

Endbericht des internen Projektes

Climate Data Rescue 2011

Projektantrag: 29. Juni 2010

Projektleitung: Mag. Anita JURKOVIĆ

Projektteam: Mag. Silke ADLER, Joachim ALBENBERGER, Hermann GALAVICS,
Dr. Wolfgang LIPA, Wolfgang LECHNER, Roland SWIETLI, Iona VOSSBERG, Silvia
ZEHEITNER

Laufzeit des Projektes: 12 Monate

Meteorologische Beobachtungen.

Beobachtungsstation: *Salzburg*
Beobachter: *...*

Stationen-Nr. *...*
Wahrungs-Nr. *...*

Wahrungs-Nr.	Beobachtungen												Wahrungs-Nr.	
	Temperatur			Luftfeuchtigkeit			Wind			Niederschlag				
Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.	Wahrungs-Nr.
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

Handwritten notes and corrections are visible throughout the table, particularly in the right-hand column and at the bottom of the page.

Wien, 30. Dezember 2011

Vorgängerprojekte:

Austrian Daily Climate Data Rescue 1872 – 1983 I
Erstantrag, 23.März 2007
Zwischenbericht: 15.Mai 2008
Endbericht des ersten Projektjahres: 19. Dezember 2008
Projektleitung: Dr. Wolfgang LIPA

Austrian Daily Climate Data Rescue 1872 – 1983 II
Projektantrag: 15.05.2008
2. Zwischenbericht 31.5.2009
Endbericht des zweiten Projektjahres: 28. Dezember 2009
Projektleitung: Dr. Wolfgang LIPA

Climate Data Rescue 2010
Projektantrag: 10.Juni 2009
Endbericht des dritten Projektjahres: 20. Dezember 2010
Projektleitung: Mag. Anita JURKOVIĆ

In Folge als Strukturprojekt CLIMATE DATA RESCUE genehmigt

Finanziert wird das Projekt intern durch das
Bundesministerium für Wissenschaft und Forschung.



Bundesministerium für Wissenschaft und Forschung

Eine kurze Projektbeschreibung ist auch auf der Homepage unseres Institutes unter http://zamg.ac.at/wir_ueber_uns/messnetz/datenpflege/ zu finden.

Inhalt

1. Motivation	4
2. Projektziele.....	5
3. Projektverlauf	6
4. Überblick.....	8
5. Archivierung.....	9
5.1. Fotografische Archivierung	9
5.2. Klimabogendatenbank	9
6. Digitalisierung, Prüfung und Korrektur	10
6.1. Datenbestand.....	10
6.2. Digitalisierung (WP1).....	11
6.3. Prüfung und Korrektur (WP2, WP3)	11
7. Ergebnisse	12
7.1. Datenbestand und -qualität	12
7.2. Meilensteine.....	13
8. Zusammenfassung und Ausblick	14
8.1. Fazit	14
8.2. Ausblick.....	14

1. Motivation

Das Ziel von Climate Data Rescue Projekten stellt die Digitalisierung, Prüfung und Korrektur von historischen Klimaaufzeichnungen (Papierdaten), welche in meist längst vergessenen geglaubten Archiven gelagert sind, dar. Solche Projekte werden auf nationaler (Austrian Climate Data Rescue bzw. das Schweizer DigiHom Projekt) sowie internationaler Ebene (bspw. WMO Projekt MEDARE Initiative Composition WG3: Developing best practices on quality control and homogenization of climate data) stark gefördert.

Durch die Unterstützung des Bundesministeriums für Wissenschaft und Forschung und den zusätzlich zur Verfügung stehenden Geldmitteln war 2007 erstmals die Möglichkeit gegeben, im Rahmen des internen Projektes Austrian Daily Climate Data Rescue 1872 – 1983 I samt Nachfolgeprojekten Austrian Daily Climate Data Rescue 1872 – 1983 II und Climate Data Rescue 2010, mit der Nacherfassung der historisch sehr wertvollen Datensätze zu beginnen.

