

---

Unterschrift des Betreuers



TECHNISCHE  
UNIVERSITÄT  
WIEN

Vienna University of Technology

# DIPLOMARBEIT

## COMPARING GRACE WITH CCI SOIL MOISTURE DATA ON A GLOBAL SCALE

Ausgeführt am  
Department für Geodäsie und Geoinformation  
der Technischen Universität Wien

unter der Anleitung von

**Univ.-Prof. Dipl.-Ing. Dr.techn. Johannes Böhm**

Forschungsgruppe Höhere Geodäsie

und

**Senior Scientist Dr.rer.nat. Wouter Arnoud Dorigo, MSc**

Forschungsgruppe Fernerkundung

durch

**Valerie Nusime**

Matr. Nr.: 1026552

Fellingersiedlung 38

4400 Steyr

Wien, Mai 2016

---

Valerie Nusime

## Abstract

Water storage and water storage changes have become a topic of increasing interest, because they can be used as an indicator of climate change impacts. *Total water storage* (TWS) is understood as the sum of groundwater, *soil moisture* (SM), surface water and snow. *Changes in TWS* (TWSC) can be detected using spaceborne gravimetric measurements. The *Gravity Recovery and Climate Experiment* (GRACE) Tellus Land dataset provides estimates of TWSC over land. Surface SM estimates can be derived from microwave remote sensing. The *Climate Change Initiative Soil Moisture* (CCI SM) project combines different microwave remote sensing products to one SM dataset mapping the whole world.

The aim of this thesis is to evaluate the correspondence between the CCI SM and the GRACE Tellus dataset using the Pearson correlation coefficient on a global scale. Precipitation datasets (GPCC and GPCP), the ERA-Interim/Land dataset and a map of area equipped for irrigation will help to understand the results of the comparison.

As expected both datasets (CCI SM and GRACE Tellus) match each other well, indicated by positive correlation coefficients. Most of the values are between 0.6 and 0.8 with no lag between the datasets. A lag of one month results in higher values (between 0.8 and 1). In the very north of the northern hemisphere and Saudi Arabia negative correlation coefficients are predominant (values between  $-0.6$  and  $-0.3$ ). For the northern hemisphere this can be explained by the fact that the CCI SM dataset only indicates liquid water. During colder periods snow and ice lead to an increase in TWS, but a decrease in SM. The decrease in SM is caused by the fact that ice and snow are masked in the CCI SM datasets.

The area surrounding Saudi Arabia shows some contradictory results, considering the different datasets used. Especially the fact that TWS increases during warmer periods and decreases in colder periods needs to be reviewed critically.

Concluding the CCI SM dataset can help to understand the water cycle, especially in combination with the GRACE Tellus dataset on a global scale. It is important to take into account that the mentioned areas need some further research to understand

the main drivers for the resulting negative correlation coefficients.

## Kurzfassung

Das Interesse am globalen Wasserspeicher und dessen Änderungen steigt immer mehr an, da der Wasserspeicher als Indikator für klimatische Veränderung verwendet werden kann. Der totale Wasserspeicher (TWS) ist die Summe aus Grundwasser, Bodenfeuchte, Oberflächenwasser und Schnee. Die Änderungen des TWS (= *changes in TWS* TWSC) können von gravimetrischen Satellitenmessungen abgeleitet werden. Der *Gravity Recovery and Climate Experiment* (GRACE) Tellus Land Datensatz stellt die TWSC der Kontinente dar. Die Bodenfeuchtigkeit kann hingegen mittels Mikrowellenfernerkung festgestellt werden. Das *Climate Change Initiative Soil Moisture* (CCI SM) Projekt erzeugt einen Datensatz, welcher die Kombination aus unterschiedlichen Mikrowellenfernerkundungsprodukten ist.

Das Ziel dieser Arbeit ist es, den Zusammenhang zwischen dem CCI SM Datensatz und dem GRACE Tellus Datensatz zu evaluieren. Für diesen Zweck werden die Pearson Korrelationskoeffizienten berechnet. Niederschlagsdatensätze (GPCC und GPCP), der ERA-Interim/Land Datensatz sowie ein Datensatz, der die Fläche, welche für Bewässerung ausgerüstet ist, darstellt, werden für ein besseres Verständnis der Ergebnisse herangezogen.

Die beiden Datensätze (CCI SM und GRACE Tellus) stimmen in allen Region der Welt gut überein, dies lässt sich aus den positiven Korrelationskoeffizienten ableiten. Die Werte liegen meist zwischen 0.6 und 0.8, ohne Verzögerung. Die Verzögerung des GRACE Tellus Datensatzes um 1 Monat führt zu höheren Werten (zwischen 0.8 und 1). Im nördliche Teil der nördlichen Halbkugel und Saudi Arabien herrschen negative Korrelationskoeffizienten vor (zwischen  $-0.6$  und  $-0.3$ ). Auf der nördlichen Halbkugel lässt sich dies durch die Tatsache erklären, dass die CCI SM Datensätze nur flüssiges Wasser im Boden anzeigen. In kälteren Perioden führen Schnee und Eis zu einem Anstieg im TWS, wohingegen die Bodenfeuchte sinkt. Der Grund des Absinkens der SM ist, dass Eis und Schnee im CCI SM Datensatz nicht dargestellt werden.

Zieht man alle verwendeten Datensätze hinzu, ergeben sich für Saudi Arabien widersprüchliche Ergebnisse. Besonders die Tatsache, dass der TWS während der warmen

Jahreszeit ansteigt und in der kühleren Jahreszeit absinkt, muss noch näher untersucht werden.

Abschließend ist zu sagen, dass der CCI SM Datensatz helfen kann, den Wasserkreislauf über den Kontinenten zu verstehen, besonders wenn er mit dem GRACE Tellus Datensatz verbunden wird. Es muss aber berücksichtigt werden, dass jene Gebiete, die negative Korrelationskoeffizienten aufweisen, noch näheren Untersuchungen unterzogen werden müssen um den TWS besser zu verstehen.

# Acknowledgement

I would like to express my gratitude to Univ.Prof. Dipl.-Ing. Dr.techn. Johannes Böhm and Senior Scientist Dr. rer.nat. Wouter Dorigo, MSc. for all their support and comments, throughout the whole thesis and that they always found a moment to help me with my concerns.

I want to sincerely thank my parents for always having a sympathetic ear at any day- or nighttime and for their attempt to find a solution whenever I am clueless. Without their support -in every sense- my years of studying would not have been as enjoyable and carefree. I also want to thank my sister who gave me a feeling of secureness, from my first day on in Vienna.

Last but not least I want to sincerely thank my friends and fellow students who made this time so special to me.

# List of Figures

1.2	The main processes related to the water cycle: evapotranspiration, surface runoff, precipitation, groundwater flow and evaporation; found on the NASA website . . . . .	3
1.3	This illustration displays the relation between TWS, precipitation and irrigation. . . . .	3
2.1	GRACE sensors/instruments, including the <i>star camera assembly</i> (SCA), <i>Global Positioning System</i> (GPS) instruments, the <i>K-Band-Ranging system</i> (KBR) and the <i>SuperSTAR accelerometer</i> (ACC); found on the CSR website . . . . .	8
2.2	GRACE satellite mission concept: the two twin satellites (GRACE 1 and GRACE 2) using GPS satellites for positioning. Between the satellites the a K- and Ka-Band distance measurement technique is symbolized as a red dashed line; found on the CRS website . . . . .	9
2.3	The GRACE Tellus data for April 2002 in [cm]. . . . .	15
2.4	Satellites and sensors, which are used for collecting the CCI SM data. SMAP and Aquarius are not included yet, but the possibility to do that exists; from Dorigo et al. [2012] . . . . .	17
2.5	This is the CCI SM dataset for April 2002, with a spatial resolution of 1° in longitude and latitude direction. . . . .	18
2.6	ERA-Interim/Land data for April 2002 for the first layer (0.00-0.07m), on a spatial 1°x1° grid. . . . .	21
2.7	The GPCC dataset for April 2002, as it can be downloaded on the DWD website. . . . .	23
2.8	The GPCP anomalies for April 2002. The daily values are merged to gain a temporal resolution of one month. Afterwards the mean over the timespan from January 2004 to December 2009 is subtracted. . .	25
2.9	Schematic representation of the methodology used for mapping AEI; by (Siebert et al. [2013]) . . . . .	26

2.10	Irrigation for the year 2005 for the whole world in AEI in [%] of total area . . . . .	27
4.1	correlation of GRACE vs CCI with no lag ( $k = 0$ ) . . . . .	31
4.2	The correlation coefficients of the GRACE Tellus dataset compared to the GPCC Full Data Reanalysis dataset with a spatial resolution of 1 degree. Between the datasets no lag was applied ( $k = 0$ ). . . . .	32
4.3	The correlation of CCI SM dataset compared to GPCC Full Data Reanalysis dataset with a spatial resolution of 1 degree. Between the datasets there is no lag ( $k = 0$ ) . . . . .	33
4.4	The correlation coefficients of the GRACE Tellus dataset compared to the GPCP 1DD dataset, with no lag ( $k = 0$ ) . . . . .	34
4.5	The correlation coefficients of the comparison between the CCI SM dataset and the GPCP 1DD dataset, with no lag ( $k = 0$ ) . . . . .	35
4.6	Correlation coefficients between the GRACE Tellus dataset and the ERA-Interim/Land dataset for all layers. No lag ( $k = 0$ ) was added. . . . .	36
4.7	Correlation coefficients of the GRACE Tellus dataset compared to the ERA-Interim/Land dataset for Layer 1 (0.00-0.07) only. The mean over the timespan 2004-2009 was subtracted from the ERA-Interim/Land dataset. The lag between the datasets was set to zero. . . . .	37
4.8	The correlation of CCI SM dataset compared to ERA-Interim/Land dataset. Only Layer 1 (0.00-0.07m) was utilized to calculate the correlation. Between the datasets there is no lag ( $k = 0$ ) . . . . .	38
5.1	A comparison of the GRACE Tellus TWS in [ $cm$ ] and the CCI SM in [ $m^3/m^3$ ] for a longitude of $11^\circ$ degrees and a latitude of $8^\circ$ (in the east of Nigeria). The time period ranges from April 2002 to December 2014. This site shows a high positive correlation. The red line stand for the CCI SM dataset and the blue one for the GRACE Tellus (TWS) dataset. . . . .	40
5.2	The correlation coefficients between the GRACE Tellus and the CCI SM datasets. The GRACE Tellus dataset lags one month behind CCI SM ( $k = 1$ in equation 3.3). . . . .	41

5.3	A comparison of the GRACE Tellus TWS [ $cm$ ] and the CCI SM [ $m^3/m^3$ ] for a longitude of $21^\circ$ degrees and a latitude of $66^\circ$ (in the North of Sweden) from April 2002 to the end of 2014. At this specific location the correlation is somewhere close $-1$ for zero lag. The red line stands for the CCI SM dataset and the blue one for the GRACE TWS. . . . .	42
5.4	A comparison of the GRACE Tellus TWS [ $cm$ ] and the CCI SM [ $m^3/m^3$ ], located at for a longitude of $21^\circ$ degrees and a latitude of $66^\circ$ (in the North of Sweden) for the period January 2007 to December 2008. A negative correlation dominates at this region of interest. The red line stands for the CCI SM dataset and the blue one for the GRACE Tellus dataset. The CCI SM is given as anomalies of SM. . .	43
5.5	Juxtaposition of the GPCP and GPCC datasets to the CCI SM dataset for the period January 2007 to December 2008. The CCI SM dataset is depicted in blue, the GPCP dataset in red and the GPCC dataset in green. . . . .	44
5.6	Juxtaposition of the GPCP and GPCC datasets to the GRACE Tellus dataset for the period January 2007 to December 2008. The GRACE Tellus dataset is shown in blue, the GPCP dataset in red and the GPCC dataset in green. All three datasets depict anomalies based on the mean over the period January 2004 to December 2009. . . . .	45
5.7	Comparison of he GRACE Tellus (red) dataset and the maximum temperature [blue] for the site in the northern part of the northern hemisphere for the years 2007 and 2008. . . . .	46
5.8	Comparison of the precipitation datasets (GPCP - red, GPCC - green) and the ERA-Interim/Land top layer (blue) for a latitude of $66^\circ$ and a longitude of $21^\circ$ . . . . .	47
5.9	Comparison of the CCI SM dataset and the GRACE Tellus dataset, for January 2007 to December 2008. The region of interest is in Saudi Arabia, at a latitude of $20^\circ$ and a longitude of $41^\circ$ . GRACE Tellus is shown in blue and the anomalies of the CCI SM dataset are shown in red. . . . .	48

5.10	A comparison of the precipitation datasets and the GRACE Tellus dataset is given for January 2007 to December 2008. The region of interest is in Saudi Arabia, at a latitude of $20^\circ$ and a longitude of $41^\circ$ . The GPCP (red) and the GPCC (green) dataset displayed anomalies, referenced to the timespan January 2004 to December 2009. The GRACE Tellus dataset is symbolized in blue. . . . .	49
5.11	A confrontation of the GPCC (green) and GPCP (red) dataset against the CCI SM (blue) dataset for January 2007 to December 2008. The point of interest is located at a latitude of $20^\circ$ and a longitude of $41^\circ$ . . . . .	49
5.12	The trend of the TWSC (red) is increasing over the timespan 2002 to 2014. The blue line symbolizes the GRACE Tellus TWS changes. . . . .	51
5.13	The trend of the CCI SM trend (red) is decreasing over the timespan 2002 to 2014. The blue line symbolizes the SM (CCI SM). . . . .	51
5.14	The trend of the TWSC mapped for the whole world. A red area means that the TWSC is increasing, and a dark blue one means that it is decreasing. If the TWSC is uniformly than it is shown in green. . . . .	52
5.15	The relationship between the trend of the TWSC and the areas irrigated with groundwater ([%] of AEI). Only areas with more than 40% are selected. A positive trend is coloured red, a negative blue. The areas irrigated with groundwater are symbolized as the brighter areas over land. . . . .	53
5.16	On this map the correlations between the CCI SM and GRACE Tellus datasets are overlaid by an irrigation dataset, showing each AEI where more than 40%] of total area are equipped. . . . .	53
A.1	The correlation coefficients between the CCI SM and GRACE Tellus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	62
A.2	The correlation coefficients between the CCI SM and GRACE Tellus datasets for different lags. The CCI SM dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	63
A.3	The correlation coefficients between the GPCC and GRACE Tellus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	64

A.4	The correlation coefficients between the GPCC and GRACE Tellus datasets for different lags. The GPCC dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	65
A.5	The correlation coefficients between the GPCC and CCI SM datasets for different lags. The CCI SM dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	66
A.6	The correlation coefficients between the GPCC and CCI SM datasets for different lags. The GPCC dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	67
A.7	The correlation coefficients between the GPCP and GRACE Tellus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	68
A.8	The correlation coefficients between the GPCP and GRACE Tellus datasets for different lags. The GPCP dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	69
A.9	The correlation coefficients between the GPCP and CCI SM datasets for different lags. The CCI SM dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	70
A.10	The correlation coefficients between the GPCP and CCI SM datasets for different lags. The GPCP dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	71
A.11	The correlation coefficients between the ERA-Interim/Land all Layers and GRACE Tellus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . .	72
A.12	The correlation coefficients between the ERA-Interim/Land all Layers and GRACE Tellus datasets for different lags. The ERA-Interim/Land dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . .	73
A.13	The correlation coefficients between the ERA-Interim/Land Layer 1 and GRACE Tellus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . .	74
A.14	The correlation coefficients between the ERA-Interim/Land Layer 1 and GRACE Tellus datasets for different lags. The ERA-Interim/Land dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . .	75
A.15	The correlation coefficients between the ERA-Interim/Land Layer 1 and CCI SM datasets for different lags. The CCI SM dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	76

A.16	The correlation coefficients between the ERA-Interim/Land Layer 1 and CCI SM datasets for different lags. The ERA-Interim/Land dataset lags corresponding to the given $k$ -values ( $k = 1, 2, \dots, 6$ ). . . . .	77
B.1	A comparison of both precipitation datasets (GPCC and GPCP) and the CCI SM dataset. The site is located at a latitude of $8^\circ$ and a longitude of $11^\circ$ . The precipitation datasets are coloured in red (GPCP) and green (GPCC) in units of $[mm/month]$ . The CCI SM dataset is shown in blue in $[m_3/m_3]$ . . . . .	78
B.2	A comparison of both precipitation datasets (GPCC and GPCP) and the GRACE Tellus dataset. The site is located at a latitude of $8^\circ$ and a longitude of $11^\circ$ . The precipitation datasets are coloured in red (GPCP) and green (GPCC) in units of $[mm/month]$ . The GRACE Tellus dataset is shown in blue in $[cm]$ . . . . .	79
B.3	A comparison of both precipitation datasets (GPCC and GPCP) and the GRACE Tellus dataset, for loaction in Saudi Arabia (latitude of $20^\circ$ and longitude of $41^\circ$ ). The anomalies of the precipitation datasets are coloured in red (GPCP) and green (GPCC) in units of $[mm/month]$ . The GRACE Tellus dataset is shown in blue in $[cm]$ . . . . .	79
C.1	The AEI for Saudi Arabia. The red marker indicates the site with a latitue of $20^\circ$ and a longitude of $41^\circ$ , which chosen for the comparison the the different datasets (5). The figure was generated using Google Maps and GMIA. The AEI are the coloured areas, mostly in the southern part auf Saudi Arabia. . . . .	80
C.2	The table represents the area actually irrigated per month in [%]. From January to May and in December the number is increasing correlated to the other month. This table can be found on the FAO website <sup>1</sup> . . . . .	81
C.3	The correlation coefficients between CCI SM and GRACE Tellus for Saudi Arabia are overlaid by a dataset showing areas using groundwater for irrigation. . . . .	82

# List of Tables

2.2	Current orbital parameters for the GRACE satellites (29-April 2016 07:37am) and the orbital parameters at the mission start in 2002 . . .	7
-----	---	---

# Contents

<b>List of Figures</b>	<b>vi</b>
<b>List of Tables</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The hydrological cycle in a changing climate . . . . .	1
1.2 Monitoring hydrological changes from space . . . . .	2
1.3 Structure of Thesis . . . . .	5
<b>2 Data &amp; Processing</b>	<b>6</b>
2.1 GRACE . . . . .	6
2.1.1 Previously performed processing steps . . . . .	10
2.1.2 GRACE Tellus . . . . .	14
2.2 ESA CCI Soil Moisture . . . . .	16
2.2.1 Previously performed processing steps . . . . .	17
2.2.2 CCI SM v02.2 . . . . .	18
2.3 ERA-Interim . . . . .	19
2.3.1 Previously performed processing steps . . . . .	19
2.3.2 ERA-Interim/Land . . . . .	20
2.4 Precipitation products . . . . .	20
2.4.1 GPCC . . . . .	21
2.4.2 GPCP . . . . .	22
2.5 GMIA . . . . .	24
2.5.1 Previously performed processing steps . . . . .	26
2.5.2 GMIA Version 5 . . . . .	27
<b>3 Correlation function</b>	<b>28</b>
<b>4 Results</b>	<b>30</b>
4.1 GRACE Tellus compared to CCI SM . . . . .	30

---

4.2	GPCC compared to GRACE Tellus & CCI SM . . . . .	31
4.2.1	GPCC compared to GRACE Tellus . . . . .	32
4.2.2	GPCC compared to CCI SM . . . . .	33
4.3	GPCP compared to GRACE Tellus & CCI SM . . . . .	33
4.3.1	GPCP compared to GRACE Tellus . . . . .	34
4.3.2	GPCP compared to CCI SM . . . . .	34
4.4	ERA-Interim/Land compared to GRACE Tellus & CCI SM . . . . .	35
4.4.1	ERA-Interim/Land compared to GRACE Tellus . . . . .	36
4.4.2	ERA-Interim/Land compared to CCI SM . . . . .	38
<b>5</b>	<b>Discussion</b>	<b>39</b>
<b>6</b>	<b>Summary &amp; Outlook</b>	<b>55</b>
6.1	Summary . . . . .	55
6.2	Outlook . . . . .	56
	<b>Bibliography</b>	<b>57</b>
<b>A</b>	<b>Map of correlation coefficients</b>	<b>62</b>
<b>B</b>	<b>Comparison of the datasets</b>	<b>78</b>
<b>C</b>	<b>Irrigation</b>	<b>80</b>

# Chapter 1

## Introduction

### 1.1 The hydrological cycle in a changing climate

In the last decades climate change has become a topic of increasing interest, not only in different fields of science, but to the public in general. As a result of climate change extreme weather phenomena like droughts and floods have become common all over the world (Cheng et al. [2012], Leitner [2015]). This affects the life of all creatures be it man or animal, in some regions to a small extent, in others with wide influence.

In this context, mapping water storage and especially the changes of water storage, as an indicator of climate change constitute an important issue. A lot of different measurement techniques have been developed to derive the physical parameters, which can subsequently be related to water storage. Some of those techniques are in-situ measurements, which give a good understanding on a smaller scale, some are airborne or from outer space, with the advantage that the whole world, even regions which might be difficult to reach, can be mapped. A very obvious spaceborne technique to map soil moisture and precipitation is used in projects like the *Climate Change Initiative Soil Moisture* (CCI SM), where microwave instruments yield those physical parameters.

Another term often related to soil moisture is *total water storage* (TWS). TWS includes groundwater, soil moisture, surface water, snow and ice. Groundwater is defined as the water which is positioned in aquifer, whereas soil moisture is located in near surface soil layers. Surface water is understood as water at the surface, like lakes and rivers (Figure 1.2).

It is a well-known fact that TWS is influenced by changes in climate and the resulting weather and the other way round as explained in Arnell [1999]. The great water cycle is shown in Figure 1.1, which includes many sub-cycles and processes, like

evaporation, precipitation, evapotranspiration or groundwater flow. In this context evaporation is any water of lakes, rivers, seas or the soil surface that enters the atmosphere as vapour. Evapotranspiration is the sum of evaporation from soil surface and transpiration from plant leaves.

The *changes in TWS* (TWSC) are made obvious by using the satellite mission *Gravity Recovery and Climate Experiment* (GRACE).

Irrigation is an issue of considerable interest, as it mostly influences top soil layers only. Depending on the manner of irrigation used it may, however, also influence groundwater reservoirs if those are used as a water dispenser. Concerning the relationship between TWS and SM, both are often influenced by similar processes.

Figure 1.2 illustrates how TWS, precipitation and irrigation are connected and influenced by each other. SM can be differentiated in surface SM, which can be sensed by microwave satellites and root-zone SM, which cannot be sensed by microwave satellites. In case of precipitation soil moisture is increasing. This leads to an increase in TWS as well. Usually, evaporation concerns the top soil layers only and causes a decrease of soil moisture and as a consequence TWS decreases as well. Irrigation influences the soil layers and the source of water. The source of water may be surface water (e.g. lakes, rivers), groundwater (aquifer, wells) or non-conventional sources (e.g. desalinated water) as described in Siebert et al. [2010]. The irrigated area will have an increased soil moisture. Depending on the chosen source of water (groundwater), it will probably reduce the stored water. This means that water from aquifers used for irrigation will be “lost” due to evaporation and evapotranspiration into the atmosphere.

If groundwater is the wanted parameter derived from TWS, all the other components like snow, surface water and SM need to be removed, as described in Rodell and Famiglietti [2002] and Rodell and Famiglietti [2001]. Rodell and Famiglietti [2001] show that, compared to surface water and snow, only SM and groundwater have a significant impact on TWS.

