Sofia University St. Kliment Ohridski
Faculty of Physics
Department Meteorology and Geophysics

Sofia University Atmospheric Data Archive

Master of Sciences Thesis

Tzvetan Simeonov Simeonov
f.n. 180046

Thesis supervisor:

/head assist. prof. G. Guerova/

Head of department:

/assoc. prof. N. Rachev/

Sofia April 2013

СОФИЙСКИ УНИВЕРСИТЕТ
„СВ. КЛ. ОХРИДСКИ“
ФИЗИЧЕСКИ ФАКУЛТЕТ
КАТЕДРА „МЕТЕОРОЛОГИЯ И ГЕОФИЗИКА“

Архив за атмосферни наблюдения на СУ „Св. Климент Охридски“

МАГИСТЪРСКА ДИПЛОМНА РАБОТА

на

Цветан Симеонов Симеонов

ф.н. 180046

Научен ръководител:

/гл. ас. д-р Г. Герова/

Ръководител катедра:

/доц. д-р Н. Рачев/

София Април 2013

Contents

List of Figures

List of Tables

Резюме

Архивът за атмосферни наблюдения на Софийския Университет (АСУ) е създаден, за да съхранява данни за интегрираната по височина водна пара (ИВП, IWV), както и за други атмосферни наблюдения. До момента в базата са заредени над 12 милиона идивидуални наблюдения за 32 станции на територията на България, обхващащи периода от 1997 година до настоящия момент.

С нарастването на количеството генерирани научни данни проблемът за тяхното съхранение става все по-належащ. Базите данни предлагат оптимизирано решение за съхранение на данни, комбинирано с улеснен достъп и единна структура на отделните записи. Проблемът на големия обем данни е особено очевиден при изследването на атмосферните процеси. За това базите данни често са използвани за съхранение на метеорологична информация.

Териториално различаваме два типа бази данни: глобални и регионални.

Като пример за глобална база данни е представен порталът за климатични промени на Световната банка (wba). Той съдържа информация за приземни температура и налягане, валеж, температура на повърхността на океана и други метеорологични и климатични величини. Базата съдържа записи от началото на XX век до наши дни за множество станции, разположени по цялото земно кълбо. Данните могат да бъдат извеждани на потребителите не само в записен формат, но и като фигури и карти на съответно избраните полета.

Една от разгледаните регионални бази данни е базата STARTWAVE на Университета в Берн, Швейцария, (*Morland et al.*, 2006). Тя съдържа не само данни за приземната температура и налягането за територията на Швейцария, но и измервания по метода ГНСС метеорология, както и наблюдения от различни инструменти. Данните са записани в база чрез стандартни директни заявки на MySQL и чрез програми, писани на MATLAB.

АСУ е създадена по подобие на Швейцарската база STARTWAVE и е предвидена да съхранява както синоптични, така и данни, получени

по метода ГНСС (Глобални Навигационни Спътникови Системи) метеорология. Това е първата база данни от този тип не само в България но и в цяла Югоизточна Европа.

В глава 2 на дипломната е представена цялостната концепция на АСУ.

Глава 2.1 са разгледани физичните величини, записани в базата данни. Таблица 2.1 представя физичните величини, точностите на измерване, както и измерителните единици на всяка една величина, записана в таблиците на базата.

В глава 2.2 е представена структурата на базата данни (фигура 2.2) както и са описани всички таблици, съществуващи в базата (в таблица 2.3). Базата данни е създадена на основата на MySQL, като за комуникация с нея се използват php, bash и MATLAB програми.

В базата съществуват четири вида таблици. Първият вид таблици са информационните таблици (глава 2.2.2). В тях се съхранява информация за станциите, тяхните координати и видовете източници на информация. Връзките за това коя станция получава данни от кой източник също е заредена тук.

Втория вид таблици са така наречените първични (таблици, описани в глава 2.2.3). В тях се съхраняват необработените данни от синоптичните и ГНСС станциите, както и данните от аерологичния сондаж. Вторичните таблици (глава 2.2.4) съдържат обработените вече данни от първичните таблици за количество водна пара в атмосферата, както и за други атмосферни параметри, пресметнати по метода ГНСС метеорология и чрез аерологичния сондаж.

Таблиците, необходими за достъпа до информация от страницата на базата данни в интернет са представени в глава 2.3 . Тяхната функция е да съхраняват информация за потребителите за по-достъпно използване на онлайн портала на базата данни. В същата глава е описан и интернет порталът на АСУ, който включва страници, обясняващи всеки един от методите, използвани при обработката на данни; новини, свързани с базата; карти на достъпните за сваляне данни и други.

В глава 2.4 и 2.5 са представени начините на запълване на таблиците и за обработка на наблюденията съответно. Процедурите се извършват с bash скриптове, както и с MATLAB програми за обработката. В приложение са добавени примерени програми.

В глава 3 е представен методът ГНСС метеорология за пресмятане на ИВП. При преминаването си през атмосферата сигналът на ГНСС се забавя и пречупва. Именно това забавяне и пречупване се използва за измерване на водната пара в атмосферата. В глави 3.2, 3.3, 3.4 са представени използваните източници на първични ГНСС наблюдения.

В глава 4 са представени източниците на синоптични данни и данни от аерологичния сондаж. АСУ съхранява наблюдения от аерологичния сондаж за станция София. Времевият ред е от 1995 година, а измерванията са извършени от Националния Институт по Метеорология и Хидрология при БАН. Синоптичните наблюдения са също от мрежата на НИМХ-БАН за територията на България. Данните за станциите извън България са измерени от националните метеорологични служби на съответните държави и са придобити от световните бази данни. В глава 4.3 е представен и численият модел за прогноза на времето WRF. В бъдеще се предвижда моделни данни да бъдат използвани за определяне на температурата и налягането в станциите, разположени на територията на България.

В глава 5 са представени изследвания, при които са използвани данни от АСУ. Изучаване на топлинната вълна над България между 19 и 25 юли 2007 година е представено в глава 5.1. Наличието на топлинни вълни също така се проявява и при годишния ход на водната пара. Така в най-топлите месеци водната пара над България е под нормите с около 6 %, докато температурата надвишава с 20 % средномесечните стойности. Също така е разгледан денонощния ход на температурата и ИВП във всяка от изучаваните станции. В крайбрежните станции е установен следобеден пик в ИВП, който се дължи на бризовата циркулация (*Simeonov et al.*, 2013).

В глава 5.2 е представено сравнение на ИВП получена по три различни модела за обработка. Резултатите показват до 3% разлика, най-вече през лятото (Юни, Юли и Август).

Abstract

The Sofia University Atmospheric Data Archive (SUADA) is developed to facilitate the use of vertically Integrated Water Vapour (IWV) data for meteorologic and climatic studies in Bulgaria/Southeast Europe. The SUADA database includes:

- Ground-based Global Navigation Satellite Systems (GNSS) observations from:
  - EUREF post-processed IWV (total 6 539 observations) for station Sofia (SOFI 04.2001-11.2004)
  - IGS reprocessed IWV (total 19 648 observations) for station Sofia (SOFI 1997-2012)
  - Zenith Total Delay data (total 11 473 034 observations) from 30 stations from ZenitGEO company network in Bulgaria (since 11.2011)
  - ZenitGEO processed IWV (total 23 233 observations) from 11 of the 30 ZenitGEO stations
  - IWV (total 763 observations) from 8 stations from IGS network on the Balkan Peninsula (for 19-25 July 2007 heatwave)
- Radiosonde IWV data (total 6 376 observations) for station Sofia (1997-2012)

The first application of the SUADA data is a study of water vapour dynamics during the 2007 heat wave in Bulgaria. IWV was derived from 8 GNSS stations in Southeast Europe for 19-25 July 2007 heat wave period. At the Black sea coastal stations Constanta and Varna, the peak of IWV is registered at 15 UTC, which is 3 hours after the temperature peak. Similarly, IWV peak at 15 UTC is characteristic for the Mediterranean sea station Athens and Adriatic sea station Dubrovnik. The IWV peak at 15 UTC can be explained with the peak of the sea breeze circulation that brings humid air from the sea inland. In contrast, at the inland stations Bucharest and Craiova the peak of IWV is between 6-9 UTC. The differences in diurnal cycle of inland and coastal stations show the high sensitivity of GNSS derived water vapour for capturing the small-scale local atmospheric circulation.

The second application of SUADA data is long term comparison of IWV processing methods. Three diffrent processing are compared for a 3 year period 2001-2004. A differnce is up to 3 % was found mostly for the summer period (June, July and August).

Chapter 1

Introduction

With the increase of generated scientific information, it becomes more important to optimize data storage and use. The best possible optimization in terms of structure, ease of access and processing is achieved by using databases. The field of atmospheric physics is especially affected by the problem because of the large amount of collected data, which has to be simultaneously processed in a restricted amount of time in order to produce weather forecasts.

