sind die Angaben in hundertstel Grad. Die Umrechnung von gerundeten Minuten auf Grad ergibt eine Ungenauigkeit von 1.66 hundertstel Grad, was auf 2 hundertstel Grad gerundet wird. Das Lademodul sucht aber nur nach Vergleichsstationen in einer Entfernung von einem hundertstel Grad je Richtung². Damit wird allein durch die Koordinatenunschärfe in der FAO-Datenbank und deren Umrechnung in Dezimalgrad in vielen Fällen nicht die passende Vergleichsstation gefunden. Dieses generelle Problem der FAO-Datenbank, das auch an anderen Stellen in Erscheinung treten wird, kann im Fall vorhandener WMO-Nummer entschärft werden, indem in der WZN-Datenbank nach der Station gesucht wird, die die gleiche WMO-Nummer hat. Falls diese Vergleichsstation existiert und höchstens 3 hundertstel Grad in jede Richtung entfernt liegt, werden die Koordinaten dieser Vergleichsstation übernommen und der Fall ist geklärt.

Falls die Stationen mit gleicher WMO-Nr weiter entfernt voneinander liegen, wird der Stationskatalog der WMO (WMO Volume A) mit in die Untersuchung einbezogen. Auch hier müssen wieder verschiedene Fälle unterschieden werden:

- Falls die Koordinaten im WMO-Lexikon mit denen in der WZN-Datenbank übereinstimmen und falls die Stationsnamen in allen drei Lexika gleich sind (bis auf unterschiedliche Schreibweise und Tippfehler), werden die Koordinaten der FAO-Datenbank als falsch angenommen und entsprechend der WZN-Datenbank verändert.
- Falls die Koordinaten im WMO-Lexikon nicht mit denen der WZN-Datenbank übereinstimmen, letztere aber auch nicht mit der Liste der bekannten WMO-Stationen mit falschen Ortsangaben, so müssen die Einträge in der WZN-Datenbank entsprechend verändert werden.
- Falls die Station der FAO-Datenbank einen signifikant anderen Namen und andere Koordinaten hat als die beiden anderen Datensätze, wird in der FAO-Datenbank nach Stationen mit ähnlichem Namen und nach Stationen mit ähnlichen Koordinaten gesucht. Auf diese Weise werden Fehler in der WMO-Nr. (Zahlendreher oder Tippfehler) gefunden und eliminiert.
- Falls die WMO-Nr der FAO-Station nicht im aktuellen Stationskatalog der WMO vorkommt, die Namen der FAO-Station und der WZN-Vergleichsstation aber quasi gleich sind und die Koordinaten sich auf eine leicht als Übertragungs- bzw. Aufzeichnungsfehler zu identifizierende Art unterscheiden, werden die Koordinaten der FAO-Station auf jene der WZN-Vergleichsstation gesetzt. Mit einer leicht als Übertragungs- bzw. Aufzeichnungsfehler zu identifizierende Art sind hier Zahlendreher, Fehler in nur genau einer Stelle, ähnliche Ziffern (z.B. 3 statt 8) und Fälle gemeint, bei denen die Grad-Nachkommastellen sich um den Faktor 6/10 unterscheiden.

Auf diese Weise konnten alle unklaren Fälle dieser Art geklärt werden. Die Station Rengat Japuri mit der WMO-Nr. 96171 hat allerdings laut der Korrekturliste des WZN die Koordinaten (-43, 10245). In der WZN-Datenbank sind jedoch die Koordinaten (-33, 10232) eingetragen.

² 0.01° entsprechen 1.111 km in Nord-Südrichtung. In Ost-Westrichtung beträgt dies am Äquator das Gleiche, nimmt aber zu den Polen hin auf 0 ab. In 30° Breite beträgt die Ost-Westentfernung noch 962m, in 50° Breite 714m. In 60° Breite darf der Abstand zwischen der FAO-Station und einer Vergleichsstation in Ost-Westrichtung 556m nicht mehr überschreiten.