Genau hier setzt das Projekt Climate Data Rescue 2011 fort. Mit der Digitalisierung von Klimabögen und der Nachbearbeitung bzw. Endkontrolle historischer sowie aktueller Datensätze sollen alle Klimadaten - welche die Basis für nahezu alle klima- und umweltrelevanten Projekte darstellen - in hoher zeitlicher Auflösung, zumindest auf Tagesbasis, digital verfügbar und auf denselben einheitlichen Qualitätsstandard gebracht werden. Die geprüften Datensätze liefern u.a. die Grundlage zur Kartenerstellung und statistischen Auswertung von ausgewählten Zeitreihen (Projekt Kärntner Klimaatlas, 3PClim).

2. Projektziele

- Nachbearbeitung aller langjährigen Zeitreihen der Periode 1872 – 1983, d.h. ca. 75 – 80 % aller verfügbaren Zeitreihen. Die ca. 20 – 25 % der nicht erfassten Kurzzeitreihen bleiben in Papierform und als fotografierte Kopien im Datenarchiv
- Bereitstellung aller wesentlichen ZAMG-Tagesdaten mit bestmöglicher und einheitlicher Qualität. Erst mit diesem Datensatz können Fragestellungen wie Dauer von Hitze-, Kälte-, Starkniederschlagsperioden hoch auflösend untersucht und beantwortet werden. Dieses Projekt ist somit das Vorprojekt für alle jene wissenschaftlichen Projekte, welche nicht mit Monatswerten das Auslangen finden, sondern auf Grundlage von Tagesdaten arbeiten müssen
- Weitergabe von bisherigen Erfahrungswerten in Bereichen der Datennacherfassung, Prüfung sowie Korrektur von Tagesdatensätzen
- Fortentwicklung bestehender Prüfroutinen und Algorithmen
- enge Zusammenarbeit mit der Regionalstelle Tirol und der Abteilung Klimaforschung im Rahmen des INTERREG Projektes 3PClim
- Kompetenzsteigerung der ZAMG im Sektor Qualitätsprüfung auf nationaler sowie internationaler Ebene

3. Projektverlauf

Nach einer 50% Kürzung der beantragten Summe, musste der Projektplan im Dezember 2010 neu ausgearbeitet bzw. angepasst werden. Dabei wurden die Projektziele – adäquat zur budgetären Kürzung – neu definiert (vgl. *Version 2*). Als Folge waren personelle Restrukturierungen bzw. Einsparungen und eine Vergabe von Werkverträgen außer Haus notwendig. Die Anzahl der Arbeitspakete konnte beibehalten werden. Die Inhalte und Ziele der Arbeitspakete sind im Folgenden aufgelistet.

- **WP1: Evaluierung der Datenqualität der noch nicht digitalisierten Papierdaten sowie deren Nacherfassung**

Projektziel für 2011:

- *Version 1*: 150 Stationsjahre
- *Version 2*: 75 Stationsjahre

- **WP2: Prüfung der digitalisierten, aber noch nicht endgeprüften Tagesdaten**

Projektziel für 2011:

- *Version 1*: 300 Stationsjahre
- *Version 2*: 150 Stationsjahre

- **WP3: regionale Prüfung der Datensätze mittels ProClim.db**

Projektziel für 2011:

- *Version 1*: Prüfung und Korrektur der Tiroler Messstellen
- *Version 2*: Abschlussarbeiten der Region Süd (Kärntner Stationen).
Nach Abschluss der kompletten Nacherfassung und Endprüfung des Bereiches West (Bundesland Tirol und Umgebung) Start der regionalen Prüfung.

Das Projektmonitoring wurde von der Projektleitung und dem Controller durchgeführt. Mittels monatlicher Übermittlung eines Projektcontrolling – Excelfiles wurden Projektablauf und Einsatz von Ressourcen (Stichwort Kostenstellenrechnung) dokumentiert. Projektablauf und -status sind Abb. 1 zu entnehmen.