Previous researches show the advantages of storing atmospheric data from different sources in a single database. Meteorology and climatology databases can be classified by the area of covered territories into:

- global - a database, which collects data from numerous stations around the globe
- regional - a database, which stores data from stations in a specific region.

A typical example for a global database is the World Bank Climate Change Data Portal *(wba, World Bank)*. It combines surface temperature and pressure data together with precipitation, ocean temperature and other climate parameters. The World Bank database stores continuous data, ranging from the beginning of the XX century until the present day. This database enables users to make atmospheric research not only in climatic timescales, but also to make case studies for synoptic events. The data can be extracted from the database in a text format or being visualized graphically on a map by a special function on the web site.

Another global database is the OGIMET *(ogi)* database. It stores meteorological and climate information from National Weather Services. This database is a scientific collaboration between the Spanish Weather Service and the University of Granada. The database web site also provides satellite images. Weather forecasts from several Numeric Weather Prediction models and synoptic charts are available in the OGIMET database.

The University of Wyoming's Department of Atmospheric Science database *(wyo)* is a global database, which stores meteorological and climate data. The database web site has sections, developed specifically for weather forecasts over the USA with a selection of available parameters. The web site also provides synoptic charts. The database also contains radiosonde measurements of the atmosphere in the "Upper Air Observations" section. Polar meteorology data and forecasts are also available.

The EGVAP database *(egv, EGVAP)* represents a regional database. Data from weather stations around Europe is stored in it. EGVAP contains data not only from synoptic stations (temperature and pressure), but also GNSS meteorology data (Global Navigational Satellites System meteorology method described in chapter 3). For the moment the EGVAP network consists of more than 1600 GNSS sites. The project is mainly focused on near-real time GNSS data for the needs ot the operational Numerical Wearther Prediction applications in Europe. The database is part of the EUMETNET Network of the European Meteorological Services.

The STARTWAVE (*Morland et al.*, 2006) database is another example of a regional database. In the database are archived observations of intergated water vapour as well as water vapour profiles made by various techniques in Switzerland and the neightbouring countries. One of the data-set in STARTWAVE is atmospheric water vapour derived from the ground-based Swiss GNSS network AGNES. In STARTWAVE the data is stored using the MySQL platform and can be retrieved either through a php-based web interface, or with MySQL queries, which are inserted into MATLAB programs (*Morland*). This combination of sources allows easier comparison of data from the different sources and faster data operations. Larger volumes of data can be accessed, manipulated and calculated much easier, when stored in a single organized structure.

The Sofia University Atmospheric Data Archive (SUADA) is a new regional database aiming at: 1) achiving atmospheric water vapour observations from diferent techniques and 2) using the data for meteorologic and climatic studies in Bulgaria/Southeast Europe. Collected and archived is data not only from stations on the territory of Bulgaria, but also from stations on the Balkan Peninsula. The main focus is to observe meteorological and climatic events, using the GNSS meteorology method. Similar to STARTWAVE databases, SUADA is designed to enhance and facilitate the atmospheric research in the Sofia University, but also to provide online data access, via e web portal, for interested researchers in Bulgaria and the neightbouring countries.

The SUADA is designed to incorporate ground-based GNSS data together with synoptic, radiosonde and model data. The structure of the database and the information flow is presented in chapter 2. The GNSS meteorology method and the sources of GNSS data are presented in chapter 3. The radiosonde and synoptic data and processing are presented in chapter 4. In chapter 5 two case studies, using SUADA data are presented.

Chapter 2

SUADA design and structure

In this chapter the technical parameters of the database are discussed. This includes the structure of the database, the uploading, calculation of data and the SUADA web portal.

2.1 Atmospheric parameters in SUADA

In table 2.1 are presented the archived atmospheric parameters, their format, range and accuracy. For example Integrated Water Vapour (IWV) is recorded in a 'double' format; the data sources are IGS, EUREF and Zenithgeo (see chapters 3.2.1, 3.2.2, 3.3, 3.4); it is in the range between 0 and 50 mm (a typical summer average is $\sim$ 25 mm and a typical winter values are $\sim$ 10 mm); the precision of measurement is 0,1 mm and the unit of measurement is millimeters.

Table field	Data type	source	range	accuracy	units
Altitude	float		0-9000	0,1	m [meters]
Longitude	float		0-360	0,01	decimal degrees
Latitude	float		-90 - +90	0,01	decimal degrees

IWV	double	chapters 3.2.1, 3.2.2, 3.3, 3.4	0-50	0,01	mm [millimeters]
IWV standard deviation	double	chapters 3.2.1, 3.2.2, 3.3, 3.4	0-2	0,01	mm [millimeters]
IWV north and east gradients	double	chapters 3.2.1, 3.2.2, 3.3, 3.4	-0,01 - +0,01	0,0001	mm [millimeters]
Pressure	float	chapters 4.1, 4.2	850-1050	0,1	hPa [hecto Pascals]
Temperature	float	chapters 4.1, 4.2	-80 - +50	0,1	°C [degrees Celsius]
ZTD	double	chapters 3.2.1, 3.2.2, 3.3, 3.4	0-3	0,0001	m [meters]
ZHD	double		0-0,4	0,0001	m [meters]
ZWD	double		0-3	0,0001	m [meters]
Height	float	chapter 4.1	0-9000	0,1	m [meters]
Dew_ Point	float	chapter 4.1	-80 - +50	0,1	°C [degrees Celsius]
Wind_ Dir	float	chapter 4.1	0-360	1	decimal degrees
Wind_ Speed	float	chapter 4.1	0-100	0,1	m/s [meters per second]
Rel_ Hum	float	chapter 4.1	0-108	1	%
MixR	double	chapter 4.1	0-40	0,01	g/kg

Table 2.1: SUADA atmospheric parameters format, range, accuracy and units.

The abbreviations used in table 2.1 are listed in table 2.2.

Abbreviation	Description
IWV	Integrated Water Vapour
ZTD	Zenith Total Delay
ZHD	Zenith Hydrostatic Delay
ZWD	Zenith Wet Delay
Rel_ Hum	Relative Humidity
MixR	Mixing ratio

Table 2.2: Abbreviations of SUADA parameters.

2.2 SUADA data structure and tables

SUADA is developed using the Structured Query Language (SQL) for relational database management system. The relational method of database management was proposed by *Codd* (1970). The method structures sequences of elements (tuples). Every single tuple consists of several elements, also called domains. The domains are organized by named relational definitions, also called attributes. The relation is a structured set of attributes and sequences, which can be represented visually as a table (see figure 2.1).

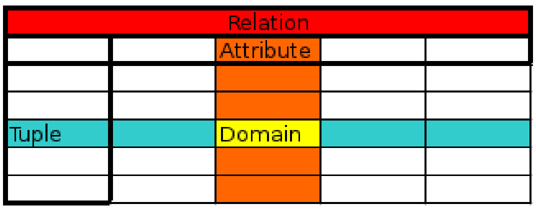

Figure 2.1: Relational database structure example.

Codd (1970) proposes several schemes of database structuring. The SUADA tables are structured as peers with additional relations between them as shown in figure 2.2 (see scheme 5, proposed by *Codd* (1970)).

A set of constraints is applied in the SUADA in order to avoid repeating data or lines. It is not desirable to have two recordings from the same source, station and time. In order to prevent this PRIMARY KEY's are set for every single table in the database. A primary key is a constraint in MySQL, which prohibits the repeating of equal values for certain, defined by the user, set of attributes. For the purposes of SUADA this means, that for a single station, source, day and time only one tuple can be recorded in the database.

It is also not desirable to have data from a non-existing source of information. Thus a set of FOREIGN KEY's are defined between the tables in SUADA. These constraints guarantee that data can be uploaded and processed only for predefined combinations of stations and sources, as every source provides data for unique set of stations. On the other hand data for a particular station can be acquired from multiple sources.

2.2.1 SUADA structure

Figure 2.2 shows the SUADA data stucture and flow.

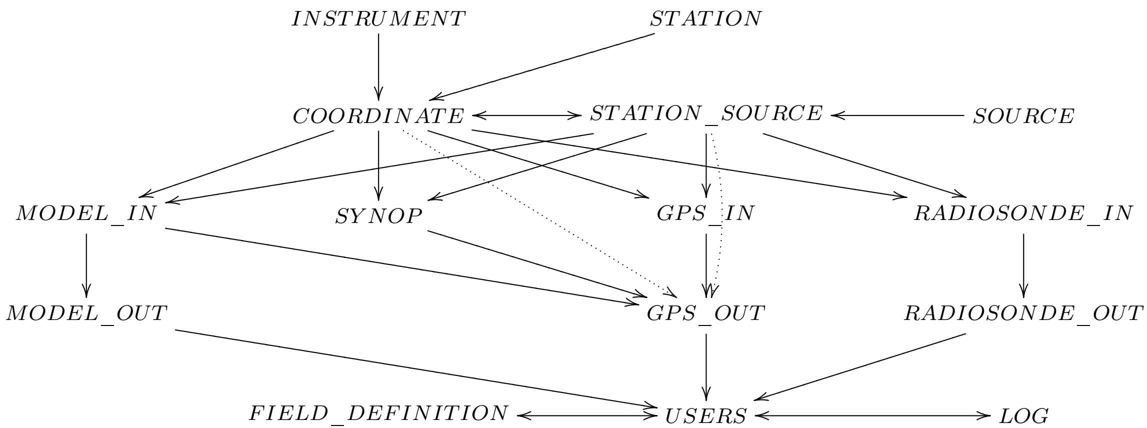

Figure 2.2: SUADA data structure and data flow.