2.4 Es existiert eine Vergleichsstation mit einem anderen Namen

Stationsnamen, die im Original in anderen als lateinischen Schriftzeichen formuliert werden, können im allgemeinen auf verschiedene Weise übersetzt werden. Dies führt zu kleinen aber nur schwer automatisch erkennbaren Unterschieden in den verschiedenen Datenbanken. Weitere Unterschiede treten dadurch auf, dass Abkürzungen unterschiedlich verwendet werden (z.B. für Mount, Sankt, Postamt, Flughafen usw.). Solche Abkürzungen finden sich insbesondere in Zusatzinformationen, die den Ort näher spezifizieren und oft notwendig sind, um die Eindeutigkeit zu gewährleisten. Zusätzlich können Ortsnamen in verschiedenen Datenbanken in verschiedenen Sprachen angegeben sein (z.B. *Swan Island* als *Islas del Cisne*). Auch können Orte durch Namen oder Eigenschaften gekennzeichnet sein (z.B. *Dum Dum* oder *Calcutta Airport*). Weitere Unterschiede in den Stationsnamen können durch unterschiedlich explizite nähere Beschreibungen, durch die Verwendung von Klammern, Kommas und Bindestrichen (z.B. *Calcutta (Airport)*, *Calcutta, Airport* oder *Calcutta – Airport*) sowie durch unterschiedliche Reihenfolge der näheren Beschreibung (z.B. *Airport Calcutta* oder *Calcutta Airport*) entstehen.

Aus all diesen Gründen treten zahlreiche Stationen auf, für die zwar die WMO-Nummer, die Länderkennung und die Koordinaten in der FAO-Datenbank identisch sind mit jenen aus der WZN-Datenbank, nicht aber der Name. In diesen Fällen, in denen außer den Namen alle Informationen vorhanden und eindeutig sind, wird der Stationsname in der FAO-Datenbank durch den der WZN-Datenbank ersetzt.

Falls keine WMO-Nummer als Vergleichsparameter vorliegt, werden die Stationsnamen der FAO-Stationen visuell mit den Namen der gefundenen Vergleichsstationen der WZN-Datenbank verglichen. Damit können alle Namensunterschiede, die auf den oben angegebenen Gründen beruhen, manuell eliminiert werden. Auch fehlende Länderkennungen werden in diesem Schritt übernommen.

In einigen Fällen hat die FAO-Station eine WMO-Nr., die Vergleichsstation der WZN-Datenbank hingegen nicht. In diesen Fällen wurden die FAO-Stationen anhand des WMO Volume A überprüft. Sie werden aber weiterhin vom Lademodul als unklare Fälle angesehen.

Da die Unschärfe in den Ortskoordinaten der FAO-Stationen größer ist als der Suchradius für Vergleichsstationen, wird bei den verbleibenden Stationen in einem größeren Umkreis von $.03^\circ$ nach Vergleichsstationen mit gleichem Namen gesucht. Dadurch konnten einige weitere Fälle geklärt werden.

Die nach dieser Vorgehensweise übrig bleibenden Stationen haben einen Namen, der nicht zum Namen der Vergleichsstation passt und keine WMO-Nummer, die zur Klärung beitragen könnte.

2.5 Die Vergleichsstation hat eine andere WMO-Nummer

Es gibt verschiedene Gründe, weshalb die WMO-Nr. einer FAO-Station sich von der WMO-Nr. der Vergleichsstation der WZN-Datenbank unterscheiden kann. In manchen Fällen hat sich die WMO-Nr. im Laufe der Zeit geändert, ohne dass dies in der FAO-Datenbank nachgetragen wurde. In diesem Fall können die Änderungen einfach von Hand vorgenommen werden.

Für die weiteren Untersuchungen wird erneut das WMO Volume A zu Hilfe genommen.

Falls die WMO-Nr. der FAO-Station in keiner der anderen Stationslexika (WZN-Datenbank oder WMO Volume A) vorkommt und falls der Name der FAO-Station gleich dem Namen der Vergleichsstation ist, wird die WMO-Nr. der FAO-Station als falsch angesehen und durch die WMO-Nr. der Vergleichsstation in der WZN-Datenbank ersetzt.

