Grant agreement no. 243964

QWeCI

Quantifying Weather and Climate Impacts on Health in Developing Countries

D1.2a: Database with gridded climate and remote sensed data and observed meteorological data

Start date of project: 1st February 2010 Duration: 42 months

Lead contractor: UoC
Coordinator of deliverable: Prof. Dr. Andreas H. Fink
Dr. Volker Ermert

Evolution of deliverable

Due date :	M15
Date of first draft :	M15
Start of review :	M16
Deliverable accepted :	M18

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	PU
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Introduction

The QWeCI project aims to identify relationships between atmospheric variables and the occurrence of diseases like malaria and the Rift Valley Fever. This deliverable reports on the construction of an atmospheric database, which will provide access to climate-driver variables for various work packages of the QWeCI project. Key variables of the investigated diseases including rainfall amounts, temperature, and humidity are entered into the database.

The status at the ground level of the Earth's atmosphere is usually represented by automatic and manual observations. These observations are carried out regularly, for example, every three hours or even every hour by weather stations (usually called a synoptic station) typically measuring a number of characteristics of the atmosphere. These include observations from instruments for example thermometers or barometers as well as visual observations of clouds, the visibility, the state of the ground, and the present and past weather conditions. According to the **W**orld **M**eteorological **O**rganization (**WMO**) Resolution 40 (Cg-XII) of the World Weather Watch Program atmospheric observations are exchanged worldwide via the **G**lobal **T**elecommunication **S**ystem (**GTS**) by means of predefined reports. These reports include, for example, 3-hourly SYNOP (code format 12 of the WMO) and monthly CLIMAT (code format 71) messages, and hourly METAR (code format 15) reports from airports.

Besides surface station datasets, quality-controlled satellite observations are available since the late 1970s. Satellites are, for instance, used to quantify the distribution of precipitation around the globe over many years. Such products are based on mircrowave or infrared data, which are calibrated by using precipitation amounts from gauges.

Atmospheric re-analyses are also added to the database. Re-analyses improve operational analyses and provide a consistent analysis of the atmosphere. Such systems are based on data assimilation techniques using a frozen version of an atmospheric model and use as much as possible original observations. For example, in addition to standard observations, the **E**uropean **C**entre for **M**edium-range **W**eather **F**orecasts (**ECMWF**) **40** Year **R**e-**A**nalysis (**ERA-40**) product made use of data from past field experiments such as the 1974 **A**tlantic **T**ropical **E**xperiment of the **G**lobal Atmospheric Research Program (**GATE**). The **E**CMWF **Interim R**e-**A**nalysis (**ERA-Interim**) is the latest global atmospheric reanalysis produced by the ECMWF.

Structure of the atmospheric database

The atmospheric database of the QWeCI project can be entered via a web portal (see http://qweci.uni-koeln.de; Figure 1), which is integrated into the web-based Java framework of work package 5.1. In the first step, short profiles are presented of different datasets, which reveals information in terms of the name and the originator of the data, the data content, as well as regarding the covered period (Figure 2). Subsequently, the user can either enter particular metadata, visit the original provider of the data, or is able to directly download data files (Figures 3-5).