## 1.2 Monitoring hydrological changes from space

Rodell and Famiglietti [2002] studied the potential of GRACE to estimate changes in groundwater, the High Plains aquifer being the region of interest. On the other hand, Rodell and Famiglietti [2001] compared averaged TWSC from ground measurements to uncertainty estimates related to the TWSC detected by GRACE. Averaged TWS is understood as a combination of four different datasets: one for groundwater, one

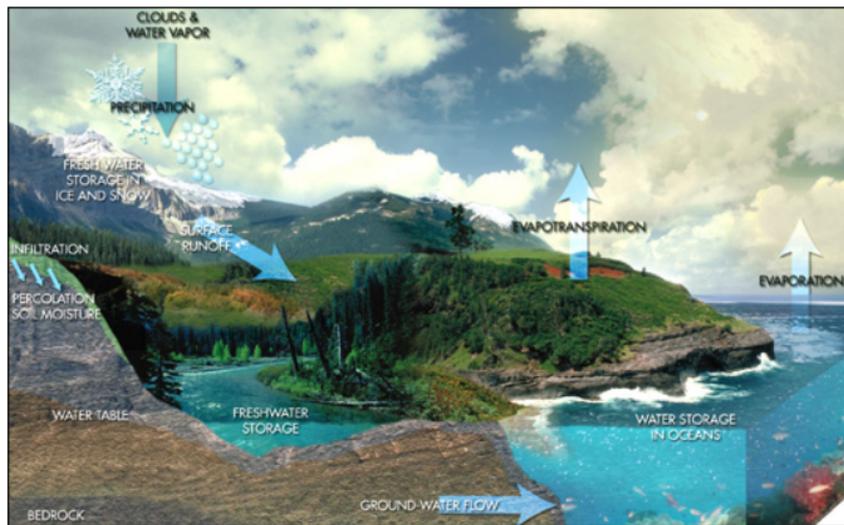


Figure 1.1: The main processes related to the water cycle: evapotranspiration, surface runoff, precipitation, groundwater flow and evaporation; found on the NASA website<sup>a</sup>.

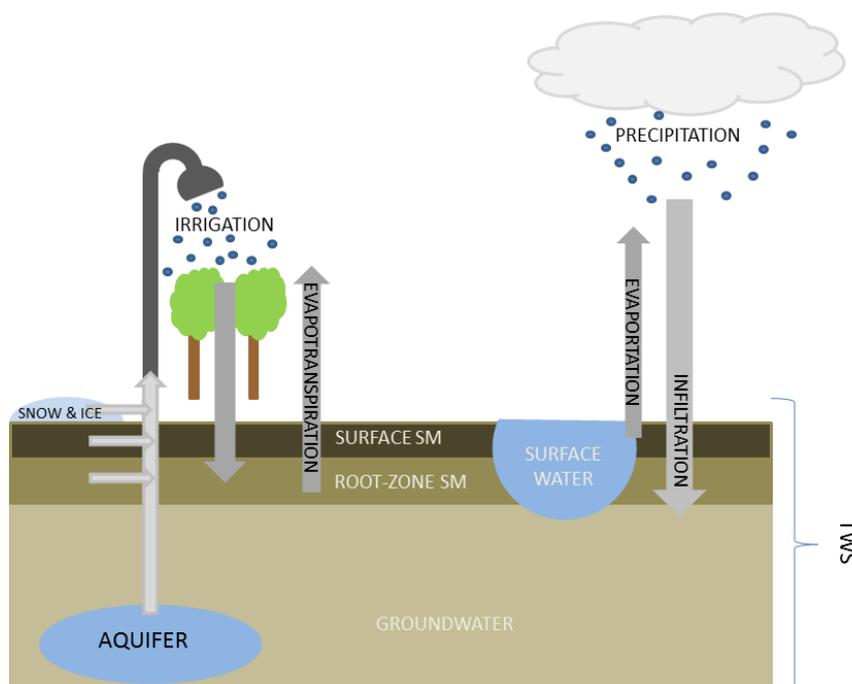


Figure 1.2: This illustration displays the relation between TWS, precipitation and irrigation.

<sup>a</sup><http://earthobservatory.nasa.gov/Features/WeighingWater/weighing2.php>

for SM, one for snow depth, snow fall and precipitation and one for the observations of reservoirs.

Several studies have already linked TWSC and SM, for example the comparison of SM and TWS in the La Plata Basin by Abelen et al. [2015]. In this study different SM datasets were checked against each other. Furthermore, GRACE TWSC were compared to the satellite SM datasets *Advanced Scatterometer* (ASCAT) and *Advanced Microwave Scanning Radiometer - EOS* (AMSR-E), which reflect only the near-surface SM. In contrast, the *WaterGAP Global Hydrology Model* (WGHM) which maps the SM in the root zone was used in this study. Additionally, the *International Disaster Database* (EM-DAT) was utilized to show the connection between hydrologic extremes and natural disasters. Since the datasets have different units (partly) representing different physical parameters, the comparison between those datasets was drawn utilizing the cross-correlation function. The results of the study by Abelen et al. [2015] reveal that the SM datasets fit quite well to each other as well as SM to precipitation in most parts of the La Plata Basin with a few exceptions only. TWS reacts with a lag of some months to droughts and floods, as indicated by the SM datasets and as well as by the EM-DAT. However, it was not possible to relate TWS to natural disasters recorded by the EM-DAT, whereas it was feasible to map those disasters in the SM datasets. In this study it is also pointed out that a high correlation coefficient does not necessarily mean that soil moisture is the dominant part of TWS, but it may be linked closely.

In another study by Allam et al. [2016] different SM datasets as well as TWSC were compared for the assessment of regional evapotranspiration. The region of interest was the Blue Upper Nile Basin, which was surveyed over a longer period of time from 2002 to 2013. In this study GRACE TWSC were compared to the CCI dataset. The result shows that there is a lag of about two months. Allam et al. [2016] explained that the reason for that lag is that the CCI SM dataset shows the SM for the top layers only and that is why the peaks follow the rainy season.

Abelen and Seitz [2013] compared SM products (derived from remote sensing and hydrological modelling) and TWSC (derived from GRACE) on a global scale with the aim of providing possibilities and benefits of comparing the datasets. They used SM products from ASCAT satellites and the WGHM model. A data harmonization was applied to all datasets and an analysis was made, whether this harmonization has an impact on the correlation coefficient results. This was followed by calculating the correlation coefficients between the different datasets. The result of this study is that high correlation coefficients between SM and TWS can be expected in regions

with a high precipitation rate, whereas arid regions show a correlation coefficient lower than 0.5, including negative values. Moreover, the study demonstrates that GRACE data can help to understand the TWS and its relation to SM, and that it can be used for validating SM products. What it excludes, however, are data points which might be affected by post-glacial rebound.

The aim of this thesis is a better understanding of the interaction between SM and TWS or its components. For this reason the GRACE Tellus and the CCI SM dataset are compared by an correlation analysis on a global scale, as well as to find possible reasons why the data may not be appropriate in some regions. Most of the previous studies concentrate just on a specific region of interest. To further pinpoint the causes of agreement and disagreement between the datasets, *Global Precipitation Climatology Centre* (GPCC), *Global Precipitation Climatology Project* (GPCP) and ERA-Interim/Land datasets as well as a dataset related to irrigation are added and compared to GRACE Tellus, CCI SM or the correlation coefficients between those two.

### 1.3 Structure of Thesis

Chapter 2 gives an overview of the datasets used. Information about the satellite missions and the required processing steps which have to be applied to retrieve either SM or TWSC are given. This is followed by an explanation of the products, that are actually utilized in this thesis.

The mathematical and statistical background for calculating the correlation coefficients and filtering those are given in Chapter 3.

Chapter 4 then presents the results of the correlation analysis. In Chapter 5 a deeper understanding and a discussion on the results is given. The thesis is completed by a short summary (Section 6.1) and an outlook for further research topics (Section 6.2).

# Chapter 2

## Data & Processing

This chapter gives an outline of the different datasets used in this thesis. Each section will contain an overview of the satellite mission, the configuration and the aim of the mission, followed by a description of the processing steps and the handling of the used datasets.

### 2.1 GRACE

The aim of the *Gravity Recovery and Climate Experiment* (GRACE) mission is to accurately map the variable Earth gravity field - with a spatial resolution of 400 km to 4,000 km as mentioned by Tapley et al. [2004]. The mission is a cooperation between the *National Aeronautics and Space Administration* (NASA) and the *Deutsche Forschungsanstalt für Luft und Raumfahrt* (DLR). Two identical satellites are flying tandem, which are separated by a distance of about 200 km. Changes in the Earth's gravity field (depending on the solid and fluid Earth) as well as the atmosphere have an effect on the distance between the satellites. The fact that one satellite follows the other, incurred them the nicknames "Tom" and "Jerry".

The satellites are at an altitude of around 450km and the orbits have an inclination of  $89.5^\circ$  (near-circular orbits). A slow decay (around  $30m/day$ ) of the satellites is caused by the atmospheric drag on them. As a result the current orbital parameters differ from those given at the mission start. The current orbital parameters can be found on the *Center for Space Research Texas* (CSR) website<sup>1</sup>. In Table 2.1 the current orbital parameters (29-April 2016 07:37am) are compared with those given at the mission start.

The period of circulation is about 90 minutes, this leads to a sampling time for the whole Earth of about one month. This accounts for the circumstance that changes in

---

<sup>1</sup><http://www.csr.utexas.edu/grace/ground/>

Table 2.1: Current orbital parameters for the GRACE satellites (29-April 2016 07:37am) and the orbital parameters at the mission start in 2002

<b>Orbital parameters</b>	<b>current</b>	<b>mission start</b>
semi-major axis [km]	6740.132501	6873
eccentricity	0.000926	0.00225
inclination [°]	59.035	89
argument of perigee [°]	96.767	96.2
right ascension of ascending node [°]	20.660	277.26
true anomaly [°]	299.663	156.5
mean anomaly [°]	299.755	456.4

the gravity field with a variation of less than one month cannot be detected properly. Due to the altitude decay the twin satellites have no fixed ground repeat rate.

Tapley et al. [2004] describe that for measuring the changes of the distance between those satellites a one way K- and Ka-band microwave ranging system (24 and 32 GHz) is used - this constellation is called "low-low satellite tracking". Using two frequencies leads to an ionosphere-free, high-precise distance measurement. It is possible to measure changes in distance with an accuracy as small as  $10\mu m$ . Both satellites are equipped with the same ranging system, except for the frequencies. To avoid mixing-up the transmitted and received signal, the frequencies are shifted by 500KHz, described in Mayer-Gürr et al. [2006]. The reference frequency is generated by the *Ultra Stable Oscillator* (USO). The microwave assembly is used for converting the frequency (up-converting the reference frequency, down-converting the received one). For transmitting and receiving the signal a horn antenna is mounted on each of the twin satellites. The horn antennas are pointing toward the other satellite. This means that the leading satellite is flying backwards, while the trailing satellite is flying forwards. This causes different signs of wear and tear and because of that the satellites switched their position in 2005. The satellite, which was the leading one until 2005, has become the trailing one.

For retrieving accurate positions and for time tagging *Global Positioning System* (GPS) receivers, known as "high-low satellite tracking", are mounted on the satellites. Mayer-Gürr et al. [2006] point out that with the help of the GPS receivers atmospheric and ionospheric profiling is possible. *Star cameras* (SCA) deliver the orientation of the satellites relative to stars. If the Sun or the Moon reaches the field of view (of one star camera), measurements cannot be obtained and the data is collected using the other star camera. These instruments (GPS and SCA) as well

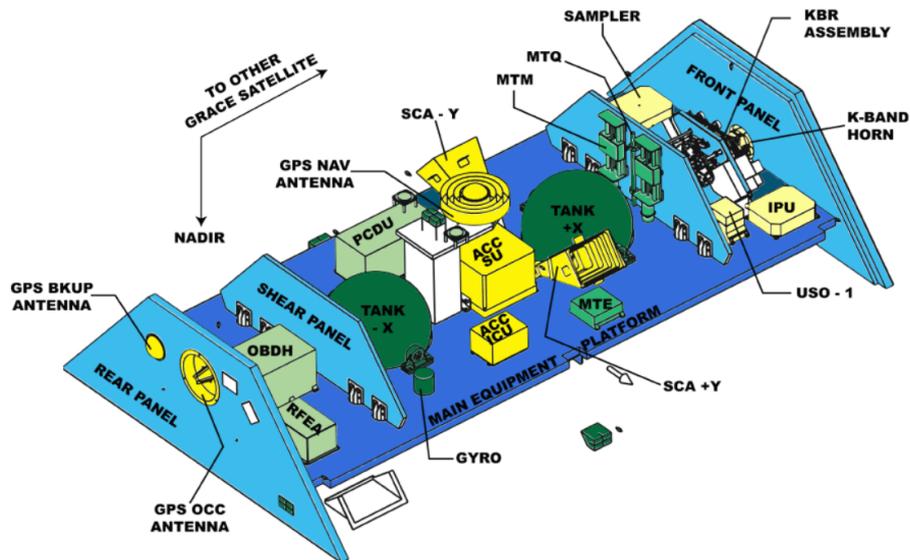


Figure 2.1: GRACE sensors/instruments, including the *star camera assembly* (SCA), *Global Positioning System* (GPS) instruments, the *K-Band-Ranging system* (KBR) and the *SuperSTAR accelerometer* (ACC); found on the CSR website<sup>2</sup>

as the *Coarse Earth and Sun Sensors* (CESS), a three-axis boom-mounted Förster magnetometer and a three-axis Inertial Reference Unit are the sensors of the *Attitude and Orbit Control System* (AOCS). AOCS provides information about the spacecraft attitude and is used for the orbital control of the satellites. The CESS is used for an omni-directional, robust, but coarse Sun and Earth tracking, utilized during the “safe mode”.

Furthermore, a Laser-Corner-Cube-Reflector Assembly is used to obtain precise orbit information and to calibrate the on-board GPS receivers. Additionally, use was made of *SuperSTAR accelerometer* (ACC) for measuring the non-gravitational accelerations, for example solar radiation and air drag. The measurement of the proof mass motion in the *Sensor Unit* (SU) indicates the non-gravitational accelerations. The accelerometer is located at the satellite’s mass centre in order to avoid acceleration offsets. The *centre of mass trim assembly* (CMT) is able to measure and correct the offset between the satellite centre of mass and the accelerometer proof mass. Via an S-Band Boom, mounted on the bottom side (looking at the Earth), the recorded data is sent back to Earth for processing.

Figure 2.1 illustrates an internal view of the GRACE satellite sensors. The outer shells of the satellites are covered with silicon-solar cell for generating power.

<sup>2</sup><http://www.csr.utexas.edu/grace/spacecraft/a1.html>

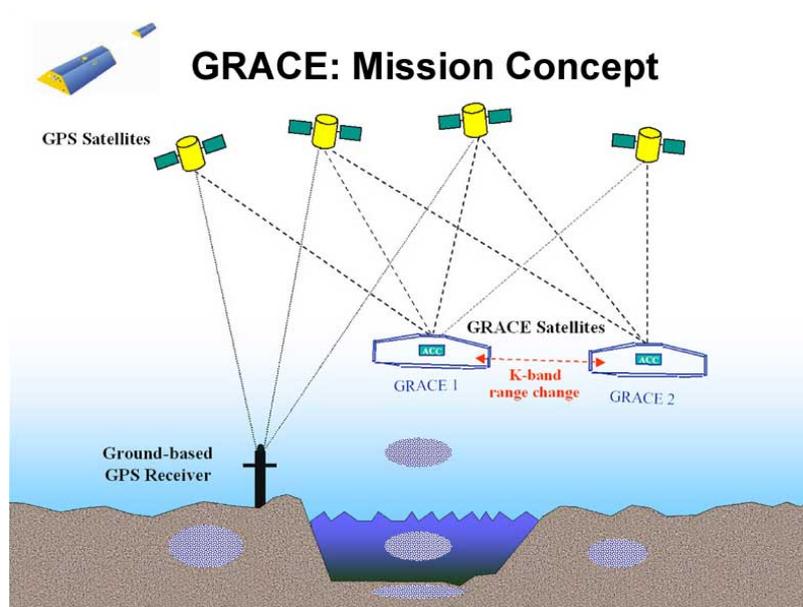


Figure 2.2: GRACE satellite mission concept: the two twin satellites (GRACE 1 and GRACE 2) using GPS satellites for positioning. Between the satellites the a K- and Ka-Band distance measurement technique is symbolized as a red dashed line; found on the CRS website <sup>3</sup>

Figure 2.2 displays the mission concept, including the twin satellites, the GPS satellites and some mass changes within the Earth (blue ellipses). The distance measurement technique is symbolized in red. The twin satellites were launched in March 2002 with a planned mission duration of around 5 years. In the last few years data gaps occur due to low battery<sup>4</sup>.

Because of the great achievements a *GRACE-follow on* (GRACE-FO) mission is planned with a launch in 2017. The new mission will use *Laser-Range-Interferometry* (LRI) to measure the variations in the distance between the satellites more accurately, as described in Sheard et al. [2012].

The collected data are provided as spherical harmonics coefficients by the *Geoforschungszentrum Potsdam* (GFZ), the CSR and the *Jet Propulsion Laboratory* (JPL), which all use different models for processing the raw data. The data is divided into various levels of processing: the raw data (Level-1), monthly gravity field estimations (Level-2) and mass anomalies (Level-3).

In this thesis the Level-3 *GRACE Tellus Land* (GRACE Tellus) dataset was used, which is described in Section 2.1.2.

<sup>2</sup><http://www.csr.utexas.edu/grace/publications/presentations/HPC2001.html>

<sup>4</sup>The mission status can be found on the CRS website: [http://www.csr.utexas.edu/grace/operations/mission\\_status/](http://www.csr.utexas.edu/grace/operations/mission_status/), last visit : April 2016

### 2.1.1 Previously performed processing steps

To obtain the Earth's gravity field from the measured changes in distance between the satellites, short term mass variations in the atmosphere and in the oceans need to be considered and removed. These mass changes cause variations in the distance between the satellites and need to be removed if the Earth's gravity field is wanted. Flechtner [2007] mentions two possible methods to reduce the effects of the atmospheric mass variations. This processing step is known as the atmospheric de-aliasing. The easiest technique is the *thin layer approximation* (TL). A more complex technique is the *vertical integration of atmospheric column* (VI). Three numerical weather services deliver the input parameters: the *Deutscher Wetter Dienst* (DWD), the *National Centre for Environmental Prediction* (NCEP) and the *European Centre for Medium-range Weather Forecasts* (ECMWF).

Assuming that most of the atmospheric mass variations are located in the lower 10km, the TL concentrates the atmospheric mass in one thin layer. So the vertical distribution is neglected as declared in Karbon et al. [2013]. Surface loads indicate density changes in the atmosphere and they are defined as mass per surface element  $q$ . The Earth's global gravity field (including solid and fluid Earth) is usually given in terms of spherical harmonic coefficients. The changes of the gravity field are

$$\begin{aligned}\Delta C_{lm} &= \frac{a^2}{(2l+1)M} \iint_{Earth} q \tilde{P}_{lm}(\cos\theta) \cos(m\phi) dS \\ \Delta S_{lm} &= \frac{a^2}{(2l+1)M} \iint_{Earth} q \tilde{P}_{lm}(\cos\theta) \sin(m\phi) dS\end{aligned}\tag{2.1}$$

with

$$dS = \sin\theta d\theta d\phi\tag{2.2}$$

In Equation (2.1)  $a$  states the radius of the sphere (on the surface it is equivalent to the Earth radius  $r$ ), and  $\theta$  and  $\phi$  the colatitude and east longitude.  $M$  the Earth mass,  $\tilde{P}_{lm}$  the normalized associated Legendre functions of degree  $l$  and order  $m$  and  $dS$  the surface element.

If  $q$  is defined by using the surface pressure  $p_s$

$$q = \frac{p_s}{g}\tag{2.3}$$

where  $g$  is the mean gravity acceleration and the elastic deformation is taken into account (using the Load Love numbers  $k_l$ ), Equation (2.1) changes to

$$\begin{aligned}\Delta C_{lm} &= \frac{a^2(1-k_l)}{(2l+1)Mg} \iint_{Earth} p_s \tilde{P}_{lm}(\cos\theta) \cos(m\phi) dS \\ \Delta S_{lm} &= \frac{a^2(1-k_l)}{(2l+1)Mg} \iint_{Earth} p_s \tilde{P}_{lm}(\cos\theta) \sin(m\phi) dS.\end{aligned}\quad (2.4)$$

In order to analyse gravity variations caused by the atmosphere it is important to calculate the mean pressure field (at least over one year)  $\bar{p}_s$ . Residual pressure data is obtained by subtracting  $\bar{p}_s$  from  $p_s$ . This leads to

$$\begin{aligned}\Delta C_{lm} &= \frac{a^2(1-k_l)}{(2l+1)Mg} \iint_{Earth} (p_s - \bar{p}_s) \tilde{P}_{lm}(\cos\theta) \cos(m\phi) dS \\ \Delta S_{lm} &= \frac{a^2(1-k_l)}{(2l+1)Mg} \iint_{Earth} (p_s - \bar{p}_s) \tilde{P}_{lm}(\cos\theta) \sin(m\phi) dS.\end{aligned}\quad (2.5)$$

TL only includes atmospheric mass changes, but for the GRACE mission changes of the centre of mass of the atmosphere are important as well. The VI considers the vertical structure of the atmosphere. Taking the volume element  $dV$

$$dM = \rho r^2 dr \sin\theta d\theta d\phi = \rho dV \quad (2.6)$$

into account the spherical harmonic coefficients can be calculated by

$$\begin{aligned}\Delta C_{lm} &= -\frac{1}{(2l+1)M^l} \iint_{Earth} \left[ \int_r^\infty r^{l+2} \rho dr \right] \tilde{P}_{lm}(\cos\theta) \cos(m\phi) \sin\theta d\theta d\phi \\ \Delta S_{lm} &= -\frac{1}{(2l+1)M^l} \iint_{Earth} \left[ \int_r^\infty r^{l+2} \rho dr \right] \tilde{P}_{lm}(\cos\theta) \sin(m\phi) \sin\theta d\theta d\phi.\end{aligned}\quad (2.7)$$

In Equation (2.6) and Equation (2.7)  $\rho$  depicts the density and  $r$  the Earth radius. Using the hydrostatic equation

$$\rho dr = -\frac{dP}{g_r} \quad (2.8)$$

and approximating the gravity acceleration in height  $r$   $g_r$  with

$$g_r = g \frac{a^3}{r} \quad (2.9)$$

the spherical harmonic coefficients can be calculated using

$$\begin{aligned}\Delta C_{lm} &= -\frac{1}{(2l+1)M^l} \iint_{Earth} \left[ \int_{p_s}^0 r^{l+4} dp \right] \tilde{P}_{lm}(\cos \theta) \cos(m\phi) \sin \theta d\theta d\phi \\ \Delta S_{lm} &= -\frac{1}{(2l+1)M^l} \iint_{Earth} \left[ \int_{p_s}^0 r^{l+4} dp \right] \tilde{P}_{lm}(\cos \theta) \sin(m\phi) \sin \theta d\theta d\phi.\end{aligned}\quad (2.10)$$

Again to analyse gravity variations caused by the atmosphere a mean pressure field (at least over one year)  $\overline{p_{VI}}$  needs to be subtracted.

$$\begin{aligned}\Delta C_{lm} &= -\frac{1}{(2l+1)M^l} \iint_{Earth} \left( \left[ \int_{p_s}^0 r^{l+4} dp \right] - \overline{p_{VI}} \right) \tilde{P}_{lm}(\cos \theta) \cos(m\phi) \sin \theta d\theta d\phi \\ \Delta S_{lm} &= -\frac{1}{(2l+1)M^l} \iint_{Earth} \left( \left[ \int_{p_s}^0 r^{l+4} dp \right] - \overline{p_{VI}} \right) \tilde{P}_{lm}(\cos \theta) \sin(m\phi) \sin \theta d\theta d\phi.\end{aligned}\quad (2.11)$$

In the *Atmosphere and Ocean De-aliasing Level-1B* (AODB1B) product the atmospheric and oceanic mass variation are already taken into account. This means that the sum of the global atmospheric variability and the oceanic pressure variability are applied (e.g. by linear or spline interpolation) to the static background mean gravity field during processing.