Kommt die WMO-Nr. der FAO-Station in einer der anderen Lexika unter ähnlichem Namen vor, so wird unter Berücksichtigung der Unschärfe in den horizontalen Koordinaten und unter Zuhilfenahme der Stationshöhe als weiteres Kriterium, eine eindeutige Zuordnung erzielt. Dazu werden entweder die Koordinaten im Rahmen der Unschärfe verändert, oder aber die als falsch erkannte WMO-Nr.

Falls die WMO-Nr. der FAO-Station in einer der beiden anderen Stationslexika an einem anderen Ort vorkommt und falls der Name der FAO-Station dem der Vergleichsstation gleicht, wird die WMO-Nr. der FAO-Station als falsch angenommen und durch die WMO-Nr. des anderen Stationslexikons ersetzt.

Kommt die WMO-Nr. der FAO-Station im WMO-Stationslexikon vor, die WMO-Nr. der Vergleichsstation aus der WZN-Datenbank hingegen nicht, und ähneln sich darüber hinaus die Stationsnamen, werden die WMO-Nummern in der WZN-Datenbank angezweifelt (es könnte sich z.B. um veraltete Kennungen handeln). In diesem Fall werden die WMO-Nr. der Stationen der FAO-Datenbank beibehalten und es wird empfohlen, jene aus der WZN-Datenbank daran anzupassen. Dies tritt bei folgenden 3 Stationen auf:

- 41940 → 41977 (Chittagong (Ambagan)),
- 84633 → 84630 (Huayao),
- 41414 → 41367 (Thamud).

2.6 Es existieren mehrere Vergleichsstationen mit teils passender Information

Da in diesen Fällen unter der Annahme richtig vorgegebener Koordinaten während des Ladeversuchs in die Datenbank unvermeidbar mehrere Vergleichsstationen gefunden werden, können diese Fälle während des Ladens nicht unterbunden werden. Jedoch zeigt ein Blick auf die Liste der betroffenen Stationen und deren Vergleichsstationen, dass unter Zuhilfenahme der Stationshöhe und Berücksichtigung der Unschärfe in der Namensschreibweise und der Ortskoordinaten diese Fälle während des Ladens geklärt werden können.

2.7 Die FAO-Station wird als neue Station erkannt

Hier muss zunächst geprüft werden, ob die Unschärfe in den Ortskoordinaten fälschlicher Weise zu diesen neuen Stationen geführt hat.

Dies geschieht indem nach in der Nähe liegenden Stationen (Suchumgebung $\pm 0.03^\circ$) mit vergleichbarem Namen und gleicher Höhe gesucht wird. Wird auf diese Weise eine eindeutige Vergleichsstation gefunden, so werden deren Koordinaten zur Klärung des Falles übernommen.

In einem weiteren Schritt wird bei den verbleibenden als neu eingestuft Stationen geprüft, ob ihre Koordinaten einen Ort markieren der innerhalb des durch die Länderkennung vorgegebenen Landes liegt.

2.8 Abschliessende Überprüfung aller unklaren Fälle und neuen Stationen

Unklare Fälle oder neue Stationen im Sinne der WZN-Datenbank können dadurch auftreten, dass die Koordinaten der FAO-Stationen teilweise falsch sind. Falls die einzulesenden Stationen keine WMO-Nr. haben, kann das Lademodul des WZN diesem Problem nicht begegnen. Hier wird ihm aber eine große Bedeutung beigemessen. Daher wird ihm mit zwei Suchstrategien begegnet.