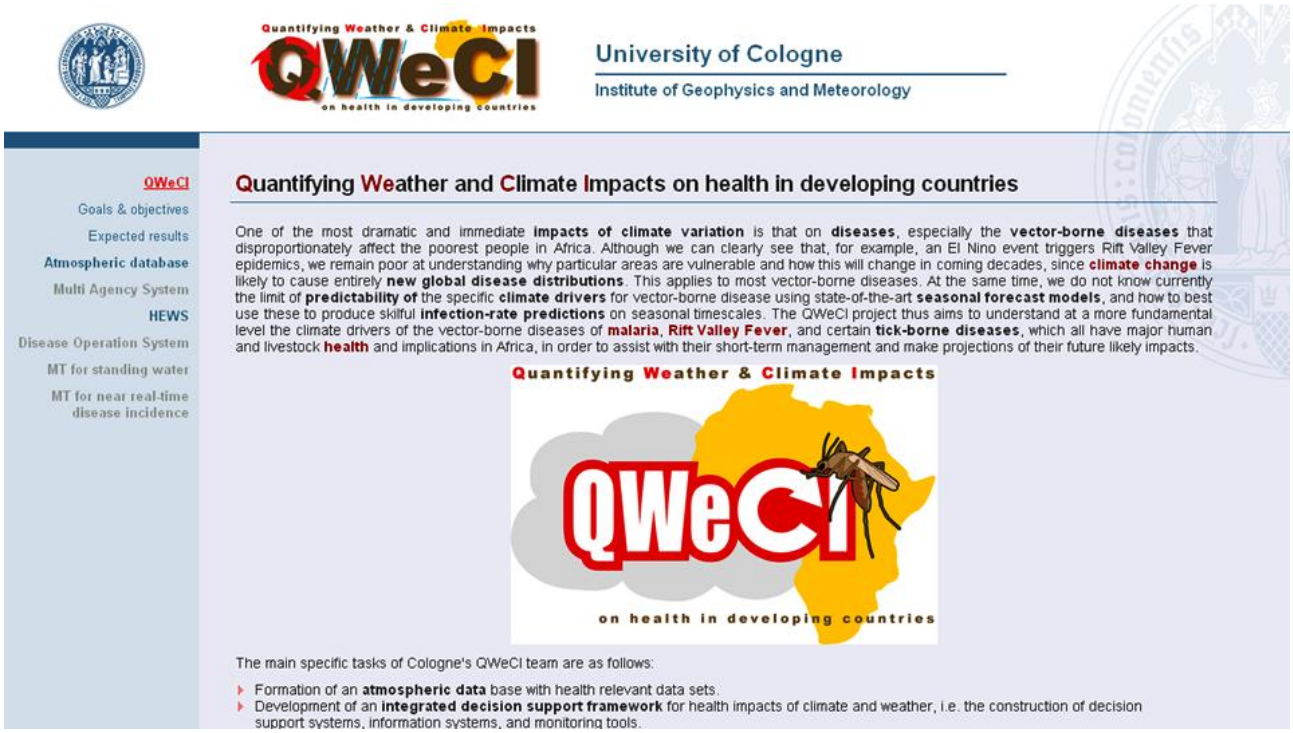


Figure 1: Web portal of the UoC regarding QWeCI available at http://qweci.uni-koeln.de.

Detailed metadata is provided for each dataset according to the **ISO 19139** geographic information metadata e**X**tensible **M**arkup **L**anguage (**XML**) schema implementation. This catalog is divided into various categories consisting amongst others of an abstract, resource overview, content information, graphical overview, lineage, resource constraints, data quality, and distribution information (Figure 4 & 5). All this information is provided by the web portal and is in addition available by an XML file.

A link is provided for the originator of the datasets. Here the user might be able to access the original data, download recent data updates, to find different versions of the dataset and additive information. Note that the original data files have been frequently processed in order to provide end-user friendly file formats such as **C**omma-**S**eparated **V**alues (**CSV**) or the **Net**work **C**ommon **D**ata **F**ormat (**NetCDF**).

The datasets of the QWeCI atmospheric database can be either downloaded from a public accessible **FTP** (**F**ile **T**ransmitting **P**rotocol) server of the UoC computer centre (free access via ftp://ftp.uni-koeln.de/institute/qweci) or via a password protected FTP service of the UoC QWeCI server (restricted access via ftp://qweci.uni-koeln.de). Downloadable are the data files, example graphics of the data, the XML file of the dataset, as well as supporting files (Figure 3).