The AODB1B product has already been applied in the Level-2 product and this processing level (Level-2) is the starting point for obtaining the surface mass from the Earth's gravity field. Later on surface mass can be related to TWS.

Wahr et al. [1998] described that the geoid shape  $N$  is stated as

$$N(\theta, \phi) = r \sum_{l=0}^{\infty} \sum_{m=0}^l \tilde{P}_{lm}(\cos \theta) (C_{lm} \cos(m\phi) + S_{lm} \sin(m\phi)). \quad (2.12)$$

The dimensionless spherical harmonic geoid coefficients  $C_{lm}$  and  $S_{lm}$  are retrieved from the satellite mission up to degree and order around 100. *Changes in  $N$*  ( $\Delta N$ ) may be caused by *changes in  $C_{lm}$*  and  *$S_{lm}$*  ( $\Delta C_{lm}$  and  $\Delta S_{lm}$ ). This leads to

$$\Delta N(\theta, \phi) = r \sum_{l=0}^{\infty} \sum_{m=0}^l \tilde{P}_{lm}(\cos \theta) (\Delta C_{lm} \cos(m\phi) + \Delta S_{lm} \sin(m\phi)). \quad (2.13)$$

$\Delta N$  may be caused by a density redistribution  $\Delta \rho(r, \theta, \phi)$ , which is reduced to a layer of thickness  $H$ .  $H$  is mostly around 10-15km, deduced by including those parts that cause mass fluctuation (atmosphere, oceans, ice caps and below-ground-water storage). “[Wahr et al. [1998]] define the change in surface density (i.e., mass/area),

$\Delta\sigma$ , as the radial integral of  $\Delta\rho$  through this layer:"

$$\Delta\sigma(\theta, \phi) = \int_{thinlayer} \Delta\rho(r, \theta, \phi) dr \quad (2.14)$$

With the help of Equation (2.15)  $\Delta\sigma$  (the TWS) can be retrieved from  $\Delta C_{lm}$  and  $\Delta S_{lm}$ .

$$\Delta\sigma(\theta, \phi) = \frac{r\rho_{ave}}{3} \sum_{l=0}^{\infty} \sum_{m=0}^l \tilde{P}_{lm}(\cos\theta) \frac{2l+1}{1+k_l} (\Delta C_{lm} \cos(m\phi) + \Delta S_{lm} \sin(m\phi)) \quad (2.15)$$

In Equation (2.15)  $\rho_{ave}$  is defined as the average density of the Earth ( $\rho_{ave} = 5517\text{kg}/\text{m}^3$ ).

The terms with  $l = 0$  and  $m = 0$  are proportional to the total mass of the Earth and do not change within time, though the  $\Delta C_{00}$  term dissipates and  $k_0$  becomes 0 as well. "[...] the position of the Earth's centre of mass relative to the centre of the coordinate system" (Wahr et al. [1998]) correspond to those terms where  $l = 1$ .

It is a well-known fact that changes in water storage cause changes in the Earth gravity field. As explained by Wahr et al. [1998], the errors in the GRACE results increase as  $l$  increases, which negatively affect on the results made available when using Equation (2.15). However, using Equation (2.16)

$$\overline{\Delta\sigma}(\theta, \phi) = \int \sin\theta' d\theta' d\phi' \Delta\sigma(\theta', \phi') W(\theta, \phi, \theta', \phi') \quad (2.16)$$

achieves useful results referred to the averages of the surface mass density. Combining and manipulating Equation (2.15) and Equation (2.16) leads to Equation (2.17)

$$\overline{\Delta\sigma}(\theta, \phi) = \frac{2r\rho_{ave}\pi}{3} \sum_{l,m} \frac{2l+1}{1+k_l} W_l \tilde{P}_{lm}(\cos\theta) [\Delta C_{lm} \cos(m\phi) + \Delta S_{lm} \sin(m\phi)] \quad (2.17)$$

with

$$W_l = \int_0^\pi W(\alpha) P_l(\cos\alpha) \sin\alpha d\alpha \quad (2.18)$$

$$P_l = \frac{\tilde{P}_{lm=0}}{\sqrt{2l+1}}.$$

$\alpha$  indicates the angle between  $(\theta, \phi)$  and  $(\theta', \phi')$ .

In Equation (2.18)  $W_l$  states a spatial averaging function, which is based on the idea by Jekeli [1981]. Jekeli [1981] illustrates a recursion relation for  $W$ , as it is shown in

Equation (2.19)

$$\begin{aligned}
 W_0 &= \frac{1}{2\pi} \\
 W_1 &= \frac{1}{2\pi} \left[ \frac{1 + \exp^{-2b}}{1 - \exp^{-2b}} - \frac{1}{b} \right] \\
 W_{l+1} &= -\frac{2l+1}{b} W_l + W_{l-1},
 \end{aligned} \tag{2.19}$$

where  $b$  is clarified in Equation (2.20)

$$b = \frac{\ln(2)}{1 - \cos(r_{ave}/r)}. \tag{2.20}$$

In Equation (2.20)  $r_{ave}$  states the averaging radius.

### 2.1.2 GRACE Tellus

In this thesis the GRACE Tellus Land (GRACE Tellus) dataset is used, which is processing level 3. The results are monthly TWSC for all continents in a  $1^\circ \times 1^\circ$  grid. TWS is understood as the sum of ground water, snow, surface water and soil moisture.

In order to consolidate all four components the TWSC are given in *liquid water equivalent* (LWE), which can be measured in *cm*.

Figure 2.3 displays an example of the GRACE Tellus data, downloaded from the NASA *Physical Oceanography Distributed Active Archive Center* (PO.DAAC) website<sup>5</sup>.

For the GRACE Tellus dataset three solutions are available, based on the RL05 spherical harmonic coefficients provided by GFZ, JPL and CSU, respectively.

In the data processing the  $C_{20}$  coefficients were replaced with *Satellite Laser Ranging* (SLR) solutions. For the estimation of the coefficients of degree 1 the method as described by Swenson et al. [2008] was used. Geruo et al. [2013] describe a *glacial isostatic adjustment* (GIA) used for GRACE Tellus, to consider the post-glacial rebound effect. If the spherical harmonic coefficients are converted to the gravitational potential and then plotted, the plot contains characteristic stripes from north to south (meridial stripes). In order to reduce them, a de-stripping filter has been applied in the GRACE Tellus processing. For smoothing the data a Gaussian filter with a radius of 300 km was utilized as described in Subsection 2.1.1 (Equation (2.19))

<sup>5</sup>[http://podaac-ftp.jpl.nasa.gov/allData/tellus/L3/land\\_mass/](http://podaac-ftp.jpl.nasa.gov/allData/tellus/L3/land_mass/)

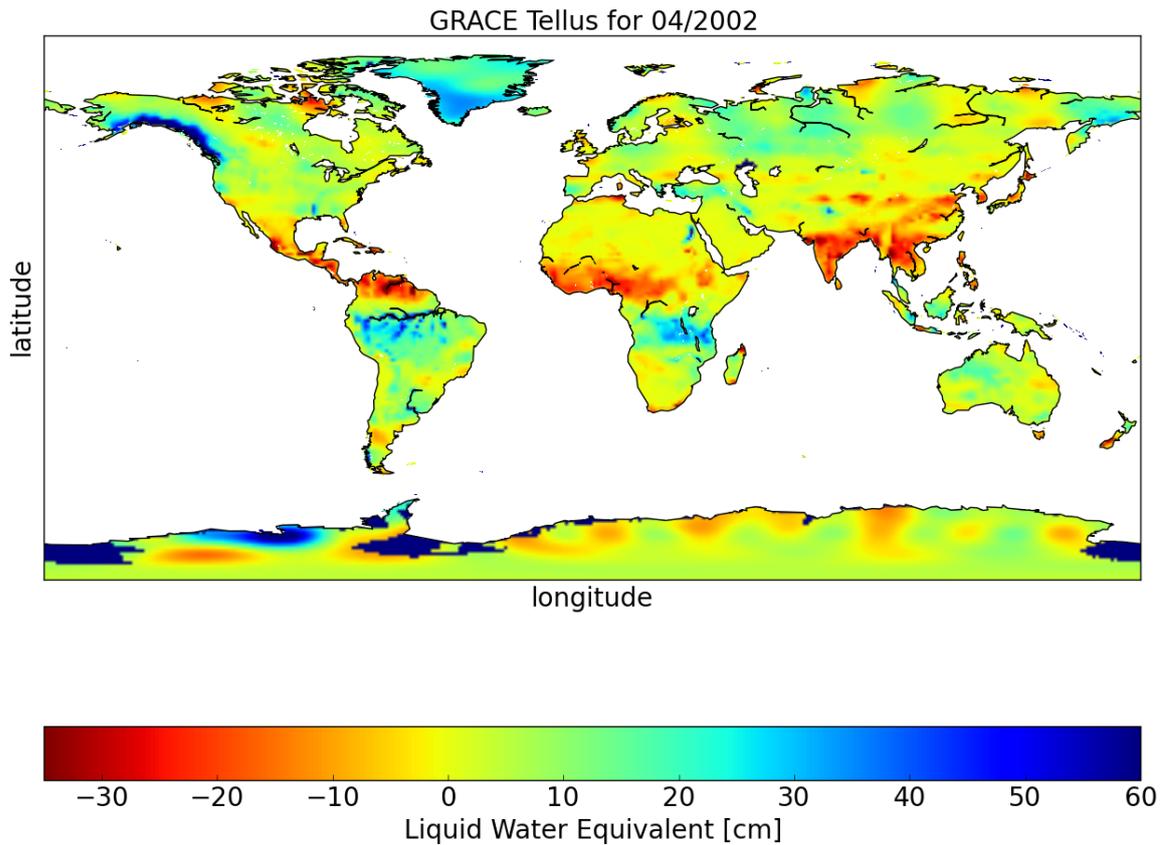


Figure 2.3: The GRACE Tellus data for April 2002 in [cm].

and Equation (2.20)). For receiving TWSC, a mean from January 2004 to December 2009 was subtracted from the obtained data.

Due to the processing (de-striping and Gaussian filter) energy is removed. By multiplying the GRACE Tellus data with the provided grid scaling, energy can be restored. The grid scaling contains “a set of scaling coefficients, one for each 1 degree bin of the land grids” (NASA website for the GRACE mission<sup>6</sup>).

As the GRACE missions started in 2002, first measurement results are available for April 2002. Unfortunately in the GRACE Tellus dataset data is not available for some months. These months need to be skipped in all other datasets as well (when comparing to GRACE Tellus), to get reliable results.

<sup>6</sup><http://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/>

## 2.2 ESA CCI Soil Moisture

The *Climate Change Initiative* (CCI) is a programme of the *European Space Agency* (ESA) to construct long-term datasets of various *Essential Climate Variables* (ECV). The *CCI for Soil Moisture* (CCI SM) project started in 2010 and it was set up for a period of 6 years with the aim to construct a consistent long-term dataset of satellite-based soil moisture. The data is derived from active and passive microwave spaceborne instruments over the timespan 1978 to 2014.

Active microwave instruments include scatterometers, synthetic aperture radars (SAR) and radar altimeters, which send electromagnetic energy to the Earth and measure the backscattering coefficient  $\sigma^0$  in Decibel [dB]. The active microwave instruments used in CCI SM are the C-band scatterometer instruments, namely the *Active Microwave Instrument Wind Scatterometer* (AMI-WS) and the *Advanced Scatterometer* (ASCAT) (Figure 2.4). Soil moisture is derived from these scatterometer measurements at the Vienna University of Technology using the WARP change detection algorithm.

Passive microwave remote sensing measures the energy emitted from the Earth surface in the microwave domain. Measurements are provided as brightness temperature  $T_B$  in Kelvin [K]. The passive microwave instruments used in CCI SM are the *Scanning Multichannel Microwave Radiometer* (SMMR), the *Special Sensor Microwave Imager* (SSM/I), the *Tropical Rainfall Measuring Mission's* (TRMM) *Microwave Imager* (TMI), the *Advanced Microwave Scanning Radiometer - EOS* (AMSR-E), WindSat and AMSR2 (Figure 2.4). Soil moisture is retrieved from the brightness temperatures by the Vrije Univeriteit Amsterdam/NASA *Land Parameter Retrieval Model* (LPRM) algorithm. The *Soil Moisture Ocean Salinity* (SMOS), *Soil Moisture Active Passive* (SMAP) and Aquarius missions are not included so far, but it might be feasible to include them later on.

The swath of the satellites on the surface is between 500 to 1400km as described in detail by Kidd and Haas [2015].

All active and passive microwave soil moisture products are first merged into a merged active and a merged passive dataset, respectively. These datasets are then merged into a combined active and passive dataset, which is the dataset used in this thesis. The data images on a  $0.25^\circ \times 0.25^\circ$  grid.

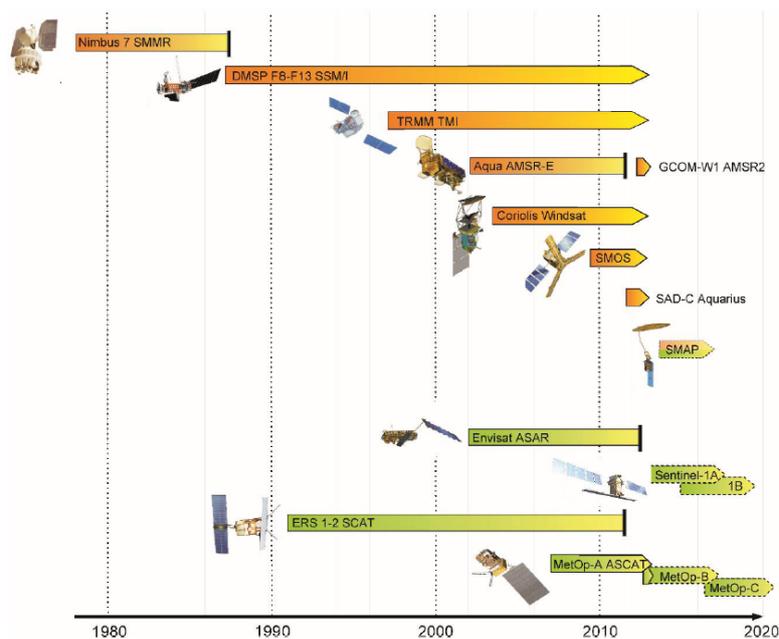


Figure 2.4: Satellites and sensors, which are used for collecting the CCI SM data. SMAP and Aquarius are not included yet, but the possibility to do that exists; from Dorigo et al. [2012]

### 2.2.1 Previously performed processing steps

The processing is divided into five different steps: Levels 0 to 4.

In Level 0, the collected raw data from ground stations is reconstructed and recorded at a data centre. Further details are described by JCGM [2008] and JCGM [2012].

The next processing step (Level 1) is the conversion of these raw data into calibrated measurements  $\sigma_0$  or  $T_B$ , depending on whether the raw data was collected by (active or passive microwave instruments. If the data are collected by several antennas or satellites, they need to be merged to one product. It is important to notice that Level 1 data are for single overpasses only.

The geophysical parameters indicating surface SM are retrieved in Level 2, in units of [%] or [ $m^3/m^3$ ], again for single overpass only. As SM cannot be directly measured, this step in the processing chain is called retrieval. The mathematical-physical background for the retrieval is, the existence of a relationship between the SM parameters and the obtained observations. However, model parameters or long-term reference data are necessary for the retrieval.

In Level 3 the geophysical product is displayed on a reference grid (space-time grid of longitude, latitude and time), having a spatial resolution of  $0.25^\circ$  for every day. This is the global SM ECV Data Product. Until Level 2 the data have only been

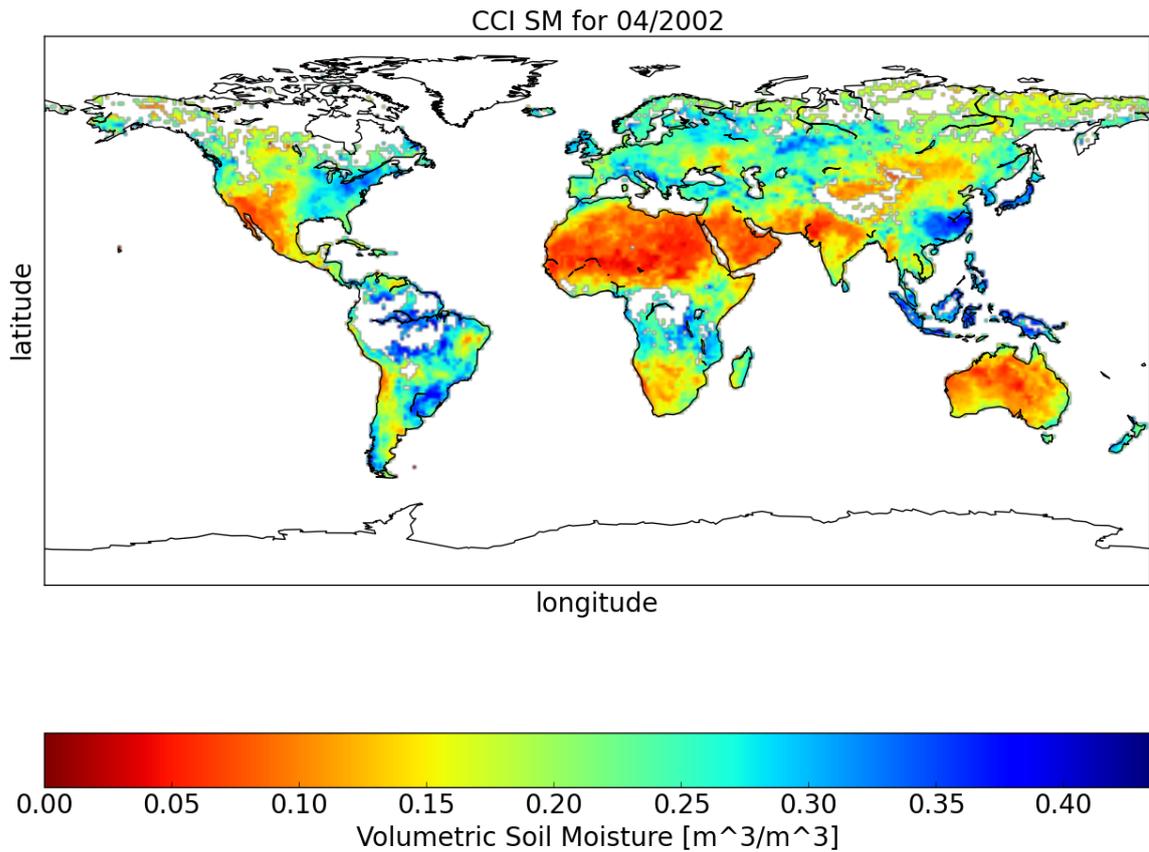


Figure 2.5: This is the CCI SM dataset for April 2002, with a spatial resolution of  $1^\circ$  in longitude and latitude direction.

available for single overpasses, a time series joining all single overpasses (over many years including different satellites) is available additionally for Level 3 and Level 4. Level 4 is used to indicate a product derived from original surface soil measurements (e.g. root-zone SM).

A deeper explanation of the processing steps can be found in Wagner et al. [2012], Liu et al. [2011], Liu et al. [2012] and Kidd and Haas [2015].

### 2.2.2 CCI SM v02.2

In this thesis the CCI SM data was resampled to a  $1^\circ \times 1^\circ$  grid to allow an easy comparison to the GRACE Tellus dataset. This was done by calculating the mean over 16 cells belonging to the one and same grid cell (4 cells in longitude direction and 4 in latitude direction). The data for April 2002 is shown in Figure 2.5. The data is given as volumetric soil moisture [ $m^3/m^3$ ], which means cubic meter water per cubic meter soil. The CCI SM dataset only contains values if the soil contains

liquid water. In other words no data can be obtained for snow or ice. Whenever a data point is absent in the CCI SM data due to snow or ice, it is likewise masked in all other datasets. At the beginning of this thesis (November 2015) the CCI SM data was only available until December 2014. However, for an accurate comparison of GRACE Tellus and CCI SM the mean over the timespan from January 2004 to December 2009 (the same period as for the GRACE data) has to be subtracted from the merged CCI SM data.

More information about the CCI SM project and the datasets can be found on the ESA homepage for the CCI project.<sup>7</sup>

## 2.3 ERA-Interim

The *European Centre for Medium-Range Weather Forecasts* (ECMWF) *Interim reanalysis* (ERA-Interim) is a climate reanalysis dataset from 1979 until present as described in Dee et al. [2011]. Climate reanalysis is defined as the numerical description of climate, which is derived from a combination of forecast models and observations of various types. The global data has a spatial resolution of around 80km and a temporal resolution of 3 hours.

### 2.3.1 Previously performed processing steps

The forecast model is an important factor in data assimilation, because smaller adjustments between observations and forecast can be achieved if a forecast model of high quality is used. The forecast model computes different variables, like soil moisture, soil temperature and atmospheric fields. The ECMWF *Integrated Forecast System* (IFS), release *Cy31r2* (atmospheric model), is used as the forecast model. It assimilates three components: atmosphere, land surface and ocean waves. A detailed description of changes to model physics is given by Dee et al. [2011].

Dee et al. [2011] describe the increasing number of observation since the beginning of the project, in the year 2010 about  $10^7$  observations per day were used. A simple reason for this is the increase in satellites, which produce the data predominantly used in this analysis. Polar-orbiting and geostationary satellites and other spaceborne sensors provide the data, which is completed by in-situ measurements (using pilot balloons, aircraft, ships, drifting buoys,...), measuring for example wind, upper-

---

<sup>7</sup>information: <http://www.esa-soilmoisture-cci.org/node/93>;

data : <http://www.esa-soilmoisture-cci.org/node/145> - a registration is necessary

air temperature and specific humidity. For a deeper understanding which datasets are used, especially compared to the early ERA-40 data, Dee et al. [2011] is recommended.

### 2.3.2 ERA-Interim/Land

In this thesis the ERA-Interim/Land dataset is used. Balsamo et al. [2015] describe the ERA-Interim/Land dataset as a land surface reanalysis dataset, which is given on a global scale. It is a simulation using the latest ECMWF land surface model (Balsamo et al. [2015]) and has a spatial resolution of around 80km. The data includes both atmospheric (air temperature, wind at different altitudes,..) and surface parameters (soil moisture, rainfall,..) and describes the development of global SM and snow water equivalent, from January 1979 until December 2014, with a temporal resolution of 6 hours. Compared to the ERA-Interim dataset it includes parametrization improvements of the land surface model as described by Balsamo et al. [2015], which leads to more suitable data for climate studies. Ground-based and remote sensing observations are used to obtain a higher quality of the ERA-Interim/Land dataset. Land parametrization revisions cause an improvement in the northern hemisphere while an improvement in the southern hemisphere is related to precipitation corrections. The data is stored on a reduced two dimensional (latitude-longitude) Gaussian grid or on the ECMWF T255 spectral grid. T255 means that the spherical harmonic coefficients are cut at wave number 255. Both can be presented as  $0.75^\circ$  grid, utilizing the *World Geodetic System 84* (WGS84).