The SUADA tables are presented in groups. The first group of tables are the "SUADA information tables", presented in section 2.2.2. They are: INSTRUMENT, STATION, COORDINATE, STATION_ SOURCE and SOURCE (lines 1 and 2 on figure 2.2). The second group of tables are the "Primary tables", presented in section 2.2.3. They are: MODEL_ IN, SYNOP, GPS_ IN and RADIOSONDE_ IN (line 3 on figure 2.2). The third group of tables are the "Secondary tables", presented in section 2.2.4. They are: MODEL_ OUT, GPS_ OUT and RADIOSONDE_ OUT (line 4 on figure 2.2). The group of tables "Information tables for the web portal" is presented in section 2.3.1. They are: FIELD_ DEFINITION, USERS and LOG (line 5 on figure 2.2).

The list of tables is presented bellow with a short summary for each of them.

Table name	Summary
INSTRUMENT	Instrument name and identification number
STATION	Station name
COORDINATE	Coordinates of the GNSS, synop and radiosonde stations
STATION_ SOURCE	Station source information (either instrument or method)
SOURCE	Contact information of SUADA data providers (name, institution, telephones, etc.)
MODEL_ IN	Numerical Weather Prediction (NWP) model data
SYNOP	Surface observations from the network of the National Institute of Meteorology and Hydrology
GPS_ IN	Tropspheric products from ground-based GNSS networks or individual station
RADIOSONDE_ IN	Data from the radiosonde network or individual station
MODEL_ OUT	Processed NWP model data (IWV or other)
GPS_ OUT	Processed GNSS data (IWV, ZHD or other)
RADIOSONDE_ OUT	Processed radiosonde data (IWV)
FIELD_ DEFINITION	List of abbreviations used in the SUADA tables

USERS	Contact information about SUADA data users (external and internal)
LOG	User log in history.

Table 2.3: SUADA table structure.

2.2.2 SUADA information tables

The information tables are the tables, which store the information about instruments, stations, station coordinates, data sources and data provides as well as station-source connections.

The INSTRUMENT table 2.4 stores information about the measuring instrument.

Table field	Data type	Summary
ID	int(11)	Unique ID number, generated for each instrument type.
Name	varchar(128)	Name of the instrument measurement technique, for example "SYNOP"
TableName	varchar(48)	Name of the database table, where data from the instrument is stored
AccessLevel	int(11)	Selects which instrument data is available for non-authorized users and which only for authorized users

Table 2.4: SUADA INSTRUMENT table field definition.

The STATION table 2.5 defines the name of the nearest town or city with the ID number of the nearby set of stations.

Table field	Data type	Summary

ID	int(11)	Unique number for individual station or set of collocated stations. Example for a set of collocated stations is station Sofia with collocate observations from the synoptic, radiosonde and GNSS networks. Each of them is referred with the same station number ID but different instrument source. Locked with a FOREIGN KEY to StationID in COORDINATE table.
Name	varchar(128)	The name of the city or village at which the station is located
Note	varchar(10000)	Field for special notes

Table 2.5: SUADA STATION table field definition.

The COORDINATE table 2.6 is a key-table for storing the coordinates of a station like latitude, longitude and altitudes. If the station coordinates are not available the station can not be processed.

Table field	Data type	Summary
StationID	int(11)	Unique number for individual station or set of collocated stations. Example for a set of collocated stations is station Sofia with collocate observations from the synoptic, radiosonde and GNSS networks. Each of them is referred with the same station number ID but different instrument source.
InstrumentID	int(11)	Unique ID number for type of measuring instrument of the station. Linked with FOREIGN KEY with ID from table INSTRUMENT

ID	int(11)	Unique number for combination of StationID and InstrumentID in order to identify the type of station
Altitude	float	Station altitude in meter above sea level
Longitude	float	Station longitude in decimal degrees
Latitude	float	Station latitude in decimal degrees
Note	varchar(10000)	Field for special notes

Table 2.6: SUADA COORDINATE table field definition.

The STATION_ SOURCE table 2.7 is a key-table for the database. Its purpose of this table is to account for different data sources for a single station. For example station Sofia with StationID number 31 uses multiple sources - 1, 2 and 3. In order to take into account that station 31 can be used with data sources 1, 2 and 3, three lines in the STATION_ SOURCE table are made: line 1 for station 31 with data source 1, line two for station 31 with data source 2 and line 3 for station 31 with data source 3.

Table field	Data type	Summary
ID	int(11)	Unique number, generated for each combination of StationID and SourceID.
StationID	int(11)	Unique ID number for individual station or set of collocated stations. Linked with FOREIGN KEY with StationID from table COORDINATE
SourceID	int(11)	Unique ID number of the data provider. Linked with FOREIGN KEY with ID from table SOURCE

Table 2.7: STATION_ SOURCE table field definition.

Table SOURCE table 2.8 holds data and contact information of data partners/providers.

Table field	Data type	Summary
ID	int(11)	Unique ID number of the data provider.
Name	varchar(128)	Name of the organization
e-mail	varchar(64)	E-mail of the organization
Phone	varchar(64)	Organization phone number
Note	varchar(10000)	Field for special notes

Table 2.8: SOURCE table field definition.

2.2.3 Primary tables

The SUADA primary tables are the tables, where the observed parameters are stored.

The MODEL_ IN table 2.9 is intended to store Numerical Weather Prediction model data, such as the WRF (more about WRF in chapter 4.3).

Table field	Data type	Summary
Station_ SourceID	int(11)	Location identification number as defined in table STATION_ SOURCE 2.7. Linked with a FOREIGN KEY with ID from table STATION_ SOURCE.
Datetime	datetime	Date and time
Pressure	float	Pressure in hPa
Temperature	float	Temperature in °C
Note	varchar(10000)	Field for special notes

Table 2.9: MODEL_ IN table field definition.

The SYNOP table 2.10 table contains unprocessed meteorological data from the National Institute of Meteorology and Hydrology and other national Meteorological Offices, who provide data for the World Meteorological Organization.

Table field	Data type	Summary
Station_ SourceID	int(11)	Location identification number as defined in table STATION_ SOURCE 2.7. Linked with a FOREIGN KEY with ID from table STATION_ SOURCE._ SOURCE
Datetime	datetime	Observation date and time
Pressure	float	Pressure in hPa
Temperature	float	Temperature in °C
Humidity	float	Relative humidity in %
Note	varchar(10000)	Field for special notes

Table 2.10: SYNOP table field definition.

The GPS_ IN table 2.11 stores observed GNSS data from different sources as described in chapter 3.

Table field	Data type	Summary
Station_ SourceID	int(11)	Location identification number as defined in table STATION_ SOURCE 2.7. Linked with a FOREIGN KEY with ID from table STATION_ SOURCE.
Datetime	datetime	Observation date and time
ZTD	double	Zenith Total Delay
Sigma_ ZTD	double	Zenith Total Delay error
Gradient_ N	double	ZTD gradient in direction North
Gradient_ E	double	ZTD gradient in direction East
Sigma_ Grad_ N	double	ZTD gradient error in direction North
Sigma_ Grad_ E	double	ZTD gradient error in direction East
Note	varchar(10000)	Field for special notes

Table 2.11: GPS_ IN table field definition.

The RADIOSONDE_ IN table 2.12 stores observed radiosonde parameters. Each observation is stored in the table in several lines from different heights (vertical profiles).

Table field	Data type	Summary
Station_ SourceID	int(11)	Location identification number as defined in table STATION_ SOURCE 2.7. Linked with a FOREIGN KEY with ID from table STATION_ SOURCE.
Datetime	datetime	Observation date and time
Pressure	float	Pressure in hPa
Temperature	float	Temperature in °C
Height	float	Height in m asl.
Dew_ Point	float	Dew point in °C
Wind_ Dir	float	Wind direction
Wind_ Speed	float	Wind speed
Rel_ Hum	float	Relative humidity in %
MixR	double	Mixing ratio g/kg
Note	varchar(10000)	Field for special notes

Table 2.12: RADIOSONDE_ IN table field definition.

2.2.4 Secondary tables

The so-called secondary tables are the tables, where the processed information is stored.

The MODEL_ OUT table 2.13 table is intended to store meteorological data from Numerical Weather Models, such as the WRF (more about WRF in chapter 4.3).