Einerseits wird für jede dieser Stationen nach Vergleichsstationen in der WZN-Datenbank gesucht, die im gleichen Land liegen und exakt den gleichen Namen haben. Andererseits wird nach Vergleichsstationen gesucht, bei denen mindestens eine der Ortskoordinaten nur um $\pm 0.03^\circ$ von den Koordinaten der FAO-Station abweicht und die Länderkennung gleich ist. Als Ergebnis dieser Suche werden 22429 Stationspaare gefunden und visuell verglichen. Falls sowohl Name als auch Land des Stationspaars identisch sind und zusätzlich die Koordinaten ähnlich bis auf Kommafehler, Vorzeichen, kleine Ungenauigkeiten, Zahlendreher und Fehler in nur einer Stelle und falls die Höhenangaben vergleichbar sind, wird die FAO-Station der entsprechenden WZN-Station zugeordnet. Falls für das Stationspaar eine der Koordinaten bis auf $.03^\circ$ genau übereinstimmt, müssen der Namen ähnlich sein (bis auf Schreibfehler, unterschiedliche Sprache bzw. Abkürzungen und Schreibweisen), die Höhe muss vergleichbar sein und die zweite Koordinate muss ähnlich sein im oben angegebenen Sinn.

Mit dieser sehr aufwendigen Prozedur konnten 864 zunächst als unklar bzw. neu eingestufte Niederschlagsstationen und 208 ebenso eingestufte Temperaturstationen eindeutig identifiziert werden.

3. Stand der FAO-Daten nach Kontrolle der Metainformationen

Durch die in den vorigen Abschnitten besprochenen Methoden konnten zahlreiche unklare Fälle behoben werden. Jedoch bleiben aus den verschiedensten Gründen noch unklare Fälle übrig. Einen Überblick über die am Ende der Untersuchung auftretenden klaren und unklaren Fälle gibt Tabelle 3. Die FAO-Station, deren WMO-Nr. in der WZN-Datenbank an einem anderen Ort vorkommt, hat gemäß der Liste von Udo Schneider in der WZN-Datenbank falsche Ortskoordinaten. Bei den drei Niederschlagsstationen und der einen Temperaturstation, bei der zwar eine Vergleichsstation gefunden wird, diese aber eine andere WMO-Nr. hat, sind die WMO-Nummern der FAO-Stationen identisch mit Vergleichsstationen aus dem WMO Volume A. Daher kann es sinnvoll sein die WMO-Nr. der WZN-Datenbank daran anzupassen. Von den 4 Niederschlagsstationen mit unterschiedlicher Länderkennung liegen 2 tatsächlich jenseits der Grenze. In den anderen beiden Fällen handelt es sich um Stationen in Alaska, bei denen in der WZN-Datenbank der Ländercode USA statt ALA angegeben ist. Diese beiden Stationen treten auch bei der Temperatur auf. Die 60+4 Fälle, bei denen eine Vergleichsstation gefunden wird, die aber einen anderen Namen aufweist, kann ad hoc keine Entscheidung gefällt werden. Bei den 440+235 Stationen mit mehreren Nachbarn, kann während des Ladens die oft eindeutig passende Station selektiert werden.

Tabelle 3: Überblick über die am Ende der Untersuchungen auftretenden klaren und unklaren Fälle.

Stationszahl	Niederschlag		Temperatur	
Im WZN-Lexikon eindeutig erkannt	10544	77.9%	5416	90.4%
Dem WZN-Lexikon neu	2478	18.3%	338	5.6%
Unklare Fälle	508	3.7%	242	4%
Anzahl der gefundenen Vergleichsstationen:				
0	2479	18.3%	338	5.6%
1	10611	78.4%	5423	90.4%
2	391	2.9%	210	3.5%
3	41		21	
4	8		4	
Details zu den unklaren Fällen:				
WMO-Nr. existiert an anderem Ort	1		0	
Eine Vergleichsstation mit anderer WMO-Nr.	3		1	
Eine Vergleichsst. Mit anderer Länderkennung	4		2	
Eine Vergl., gleiche WMO-Nr., anderer Name	0		0	
Eine Vergl., keine WMO-Nr., anderer Name	60		4	
Mehrere Vergl. Mit teils gleicher Information	440	3.3%	235	3.9%