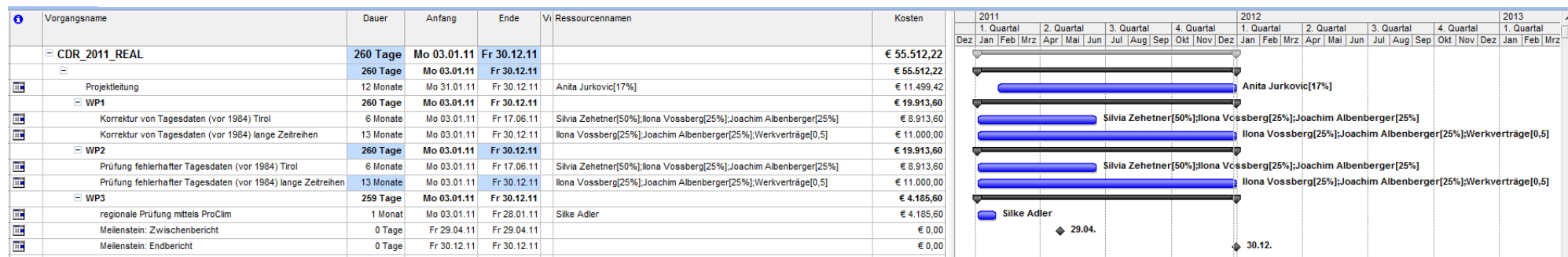


Abb. 1: Zeitliche Staffelung der Arbeitspakete (WP). Blaue Balken kennzeichnen die Dauer des Arbeitspakets. Die benötigten Ressourcen sind neben dem Balken aufgelistet. Meilensteine sind als Punkte dargestellt.

4. Überblick

Das Messnetz der ZAMG umfasst derzeit

- 246 TAWES*
- 9 Ö3*
- 6 Mini Taklis*
- 4 ACG Stationen*

An 193 Standorten finden neben automatischen Messungen auch Zusatzbeobachtungen statt. Diese Messstellen werden als Klimastationen bezeichnet und von einem Beobachter betreut, der zu den Klimaterminen Handmessungen von Niederschlag /Schnee/Neuschnee bzw. Erdbodenzustandsklassifikationen und Sichtweitenbestimmungen vornimmt.

Der beobachtete Wert wird im Regelfall ins Klimatagebuch und auf einem Klimabogen eingetragen und schlussendlich als Prüf- oder Zusatzwert in die Datenbank eingespielt. Der Weg vom beobachteten Wert bis zum Einspielen in die Klimadatenbank unterliegt vielen kleinen Einzelschritten:

- (a) Wert wird vom Beobachter im Klimatagebuch und im Klimabogen verzeichnet
- (b) Übermittlung des Klimabogens an die zuständige Dienststelle der ZAMG (jeweils am Monatsersten)
- (c) Bei Ankunft Eintrag ins Postbuch
- (d) Digitale Sicherung des Klimabogens. An den Aussendienststellen erfolgt ein Scan und die Übermittlung als pdf an die ZAMG Wien.
- (e) An der ZAMG Wien werden die Klimabögen fotografiert, gesammelt und mit den konvertierten pdf's der Regionalstellen auf einen Server gestellt.
- (f) Plausibilitätsprüfung und Korrektur der Werte und Kontakt zu den Beobachtern
- (g) Verschlüsselte Eingabe des Klimabogens in die Datenbank
- (h) Archivierung ins Klimabogenarchiv

Eine wesentliche Vereinfachung soll durch die Einführung einer automatischen Übermittlung stattfinden. Durch das web-basierte Klima Eingabe System (KSE) sollen die Punkte a bis e und g vollständig entfallen und die Daten zeitnah für die Prüfung zur Verfügung gestellt werden. Bis alle Beobachter auf dieses neue System, welches derzeit im Probetrieb läuft, umgestiegen sind, muss das Archiv laufend betreut und die aktuellen Klimabögen inventarisiert werden.

*Die Definitionen der Stationsarten sind auf unserer Homepage, unter http://www.zamg.ac.at/wir_ueber_uns/messnetz/, zu finden.

5. Archivierung

Die Grundlage jedes CDR Projektes stellt das Papierdatenarchiv dar. Erst das Wissen über den tatsächlichen Archivdatenbestand und den Qualitätsstandart historischer Aufzeichnungen konkretisiert das Ziel eines solch umfassenden Projektes und führt zu groben Abschätzungen bezüglich Projektdauer bzw. -aufwand.