Table field	Data type	Summary
Station_ SourceID	int(11)	Location identification number as defined in table STATION_ SOURCE 2.7. Linked with a FOREIGN KEY with ID from table STATION_ SOURCE.
Datetime	datetime	Simulation date and time
Pressure	float	Pressure in hPa
Temperature	float	Temperature in °C
Humidity	float	Humidity in %
IWV	float	IWV in mm
Note	varchar(10000)	Field for special notes

Table 2.13: MODEL_ IN table field definition.

The GPS_ OUT table 2.14 stores processed GNSS data. Data stored in GPS_ IN, STATION, SYNOP and MODEL tables is required for the processing. The GNSS meteorology method presented in chapter 3.1 is used for deriving IWV from the observed Zenith Total Delay. The IWV_ RH field is not being used yet. It is intended to store IWV with altitude correction described in *Morland and Matzler* (2007).

Table field	Data type	Summary
StationID	int(11)	Station ID number as defined in table STATION 2.5
SourceMetID	int(11)	Source ID number of the meteorological data as defined in table SOURCE 2.8
SourceGpsID	int(11)	Source ID number of the GNSS data as defined in table SOURCE 2.8
Datetime	datetime	Observation date and time
IWV	double	Integrated Water Vapour
IWV_ RH	double	Integrated Water Vapour with altitude correction (see *Morland et al.* (2009))
Sigma_ IWV	double	Integrated Water Vapour error

Sigma_ IWV_ RH	double	Integrated Water Vapour with altitude correction error
IWV_ N	double	Integrated Water Vapour gradient in direction North
IWV_ E	double	Integrated Water Vapour gradient in direction East
Sigma_ IWV_ N	double	Integrated Water Vapour gradient error in direction North
Sigma_ IWV_ E	double	Integrated Water Vapour gradient error in direction East
Timestamp	timestamp	The last time of entry alteration
Pressure	float	Pressure in hPa
Temperature	float	Temperature in °C
ZTD	double	Zenith Total Delay
ZHD	double	Zenith Hydrostatic Delay
ZWD	double	Zenith Wet Delay
Note	varchar(10000)	Field for special notes

Table 2.14: GPS_ OUT table field definition.

The RADIOSONDE_ OUT table 2.15 stores processed radiosonde data using data from STATION and RADIOSONDE_ IN tables. In contrast to RADIOSONDE_ IN table, in RADIOSONDE_ OUT table one line equals to one measurement.

Table field	Data type	Summary
StationID	int(11)	Station ID number as defined in table STATION 2.5
SourceMetID	int(11)	Source ID number of the meteorological data as defined in table SOURCE 2.8
SourceRadID	int(11)	Source ID number of the radiosonde data as defined in table SOURCE 2.8
Datetime	datetime	Observation date and time

IWV	double	Integrated Water Vapour
IWV_ RH	double	Integrated Water Vapour with altitude correction
Sigma_ IWV	double	Integrated Water Vapour error
Sigma_ IWV_ RH	double	Integrated Water Vapour with altitude correction error
IWV_ WAYOM	float	Integrated Water Vapour as calculated by the National Institute of Meteorology and Hydrology
Note	varchar(10000)	Field for special notes

Table 2.15: RADIOSONDE_ OUT table field definition.

2.3 SUADA Web portal

The SUADA web portal (*www.suada.phys.uni-sofia.bg*) presents the SUADA project. The web portal has five main menus with 18 sub-menus presenting different aspects of the project. In the "SUADA" sub-menu the database is accessible for external users. This is an interactive page where the user can specify the date, the station (or a specific area, defined by coordinates) and the method (GNSS or radiosonde) for which the data is extracted. All users can download up to 100 lines without registration. For unlimited data access user registration is required and is done via a web based registration form (http://suada.phys.uni-sofia.bg/?page_ id=190).

The "In Brief" sub-menu presents the different SUADA data sources. There is also a short representation of the GNSS meteorology method (also presented in chapter 3). In the "WRF Weather Prediction" sub-menu are presented the weather prediction charts of surface temperature and rainfall updated every 12 hours (see chapter 4.3). The other sub-menus of the web portal present news about the project; the people involved; list of publications, related to SUADA; and information for science contests, conferences ect. The web page was developed in WordPress content management system, which facilitates editing.

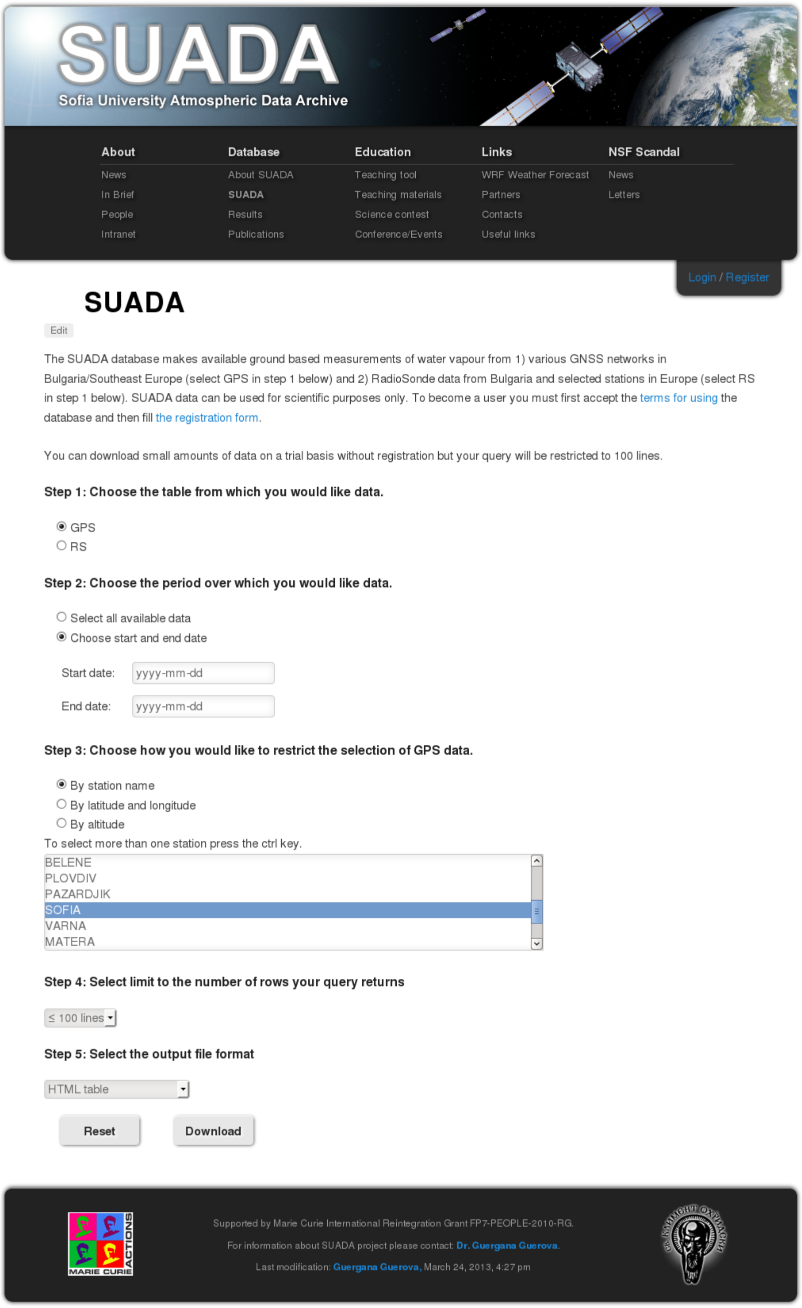

Figure 2.3: "SUADA" sub-menu as presented on the SUADA web portal.

2.3.1 Information tables for the web portal

The informational tables for the web portal are the tables with information for the web portal *http://suada.phys.uni-sofia.bg/?page_ id=964*. They are FIELD_ DEFINITION, USERS and LOG. They serve as a interface between the data users and the database.

Table FIELD_ DEFINITION table 2.16 contains information about field abbreviations in all other database tables. It is required by the users for understand the names of the fields.

Table field	Data type	Summary
FieldName	varchar(32)	Table field name as defined in the database
Description	varchar(128)	Field name description
Units	varchar(8)	Field value unit
UserVisible	tinyint(1)	Table fields visible by the users defined in USERS table 2.17

Table 2.16: FIELD_ DEFINITION table field definition.

Table USERS 2.17 contains contact information about the SUADA data users both internal and external, with data access rights defined in INSTRUMENT table 2.4. This table is connected to the "Login" section of the SUADA web portal.

Table field	Data type	Summary
Username	varchar(32)	User name, identifying the person with access rights
Firstname	varchar(32)	First name of the user
Surname	varchar(32)	Surname of the user
Email	varchar(64)	Organization e-mail
Phone	varchar(32)	User's phone number
Address	varchar(128)	User's contact address

Institution	varchar(128)	User's home institution
Access	int(11)	Shows user authorization status
Pwd	blob	User's password, coded
Timestamp	timestamp	The last time when this entry was altered
Cookie	blob	File stored on the user's side for autonatic log-in
note	varchar(10000)	Field for special notes
ID	int(11)	Unique ID number for each user

Table 2.17: USERS table field definition.