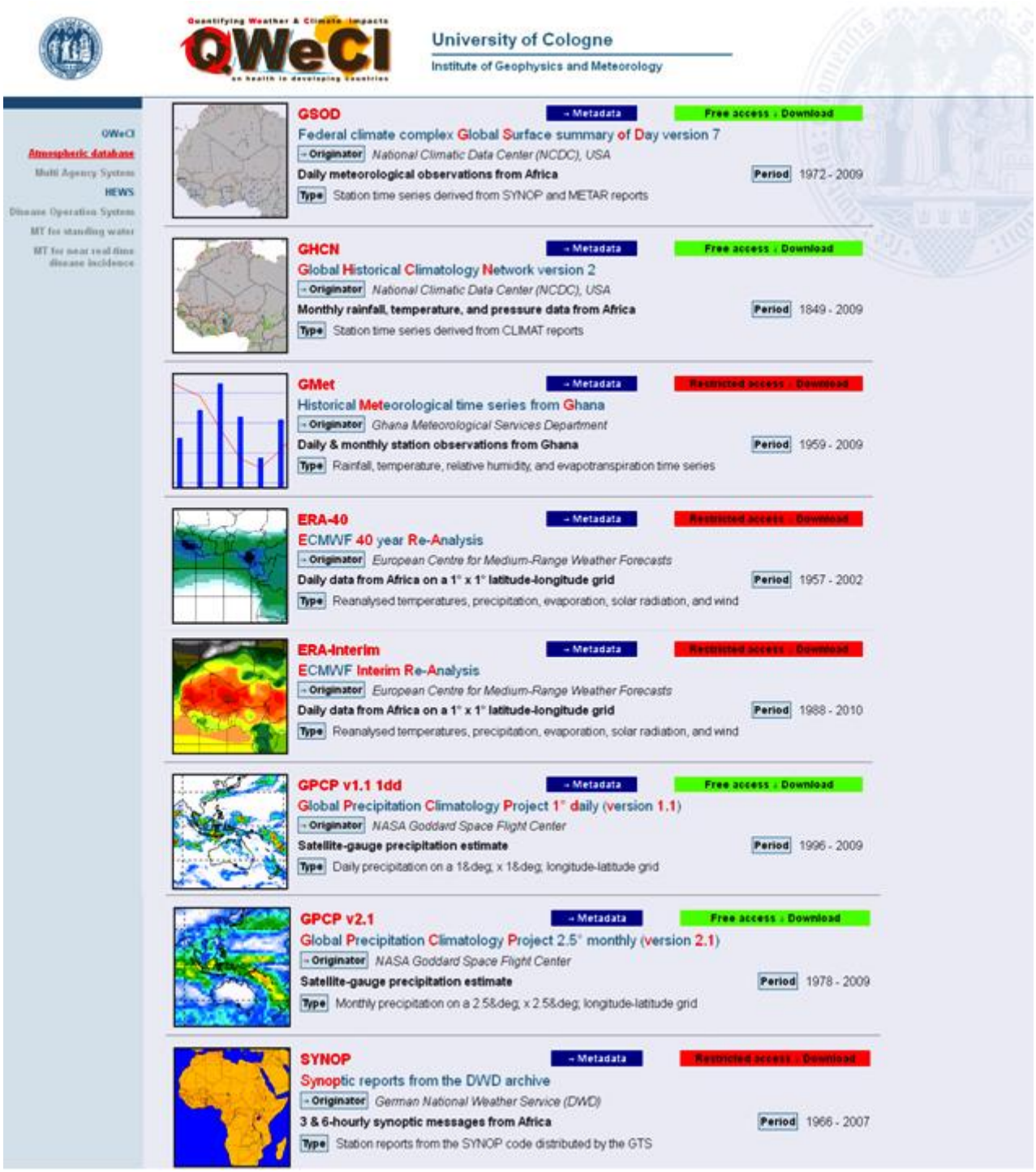


Figure 2: Short profiles of the datasets as presented by the web-based Java framework.