In the dataset, the surface is divided into 4 different layers: Layer 1 from 0.00-0.07m, Layer 2 from 0.07-0.28m, Layer 3 from 0.28-1.0m and Layer 4 from 1.0-2.89m.

The datasets can be downloaded from the ECMWF Public Dataset website <sup>8</sup>.

Figure 2.6 shows the first layer (Layer 1) of the ERA-Interim/Land dataset for April 2002.

## 2.4 Precipitation products

Precipitation is defined as “rain, snow, sleet, or hail that falls to or condenses on the ground” <sup>9</sup>.

---

<sup>8</sup><http://apps.ecmwf.int/datasets/data/interim-land/type=fc/> - a registration is necessary

<sup>9</sup><http://www.oxforddictionaries.com/definition/english/precipitation>

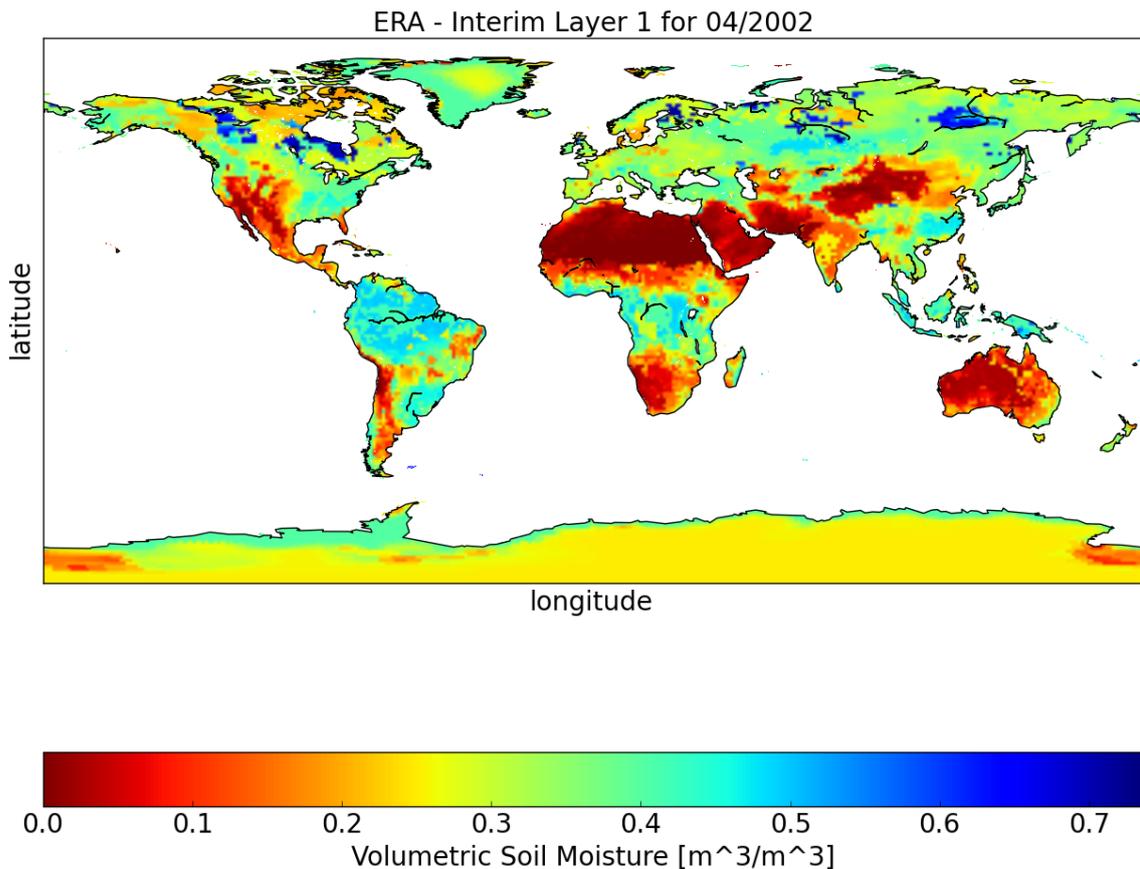


Figure 2.6: ERA-Interim/Land data for April 2002 for the first layer (0.00-0.07m), on a spatial 1°x1° grid.

### 2.4.1 GPCC

The *Global Precipitation Climatology Centre* (GPCC), which is operated by the *Deutscher Wetterdienst* (DWD), was initiated in 1989. It provides gridded precipitation products based on in-situ measurements of precipitation. A short description of the GPCC dataset, especially related to the GPCP dataset can be found in Huffman et al. [1997].

Different GPCC products like the First Guess Product, the First Guess Daily Product, the Monthly Monitoring Product (Version 5) and the Full Data Reanalysis Product (Version 7) exist. A list of all products, their descriptions and the data itself can be found on the DWD website<sup>10</sup>. In Chapter 2.4.1 an outline of the dataset used for this thesis is given.

<sup>10</sup>decription: <http://www.dwd.de/EN/ourservices/gpcc/gpcc.html>;  
data: [ftp://ftp.dwd.de/pub/data/gpcc/html/fulldata\\_v7\\_doi\\_download.html](ftp://ftp.dwd.de/pub/data/gpcc/html/fulldata_v7_doi_download.html)

### Previously performed processing steps

The in-situ data is only obtained at specific points on the surface, which means that the data must be interpolated to a regular grid of a certain resolution. This is done using the Spheremap interpolation routine. This interpolation method is basically a modification of the well-known Shepard interpolation, adapted to a spherical surrounding, as described in Adler et al. [2003]. In a nutshell, the distance between the grid point and the in-situ measurement station is considered, utilizing a weighting method, as well as the directional distribution of the in-situ measurement stations and the gradients of the data field. A more detailed description of the Spheremap interpolation can be found by Willmott et al. [1985].

### GPCC Full Data Reanalysis

In this thesis the GPCC Full Data Reanalysis dataset is used. It is available in a spatial resolution of  $0.5^\circ$ ,  $1.0^\circ$  or  $2.5^\circ$  in longitude and latitude, with a temporal resolution of one month over the time period 1901-2013. The data was derived from more than 67,000 in-situ stations, spread all over the world. As the GRACE data belongs to a  $1^\circ \times 1^\circ$  grid, the GPCC Full Data Reanalysis with a  $1^\circ$  resolution was used. Figure 2.7 shows the dataset for April 2002.

## 2.4.2 GPCP

In 1986 the *Global Precipitation Climatology Project* (GPCP) was initiated by the *World Climate Research Program* (WCRP) with the aim to provide precipitation data on global scale with a temporal resolution of one day (over the period from 1997 to present) or one month (in the time interval of 1979-present). The daily dataset is provided on a  $1^\circ \times 1^\circ$  (latitude \* longitude) grid and the monthly dataset on a  $2.5^\circ \times 2.5^\circ$  grid.

The provided dataset combines various sources of precipitation information. The in-situ measurements are analysed by the GPCC. At the era before GPCC, the *Global Historical Climate Network* (GHCN) and *Climate Assessment and Monitoring System* (CAMS) are used. The SSMI/I is used to obtain microwave data at four different frequencies (19.35, 22.235, 37 and 85,5 GHz). Two algorithms are applied to obtain the data: one for land regions (scattering) and one for ocean regions (emission). Huffman et al. [1997] describe how both estimates (land and ocean)

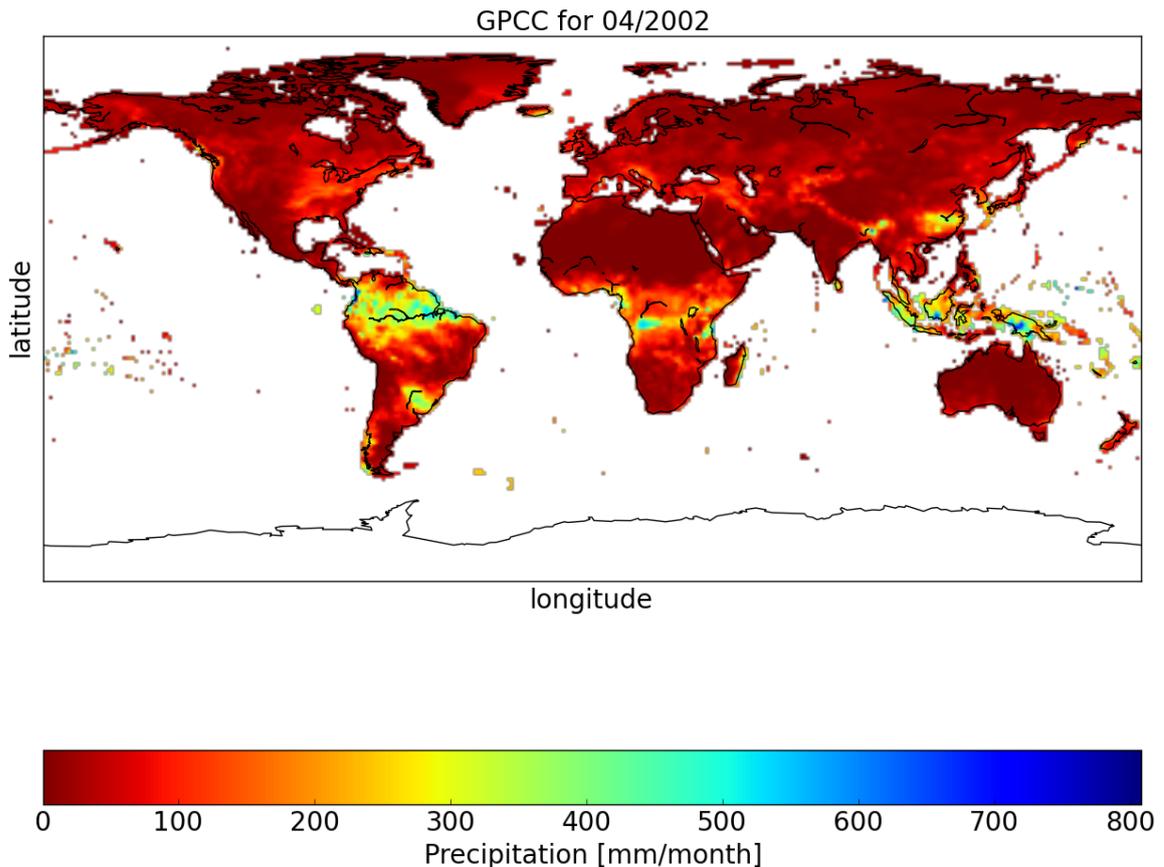


Figure 2.7: The GPCP dataset for April 2002, as it can be downloaded on the DWD website<sup>11</sup>.

are merged for near-coastal regions. The *infrared* (IR) precipitation measurements are obtained from the merged product (of different geostationary satellites) using the *Geostationary Operational Environmental Satellites* (GEOS) *precipitation index* (GPI) technique, which related cold cloud-top area to rain rate. Additionally, data is obtained from the *Television and Infrared Observation Satellites* (TIROS) *Operational Vertical Sounder* (TOVS).

Each dataset has its own assets and drawbacks, which are explained by Huffman et al. [1997]. This product was calculated by making use of the various advantages (like for example spatial resolution vs. temporal resolution) of the different sources. A detailed description of the dataset utilized to produce GPCP can be found by Huffman et al. [1997] and Adler et al. [2003].

<sup>11</sup>[ftp://ftp.dwd.de/pub/data/gpcc/html/fulldata\\_v7\\_doi\\_download.html](ftp://ftp.dwd.de/pub/data/gpcc/html/fulldata_v7_doi_download.html)

### Previously performed processing steps

An essential processing step is the merging of various datasets from different sources to one combined product. As there are no microwave-based estimates available for the period 1979-1987, two different techniques have to be applied.

The first technique, used for the period 1987-present, combines the input data by highlighting the advantages of each. It results in a single dataset that superimposes the various input datasets, like frequent time sampling and overwhelming physical basis. An outline is illustrated in Huffman et al. [1997]. For the earlier period (1979-1987) outgoing long-wave radiation (IR radiation - radiating from the Earth) was utilized to produce a monthly estimation, while no microwave-based estimates were available. Xie and Arkin [1998] introduced an *outgoing longwave radiation based precipitation index* (OPI). Applying the GPCP dataset for the period after 1979 the OPI can be linked to precipitation by developing regression coefficients. For further information the reader is referred to Xie and Arkin [1998] and Adler et al. [2003].

### GPCP 1DD

For this thesis the GPCP *1-degree-daily* (1DD) dataset is used - this implies that the data is stored in a  $1^\circ \times 1^\circ$  grid with a temporal resolution of one day. For a constructive comparison of the GPCP 1DD dataset with the GRACE Tellus dataset, the calculation of the sum of precipitation of each month, in units of [*mm/month*], is required. Furthermore, the mean of all data of the period from January 2004 to December 2009 was calculated for each grid cell and subtracted from the monthly sum.

Figure 2.8 depicts the monthly precipitation sums for April 2002.

The GPCP 1DD dataset can be found on the NASA website for precipitation analysis<sup>12</sup>.

## 2.5 GMIA

Irrigation has become a topic of increasing interest due to the cultivation of crop and features prominently in food production and food security. Irrigation influences both SM and TWS and because of that it is of great interest for this thesis. The *Digital Global Map of Irrigated Area* (GMIA) was initiated with the release of the first map (Version 1) in 1995. From that point on, four updated versions followed, the latest

---

<sup>12</sup><http://precip.gsfc.nasa.gov/>

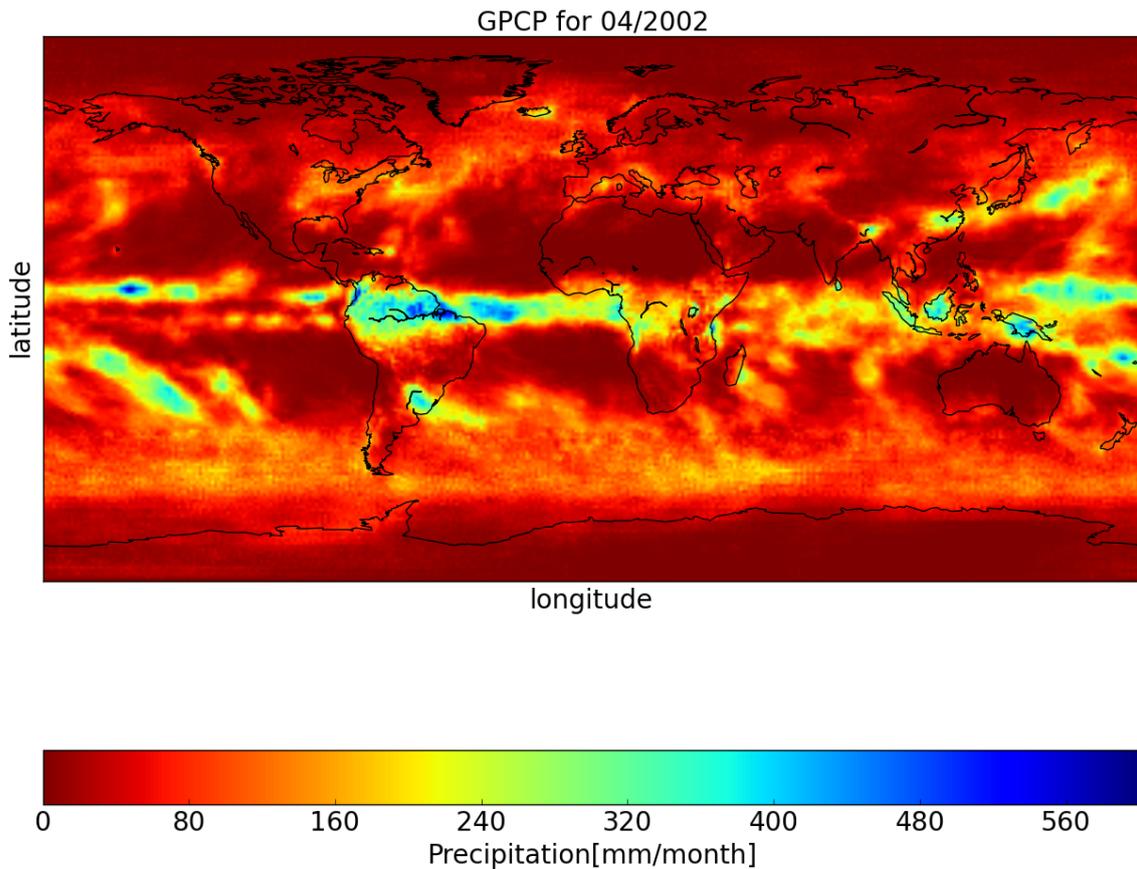


Figure 2.8: The GPCP anomalies for April 2002. The daily values are merged to gain a temporal resolution of one month. Afterwards the mean over the timespan from January 2004 to December 2009 is subtracted.

which is Version 5 is utilized in this thesis. GMIA is a cooperation between the Johann Wolfgang Goethe University (Frankfurt am Main / Germany), the Rheinische Friedrich-Wilhelms-Universität Bonn (Bonn / Germany) and the *Food and Agriculture Organization of the United Nations* (FAO). To be more specific, it was the *FAO global information system on water and agriculture* (Aquastat) that was involved in this cooperation. Aquastat consolidates data about water resources, water uses and agricultural water management per country. More detailed information can be found in FAO [2014] and Siebert et al. [2013]. GMIA is available on a global scale ever since Version 3.

The irrigation dataset includes both, irrigation density (i.e. the percentage of *area equipped for irrigation* (AEI)) as well as *the area actually irrigated* (AAI).

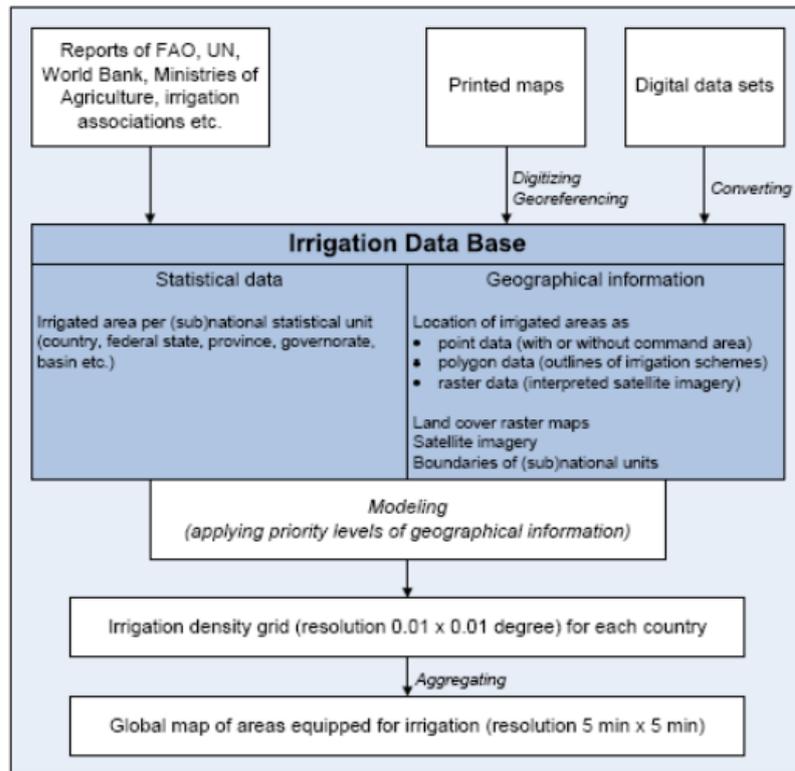


Figure 2.9: Schematic representation of the methodology used for mapping AEI; by (Siebert et al. [2013])

### 2.5.1 Previously performed processing steps

The dataset is a combination of irrigation statistics from different sources (e.g. sub-national units, national census, FAO reports...). If the values provided by Aquastat are more representative, then a scaling was implemented so that the sum of the irrigated area is equivalent to the AEI (from Aquastat). It is important to notice that the AEI and the AAI differ from each other wherein the AAI may be lower.

If for some reason only the AAI was given, then the AEI was extrapolated by using a time series of AAI in the region of interest over several years.

Siebert et al. [2013] explain that a lot of different data sources were used regarding the distribution of irrigated area within sub-national units.

Figure 2.9 shows a diagram of the methodology used to obtain the GMIA product. At the top the different input datasets are symbolized, with a short explanation of their meaning. This is followed by the modelling and the grid building.

The AEI can be distinguished depending on where the water comes from: groundwater, surface water or water derived from non-conventional sources. A detailed description of the datasets utilized and the processing steps is given by Siebert et al.

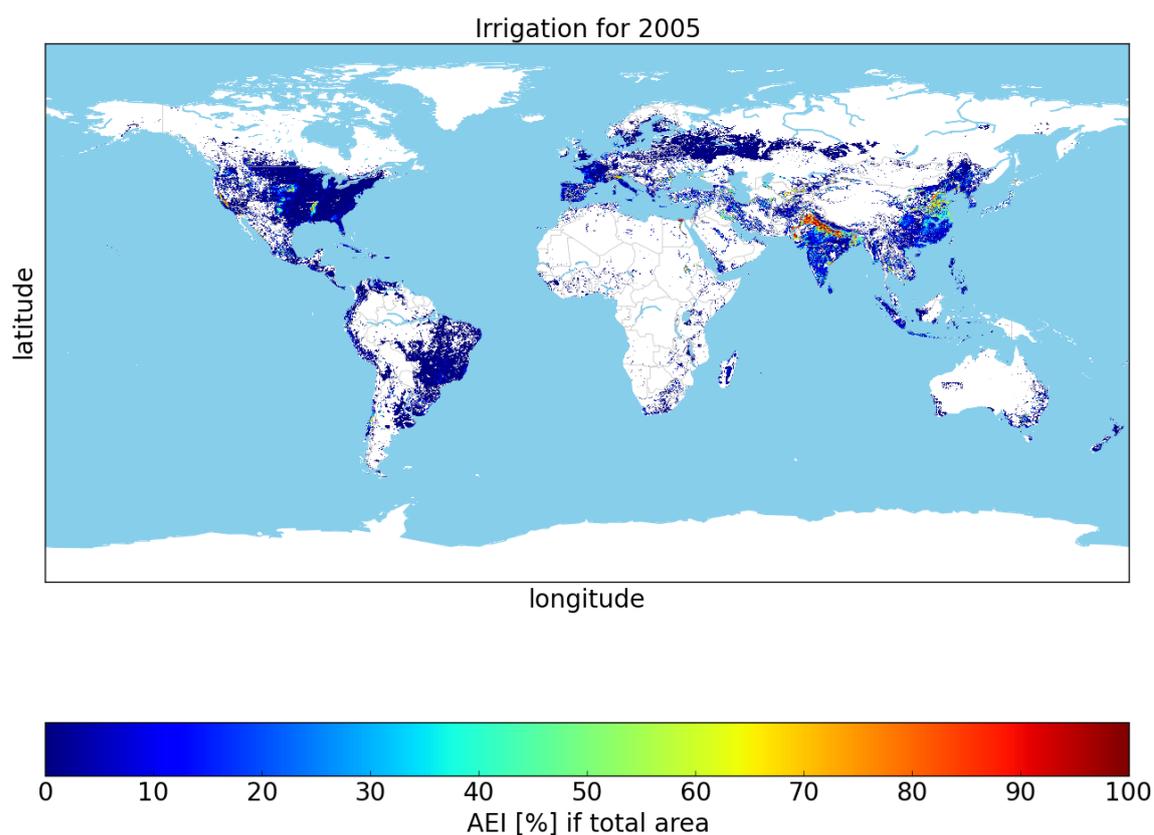


Figure 2.10: Irrigation for the year 2005 for the whole world in AEI in [%] of total area

[2013].