Table LOG 2.18 contains information about the log-in time of the sessions, initiated by users.

Table field	Data type	Summary
ID	int(11)	Session ID number
Type	int(11)	Session type
UserID	int(11)	ID number of the user. Linked with a FOREIGN KEY to table USERS
Note	blob	Note
Date	timestamp	Recording of the date and time of the beginning of the session

Table 2.18: LOG table field definition.

2.4 SUADA data upload

The SUADA database is maintained and updated mainly using scripts. For the data upload a set of bash scripts is used. Their role is to extract the necessary data files from ftp servers on the internet, download them locally on the PHYSON cluster, select data from them and upload the structured data into the database tables. The bash scripts are the backbone of the uploading. They are used for the downloading of external files and for uploading the data into SUADA via MySQL queries. The awk and sed scripts are used for selecting the necessary data from the external data files and arranging the data in a readable for MySQL form.

A simplified example of a bash script for uploading data is enclosed in the Appendix 1.

2.5 SUADA data processing

The relations between the tables, presented in chapter 2.2 suggest the existence of processing data from one table and uploading the calculated values into another. This is done by using MATLAB scripts. MATLAB is a numerical computation language, visualization and programming tool (*MathWorks*, 2011). It has a distinctive advantage over C as it does not require variable declaration and the ease of manipulating large data arrays. For the needs of SUADA the MATLAB Database toolbox is used. It comprises special functions, which allow the database access. One of the advantages of the Database toolbox are the functions, reading date and time in MySQL format (yyyy-mm-dd HH:MM:SS).

In the project MATLAB is used not only for data processing but also as powerful visualization tool producing the following figures: 5.1 5.2, 5.3, 5.4, 5.5, 5.6, 5.7.

In Appendix 2 an example of MATLAB script, used for data processing is presented.

Chapter 3

Source: GPS

3.1 GNSS Meteorology

The concept of GNSS Meteorology was suggested by Bevis (*Bevis et al.*, 1992). The propagation of the GNSS signal through the atmosphere is affected by the atmospheric gases (*Tralli and Lichten*, 1990). The magnitude of the atmospheric effects depends on several factors: on the composition of the atmosphere; on the elevation of the receiver (thus on the thickness of the atmosphere); on the elevation angle of the satellite and on the amount of water vapour, which is mostly dependent on the atmospheric conditions.

When the signal is traveling from the satellite to the receiver, it first gets delayed by the ionized gases in the upper atmosphere (100-500 km above the surface of the Earth). This delay can be measured and can be used for sounding the properties of the ionosphere. The processes in the ionosphere a very dynamic because of the sun-sensible daily ion content fluctuation. After the ionosphere the signal travels through denser, non-ionized layers of the atmosphere.

The troposphere, the lowest and densest part of the atmosphere, consists of a mixture of nitrogen, oxygen, noble gases, trace gases and water vapour. All these gases are responsible for the total tropospheric delay. The total tropospheric delay in direction zenith is called ZTD.

There are two contributing factors for the total delay. They are the hydrostatic and the wet delay. The hydrostatic delay is caused by all the gases in the atmosphere, except the water vapour. They are the main contributor to the positioning uncertainty. The hydrostatic delay in direction zenith is called ZHD and is relatively stable in a day timescale. It can be easily derived, using its dependency on the local atmospheric pressure:

$$ZHD = (2.2768 + 0.0024)\frac{p_s}{f(h,\theta)} \tag{3.1}$$

$$f(h,\theta) = 1 - 0.00266cos(2\theta) - 0.00028h \tag{3.2}$$

where p_s is local surface pressure and $f(h,\theta)$ is a factor, dependent on height h and the latitude variation of the gravitational acceleration θ.

The second contributing factor to the ZTD is the ZWD. It is caused by the water vapour in the atmosphere. The ZWD has a large temporal variation in a hour timescale. This is the reason, why the GNSS derived IWV is so valuable with its high temporal resolution. The ZWD contributes less then 10% of the ZTD. ZWD and IWV can be calculated with the use of these expressions:

$$ZWD = ZTD - ZHD. \tag{3.3}$$

$$IWV = \frac{10^6}{(k_3/T_m + k_2')R_v}ZWD \tag{3.4}$$

where k_2, k_3 and R_v are constant and T_m is the weighted mean atmospheric temperature.

For the SUADA database processing surface pressure and temperature of the atmosphere are obtained from the synoptic observing network operated by the National Institute of Meteorology and Hydrology. The GNSS data was obtained from several sources. They are the International GNSS Service, EUREF and ZenithGEO. One additional processing was made by *Keranka Vassileva* for July 2007 heat wave.

The pressure corrections between the altitudes of the meteorological and the GNSS stations are calculated, using the US Standard Atmosphere (*Sissenwine et al.*, 1962). The formula for correcting pressure between different altitudes is:

$$p_c = p_o(1 - 0,000002257(z_g - z_m)^{5,255} \tag{3.5}$$

where p is the pressure correction, p_o is the pressure at sea level (1013,5 hPa), z_m is the height of the meteorological station and z_g is the height of the GNSS station. The more accurate barometric formula (*Ahrens*, 2008) will be used:

$$p = p_o e^{(-\frac{g}{RT}(z_g - z_m))} \tag{3.6}$$

where g is the gravitation constant, R is the universal gas constant and T is average temperature of the air between the stations in Kelvin.

In addition gravitation correction is applied:

$$e_f = 1 - 0,00266cos(2\lambda) - 0,00028H \tag{3.7}$$

where λ[radians] is the GNSS station latitude and H[km] is the GNSS station altitude.

Then the pressure at the height of the GNSS station is determined as follows:

$$P = \frac{c.p_c}{e_f} \tag{3.8}$$

where c is a empirically determined constant (*Bevis et al.*, 1992).

A temperature correction due to the height difference is calculated using the dry adiabatic gradient:

$$\Delta T = 0.0065 * (z_m - z_g); \tag{3.9}$$

3.2 International GNSS Service

The International GNSS Service (IGS) is a scientific consortium of scientists from over 200 worldwide national agencies, universities and research institutions in more than 80 countries divided in 6 regional centers. It computes and maintains a global network of over 350 permanent, continuously operated GNSS sites.

IGS provides data for ZTD from its establishment in 1994 (*Beutler* *et al.*). It started processing GNSS data for station Sofia (SOFI) from 1997. The SUADA contains continuous data GNSS data for station Sofia from the beginning of the IGS processing in 1997 to 2012.

IGS dataf from the second reprocessing *(chapter3.2.2)*, uploaded into the SUADA, is data, which was processed by the Center for Orbit Determination in Europe (CODE), Astronomical Institute, University of Bern. Only the GPS system was used for the processing. Data from other centers is intended to be uploaded into the database as well.

3.2.1 IGS first reprocessing campaign

In 2008 a reprocessing campaign was initiated by IGS (*Rebischung et al.*, 2012). The campaign was part of a new IGS08 reference frame, which was a substitution for the previous IGS05. The reprocessing was needed because of the simultaneous adoption of a new set of antenna phase center calibrations for 65 out of 232 stations. Further more the reprocessing was an upgrade, which followed the newer 2008 International Terrestrial Reference Frame.

SUADA stores data for station Sofia (SOFI) from 1997 to 2007 from the first IGS reprocessing campaign. This data has been processed by the CODE center in Bern University, Switzerland.

3.2.2 IGS second reprocessing campaign

In 2011 the second reprocessing campaign was initiated by IGS (*Meindl et al.*). During the campaign data from 1993 onward was processed, data with temporal resolution of 30 seconds from 1996. The 13 to 25 % increase in precision is achieved by improved antenna calibration of the GNSS receivers, updated solar radiation model and better orbit and clock products.

Several analysis centers contributed to the campaign. They are CODE from Switzerland; ESA, GFZ and PDR from Germany; EMR from Canada and SIO, MIT, JPL and NGC from the USA. The data, which is used in the SUADA is processed by the CODE center in Bern University, Switzerland.

In addition in SUADA six European stations from the 2011 reprocessing campaign were uploaded into the database. They are Zimmerwald (ZIMM), Switzerland; Onsala (ONSA), Sweden; Ondrejov (GOPE), Czech Republic;

Medicina (MEDI), Italy; Matera (MATE), Italy; Potsdam (POTS), Germany and Sofia(SOFI), Bulgaria. Stations Zimmerwald, Onsala, Potsdam and Matera are reference stations for IGS in Europe. These stations cover several different climate zones. Moreover all the European stations are situated to the west of Bulgaria, on the way of the downwind of air masses. The main objective of selecting these stations is that they have been processed consistently for a long period of time. Station Sofia is the only GNSS station in Bulgaria, processed by IGS.