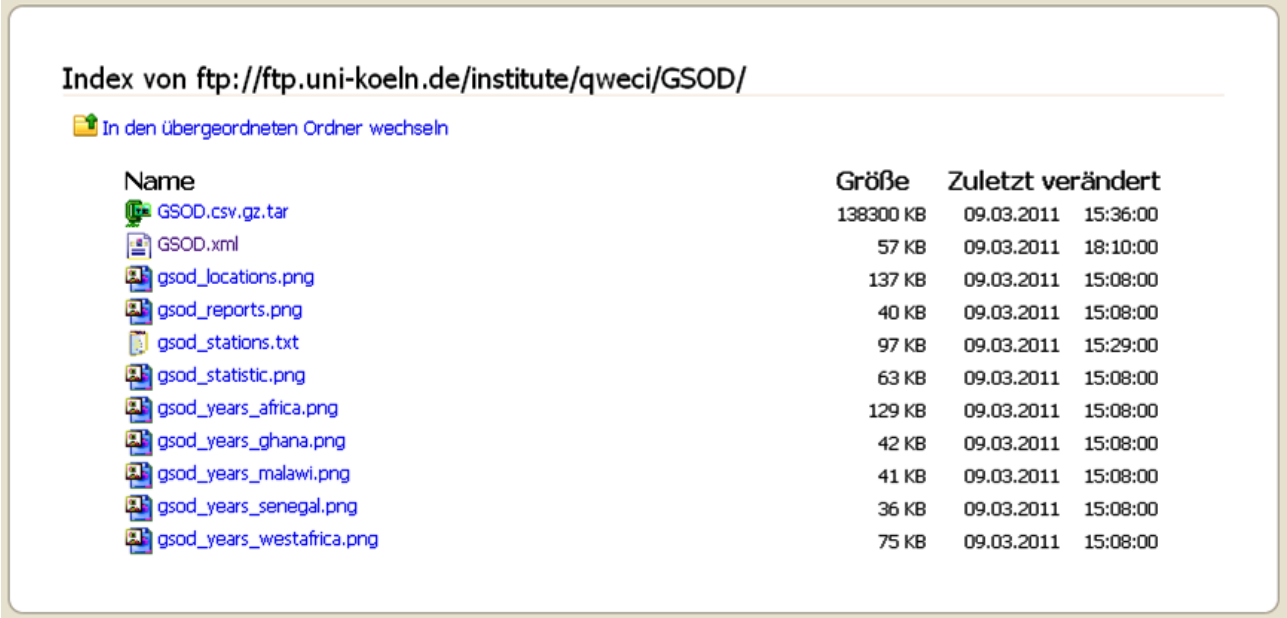
Datasets

In the following, datasets are presented, which were already entered into the atmospheric database. The alpha version of the atmospheric database includes altogether eight different datasets:

- GSOD: Federal climate complex Global Surface summary of Day version 7

The federal climate complex **G**lobal **S**urface summary **o**f **D**ay **(GSOD**) dataset is produced and regularly updated by the **N**ational **C**limatic **D**ata **C**enter (**NCDC**). In the dataset a subset of the global surface summary of day data was extracted for Africa. The daily entries of GSOD include 18 surface meteorological elements, which were derived from SYNOP and METAR reports. The historical time series are generally available for 1973 to the present. Some stations in Africa reveal data back to before 1900. The variables included in the dataset differ from station to station, they include mean temperature, mean dew point, mean sea level pressure, mean station pressure, mean visibility, mean wind speed, maximum sustained wind speed, maximum wind gust, maximum temperature, minimum temperature, precipitation amount, snow depth, as well as an indicator for the occurrence of fog, rain or drizzle, snow or ice pellets, hail, thunder and tornado/funnel clouds of a particular day.

Figure 3: Directory of the GSOD dataset on the public accessible FTP server of the UoC computer



centre (downloadable via ftp://ftp.uni-koeln.de/institute/qweci/GSOD*).*

- GHCN: Global Historical Climatology Network version 2

The global historical climatology network (GHCN) monthly data is produced and regularly updated by the NCDC. Here a subset of the global GHCN version 2 dataset was extracted for Africa. The GHCN database contains monthly historical mean, maximum and minimum temperatures, precipitation amounts, and pressure data for synoptic weather stations, which were mostly gathered from CLIMAT reports. The period of record varies from station to station; some of the stations in Africa extend back to the 19th century.

- GMet: Historical Meteorological time series from Ghana

The Ghana dataset contains both daily and monthly data of several synoptic weather stations as well as hydro-meteorological stations from Ghana. Daily data include minimum and maximum

temperatures, precipitation amounts, as well as sunshine duration. The period of record varies from station to station and includes various data gaps. For most stations data is available between 1960 and 2008.