Das Archiv der ZAMG umfasst u.a. historische, bis zu 150 Jahre alte, Klimabögen aus dem In- sowie Ausland und sichert somit Kulturgut, welches als besonders schützenswert angesehen werden muss (WMO-Richtlinie). Im Zuge der Inventarisierung des ZAMG Klimabogenarchives konnten viele in- und ausländische Datens(ch)ätze neu gefunden werden.

Nach Abschluss der Komplettinventarisierung des Klimaarchives im Dezember 2010 stellt die Wartung der Klimabogendatenbank sowie die routinemäßige Archivierung aktuell übermittelter Klimabögen die zentrale Aufgabe dar. Derzeit werden rund 223.000 Klimabögen im Archiv gelagert.

5.1. Fotografische Archivierung

Wie bereits im Überblick gebracht werden alle Klimabögen fotografiert und im pdf-Format auf einem intern zugänglichen Server gespeichert. 2011 wurden insgesamt 6831 Fotos konvertiert und nacherfasst.

5.2. Klimabogendatenbank

Herzstück der Archivierung ist die Klimabogendatenbank. In dieser sind alle Lagerplätze der historischen und aktuellen Klimabögen verzeichnet und zentral gespeichert. Dadurch kann neu eintreffenden Klimabögen automatisch ein neuer Lagerplatz zugewiesen werden. Die Bedienung und Eingabe erfolgt über einen selbstentwickelten Editor auf ACCESS Basis und liefert einen Überblick über folgende Punkte:

- pdf-Vorschau
- Bestandsliste
- Art des Klimabogens
- Bearbeitungsstatus
- Entlehnungsstatus

Der Umgang mit der Softwareapplikation wurde im Rahmen einer internen Fortbildung Anfang 2011 nähergebracht.

Schulung „Einführung in ACCESS Datenbanken- Die Klimabogendatenbank“
Teilnehmer: Albenberger, Lang, Lechner, Lipa, Manhard, Vossberg, Zehetner
Datum, Ort: 20.,21.04.2011; Lesesaal Kreil-Haus

6. Digitalisierung, Prüfung und Korrektur

6.1. Datenbestand

Zurzeit umfasst die Klimadatenbank, bestehend aus Einminuten-, Zehnminuten-, Stunden-, Tages-, Monatswerten und Phänologiedatensätzen, 450 GB an Massenspeicher. Umgerechnet entspricht dies 56250 Magnetplatten, welche in den 80er Jahren als Speichermedium verwendet wurden, oder 5.6 Milliarden Lochkarten (Datenerfassung in den 70er Jahren). Würde man die Lochkarten aneinanderreihen, ergäbe dies eine Länge von 1125 km, was in etwa der Luftlinie Wien – Paris entspricht.

Die Tagesdaten – Datengrundlage dieses Projektes – werden in der tag-Tabelle auf dem Langzeitserver SYBKLIM gespeichert. Jeder Tagesdatensatz wird mit einem Qualitätsflag versehen, wobei die Klassifizierung typisiert erfolgt.

typ 0	Papierdaten
typ 1	ungeprüfte Daten
typ 2	teilweise geprüfte Daten - Lochkartendaten
typ 3-5	vorgeprüfte Daten
typ 6	endgeprüfte Daten

Die langfristige Zielsetzung von Climate Data Rescue ist die schlussendliche Zuordnung aller Datensätze auf typ 6. Dies beinhaltet einerseits die **Digitalisierung** alter Papierdaten (typ 0) und andererseits die **Prüfung und Korrektur** von typ 1 - typ 5 Daten. Zurzeit sind 40% der rund 5 Millionen Tagesdaten als typ 6 gekennzeichnet. Der aktuelle Datenbestand der Tagesdaten ist Abb. 2 zu entnehmen.