3.2.3 IGS Balkan stations processing

Several IGS Balkan stations were also uploaded into the database. These stations are from the permanent IGS network, but are not processed using IGS software. They are Dubrovnik (DUBR), Athens (NOA1), Thessaloniki (AUT1), Craiova (CRAI), Constanta (CONS), Bucharest (BUCU) and Varna (VARN). The data, uploaded for these stations is for the period of July 19-26, 2007. This is the period of a heatwave over the Balkan Peninsula. The analysis of the data showed the short-term breeze circulation, present in the seaside stations (*Simeonov et al.*, 2013).

GPS data from 19 GNSS permanent stations (AUT1, NOA1, BUCU,COST, DUBR, GLSV, GRAZ, MATE, ORID, PENC, POLV, ROZH, SOFI, SULP, MIKL, WTZR, ZIMM, VARN, CRAI) from Central and Eastern Europe were processed with the Bernese software, version 5.0 for the period 19-25 July, 2007. Sixteen of them are IGS (International GNSS Service) and EPN (EUREF Permanent network) stations. Seven sessions of 24 hours have been created. Daily station coordinates and station troposphere zenith delays in every 1 hour of each session are estimated. The troposphere model used is Saastamoinen dry model with Niell dry mapping and tilting gradient model. Corrections to the introduced zenith values have been estimated and finally the total zenith delays have been obtained as well as gradient parameters in North and East directions. These daily tropospheric files with estimated zenith total delays have been uploaded into the database and then used for estimation of IWV.

3.3 EUREF

Figure 3.1: EUREF permanent network. Active stations in green, former stations in red.

The European Reference Frame (EUREF) is a European network operating since 1995 with objective to provide a standard precise GNSS-based reference system for Europe. As of April 15, 246 permanent GNSS tracking stations are part of EUREF permanent network. On the Balkan Peninsula there are 15 EUREF stations: 5 stations in Greece and Romania each and 1 station in Turkey, Croatia, Macedonia, Slovenia and Bulgaria, totaling 15 stations. The station Sofia (SOFI) is part of the EUREF permanent network 1997. This remains the only GNSS station in Bulgaria, which is part of the IGS and EUREF permanent networks.

EUREF coordinates several analysis centers in Europe. The data, uploaded into SUADA has been processed by the BKG (Bundesamt für Kartographie und Geodäsie) center in Germany. The data is available between 2001 and 2004.

3.4 ZenitGEO

The Bulgarian ZenitGEO network was established in 2009. It consists of 30 GNSS stations, evenly distributed over Bulgaria (figure 3.2).

The ZenitGEO process the GNSS data with the commercial Trimble software and provides tropospheric files with very high temporal resolution of 5 min (300 s). Currently, processed are 11 stations namely: Vidin, Oryahovo, Lovech, Velika Tarnovo, Ruse, Razgrad, Silistra, Shabla, Kyustendil, Pazardzhik and Sliven, from which 8 in north Bulgaria. Future processing of GNSS data using NWP models will be considered.

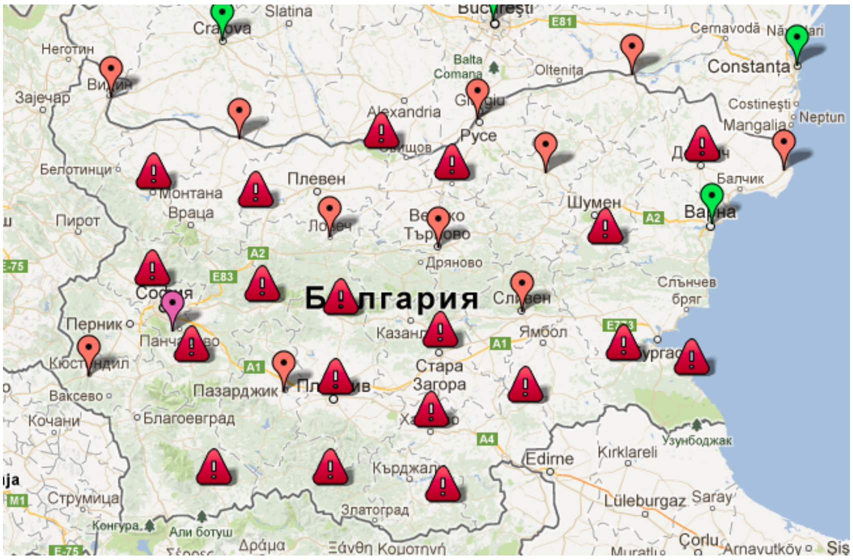

Figure 3.2: ZenitGEO network. Processed stations with red markers, unprocessed stations with red triangles, other network stations with green and purple markers.

The stations of figure 3.2, marked with red triangles (!) are ZenitGEO stations, which are uploaded into SUADA, but not processed due to lack of meteorological data. The stations, marked with a red marker are processed using synoptic data from the National Institute of Meteorology and Hydrology. The stations, marked with green markers are from the IGS08 Balkan stations reprocessing. Sofia station is not part of the ZenitGEO network.

Chapter 4

Radiosonde and SYNOP data

4.1 Radiosonde measurements and data

Radiosonde probing is an in situ measurement method of the atmosphere. The method is approved by the World Meteorological Organization (WMO) and is widely adopted for temperature, pressure, humidity, wind speed and direction, as well as other profile measurements in the atmosphere.

Sofia station of the National Institute of Meteorology and Hydrology performs daily radio sounding observations in 12 UTC. Since 2005 VAISALA RS92KL probe has been used. The temperature sensor measures from +60 °C to -90°C with resolution of 0,1°C . The barometer measures from 1080hPa to 3hPa with resolution of 0,1-1,5hPa. The relative humidity sensor is a thin-film capacitor heated twin sensor with measurement range between 0 and 100%, resolution 1% and total uncertainty in sounding 5%.

Although radiosonde producers claim 1% resolution, when the humidity in the atmosphere drops below 20%, the total uncertainty increases dramatically (*Morland et al.*, 2009)

As far as radio sounding is used as a verification method for the GNSS derived IWV in the atmosphere, the relative humidity data has to be transformed into IWV following the equation:

$$IWV = \frac{1}{\rho_w} \int_{h_0}^{h_{top}} \rho_{wv}(h) dh \tag{4.1}$$

where h_0 is the altitude of the station, where the probe is released, h_{top} is the maximum achieved height by the probe during sounding, ρ_w is the density of water, ρ_{wv} is the density of water vapour (*Guerova et al.*, 2003). IWV is measured in millimeters.

4.2 SYNOP measurements and data

As discussed in chapter 2.2.3, synoptic data is also uploaded into the database. The data is acquired from *OGIMET*, which is an online weather information server.

The surface synoptic observations (SYNOP) are formatted using the *WMO* data coding rules. They are interpreted by the scripts, presented in chapter 2.4. The measurements are done at least 8 times a day. Measurements in synoptic stations are done simultaneously in *00, 03, 06, 09, 12, 15, 18, 21 GMT*. The meteorological data, which is collected from every station is: 10 meter above surface wind direction and speed; surface pressure; 2 meter temperature; dew point and other more specific measurements. The meteorological parameters, uploaded into SUADA are 2 meter temperature, surface pressure and humidity. The other measured values are not necessary for the GNSS Meteorology method.

It is to be noted, that every national met office has its own practices in collecting meteorological information. Data from Bulgarian stations is collected manually every 3 hours. The data from Romanian stations is hourly and obtained from automatic weather stations. This is the reason for the different time resolution, provided for each processed location in the database.

4.3 Model - WRF

The Weather Research and Forecasting (WRF) model is a mesoscale model, which allows parallel computation. It has been jointly developed by the National Center for Atmospheric Research (NCAR), the Forecast Systems Laboratory and the National Centers for Environmental Prediction of the National Oceanic and Atmospheric Administration (FSL, NCEP/NOAA) and the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma. The spatial resolution of the model is between 1 and 10 km. Numerous specific models, such as the Hurricane Weather Research and Forecasting (HWRF) have been created upon WRF. From the first release in 1990 until now the model has evolved (*Michalakes et al.*, b). Additional packages have been created for interactive nesting, upgraded physics, three dimensional data assimilation and simplified parallelization (*Michalakes et al.*, a).