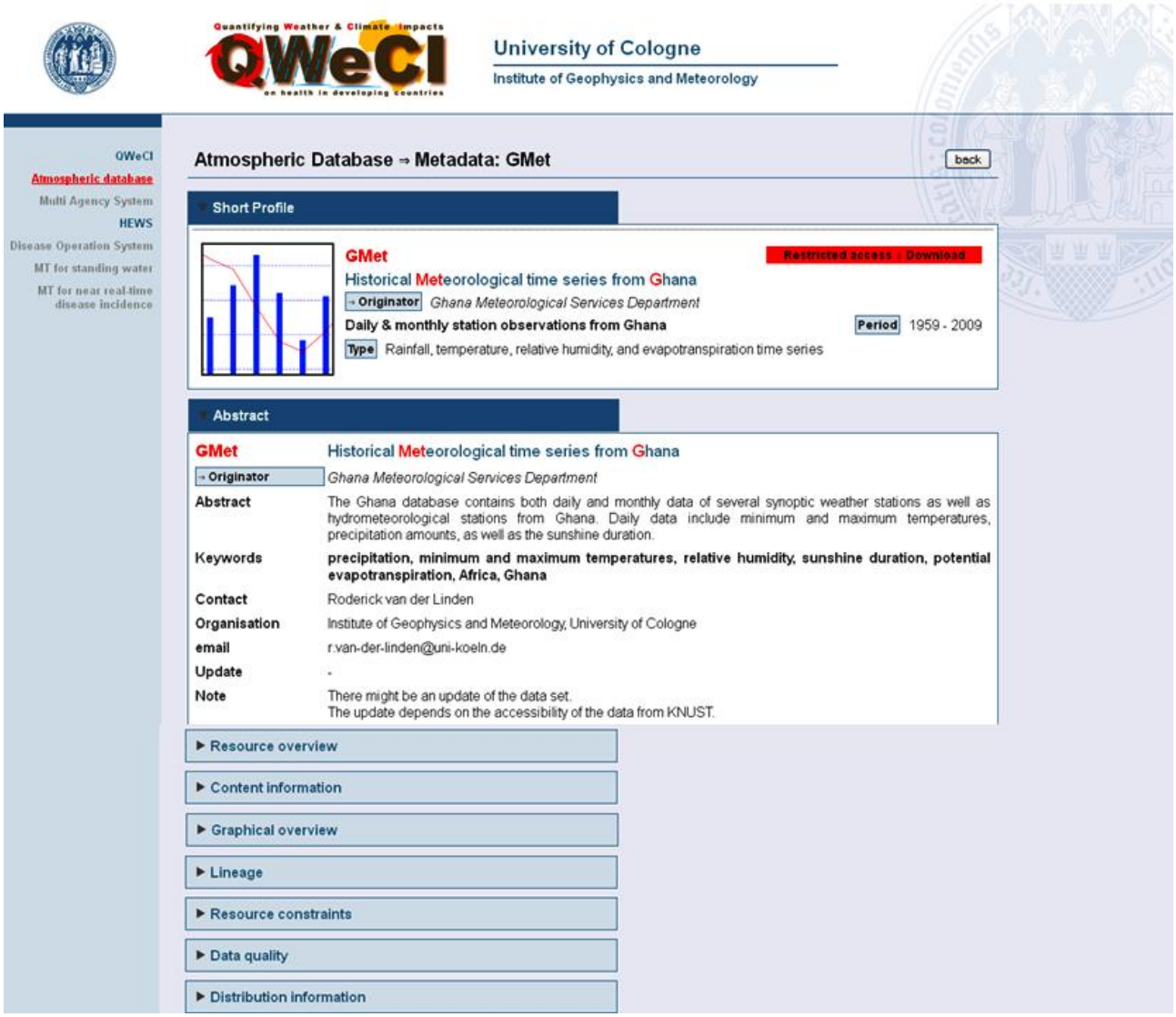


Figure 4: Metadata of the GMet dataset with the extended 'Short Profile' and 'Abstract' sections.

- ERA-40: ECMWF 40 year Re-Analysis

ERA-40 covers the period from September 1957 to August 2002. The three dimensional variational technique was applied using the T159L60 version of the Integrated Forecasting System to produce the analyses every six hours and forecasts each three hours. The analysis involved comprehensive use of satellite data, starting from the early vertical temperature profile radiometer data in 1972, then later including data from various satellites. Cloud motion winds were used from 1979 onwards. ERA-40 makes also use of data from past field experiments such as GATE. The QWeCI database contains a subset of the ERA-40 data. Only daily mean and daily accumulated surface variables were added to the database. For the height of 2 m included are the daily mean temperature, maximum and minimum daily temperature and relative humidity. The dataset furthermore comprehends the zonal and meridional wind component, skin temperature, ground temperature and soil moisture below surface, evaporation, total precipitation, and incoming solar radiation.