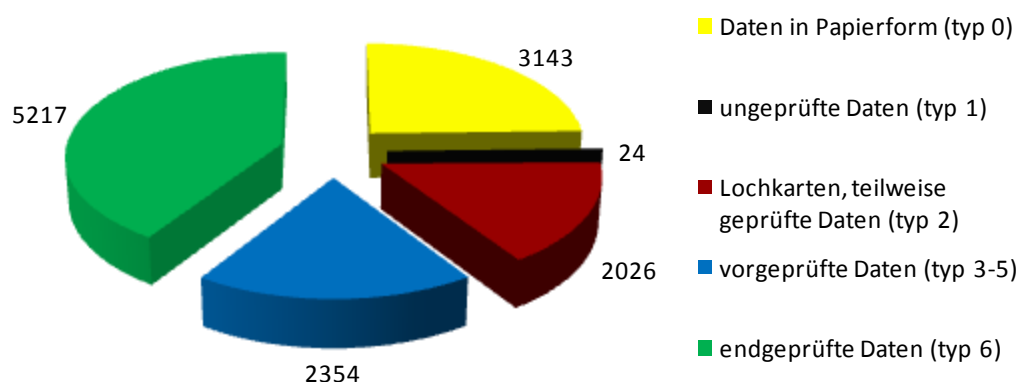


Abb. 2: Datenbestand der Tagesdaten in Stationsjahren. Die Abfrage des historischen Bestandes (Datensätze vor 1984) wurde am 30.12.2011 durchgeführt.

6.2. Digitalisierung (WP1)

Da während des zweiten Weltkrieges nahezu alle Klimabögen nach Berlin gesendet und durch einen Bombenangriff vernichtet wurden, sind zahlreiche historische Aufzeichnungen unwiederbringlich zerstört worden. Die wenigen Duplikate, welche an der Zentralanstalt und deren Regionalstellen noch auffindbar waren, und somit als absolut schützenswertes Kulturgut einzustufen sind, werden nun im Archiv der ZAMG aufbewahrt. Schon in den 70er Jahren wurde ein erster Versuch gestartet die historischen Klimabögen in Form von Lochkarten zu digitalisieren. Zu dieser Zeit erfolgte aber keine Art von Plausibilitäts- oder gar Vollständigkeitsprüfung. Erst ab 1984, mit der Einführung einer allgemein gültigen Prüfroutine, kann man von geprüften Datensätzen sprechen. Durch die Implementierung der kommerziellen Software GEKIS (geographischen Klima-Informationen-System) im Jahre 1997 konnten Stundenwerte im Routinebetrieb geprüft werden. Die letzte Erneuerung und Verbesserung der Prüfmethodik stellte 2006 die von Ernst Basler und Partner entwickelte Software QualiMet dar. Diese erfasst Fehler in den Ein- bzw. Zehnminutenwerten der TAWES Messstellen.

Ein wesentlicher und essentieller Bestandteil dieses Projektes – und somit auch als eigenes Arbeitspaket geführt – war die Nacherfassung historischer Klimabögen. In Hinblick auf die weitere Zusammenarbeit mit der Außenstelle Innsbruck im Rahmen des INTERREG Projektes 3PClim lag der Fokus auf der Digitalisierung des Bereiches Tirol.

6.3. Prüfung und Korrektur (WP2, WP3)

Für die Prüfung und Korrektur der historischen Tagesdaten wird als Basis das seit 1984 bestehende und stetig modernisierte BULL System herangezogen. Die Software prüft die innere Konsistenz der Zeitreihen, den Wertebereich ausgewählter Klimaparameter und die Plausibilität bei stark korrelierten Klimaelementen, wie bspw. Sonnenscheindauer und Niederschlag.

Nach Abschluss der Qualitätsprüfung der Tagesdaten (Setzung auf typ 6) erfolgt, so wie derzeit für den bestehenden Datenprozess der online Prüfung, als zweiter Schritt eine endgültige Qualitätsüberprüfung auf Basis von räumlichen und statistischen Testverfahren. Hierfür werden selbstentwickelte Prüfapplikationen (Schnee2010) bzw. bewährte Softwareprogramme (ProClim.db) verwendet.