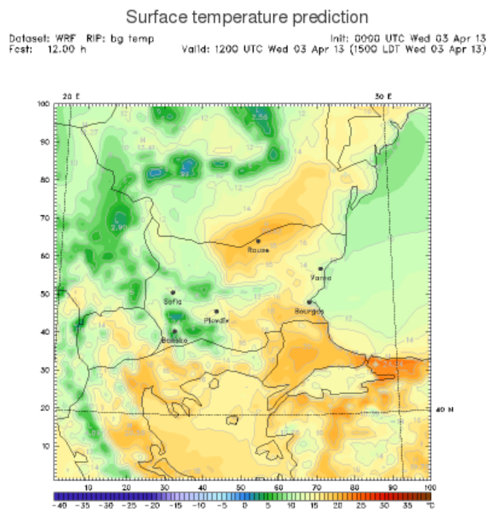

Figure 4.1: WRF generated forecast chart for temperature,

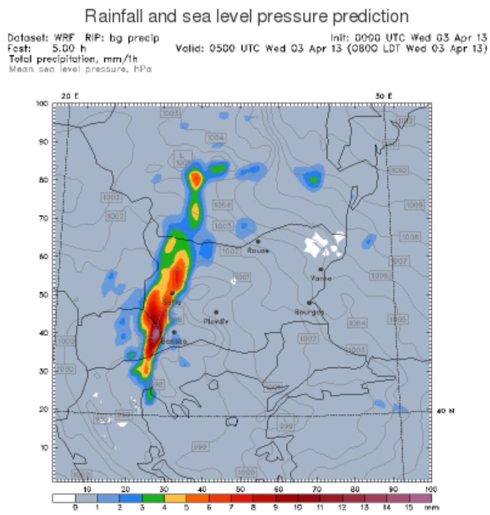

Figure 4.2: WRF generated forecast chart for rainfall and pressure.

Since February 2013 the WRF, version 3, is running on the PhysOn cluster. It is used for temperature, pressure and rainfall forecasting over Bulgaria with a spatial resolution of 9,7 km or 0,1 °.

The data from the model is intended to be included into the database, especially for computing surface temperature and pressure over Bulgaria for the needs of GNSS Meteorology processing.

The WRF model is operational and provides 24 hour weather forecast, computed two times a day. Figure 4.1 shows the temperature forecast for 12UTC on the 3rd of April 2013. The rainfall and sea level pressure forecasts are presented on figure 4.2.

Chapter 5

Case studies

5.1 Heat wave: 2007

Data, stored in SUADA was used to analyze the 2007 heat wave over Bulgaria and Southeast Europe (*Simeonov et al.*, 2013). Heat waves are severe meteorological events with temperature anomalies of over 5°C for at least five days. A temperature increase of 1°C is followed by an increase in the atmospheric water vapour by 7%.

Monthly mean IWV from GNSS station SOFI is shown in figure 5.1. When compared to 2001-2006 observations (dashed line), the following features stand out in 2007 (solid line): i.) IWV decrease in April, ii.) IWV increase in May and iii.) IWV decrease in July. The 2001-2006 mean IWV (column 2 in table 5.1) is 15.0 mm versus 14.2 mm in 2007, which is an annual decrease of -5 %. The seasonal mean IWV for December and February (DF), March to May (MAM) and June to August (JJA) are given in columns 4 to 9 of table 5.1. During the 2007 winter, the IWV decreased by -3 % (-0.3 mm). The 2001-2006 mean winter was -0.9°C colder that 2007 (fig. 5.2). The 2007 spring IWV (fig. 5.1 bottom) overall is close to the 2001-2006 mean. In the 2007 summer, the IWV decreased by 1.3 mm or by about -6 %. The 2007 summer temperature at SOFI elevation (1 119 m asl) increased by 1.8°C compared to 2001-2006 mean. In particular, July 2007 was +3.7°C hotter and with -18 % less water vapour than the 2001-2006.

Column 1-3 of table 5.1 is station name, annual mean for 2001-2006

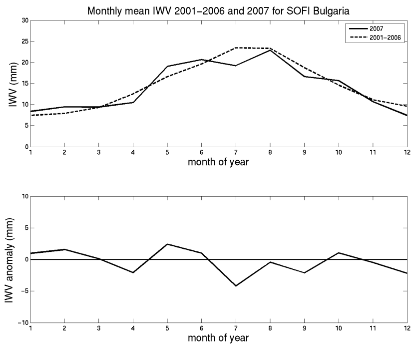

Figure 5.1: Monthly mean IWV (top) and anomaly (bottom) for SOFI, Bulgaria (thick line 2007, dashed line 2001-2006).

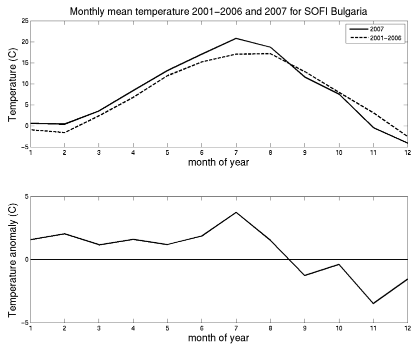

Figure 5.2: Monthly mean temperature (top) and anomaly (bottom) for SOFI, Bulgaria (thick line 2007, dashed line 2001-2006).

Met station	2001-2006	2007	2001-2006	2007	2001-2006	2007	2001-2006	2007
IWV	mm	mm	DF	DF	MAM	MAM	JJA	JJA
SOFI	15.0	14.2	8.8	8.5	12.9	13.0	22.2	20.9
Change		(-5 %)		(-3 %)		(+1 %)		(-6 %)
Temperature	[$^\circ$C C]	[$^\circ$C C]	DJF	DJF	MAM	MAM	JJA	JJA
SOFI	8.4	8.8	-2.2	-1.3	7.1	8.4	16.5	18.7
Change		(+5 %)		(+41 %)		(+18 %)		(+13 %)

Table 5.1: IWV & Temperature at station SOFI, Bulgaria.

period and 2007 accordingly; column 4-5: December and February (DF) mean for 2001-2006 and 2007; column 6-7: March, April and May (MAM) mean for 2001-2006 and 2007; column 8-9: June, July and August (JJA) mean for 2001-2006 and 2007. The 2007 departure from 2001-2006 mean is given in % in the brackets.

The diurnal cycle of the temperature and IWV during the heat wave is presented in figures 5.3 and 5.4. At the Black sea coastal stations Constanta and Varna (figure 5.3) the IWV minimum is around 00 UTC and the peak is at 15 UTC, i.e., 3 hours after the temperature peak. The IWV peak at 15 UTC can be explained with the peak of the sea breeze circulation that brings humid sea air inland. Similarly, the IWV peak at 15 UTC is also characteristic for the Mediterranean sea station Athens and Adriatic sea station Dubrovnik.

Morland et al. (2009) study the diurnal climatology from 2003 to 2007 for station Bern and conclude that the IWV diurnal cycle peak occurs about 6 hours later, than the daily temperature maximum. At the inland station Sofia (figure 5.4) the IWV minimum is at 6 UTC and a broad peak is seen between 12-18 UTC. It is to be noted that the GNSS station in Sofia located in a mountainous area outside of the city at altitude of 1120 m msl. which makes it suitable for study of local IWV circulation in mountainous regions(*Simeonov et al.*, 2013).

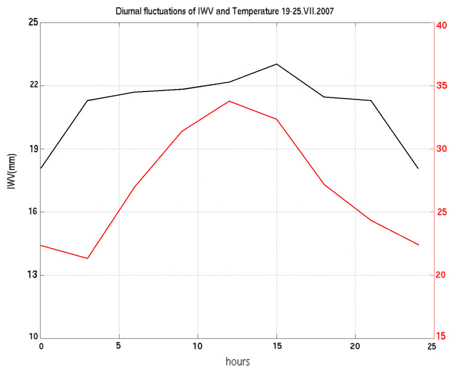

Figure 5.3: Daily IWV cycle for station Varna during July 2007 heat wave.

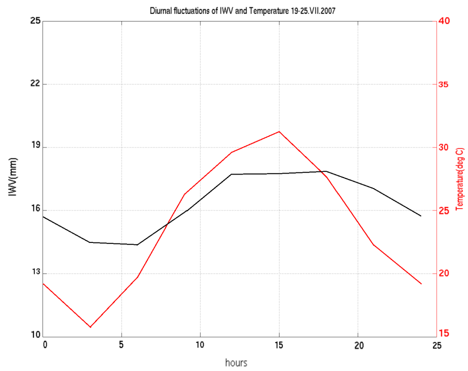

Figure 5.4: Daily IWV cycle for station Sofia during July 2007 heat wave.

5.2 Validation EUREF/IGS1/IGS2

Station Sofia is a key station for the whole database. The GNSS data, collected from it has been processed by many GNSS services in Europe. This makes it possible to compare results from the different centers in order to better establish the dynamics of the climatic and meteorological processes, relative to the station.

On figure 5.5 is presented the IWV yearly cycle, using data from EUREF (in blue color), IGS repro 1 (in green color) and IGS repro 2 (in red color). The temperature and pressure data for the three curves are from the same source, so the difference comes only from the different calculation of ZTD. The shift in the measured IWV between is due to the different processing techniques, used in the GNSS centers. The EUREF postprocessing scheme is calculating the delays 14 days after the data harvesting. IGS uses a consistent scheme of reprocessing campaigns, where large timescales are processed simultaneously, which increases the data precision.