### 2.5.2 GMIA Version 5

Version 5 contains the irrigation data for the reference year 2005 with a spatial resolution of 5 arc minutes. It consists of layers for AEI, “source of irrigation water” and “percentage of AAI” (Siebert et al. [2013]). Due to the fact that different ministries are responsible for providing the irrigation data, the irrigation data is available for different reference years and they also have different resolutions of the statistics. This is the reason why those datasets were used, which provided the highest resolution and had timestamps closest to 2005. Figure 2.10 shows the percentage of AEI (per country) for the year 2005, as it can be on the FAO website<sup>13</sup>.

<sup>13</sup><http://www.fao.org/nr/water/aquastat/irrigationmap/index10.stm>

# Chapter 3

## Correlation function

If making comparison within one time series the autocorrelation function is applicable, whereas the cross-correlation function has to be used for comparing two different datasets. In this thesis the Pearson correlation coefficient is used, which highlights the linear relationship between the input datasets. An alternative possibility would be the Spearman correlation coefficient, which is a rank correlation coefficient and therefore it is not the first choice.

The Pearson cross-correlation function is based on the assumptions that the datasets are normally distributed (Gaussian distribution) and stationary as well as that equidistant timestamps are predominant, resulting in a discrete time series. Since the autocovariance function  $\gamma_k$  is equivalent to  $\gamma_{-k}$ , and the same is true for the autocorrelation function  $\rho_k$ , both only need to be considered for a positive lag ( $k = 0, 1, 2, \dots$ ). In this thesis it is assumed that due to the short timespan (2002-2014) there is no trend or that the trend rises only minimally, at the most. This means that the process is assumed to be stationary. As a consequence of the stationary process the mean ( $\mu$ ) and the variance ( $\sigma^2$ ) are constant. As described in Horvatic et al. [2011] the absence of the detrending process can lead to misinterpreted correlation coefficients. This needs to be considered for longer timespans or if a stronger rise of the trend is expected.

The consequence of the Gaussian distribution is that the process is uniquely defined by its mean and covariance matrix.

Regarding the cross-correlation function it is important to keep in mind that  $\gamma_{xy}$  and  $\gamma_{yx}$  produce different results, but  $\gamma_{xy}(k)$  is still equivalent to  $\gamma_{yx}(-k)$ , which is why only a function for  $k = \pm 0, 1, 2, 3, \dots$  needs to be defined. This is called the cross-covariance function of a stationary bivariate process. Dividing the cross-covariance function by the covariance between  $x$  and  $y$  leads to the cross-correlation coefficient

(3.1).

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y}, \quad (3.1)$$

where  $\gamma_{xy}(k)$  is defined in Equation (3.2),  $k$  states the lag and  $x$  and  $y$  are the different time series.

$$\gamma_{xy}(k) = E[(y_t - \mu_y)(y_{t+k} - \mu_y)] \quad (3.2)$$

The estimated cross-correlation function results from differencing the original datasets  $d$  times, so that  $n = N - d$  pairs of values are available. This leads to Equation (3.3)

$$c_{xy}(k) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(y_{t+k} - \bar{y}) & k = 0, 1, 2, \dots \\ \frac{1}{n} \sum_{t=1}^{n+k} (y_t - \bar{y})(x_{t-k} - \bar{x}) & k = 0, -1, -2, \dots \end{cases} \quad (3.3)$$

in which  $c_{xy}$  is the cross-covariance coefficient. Finally the cross-correlation coefficient  $r_{xy}$  is obtained by dividing Equation (3.3) with  $c_{xx}(0)$  and  $c_{yy}(0)$ , which are common as standard deviation.  $r_{xy}$  ranges between  $-1$  and  $+1$ .  $-1$  means that dataset  $x$  increases and  $y$  decreases and  $+1$  means that both increase or decrease at the same time. A more detailed description on auto-correlation and cross-correlation function and coefficients is given by Box and Reinsel [2008].

Before drawing any conclusion it is necessary to have a cursory glance at the significance. The significance qualifies the cross-correlation coefficient, meaning that it filters if a cross-correlation occurs randomly or not. The significance can be obtained by using the p-value. The p-value displays “the probability of obtaining data at least as extreme as ours” (Blockera et al. [2006]). The idea behind the p-value for significance tests is related to testing the null hypothesis. It is recommended to choose the null hypothesis in such a way that the resulting p-value should be small for rejecting the null hypothesis. Blockera et al. [2006] give an easy to understand example using the Higgs boson: the data just appears from the background process is the null hypothesis. This means if the p-value is small, than the Higgs boson is present.

A threshold needs to be defined, which determines if the correlation coefficient is significant or not. This threshold can be chosen freely. In this thesis it was chosen to be 0.05. This means that all correlation coefficients, having a p-value greater than 0.05 are excluded. The p-value in this thesis is a 2-tailed p-value originated from the t-test probability.

# Chapter 4

## Results

Each section contains the comparison of two datasets. All the represented correlation coefficients are pre-filtered, using the p-value described in Section 3.

In the following figures red areas symbolize correlation coefficients close to +1, meaning that both datasets increase, whereas the blue coloured areas illustrate correlation coefficients close to  $-1$  (one dataset increases while the other decreases). A green colour means that the correlation coefficient is zero (no correlation between the datasets).

### 4.1 GRACE Tellus compared to CCI SM

The main aim of this thesis is to generate a global map of correlation coefficients between the CCI SM and the GRACE Tellus datasets. The datasets are processed as described in Chapter 2.1.2 and 2.2.2.

The correlation coefficients were calculated using Equation (3.3), where  $k$  was chosen to range from 0 to 6, in units of months. Since it is important to use the same timespan for all datasets the period between April 2002 and December 2014 was chosen for this thesis. This specific time interval was predefined by the availability of the different datasets.

The outcomes of those calculations (i.e. filtered correlation coefficients) with no lag ( $k = 0$  in Equation (3.3)) are shown in Figure 4.1. There is no data available over the Antarctica and Greenland as those regions are well-known for housing a lot of snow and ice. Positive correlation coefficients are predominant (mostly between 0.6 and 0.8) in most parts of the world, except for the northern part of the northern hemisphere and Saudi Arabia (values between -0.6 and -0.3).

The highest correlation can be achieved when the GRACE Tellus datasets has a lag of one month compared to the CCI SM dataset. Figure A.1 shows the correlation

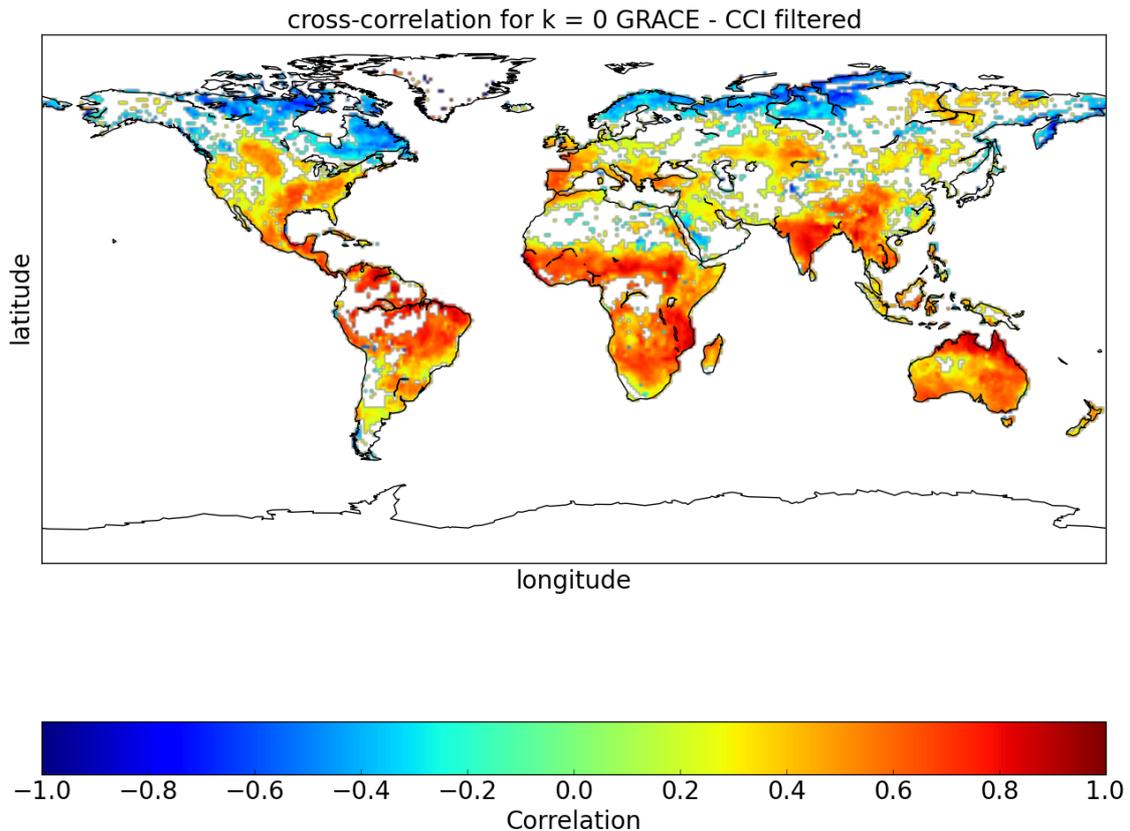


Figure 4.1: correlation of GRACE vs CCI with no lag ( $k = 0$ )

coefficients, if a lag of  $k$  ranging from 1 to 6 is assumed for the GRACE Tellus dataset. The other way round (CCI SM has a lag) is depicted in Figure A.2.

## 4.2 GPCC compared to GRACE Tellus & CCI SM

This section compares the GPCC dataset with both the GRACE Tellus and the CCI SM datasets. This comparison, as well as those in Section 4.3 and Section 4.4, are made for a deeper understanding of the results shown in Section 4.1. The GPCC dataset is only available until the end of 2013, which only permits comparisons within a shorter timespan than all the other datasets (April 2002 - December 2013).

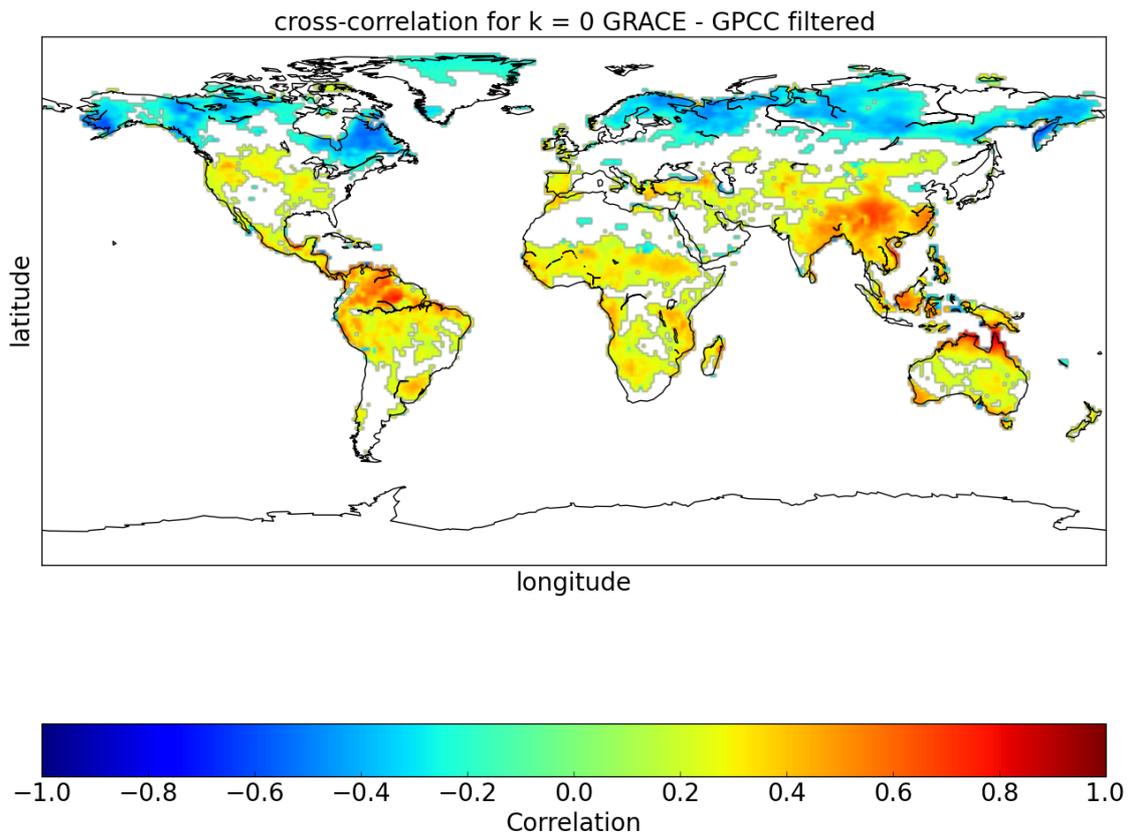


Figure 4.2: The correlation coefficients of the GRACE Tellus dataset compared to the GPCC Full Data Reanalysis dataset with a spatial resolution of 1 degree. Between the datasets no lag was applied ( $k = 0$ ).

#### 4.2.1 GPCC compared to GRACE Tellus

The anomaly of precipitation was ascertained by calculating the mean over the timespan 2004 to 2009 for each grid cell, which then was subtracted from each available data point. Again no data points are available for Antarctica, because the GPCC dataset is obtained by using in-situ measurements. Even so some data is available for some parts of Greenland. Figure 4.2 shows the correlation coefficients comparing the GRACE Tellus dataset to the anomalies of the GPCC dataset.

Figure A.3 and Figure A.4 show the correlation coefficients for different lags  $k$ . The highest correlation coefficient results, if the GRACE Tellus has a lag of 2 months (Figure A.3).

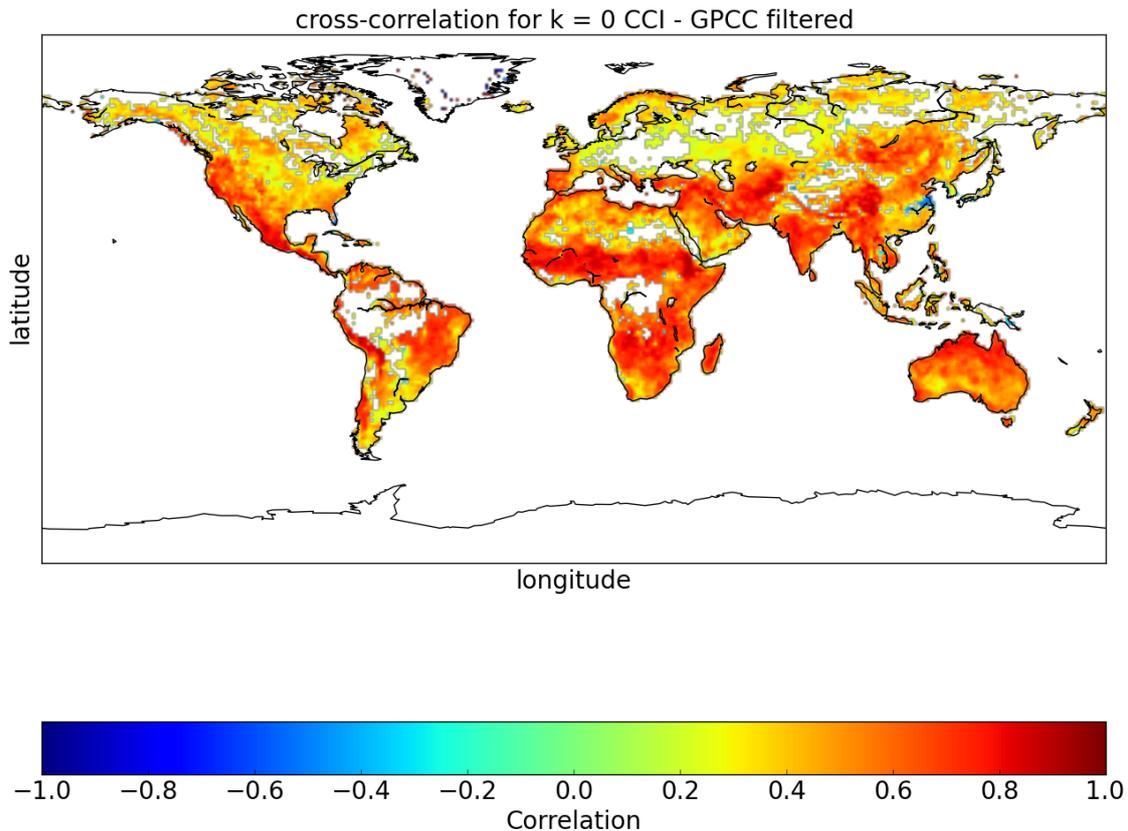


Figure 4.3: The correlation of CCI SM dataset compared to GPCP Full Data Reanalysis dataset with a spatial resolution of 1 degree. Between the datasets there is no lag ( $k = 0$ )

#### 4.2.2 GPCP compared to CCI SM

Both datasets can be compared, without subtracting the mean. Although both datasets are available with a temporal resolution of 12 hours or 1 day, the sum (mean) of each month was calculated and compared to each other. The results are shown in Figure 4.3.

The highest correlation coefficients can be found for a lag of 1 or 2 month for the GPCP dataset (see Figure A.5 and Figure A.6).

### 4.3 GPCP compared to GRACE Tellus & CCI SM

Another precipitation dataset used in this thesis is the GPCP dataset, which includes the GPCP dataset. This is why the results are very similar to those in Section 4.2.

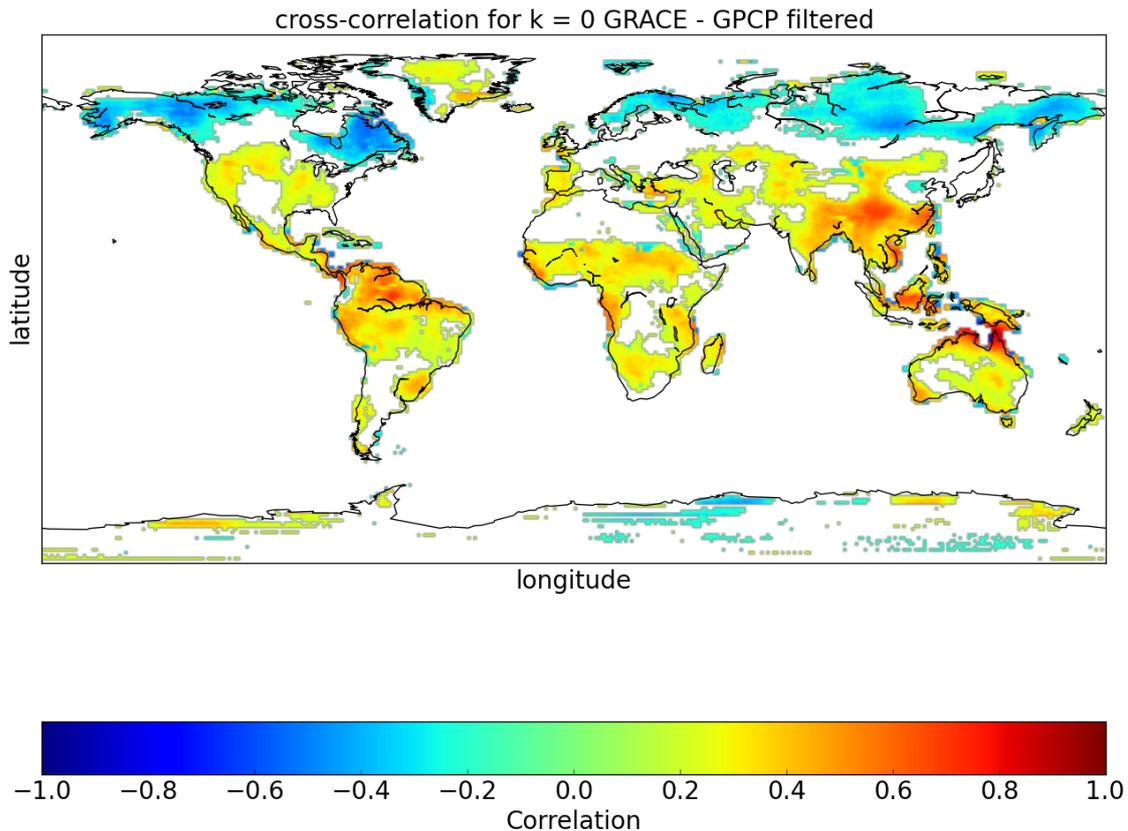


Figure 4.4: The correlation coefficients of the GRACE Tellus dataset compared to the GPCP 1DD dataset, with no lag ( $k = 0$ )

### 4.3.1 GPCP compared to GRACE Tellus

For calculating the correlation coefficients (Equation (3.3)) both datasets are processed as described earlier (Section 2.1.2 and Section 2.4.2). Regarding the calculation of correlation coefficients, it is important to skip those dates, for which no data is available for the GRACE Tellus dataset.

The outcome is shown in Figure 4.4. Because of the definition of precipitation (see 2.4) some data points are available even in the regions like Antarctica and Greenland. Using a lag on 2 months for the GRACE Tellus dataset achieves the highest correlation coefficients. This can be seen by comparing Figure A.7 and Figure A.8.

### 4.3.2 GPCP compared to CCI SM

Since the GPCP dataset analyses precipitation and the CCI SM dataset shows soil moisture, high positive correlation coefficients are to be expected. Figure 4.5 confirms this. In most parts of the world high positive correlation coefficients are predominant.

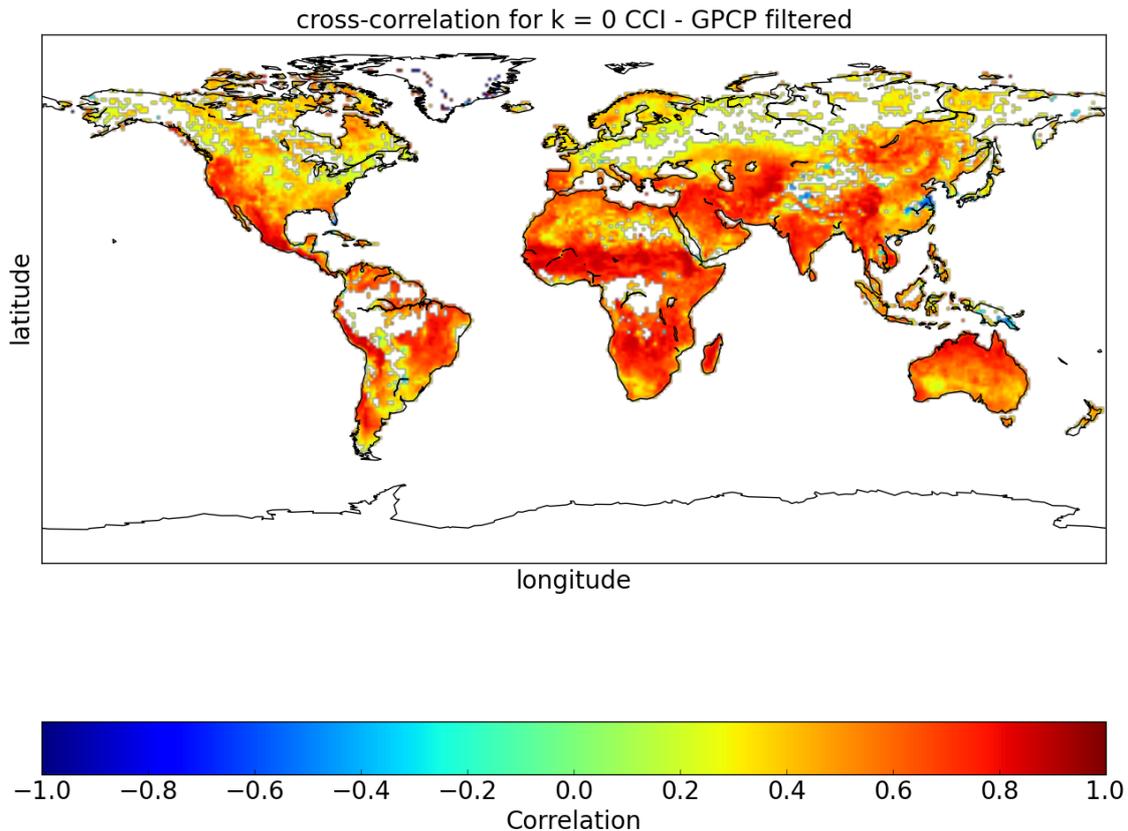


Figure 4.5: The correlation coefficients of the comparison between the CCI SM dataset and the GPCP 1DD dataset, with no lag ( $k = 0$ )

Generally, the correlation coefficients between GPCP and CCI are higher than those between GPCP and GRACE.