7. Ergebnisse

7.1. Datenbestand und -qualität

Tabelle 1 zeigt den Prozess der Datendigitalisierung und Prüfung seit dem Beginn der Projektinitiative im Dezember 2007. Die Variation der Gesamtanzahl an Datensätzen (Spalte TOTAL) bzw. die Zunahme der Anzahl an typ 0 bis typ 3 Datensätzen ist durch die stetige Bearbeitung der tag-Tabelle begründbar. Teils mussten synthetisch generierte und nicht verwertbare Datensätze bereinigt (vgl. dazu Abbildung 3 – das Jahr 2010) bzw. Werte von anderen Quellen hinzugefügt werden.

PROJEKT	STATUS	typ 0	typ 1	typ 2	typ 3-5	typ 6	TOTAL
BASIS	31.12.2007	1898	-	2032	1142	4070	9142
ACDR 1872-1983 I	31.12.2008	1550	-	1889	1733	3970	9142
ACDR 1872-1983 II	31.12.2009	1473	-	1624	1633	4412	9142
CDR 2010	31.12.2010	1293	-	1458	1368	4906	9025
PLANZIEL CDR 2011		1143	-	1408	1218	5256	9025
CDR 2011	31.12.2011	3143	24	2026	2354	5217	12764

Tab. 1: Anzahl der Tagesdaten in Stationsjahren aller Datenzustände zu Projektende aller CDR Initiativen.

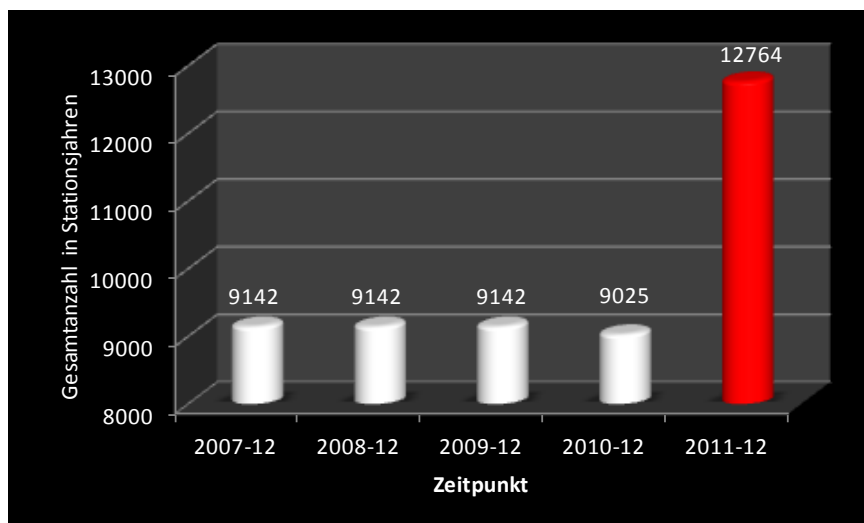


Abb. 3: Gesamtanzahl der zur Verfügung stehenden Tagesdaten (typ0-typ6) seit Projektbeginn von Austrian Climate Data Rescue 1872-1983 I im Jahr 2007.

Der sprunghafte Anstieg des Datenvolumens im Jahr 2011 ist auf die automatisierte Abfrage des Inventarisierungsstatus (Neuerfassung der Anzahl an Papierdaten, bis dato nur geschätzt) bzw. das nachträgliche Einspielen von zusätzlichen Tagesdaten zurückzuführen (bspw. „HYDRO“ Datensätze).

Die als typ3 klassifizierten „HYDRO“ Datensätze konnten im Zuge des Projektes FORALPS am HZB (Hydrographisches Zentralbüro) digitalisiert werden. Dabei handelte es sich Großteils um ZAMG/HD Parallel-Aufzeichnungen von ein und derselben Messstelle. Entgegen unserer Vorgehensweise, die Bögen im zweiten Weltkrieg nach Berlin zu übermitteln, verwahrte der hydrographische Dienst die Aufzeichnungen im Archiv.

Aufgrund dieser Tatsache nahm der absolute Prozentsatz an geprüften Datensätzen von ursprünglich 55% (Datenbestand 2010; siehe Endbericht CDR 2010) auf 40% ab.