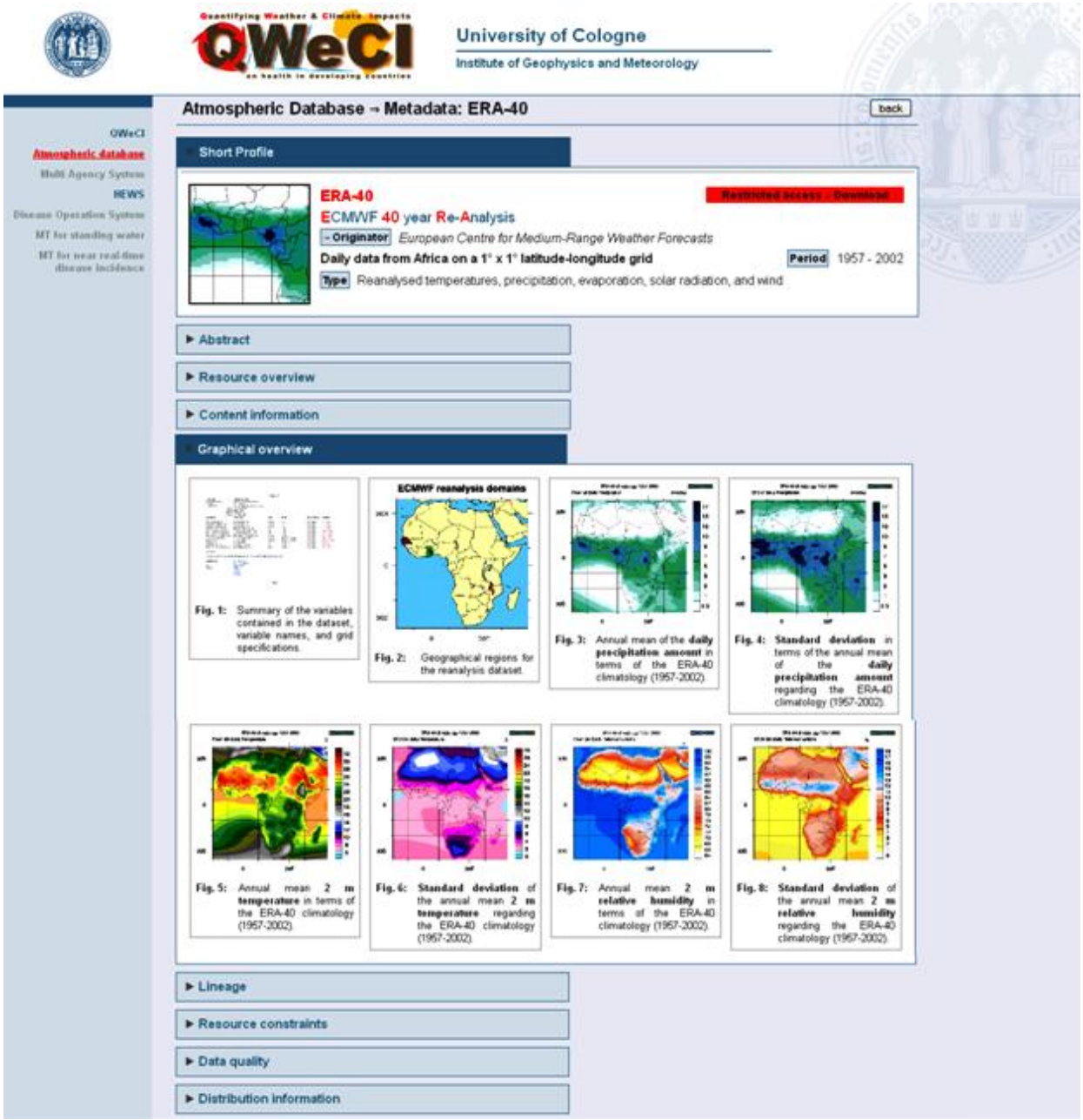


Figure 5: Metadata of the ERA-40 dataset with the extended 'Short Profile' and 'Graphical overview' sections.