The highest correlation coefficients are shown in Figure A.9, for a lag  $k$  of 1 month for the CCI SM dataset. No lag results in high correlation coefficients in most parts of the world, whereas a lag of 1 month shows high correlation coefficients for some parts only. Figure A.10 depicts the correlation coefficients for different lags of the GPCP dataset.

## 4.4 ERA-Interim/Land compared to GRACE

### Tellus & CCI SM

Since the ERA-Interim/Land dataset is a reanalysis product, it gives different results compared to the other datasets. Different layers representing various soil depths are available. This fact needs to be considered when comparing it to other datasets. In

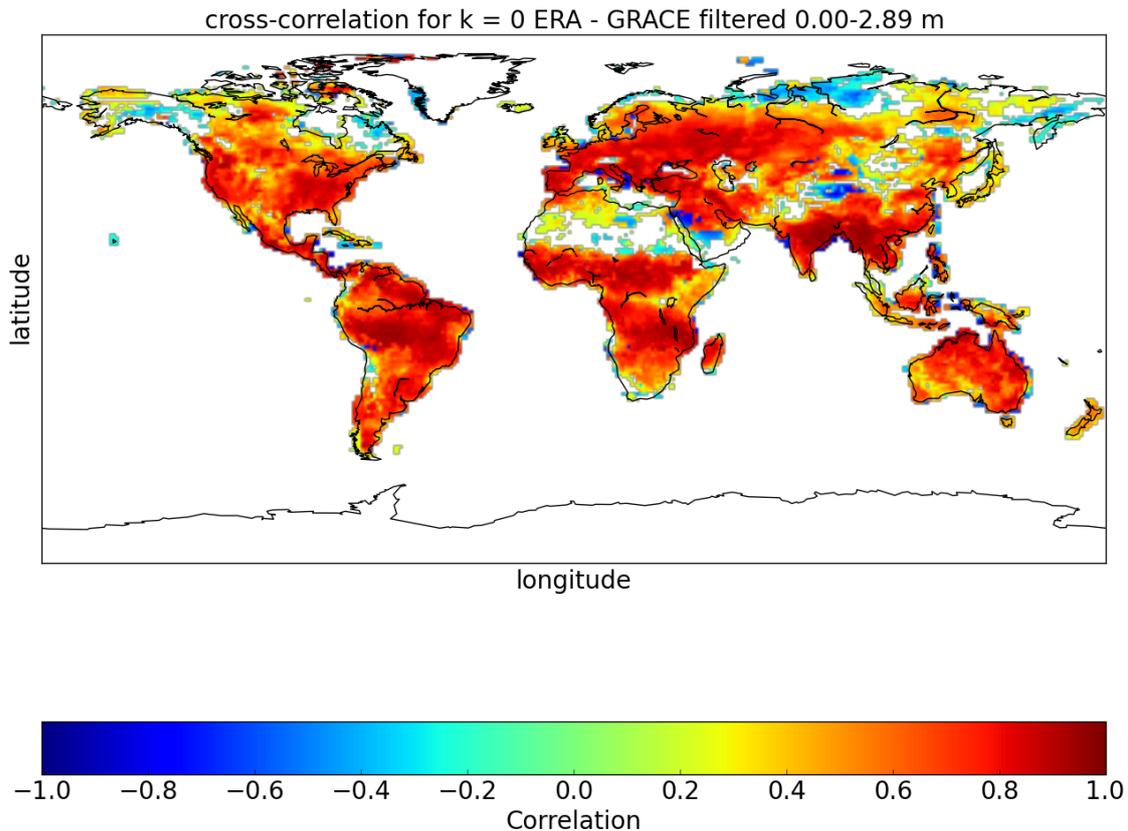


Figure 4.6: Correlation coefficients between the GRACE Tellus dataset and the ERA-Interim/Land dataset for all layers. No lag ( $k = 0$ ) was added.

order to obtain only one value per month, the mean over each month was calculated.

#### 4.4.1 ERA-Interim/Land compared to GRACE Tellus

The GRACE Tellus dataset is compared to all four layers (0.00-2.89m) of the ERA-Interim/Land dataset and only to the first one (0.00-0.07m).

##### ERA-Interim/Land Layer 1-4 (0.00 - 2.89 m)

For comparing the ERA-Interim/Land to the GRACE Tellus dataset, all layers (Layer 1 - 4) need to be merged. A weighted sum is applied, because GRACE measures changes in TWS, in which the whole ground is considered. Taking all 4 layers into account results in a soil layer depth of 2.89 m, counting from the Earth surface. For calculating the correlation coefficients between the ERA-Interim/Land and the GRACE datasets, a mean over the period 2004-2009 needs to be subtracted from

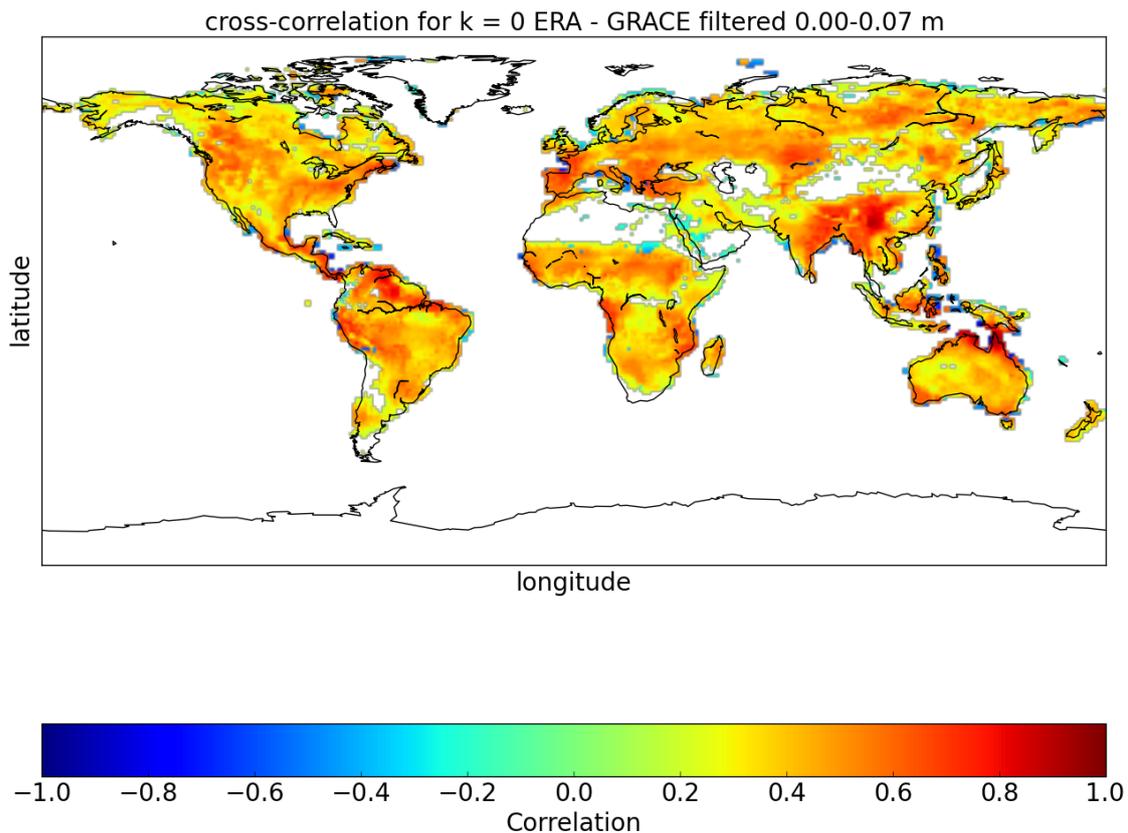


Figure 4.7: Correlation coefficients of the GRACE Tellus dataset compared to the ERA-Interim/Land dataset for Layer 1 (0.00-0.07) only. The mean over the timespan 2004-2009 was subtracted from the ERA-Interim/Land dataset. The lag between the datasets was set to zero.

the ERA-Interim/Land dataset. This leads to anomalies (in this thesis anomaly is understood as the anomaly with respect to the average value over the entire period) of volumetric soil moisture. Figure 4.6 shows the correlation coefficients of the ERA-Interim/Land dataset compared to the GRACE Tellus dataset with zero lag. Comparing the datasets for different lags (see Figure A.11 and Figure A.12) shows that the highest correlation coefficients are if no lag is applied.

#### **ERA-Interim/Land Layer 1 (0.00 - 0.07 m)**

Comparing the GRACE Tellus dataset to Layer 1 (0.00-0.07m) only, is similar to the comparison of the GRACE Tellus and the CCI SM datasets. This is because both (CCI SM and ERA-Interim/Land - Layer 1) provide only the topmost layer of soil. The mean over the period 2004 to 2009 was subtracted to retrieve anomalies

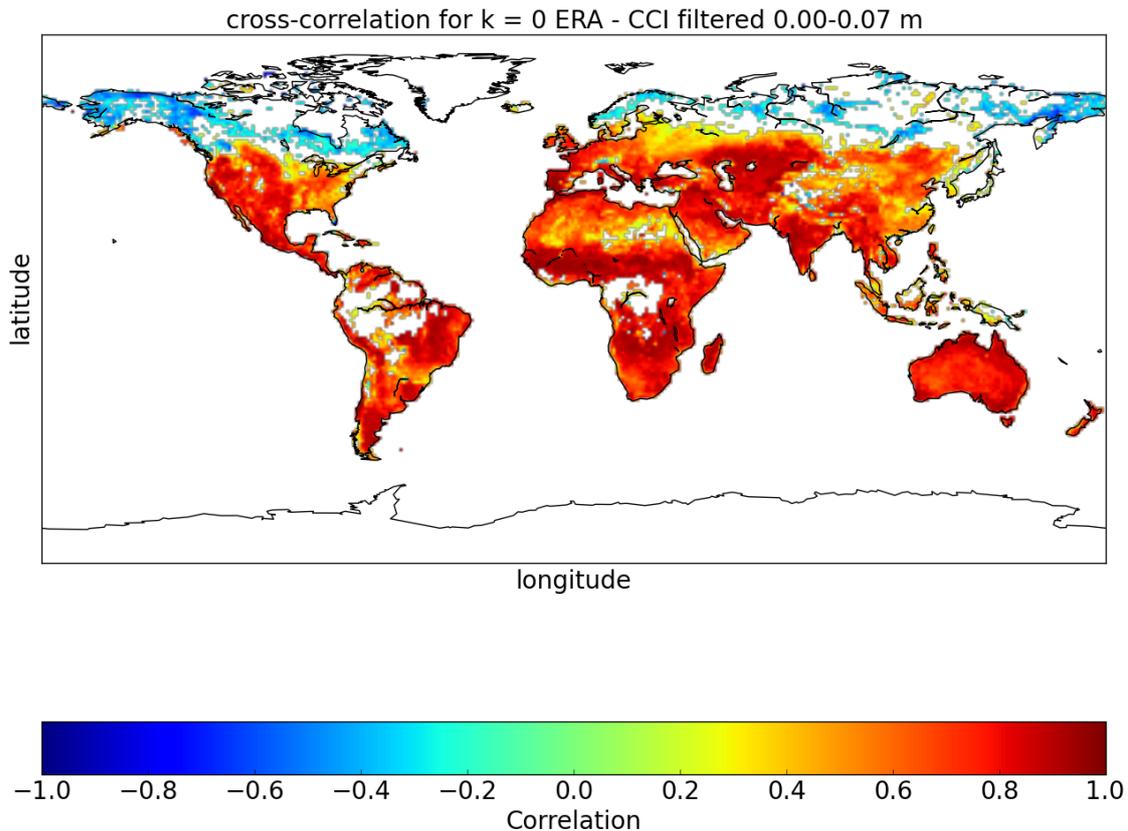


Figure 4.8: The correlation of CCI SM dataset compared to ERA-Interim/Land dataset. Only Layer 1 (0.00-0.07m) was utilized to calculate the correlation. Between the datasets there is no lag ( $k = 0$ )

of volumetric soil moisture (ERA-Interim/Land). The correlation coefficients are depicted in Figure 4.7.

The highest correlation coefficients can be achieved with a lag of one or two months, depending on the region of interest, for the GRACE Tellus dataset (see Figure A.13 and Figure A.14).

#### 4.4.2 ERA-Interim/Land compared to CCI SM

Since the CCI SM dataset shows surface soil moisture, for a comparison with the ERA-Interim/Land dataset, only Layer 1 of ERA-Interim/Land was utilized. As a consequence, the soil moisture was only compared for a depth of up to 0.07m. Figure 4.8 shows the correlation coefficients for this comparison.

No lag results in the highest correlation coefficients for most parts of the world (compare Figure A.14, Figure A.16 and Figure 4.8).

# Chapter 5

## Discussion

This chapter will provide a deeper understanding, interpretations and further information based on the results depicted in Chapter 4. The focus is on the correlation coefficients between the GRACE Tellus and the CCI SM datasets. The other datasets are used as additional information for the analysis and for a better comprehension of the results.

Having a closer look at Figure 4.1 (p. 31), the very north part of the northern hemisphere stands out as having negative correlation coefficients, whereas the remaining parts of the world are dominated by positive correlation coefficients.

Firstly an analysis for a site presenting a positive correlation coefficient is performed. Therefore Figure 5.1 displays a comparison of the GRACE Tellus dataset against the CCI SM dataset for a longitude of  $11^\circ$  and a latitude of  $8^\circ$ . This specific site is located in eastern part of Nigeria, close to the border of Cameroon. The anomalies of CCI SM (CCI SM Anomalies) are symbolized in red and the the TWSC (GRACE Tellus) are coloured in blue. It can be clearly seen that the datasets are closely related to each other. This in turns means that if the CCI SM dataset has a maximum then the GRACE Tellus dataset also has a maximum.

On closer inspection it can be seen that a small temporal lag between the different datasets exists. This is also confirmed if the correlation between those two datasets is calculated for different lags ( $k = 1, 2, \dots, 6$  see Equation 3.3). It is necessary to consider which dataset lags in reference to the other, as mentioned in Chapter 3. Especially using a lag of one month for the GRACE Tellus dataset yields a high correlation coefficient at this site. The correlation coefficients between the GRACE Tellus and the CCI SM dataset, with a lag of one month ( $k = 1$ ) for the GRACE Tellus dataset are depicted in Figure 5.2. This implies that the TWS reacts with a delay of around one month to changes in the top soil layer only. This might be caused

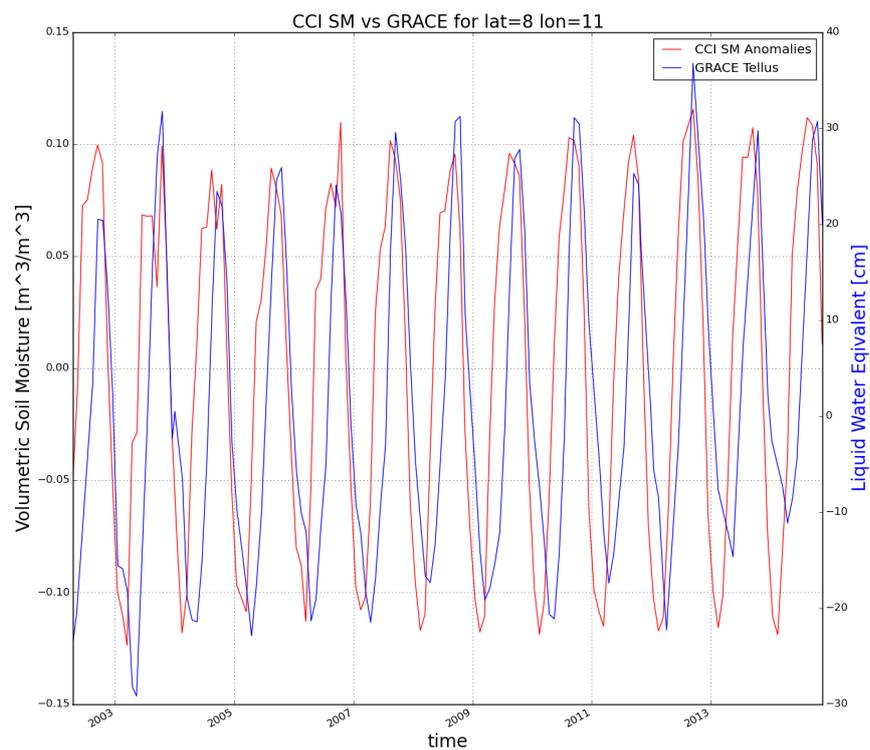


Figure 5.1: A comparison of the GRACE Tellus TWS in [ $cm$ ] and the CCI SM in [ $m^3/m^3$ ] for a longitude of  $11^\circ$  degrees and a latitude of  $8^\circ$  (in the east of Nigeria). The time period ranges from April 2002 to December 2014. This site shows a high positive correlation. The red line stand for the CCI SM dataset and the blue one for the GRACE Tellus (TWS) dataset.

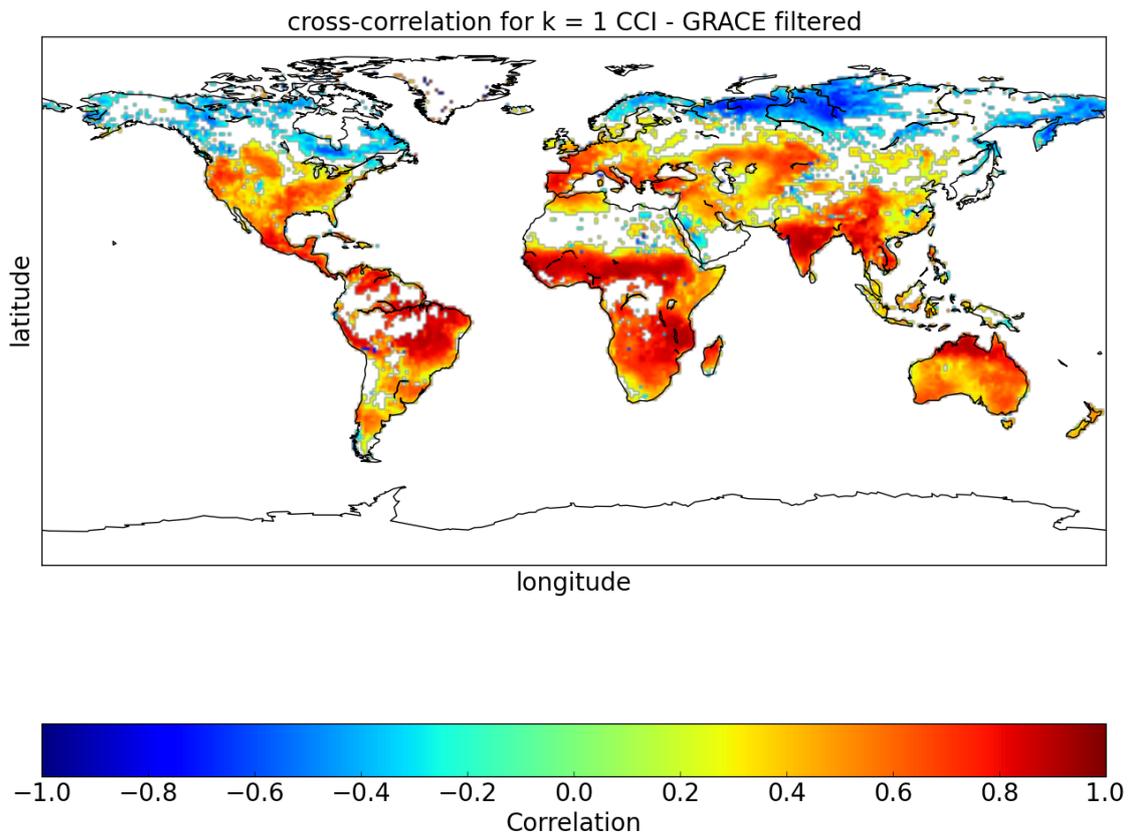


Figure 5.2: The correlation coefficients between the GRACE Tellus and the CCI SM datasets. The GRACE Tellus dataset lags one month behind CCI SM ( $k = 1$  in equation 3.3).

by the fact that CCI SM provides daily data, whereas the GRACE mission has a sampling time for the whole Earth of one month. This means that each region is only observed once per month. However, this indicates furthermore that the SM is the main driving force related to TWSC. A similar behaviour can be found when comparing the GPCP and GPCC datasets to both CCI SM and GRACE Tellus datasets, for the same region of interest (latitude of  $8^\circ$  and longitude of  $11^\circ$ ) (see Appendix B Figure B.1 and Figure B.2). In both cases an increase of the precipitation data (GPCC and GPCP) causes an increase in the GRACE Tellus TWSC and the CCI SM data. It needs to be considered that the GRACE Tellus TWSC are compared to anomalies of the GPCC and GPCP data, whereas the CCI SM dataset is compared to the precipitation falling on Earth during one month.