Die konkreten Planziele des Projektantrages – *Version 1* (vgl. Tabelle 1, blau gekennzeichnet)

- Digitalisierung von 150 Stationsjahren (typ 0)
- Bearbeitung von 18.263 Lochkartendatensätzen (typ 2)
- Prüfung und Korrektur von 54.788 vorgeprüften Tagesdaten (typ 3-5)
- Endprüfung (typ 6) von 127.838 Tagesdatensätzen (350 Stationsjahren)

konnten aufgrund von budgetären Kürzungen und Anstieg des zu bearbeiteten Datenvolumens nicht erreicht werden. Dennoch war es mit den zur Verfügung stehenden Ressourcen möglich 311 der ursprünglich geplanten 350 Stationsjahre auf typ 6 Status zu bringen, womit das neu formulierte Planziel (*Version 2*) von 175 kontrollierten Stationsjahren sogar weit übertroffen werden konnte.

7.2. Meilensteine

Meilenstein 1: Archivierung und fotografische Sicherung von Klimabögen

Komplettinventarisierung des Klimabogenarchives abgeschlossen. Laufende Nacherfassung und Archivierung von aktuellen Klimabögen. Status: 2010 alle Bereiche fertig.

Meilenstein 2: Softwareentwicklung - Datenbestand Archiv

Durch die von Hr. Roland Swietli entwickelte ACCESS Applikation konnte eine wesentliche Verbesserung der Übersicht des tatsächlichen Datenbestandes erzielt werden. Alle Lagerplätze und digitalen Sicherungen sind mittels Abfrageeditor abrufbar. Im Rahmen einer internen Fortbildung Anfang 2011 konnte den Mitarbeitern die praktische Anwendung im Routinedienst nähergebracht werden.

Meilenstein 3: Digitalisierung historischer Aufzeichnungen der Region West

Meilenstein 4: Prüfung und Korrektur von 311 Stationsjahren

Meilenstein 5: Import von zusätzlichen Tagesdatensätzen

Erhöhung des Datenvolumens in der tag_Tabelle. Bei den nachträglich eingespielten Zeitreihen handelt es sich überwiegend um historische langjährige Datensätze (1895- 1936), welche in Folge für Klimaauswertungen und –modellierungen herangezogen werden.

8. Zusammenfassung und Ausblick

8.1. Fazit

Da erst rund 40% der gemessenen und beobachteten Tagesdaten des Zeitraums 1872 – 1983 in digitalisierter und endgeprüfter Form vorliegen, wurden im Zuge der Fortsetzung von vorgängigen Projekten die verbleibenden Tagesdaten in verstärktem Maße nacherfasst und auf typ 6 (endgeprüfter Zustand) gebracht.

Infolge der Interaktion mit dem INTERREG Projekt 3PClim (Tiroler Klimatografie) lag das Hauptaugenmerk auf der Digitalisierung und Bearbeitung von Tagesdaten der Region West.

Budgetäre Kürzungen erforderten u.a. eine Neuformulierung der Projektziele und Umschichtungen bzw. Auslagerungen von ganzen Aufgabenbereichen. Dieser Umstand, verbunden mit dem unvorhersehbarem Anstieg des Datenvolumens, welcher auf den Import neuer Datensätze zurückzuführen war, führte zu einer knappen Verfehlung der anfänglich definierten Projektziele. Nach einer adäquaten Anpassung an die budgetäre Kürzung von 50% konnte das neu formulierte Planziel von 175 kontrollierten Stationsjahren weit übertroffen werden. Letztendlich konnten 311 Stationsjahre geprüft werden.

8.2. Ausblick

- Fortführung der Digitalisierung und Endprüfung aller vorhandenen Tiroler Tagesdaten
- Weitere Nacherfassung von historischen Datensätzen nach Prioritätsstufen im Zuge des Strukturprojektes Climate Data Rescue
- Weiterentwicklung und Implementierung von zusätzlichen Softwareapplikationen
- Erweiterung und Verbesserung der Prüfsoftware in Zusammenarbeit mit anderen Abteilungen
- Zusammenarbeit mit anderen Institutionen wie der Universität Wien