- ERA-Interim: ECMWF Interim Re-Analysis

ERA-Interim is the latest global atmospheric reanalysis produced by the ECMWF, covering dates from 01 January 1989 onward to present. The ERA-Interim project is conducted in part to prepare for a new atmospheric reanalysis to replace ERA-40, which will extend back to the early part of the 20th century. Like for ERA-40 only a subset of the dataset was extracted for the QWeCI atmospheric database. Analogous to ERA-40 only surface variables are included.

- GPCP v1.1 1dd: Global Precipitation Climatology Project 1° daily (version 1.1)

GPCP was established to quantify the distribution of precipitation around the globe over many years. In support of this work an international group of precipitation experts developed and produced the GPCP Version 1.1 satellite-gauge **One-D**egree **D**aily (**1DD**) combined precipitation

dataset. The 1DD product provides precipitation estimates on a 1-degree grid over the entire globe at 1-day (daily) for the period October 1996 - present. The 1DD product is consistent with the Version 2 monthly product in the sense that the 1DD approximately sum to the monthly satellite-gauge estimate. All precipitation products are produced by optimally merging estimates computed from microwave, infrared, and sounder data observed by the international constellation of precipitation-related satellites, and precipitation gauge analyses.

- GPCP v2.1: Global Precipitation Climatology Project 2.5° monthly (version 2.1)

The GPCP 2.5-degree version 2 monthly product covers the period January 1979 to the present, with a delay of two to three months for data reception and processing. All precipitation products are produced by optimally merging estimates computed from microwave, infrared, and sounder data observed by the international constellation of precipitation-related satellites, and precipitation gauge analyses. The precipitation gauge analysis used in the GPCP satellite-gauge is created by the Global Precipitation Climatology Centre.

- SYNOP: Synoptic reports from the DWD (German Weather Service) archive

The SYNOP dataset includes meteorological messages, which are based on data exchanged under the WMO World Weather Watch Program according to WMO Resolution 40 (Cg-XII). The archive includes data from SYNOP messages as distributed by weather stations. The data was provided by the archive from the German Weather Service (DWD). A subset of the SYNOP dataset available in the DWD archive was extracted for Africa for the WMO block numbers 60, 61, 62, 63, 64, and 65. Included in the synoptic reports are atmospheric variables such as air temperature, dew-point temperature, atmospheric pressure, wind speed, the precipitation amount, present and past weather, cloud observations, etc. The archive contains data from 1966 to 2010 and is most complete since about 2000. For example, on average for each weather station contained in the dataset, more than four synoptic messages are available per day in 2009.

Future prospects

In the near future, further datasets will be added to the atmospheric database. It is planned to add an additional datasets including station observations. The **M**et Office **I**ntegrated **D**ata **A**rchive **S**ystem (**MIDAS**) Land Surface Stations data will supplement the SYNOP and GSOD datasets. Satellite rainfall estimates of the **T**ropical **R**ainfall **M**easuring **M**ission (**TRMM**) in terms of Version 6 of the 3B42 and 3B43 algorithms will amend the GPCP SG precipitation estimates. Further re-analysis data will be provided in terms of the **NCEP**/NCAR Re-Analysis **1** (**NCEP-1**). NCEP-1 is a joint product from the **N**ational **C**enters for **E**nvironmental **P**rediction (**NCEP**) and the **N**ational **C**enter for **A**tmospheric **R**esearch (**NCAR**). Regarding the pilot areas of the QWeCI project, historical time series from Senegal and Malawi will be added to the database. Due to data restrictions only a small subset of the available historical time series will be accessible from the two countries.

The atmospheric database will be used for the assessment of ERA-Interim products for the three target countries (D1.2b; M24). One problem is related to the fact, that long-term historical observations are missing for Senegal and Malawi. This problem can be partly compensated by means of the SYNOP, GSOD, MIDAS, and GHCN datasets.

The atmospheric database was released on M15 as the alpha version (D1.2a). The next updated beta version of the database will contain additional datasets. It is planned to add other functionalities like a search function for further versions of the database. From a technical point of view such features are relatively easy to include since the database is incorporated into the web-based Java framework of work package 5.1. It is foreseen to complete the final version of the atmospheric database, which is ready for transfer into the AMMA database, at M26 of the QWeCI project (D1.2c).