On the other hand, a comparison between the CCI SM and the GRACE Tellus datasets for a region with negative correlation is shown in Figure 5.3. This site is

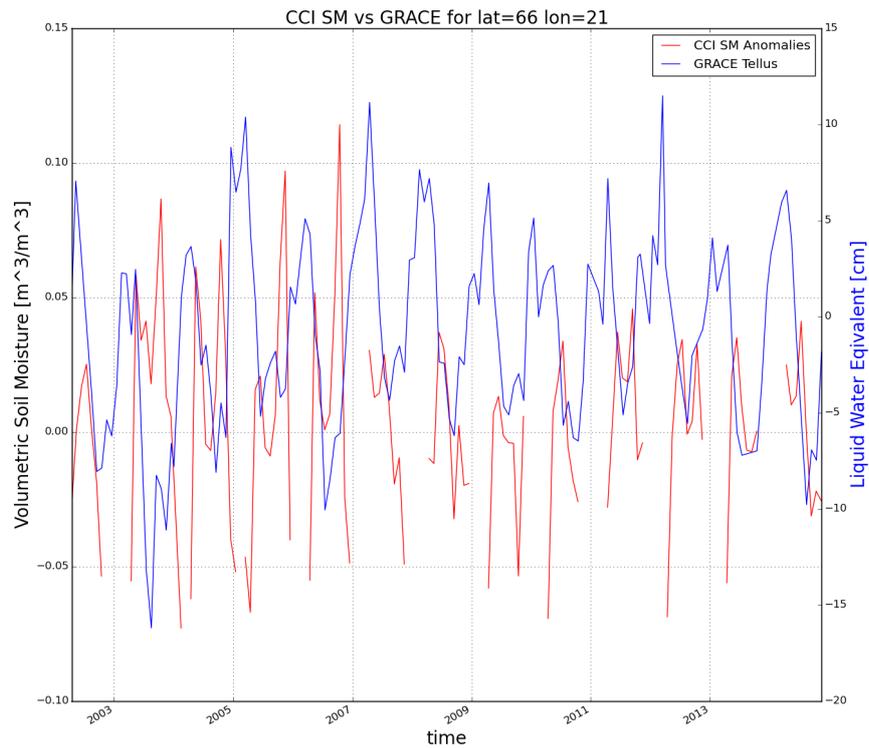


Figure 5.3: A comparison of the GRACE Tellus TWS [ $cm$ ] and the CCI SM [ $m^3/m^3$ ] for a longitude of  $21^\circ$  degrees and a latitude of  $66^\circ$  (in the North of Sweden) from April 2002 to the end of 2014. At this specific location the correlation is somewhere close  $-1$  for zero lag. The red line stands for the CCI SM dataset and the blue one for the GRACE TWS.

located in the north of Sweden at a latitude of  $66^\circ$  and a longitude of  $21^\circ$ . The red line stands for the CCI SM anomalies, while the blue one shows the TWSC. It can be seen that the CCI SM dataset regularly has data gaps. This is caused if no data is available, most likely during winter in the northern hemisphere. The fact that snow and ice are not included in the CCI SM datasets is responsible for those gaps. Having a closer look at the period from January 2007 to December 2008 (Figure 5.4) it can be recognized that the TWSC (GRACE Tellus, blue) has a maximum (minimum), while the anomalies of SM (CCI SM, red) are decreasing (increasing) or not available. Snow and ice are measured by the GRACE mission. This means that, if snow falls or liquid water freezes, the mass changes are still measurable for the GRACE mission. This assumes that precipitation increases in winter and cannot be measured related to SM using remote sensing. The CCI SM dataset only provides a measure of liquid water in the soil. As soon as it is getting colder, liquid water is decreasing and converted to ice and snow. This fact will increase the amount of snow

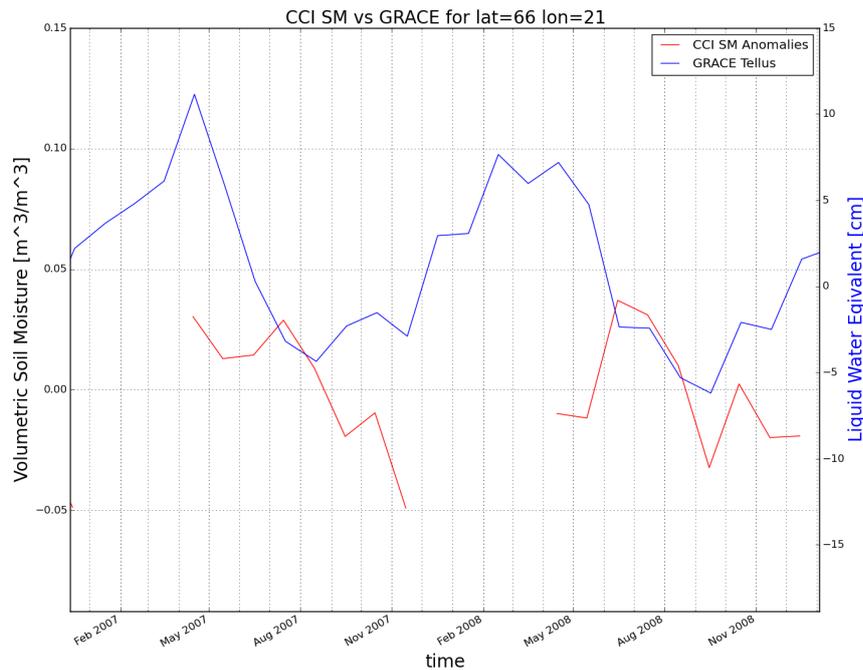


Figure 5.4: A comparison of the GRACE Tellus TWS [ $cm$ ] and the CCI SM [ $m^3/m^3$ ], located at for a longitude of  $21^\circ$  degrees and a latitude of  $66^\circ$  (in the North of Sweden) for the period January 2007 to December 2008. A negative correlation dominates at this region of interest. The red line stands for the CCI SM dataset and the blue one for the GRACE Tellus dataset. The CCI SM is given as anomalies of SM.

and ice above and under the surface. As a consequence soil moisture will decrease until it is unavailable in cold periods. In contrast during warmer periods frozen water (snow and ice) will melt and become measurable for remote sensing again. This can be explained and understood with the help of the scattering mechanisms in dry/wet soils. Pathe [2009] explains that high SM can be detected by high backscattering, caused by a high dielectric constant. The reflection and attenuation characteristics of an electromagnetic wave depend on the dielectric constant. Free water reacts as a dipole on an applied electric field. The molecules of the bounded water are immobilized. This is documented when having a look at the month from May to November in Figure 5.4. Increasing temperature leads immediately to measurable SM. In addition SM increases.

Another reason for the decrease of the SM during the cold period of the year is that less precipitation is falling. Therefore Figure 5.5 shows the comparison between the CCI SM dataset and the GPCP and GPCC datasets, for the years 2007 and 2008. The CCI SM values are displayed in blue, expressed in units of [ $m^3/m^3$ ], the

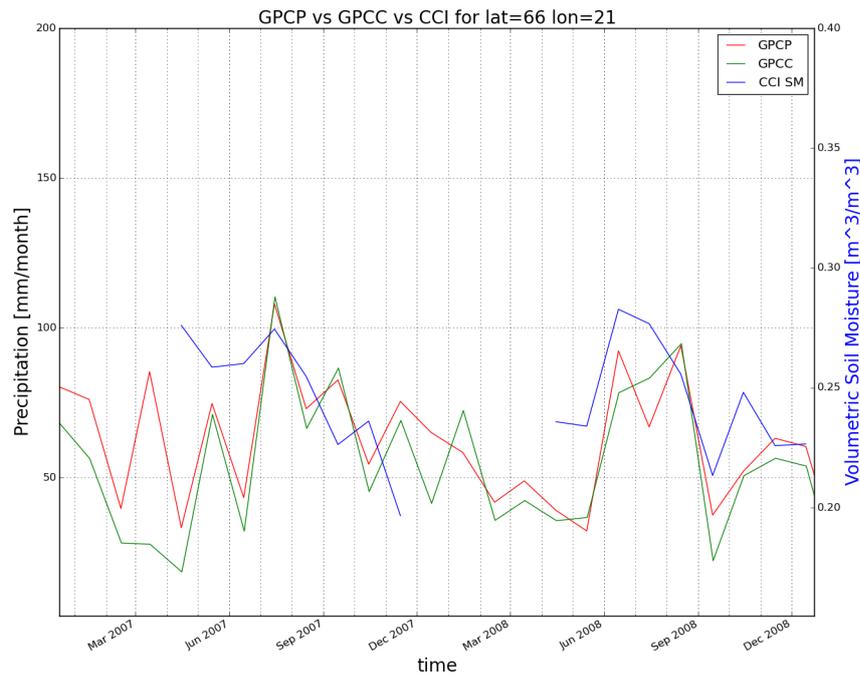


Figure 5.5: Juxtaposition of the GPCP and GPCC datasets to the CCI SM dataset for the period January 2007 to December 2008. The CCI SM dataset is depicted in blue, the GPCP dataset in red and the GPCC dataset in green.

GPCC values in green and the GPCP values in red, both in units of  $[mm/month]$ . The correspondence between the GPCC and the GPCP dataset can be easily identified. This is due to the fact that the GPCC dataset is used in the GPCP dataset. The differences between the GPCP and the GPCC dataset can be explained by the different data acquisitions. The GPCC dataset (in-situ measurements) needs to be interpolated, while the GPCP dataset compares various methods including satellite measurements, which leads to more accurate results. Additionally, snowfall is usually not captured by satellite precipitation retrievals.

The behaviour between the CCI SM dataset and the GPCC/GPCP datasets matches each other quite well during most seasons of the year. During the colder period of the year no data is available for the CCI SM dataset. As long as there is no ice and snow is present the CCI SM dataset matches the GPCP and GPCC datasets well. The comparison of the precipitation datasets (GPCC and GPCP) with the GRACE Tellus dataset is illustrated in Figure 5.6. The GRACE Tellus dataset is symbolized in blue, the GPCP dataset in red and the GPCC dataset in green. Since the GRACE Tellus dataset displays the TWSC, the anomalies of the GPCP and GPCC datasets are demonstrated in this comparison.

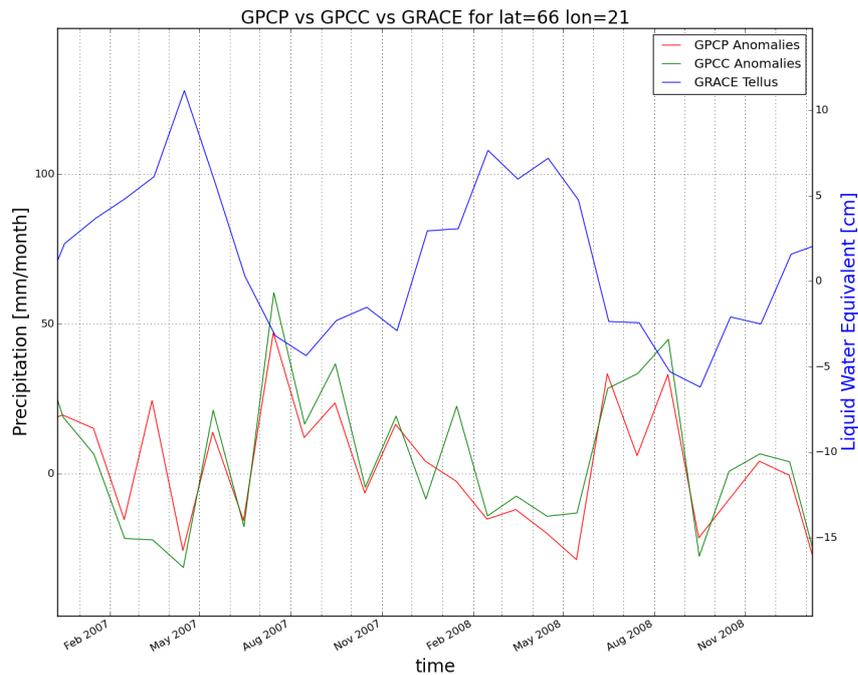


Figure 5.6: Juxtaposition of the GPCP and GPCC datasets to the GRACE Tellus dataset for the period January 2007 to December 2008. The GRACE Tellus dataset is shown in blue, the GPCP dataset in red and the GPCC dataset in green. All three datasets depict anomalies based on the mean over the period January 2004 to December 2009.

The converse behaviour of TWSC attracts attention in Figure 5.6. Over the period from January 2007 to April/May 2007 the changes in TWSC increase, but overall changes in precipitation decrease, with some fluctuation from month to month. The same phenomenon can be detected over the same months in 2008 (the increase already starts in September 2007). A reason for that circumstance might be that most of the precipitation that reaches the surface in winter will be snow, which accumulates on the surface as long as temperature remain below 0 °C. Temperatures beneath 0°C cause no water evaporation, which means that more water in the form of snow and ice will stay on or in the Earth. A comparison of TWS and the maximum temperature is given in Figure 5.7. The temperature data was downloaded from the Met Office website<sup>1</sup>. It can be seen that the maximum temperature is just above or under 0°C during colder periods. This verifies the assumption that during colder periods less water can evaporate. The lag between the datasets can be caused by the fact that in this figure air temperature is shown and the soil may probably be warmer. Another reason might be the fact that it takes some time until water is really frozen.

<sup>1</sup>[www.metoffice.gov.uk](http://www.metoffice.gov.uk)

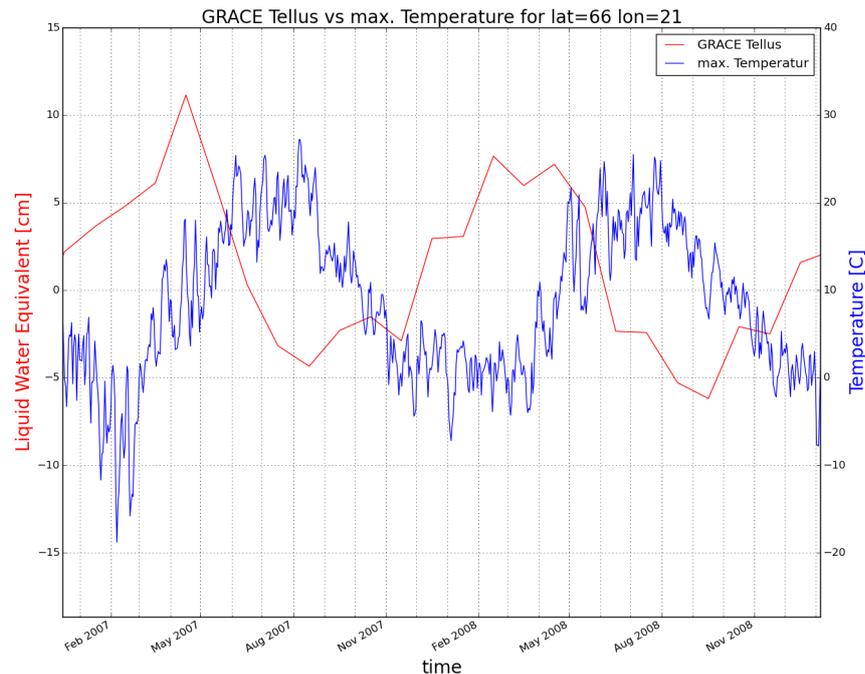


Figure 5.7: Comparison of the GRACE Tellus (red) dataset and the maximum temperature [blue] for the site in the northern part of the northern hemisphere for the years 2007 and 2008.

The reverse effect (decreasing TWS, increasing precipitation changes) can be recognized from April/May to September. In other words more water from TWS disappears than actually reaches the surface as precipitation. This might be because more water evaporates during the warmer period. This means that although more precipitation is falling, more water, including water already stored in the soil and water from precipitation, evaporates to the atmosphere and causes a decrease in the TWSC. This effect may not occur at the site located in Nigeria, because it is situated in the climatic zone of the tropics. The tropics are characterized by the same weather conditions throughout the whole year. Due to the slight variations in temperature the amount of water that evaporates will almost stay the same (more precipitation - higher SM and TWS).

Comparing the ERA-Interim/Land to both GRACE Tellus and CCI SM dataset (Figure 4.6 (p. 36) and Figure 4.8 (p. 38)) shows some slightly different results. The correlation coefficients for the CCI SM dataset and the ERA-Interim/Land dataset are positive all over the world, except for the northern part of the northern hemisphere. Again snow and ice may explain this phenomenon. The ERA-Interim/Land dataset is available for every season of the year, whereas the CCI SM dataset is unavailable during winter. Figure 5.8 depicts the comparison of the ERA-Interim/Land

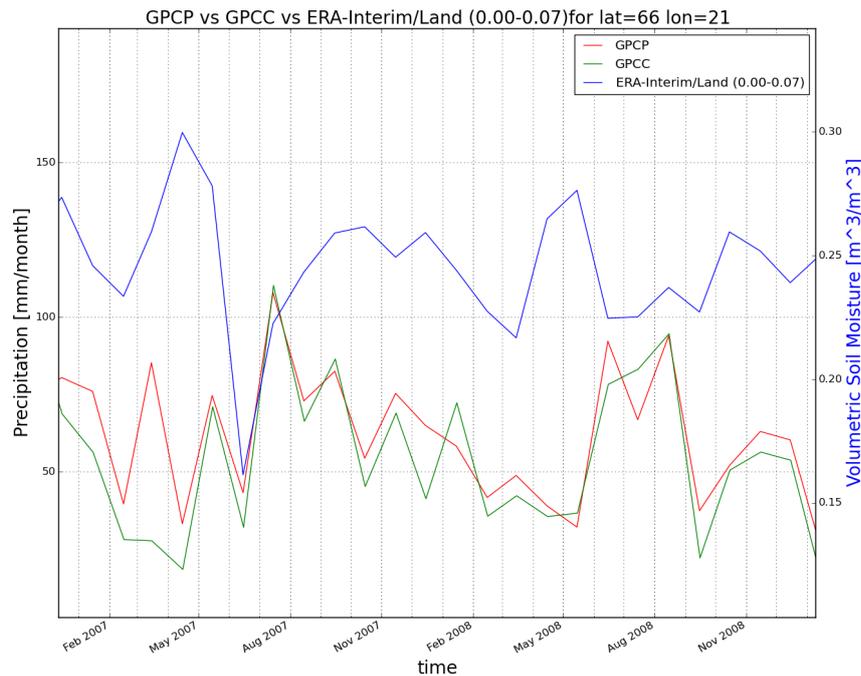


Figure 5.8: Comparison of the precipitation datasets (GPCP - red, GPCC - green) and the ERA-Interim/Land top layer (blue) for a latitude of  $66^\circ$  and a longitude of  $21^\circ$ .

Layer 1 to the GPCC and GPCP datasets for a site located in the northern part of the northern hemisphere. It can be seen that the ERA-Interim/Land dataset is available for all seasons of the year - during winter month as well. Most of the time the different datasets show similar behaviour, by means of less precipitation causes an decrease in the ERA-Interim/Land dataset as well. During the period March 2008 and September 2008 the ERA-Interim/Land shows an increase, whereas the precipitation datasets show a decrease. This might be due to the fact that the ERA-Interim/Land dataset is a reanalysis product and combines observations and forecast models. However, when comparing the ERA-Interim/Land dataset to the GRACE Tellus dataset positive correlation coefficients can be found nearly all over the world, no matter if only one or all layer are used. The correlation coefficients between Layer 1 (0.00-0.07m) and the GRACE Tellus dataset are not as high as they are for all layers and the GRACE Tellus dataset. This is due to the fact that the GRACE mission measures the TWS from the centre of the Earth to the Earth surface (in theory).

Another area of particular interest in Figure 4.1 (p. 31) is Saudi Arabia. This area has negative correlation coefficients while the areas around have mostly positive correlation coefficients, if data is available at all.

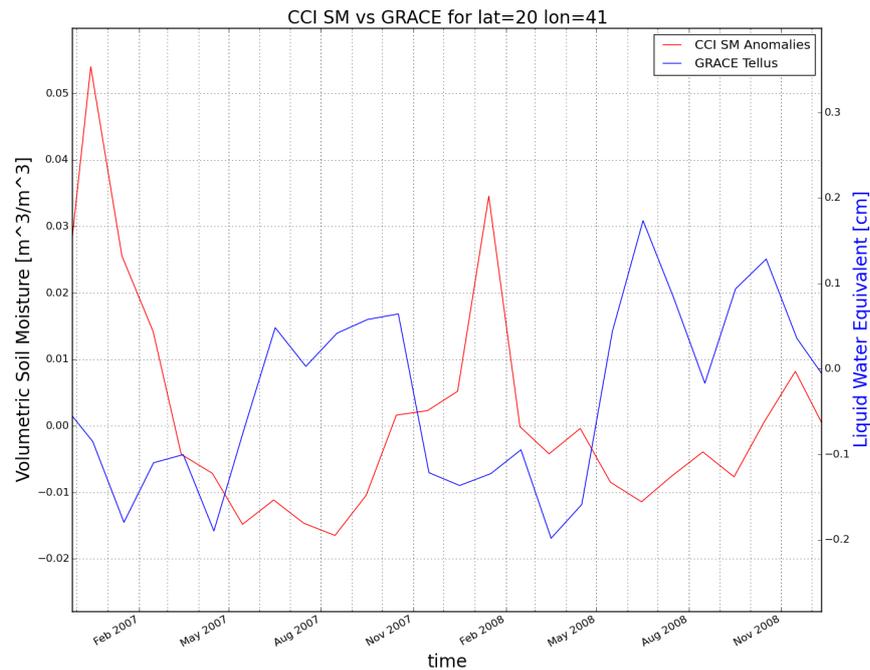


Figure 5.9: Comparison of the CCI SM dataset and the GRACE Tellus dataset, for January 2007 to December 2008. The region of interest is in Saudi Arabia, at a latitude of  $20^{\circ}$  and a longitude of  $41^{\circ}$ . GRACE Tellus is shown in blue and the anomalies of the CCI SM dataset are shown in red.

A negative correlation coefficient means that the CCI SM anomalies increase while the TWS decrease and the other way round. This can be seen in Figure 5.9, where the GRACE Tellus dataset is shown in blue and the anomalies of the CCI SM dataset are coloured in red. During warmer periods the TWS increases whereas the SM decreases. This phenomenon can probably be explained by soil conditions. In those areas the soil may be sandy and as a natural consequence water may infiltrate more easily. During heavy rainstorms the water may infiltrate through the surface to deeper layers. As a consequence, the TWS would increase, although the SM would not indicate any positive change. This can be satisfied by comparing the TWS and the SM against the precipitation data (for GRACE Tellus: Figure 5.10; for CCI SM: Figure 5.11). Whenever the precipitation, shown by the GPCP (red) and GPCC (green) datasets, increases or has a peak, the same appears in the TWS data (GRACE Tellus, blue). A decrease of SM during that period can be ascribed to the fact that water which is stored on top layers may evaporate due to high air temperatures. This means that SM decreases due to evaporation, whereas TWS increases due to the quick infiltration.

Especially during colder periods (November to February/March) SM increases and

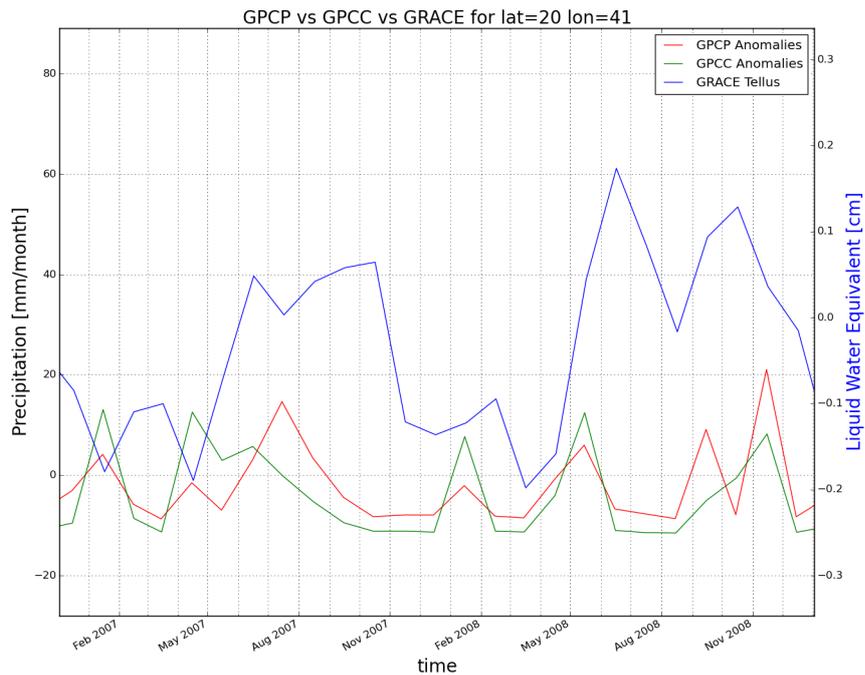


Figure 5.10: A comparison of the precipitation datasets and the GRACE Tellus dataset is given for January 2007 to December 2008. The region of interest is in Saudi Arabia, at a latitude of  $20^\circ$  and a longitude of  $41^\circ$ . The GPCP (red) and the GPCC (green) dataset displayed anomalies, referenced to the timespan January 2004 to December 2009. The GRACE Tellus dataset is symbolized in blue.