The EUREF data is processed only for the 2001-2004 period. This is

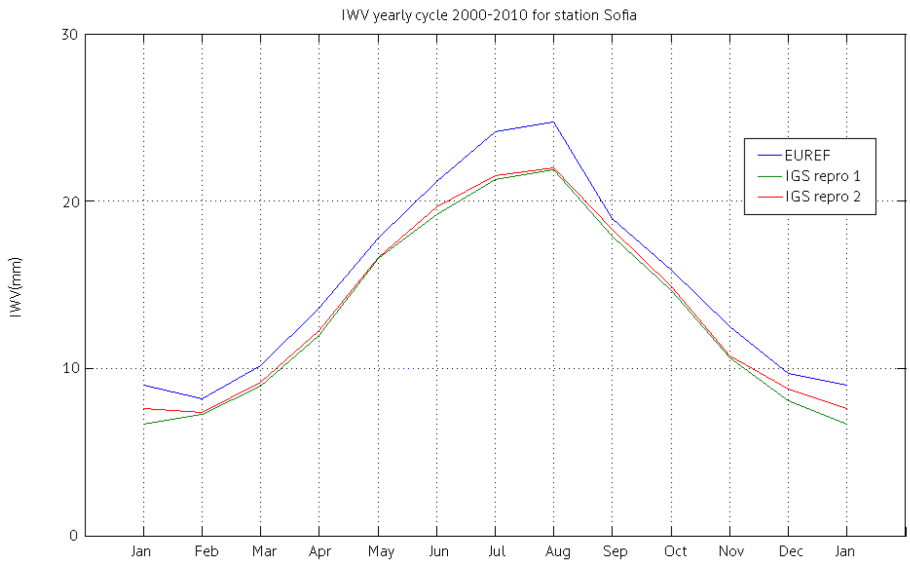

Figure 5.5: EUREF/IGS1/IGS2 comparison.

one of the reasons for the differences between the curves. IGS repro 1 data is from 2000 to 2007. The IGS repro 2 data is from 2000 to 2010.

The average IWV from the EUREF processing is 15,29mm, while from the IGS repro 1 is 14,56mm and from IGS repro 2 is 14,16mm.

On figure 5.6 the IWV yearly cycle for the 2001-2004 period is presented (in black), compared with the 2000-2007 yearly cycle for IGS first reprocessing (in red). Similarly on figure 5.7 the IWV yearly cycle for the 2001-2004 (in blue) period is compared with the 1999-2009 yearly cycle for IGS second reprocessing. For 2001-2004 period the average IWV from IGS repro 1 is 14,86mm and from IGS repro 2 is 14,97mm.

This 3 year period is too short for climatological analysis, but shows insignificant (up to 3%) differences between the processed IWV from these three sources.

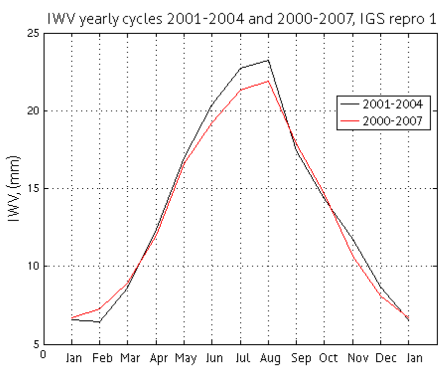

Figure 5.6: IGS repro 1 yearly 2001-2004 and 2000-2007 cycles.

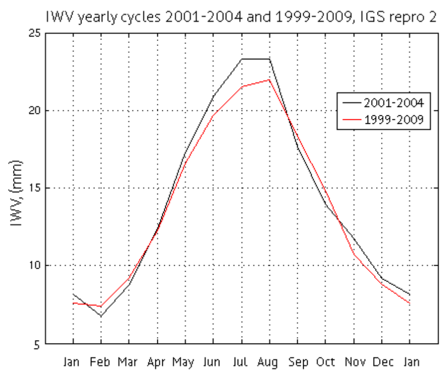

Figure 5.7: IGS repro 2 yearly 2001-2004 and 1999-2009 cycles.

Chapter 6

Conclusions

The SUADA was created to store atmospheric data from standard synoptic measurements, radiosonde measurements and GNSS derived water vapour. The database is created using MySQL. The data is structured in 16 tables, containing raw GNSS, meteorological data and processed GNSS data. The data is uploaded into the database using bash scripts. The processing of the data is done using MATLAB programs with inbuilt SQL queries. Stored information in the database is accessible through the SUADA web portal.

The data, uploaded into the database was obtained from global meteorological and GNSS databases. More than 12 million data lines have been uploaded into the database so far. The data is from 32 stations on the territory of Bulgaria, 7 stations on the Balkan peninsula and 6 other stations in Europe.

Data from SUADA has been used to analyze the 2007 July heat wave over the Balkans, as well as to create 2D charts of the IWV distribution over Bulgaria for case studies. The variety of sources of information enables further comparison between the sources.

The SUADA is in constant development. Model data are intended to be uploaded into the database in order to be validated with measurements. The WRF model data are going to be used for surface temperature and pressure estimation for stations, which lack meteorological data. A change in the pressure correction scheme is planed.

Bibliography

Egvap project, http://egvap.dmi.dk/.

Ogimet weather information service, http://www.ogimet.com.

World bank, http://iridl.ldeo.columbia.edu/maproom/.Global/.World_Bank/.Climate_Variability/index.html, climate Change Data Portal.

Wyoming atmospheric database, http://weather.uwyo.edu/.

Ahrens, D., *Meteorology Today, 9th edition*, 624 pp., Brooks Cole, 2008.

Beutler, G., M. Rothacher, T. Springer, J. Kouba, and R. Neilan, The internation gps service (igs): an interdisciplinary service in support of earth sciences, in *32nd COSPAR Scientific Assembly, Nagoya, Japan*.

Bevis, M., S. Businger, T. Herring, C. Rocken, R. Anthes, and R. Ware, Gps meteorology: Remote sensing of atmospheric water vapor using the global positioning system, *JGR*, *97*, 15 787–15 801, 1992.

Codd, E., A relational model of data for large shared data banks, *Communications of the ACM*, *13, issue 6*, 377–387, doi:10.1145/362384.362685, 1970.

Guerova, G., E. Brockmann, J. Quiby, F. Schubiger, and C. Matzler, Validation of nwp mesoscale models with swiss gps network agnes, *Journal of Applied Meteorology*, *42*, 141–150, 2003.

MathWorks (Ed.), *MATLAB 7 Getting Started Guide*, 2011.

Meindl, M., R. Dach, and Y. Jean, International gnss service technical report 2011, *Tech. rep.*, IGS Central Bureau.

Michalakes, J., J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, and W. Wang, The weather reseach and forecast model: Software architecture and performance, in *Proceedings of the 11th ECMWF Workshop on the Use of High Performance Computing In Meteorology*, a.

Michalakes, J., J. Dudhia, D. Gill, J. Klemp, and W. Skamarock, Design of a next-generation regional weather research and forecast model, *World Scientific, River Edge, New Jersey*, b.

Morland, J., Startwave database documentation, *Tech. rep.*, University of Bern.

Morland, J., and C. Matzler, Spatial interpolation of gps integrated water vapour measurements made in the swiss alps, *Meteorological applications*, *14*, 15–26, 2007.

Morland, J., B. Deuber, D. G. Feist, L. Martin, S. Nyeki, N. Kampfer, C. Matzler, P. Jeannet, and L. Vuilleumier, The startwave atmospheric water vapour database, *Atmospheric Chemistry and Physics*, *6*, 2039–2056, doi:10.1029/2008JD010110, 2006.

Morland, J., M. C. Coen, K. Hocke, P. Jeannet, and C. Matzler, Tropospheric water vapour above switzerland over the last 12 years, *Atmospheric Chemistry and Physics*, *9*, 5975–5988, 2009.

Rebischung, P., J.Griffiths, J. Ray, R. Schmid, X. Collilieux, and B.Garayt, Igs08: the igs realization of itrf2008, *GPS Solutions*, *16*, 483–494, 2012.

Simeonov, T., K. Vasileva, and G. Guerova, Application of ground-based gnss meteorology in bulgaria/southeast europe: case study 2007 heat wave., *Annual of Sofia University*, *106*, 2013.

Sissenwine, N., M. Dubin, and H. Wexler, The u.s. standard atmosphere, *Journal of Geophysical Research*, *67, issue 9*, 3627–3630, dOI: 10.1029/JZ067i009p03627, 1962.

Tralli, D. M., and S. M. Lichten, Stochastic estimation of tropospheric path delays in global positioning system geodetic measurements, *Bulletin géodésique*, *64, issue 2*, 127–152, doi:10.1007/BF02520642, 1990.

Acknowledgments

I would like to thank to:

- My project and thesis supervisor Dr. Guergana Guerova for the enthusiastic help she gave me whenever possible
- My team mates Nikola Irinchev for helping me with MySQL and Rositsa Topchiyska for her leading role in creating the SUADA web site and connecting it to the database
- Dr. Stoyan Pisov for providing the team with technical support for the PhysOn cluster, where the database and the website are stored and where the database processing is made
- Dr. Keranka Vassileva for the July 2007 heat wave GNSS processing
- ZenitGEO for the provided GNSS data