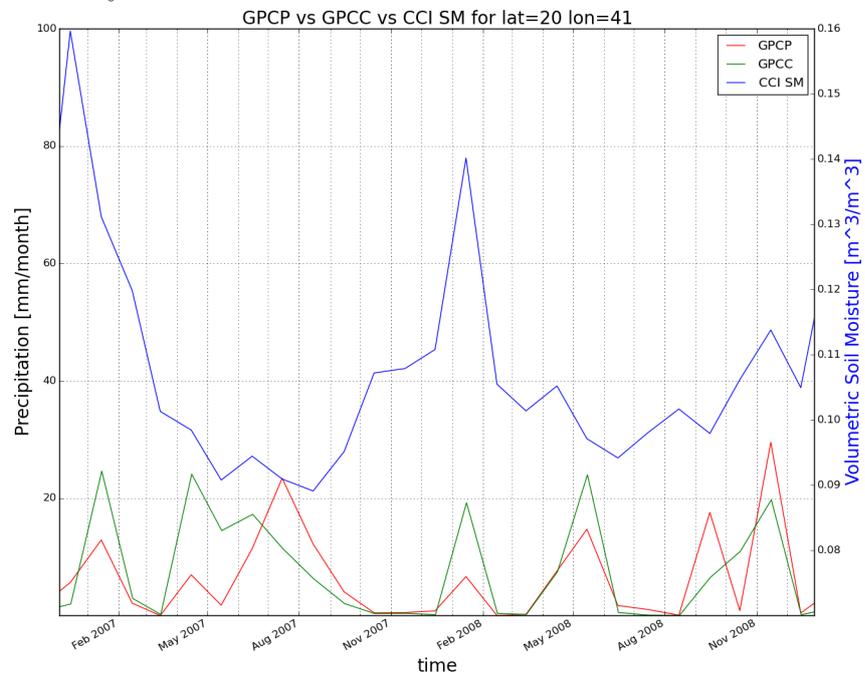


Figure 5.11: A confrontation of the GPCC (green) and GPCP (red) dataset against the CCI SM (blue) dataset for January 2007 to December 2008. The point of interest is located at a latitude of  $20^\circ$  and a longitude of  $41^\circ$ .

TWSC decreases. The increase of SM during colder periods may be explained with respect to the temperature during those months. The air temperature is colder compared to the hot summer months (up to 40° and more). This means that if the same amount of rain falls in colder months it will not evaporate as quickly as it would do during hotter months. However, this contradicts the fact that in the same period the TWS decreases. One possible explanation for this phenomenon is irrigation, where ground water or at least water from deeper soil layers, which cannot be measured by the CCI SM sensors, is being used.

Having a closer look at the GMIA shows that in the area itself no irrigation is used. However, the surroundings are irrigated. A figure displaying the AEI for Saudi Arabia in [%] of total area can be found in Appendix C, Figure C.1. A calendar depicting the total area which is actually irrigated gives even more credence to the fact that TWS decreases (Figure C.2, p. 81), because during the period from December to May the percentage of AAI increases, vis-a-vis to the period from June to November. Siebert et al. [2010] depict that around 88% of the AEI are equipped using ground water at the Arabian Peninsula. Furthermore, they mention that the percentage of *irrigation consumptive water use from ground water* (ICWU\_GW) is around 401% for the Arabian Peninsula (see Figure C.3). In this study it was found out that the Arabian Peninsula uses ground water as their main source of irrigation. This suggests that TWSC will decrease over a long period. In Figure 5.12 the opposite can be found. The trend (blue) is rising during the period 2002 to 2014. This means that the TWS is increasing as well in this specific area. This leads to the assumption that although mostly ground water is used for irrigation, aquifer recharge outpaces withdrawal.

This contradicts the results depicted in Figure 5.13, where the trend of the SM (red) shows a decrease.

However, interpreting the trend can be tricky, because the increase/decrease may also have other reasons (e.g. related to the satellites, sensors) as well. Some further research is necessary to verify these assumptions (including the reason for the increase of TWS) probably by using other datasets.

In general, no global, but only regional trends in TWSC can be found, as shown in Figure 5.14. In the red areas TWS increased in the period from 2002 to 2014, in the blue ones TWS decreased. The relationship between the trend of TWSC and irrigation using groundwater is depicted in Figure 5.15. The brighter areas in Figure 5.15 are the areas irrigated with groundwater expressed as percentage of AEI. Only areas with more than 40% AEI using groundwater are selected. It seems like there

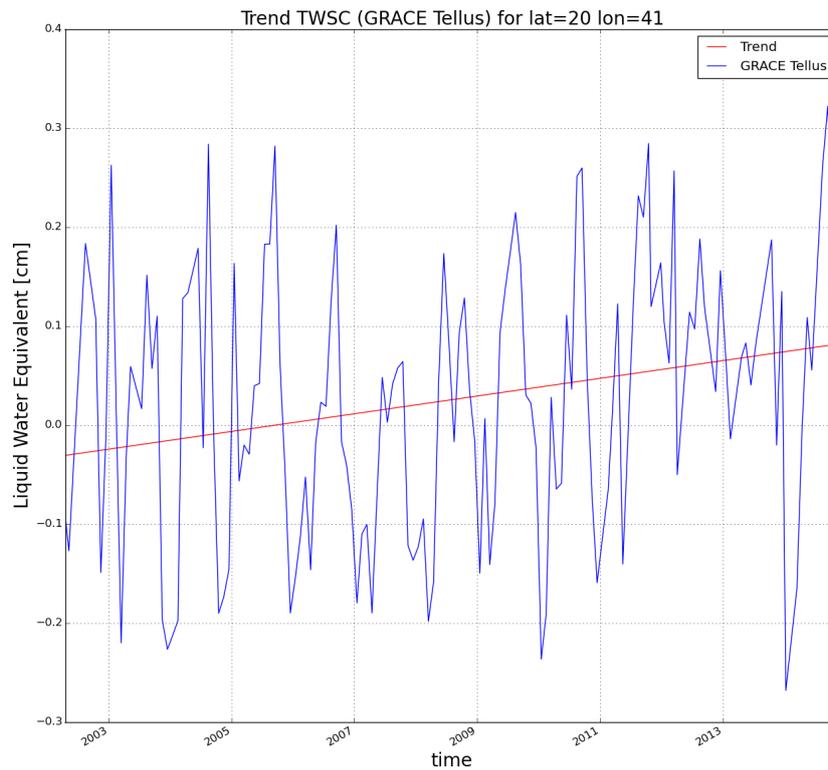


Figure 5.12: The trend of the TWSC (red) is increasing over the timespan 2002 to 2014. The blue line symbolizes the GRACE Tellus TWS changes.

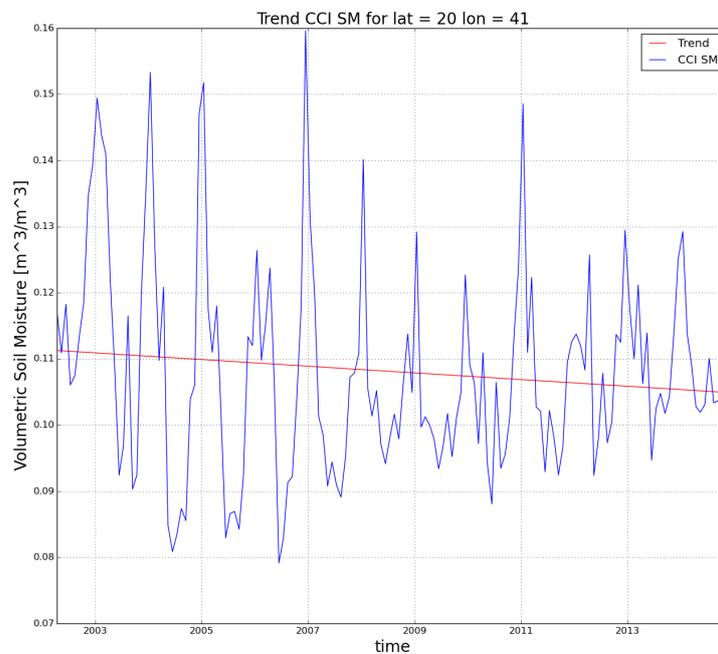


Figure 5.13: The trend of the CCI SM trend (red) is decreasing over the timespan 2002 to 2014. The blue line symbolizes the SM (CCI SM).

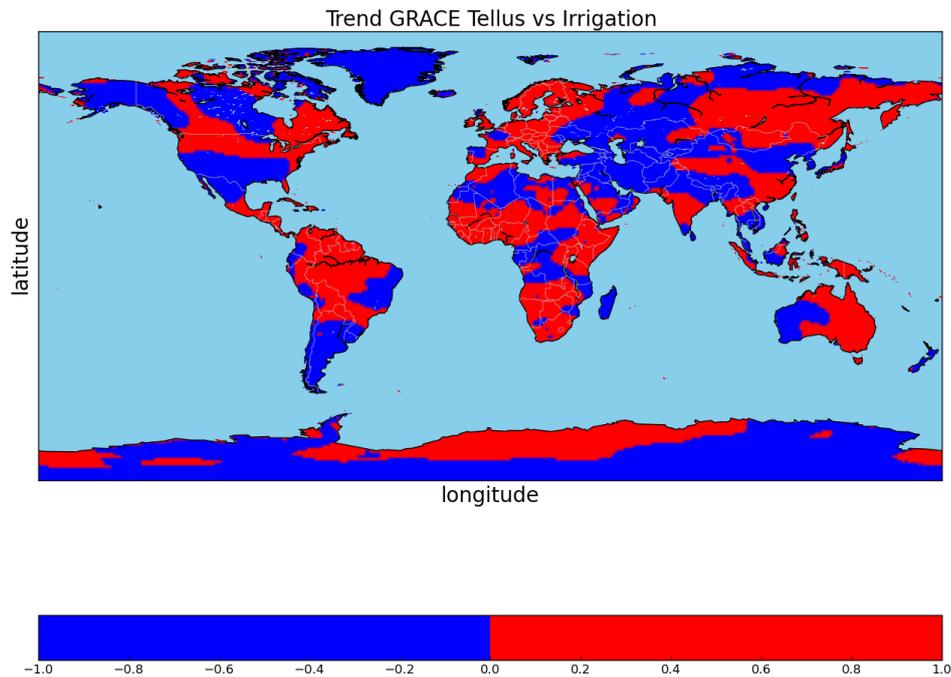


Figure 5.14: The trend of the TWSC mapped for the whole world. A red area means that the TWSC is increasing, and a dark blue one means that it is decreasing. If the TWSC is uniformly than it is shown in green.

is no relationship between irrigation and the trend of TWS. The trend of TWSC and irrigation are not linked by default. However, irrigation naturally influences SM. Therefore, the correlation between CCI SM and GRACE Tellus is compared to AEI in percentage [%] of total area as shown in Figure 5.16. A threshold of 40% was chosen assuming that areas where less than 40% of total area are AEI will not have a wide influence on SM or TWSC. This figure merges two datasets, the black areas representing the AEI, while the colored ones show the correlation coefficients between the CCI SM and the GRACE Tellus, with no lag (c.f. Figure 4.1 on p. 31). Irrigation only takes place in areas where positive correlation coefficients are predominant. Mostly irrigation is used in regions belonging to the subtropics or the temperate zone. No link between correlation coefficients and irrigation can be found, which means that no special range of correlation coefficients can be linked to AEI. However, it appears that regions with correlation coefficients ranged from 0.3 to 0.6 are more likely to be equipped for irrigation. Yet, India, Australia and especially Saudi Arabia contradict this assumption.

When comparing the ERA-Interim/Land to both CCI SM and GRACE Tellus (see

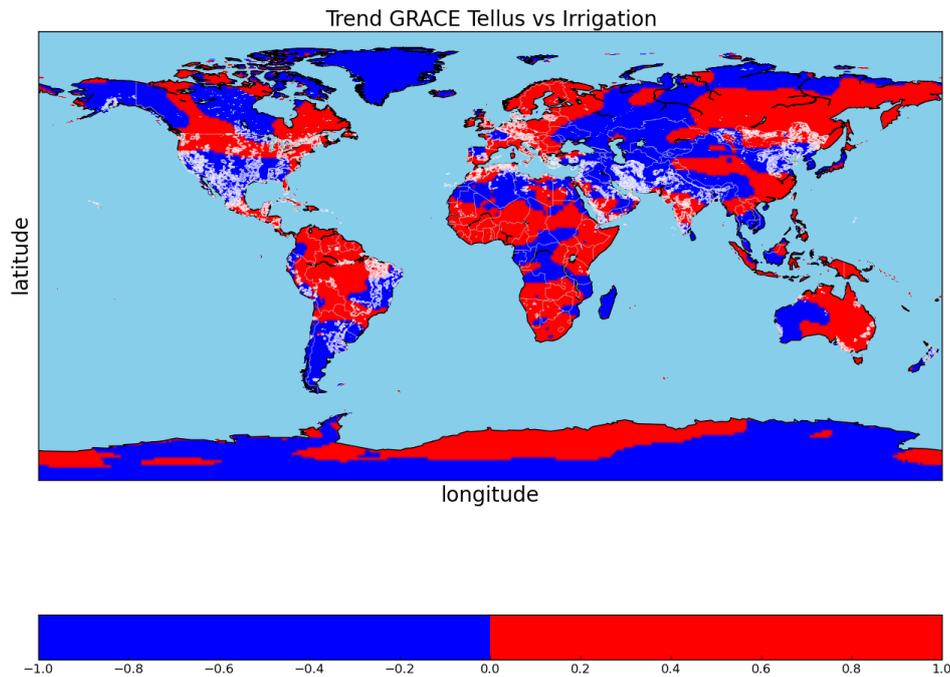


Figure 5.15: The relationship between the trend of the TWSC and the areas irrigated with groundwater ([%] of AEI). Only areas with more than 40% are selected. A positive trend is coloured red, a negative blue. The areas irrigated with groundwater are symbolized as the brighter areas over land.

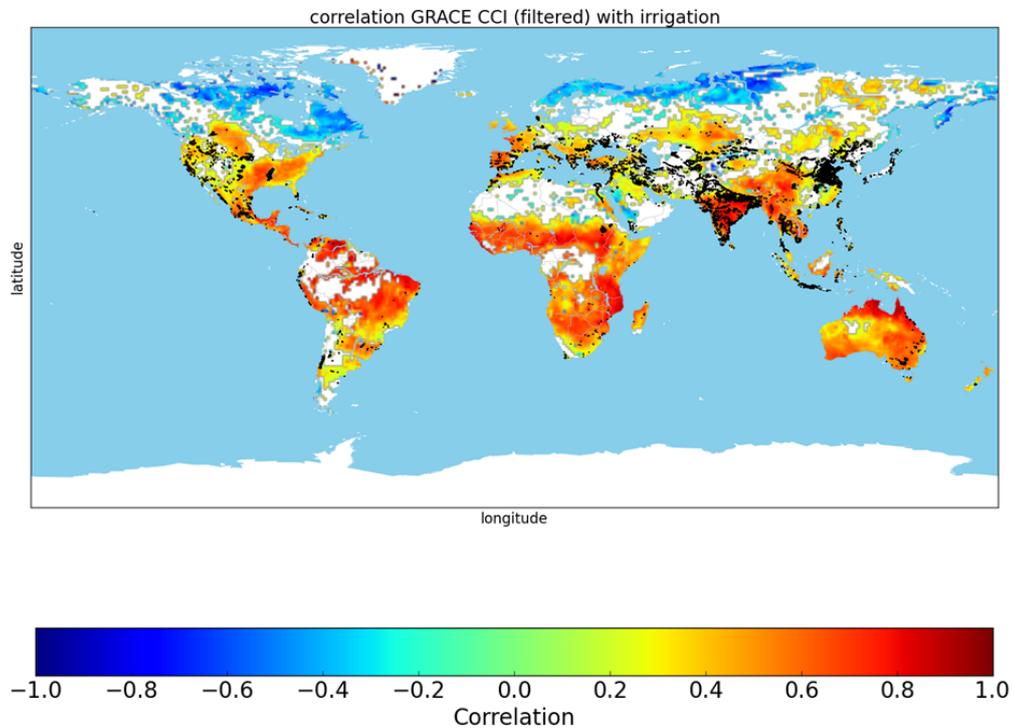


Figure 5.16: On this map the correlations between the CCI SM and GRACE Tellus datasets are overlaid by an irrigation dataset, showing each AEI where more than 40% of total area are equipped.

---

Section 4.4) the same mathematical results and corresponding interpretations are obtained. Irrespective of which layer was used, the comparison of the GRACE Tellus dataset to the ERA-Interim/Land shows negative correlation coefficients for Saudi Arabia. The comparison of the CCI dataset and the ERA-Interim/Land for Layer 1 results in positive correlation coefficients. The reason might be that those correlation coefficients are not significant and filtered because of that (p-value, see Chapter 3).

The quality of the original datasets has not been considered in this thesis. This means that it was assumed that all datasets are free from error. Nevertheless, it is important to mention that data obtained from any measurement are never completely free of errors.

# Chapter 6

## Summary & Outlook

### 6.1 Summary

The aim of this thesis was a global comparison of SM and TWS, using the CCI SM and the GRACE Tellus datasets and deducing how well these datasets match.

Correlation coefficients between the CCI SM, the GRACE Tellus, the GPCC, the GPCP and the ERA-Interim/Land datasets were calculated and mapped for the whole world. When comparing GRACE Tellus and CCI SM, positive correlation coefficients are predominant, except for the northern part of the northern hemisphere and Saudi Arabia. This leads to a more detailed investigation of those regions.

A site located in Sweden at a longitude of  $21^\circ$  and a latitude of  $66^\circ$  was chosen for more intensive research. Both CCI SM and GRACE Tellus datasets were compared to precipitation datasets (GPCC and GPCP). The opposite behaviour of SM and TWSC in this region can be explained by the impact of snow and ice. It is impossible to obtain information about frozen water using the CCI SM datasets, because they only contain data for liquid water, thus this lack of information leads automatically to a decrease in SM. However, frozen water does not cause any problems for the GRACE mission. As a consequence, TWS increases if snow accumulates on the surface. During warmer seasons frozen water will melt and is measurable again when using remote sensing - the SM will increase. One reason for the decrease of TWSC may be due to evaporation during the warmer months.

Finding a convincing explanation for the phenomenon in Saudi Arabia is quite challenging. A comparison of the different datasets was made for a site located at a latitude of  $20^\circ$  and a longitude of  $41^\circ$ . It is shown that TWS increases during warmer periods, whereas SM is decreasing. One possible reason might be related to the soil conditions in those areas. Precipitation might infiltrate easily into deeper soil layers, due to the sandy soil. This explains why TWS increases, although SM decreases.

This statement is verified by a juxtaposition of the precipitation datasets (GPCC and GPCP) and both the CCI SM and the GRACE Tellus datasets. The increase of SM during colder periods may be caused by less evaporation due to lower temperatures. On the other hand, TWS decreases during the cooler season.. One possible explanation is the use of groundwater as a source for irrigation. This is confirmed by taking GMIA into account. The area around the region of interest is irrigated more during colder months. This suggests that TWS decreases over a longer timespan, but the opposite turned out to be true.

Additionally an attempt has been made to link irrigation and the trend of TWS. However, only AEI are mapped, irrespective of which source of water was used for irrigation. In this regard no relation can be found.

## 6.2 Outlook

The region in Saudi Arabia particularly needs some further research. It is advisable to figure out how it is possible that the TWS increases, although Saudi Arabia is well-known for being one of driest countries. Using alternative datasets (e.g. Globesnow, evapotranspiration datasets, temperature,...) might help to find a convincing explanation for those phenomenons. In general, having a closer look at regions showing negative correlations by using different datasets will improve the understanding of the water cycle.

Detrending the different datasets before calculating the correlation coefficients, will deliver more accurate results for the correlation coefficients. Furthermore, the quality of the different datasets should be taken into account.

The GRACE-FO mission might measure the changes in the distance between the satellites more accurately and as a consequence TWSC may be more accurate. Making use of the GRACE-FO data will probably deliver better results. Comparing the CCI SM dataset to the combination of the GRACE and GRACE-FO datasets will improve the understanding of the relation between SM and TWS, due to the longer timespan.

# Bibliography

- Sarah Abelen and Florian Seitz. Relating satellite gravimetry data to global soil moisture products via data harmonization and correlation analysis. *Remote Sensing of Environment*, 136:89–98, 2013.
- Sarah Abelen, Florian Seitz, Rodrigo Abarca-del Rio, and Andreas Güntner. Droughts and floods in the la plata basin in soil moisture data and grace. *Remote Sensing*, 7(6):7324–7349, 2015.
- Robert F Adler, George J Huffman, Alfred Chang, Ralph Ferraro, Ping-Ping Xie, John Janowiak, Bruno Rudolf, Udo Schneider, Scott Curtis, David Bolvin, et al. The version-2 global precipitation climatology project (gpcp) monthly precipitation analysis (1979-present). *Journal of hydrometeorology*, 4(6):1147–1167, 2003.
- Mariam M Allam, Anjuli Jain Figueroa, Dennis B McLaughlin, and Elfatih AB Eltahir. Estimation of evaporation over the upper blue Nile basin by combining observations from satellites and river flow gauges. *Water Resources Research*, 2016.
- Nigel W Arnell. Climate change and global water resources. *Global environmental change*, 9:S31–S49, 1999.
- Gianpaolo Balsamo, Clement Albergel, Anton Beljaars, Souhail Boussetta, Eric Brun, Hannah Cloke, Dick Dee, Emanuel Dutra, Joaquín Muñoz-Sabater, Florian Pappenberger, et al. Era-interim/land: a global land surface reanalysis data set. *Hydrology and Earth System Sciences*, 19(1):389–407, 2015.
- Craig Blockera, John Conwayb, Luc Demortierc, Joel Heinrichd, Tom Junke, Louis Lyonsf, and Giovanni Punzig. Simple facts about p-values. 2006.
- George E. P Box and Gwilym M Jenkins; Gregory C Reinsel. *Time Series Analysis:*

- Forecasting and Control, Fourth Edition*. John Wiley Sons, 2008. ISBN 0-470-27284-8; 1-118-21087-5; 1-118-61919-6; 1-118-61906-4.
- Chad Shouquan Cheng, Heather Auld, Qian Li, and Guilong Li. Possible impacts of climate change on extreme weather events at local scale in south-central canada. *Climatic change*, 112(3-4):963–979, 2012.
- DP Dee, SM Uppala, AJ Simmons, Paul Berrisford, P Poli, S Kobayashi, U Andrae, MA Balmaseda, G Balsamo, P Bauer, et al. The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597, 2011.
- Wouter Dorigo, Wolfgang Wagner, Bernhard Bauer-Marschallinger, Daniel Chung, Richard De Jeu, Robert Parinussa, and XYi Liu. Constructing and analyzing a 32-years climate data record of remotely sensed soil moisture. In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, pages 2028–2031. IEEE, 2012.
- FAO. Understanding aquastat - fao’s global water information system. 2014.
- Frank Flechtner. Aod1b product description document for product releases 01 to 04 (rev. 3.1, april 13, 2007). *GRACE project document*, pages 327–750, 2007.
- A Geruo, John Wahr, and Shijie Zhong. Computations of the viscoelastic response of a 3-d compressible earth to surface loading: an application to glacial isostatic adjustment in antarctica and canada. *Geophysical Journal International*, 192(2):557–572, 2013.
- Davor Horvatic, H Eugene Stanley, and Boris Podobnik. Detrended cross-correlation analysis for non-stationary time series with periodic trends. *EPL (Europhysics Letters)*, 94(1):18007, 2011.
- George J Huffman, Robert F Adler, Philip Arkin, Alfred Chang, Ralph Ferraro, Arnold Gruber, John Janowiak, Alan McNab, Bruno Rudolf, and Udo Schneider. The global precipitation climatology project (gpcp) combined precipitation dataset. *Bulletin of the American Meteorological Society*, 78(1):5–20, 1997.

- JCGM. Evaluation of measurement data—guide to the expression of uncertainty in measurement. *JCGM*, 2008.
- JCGM. International vocabulary of metrology – basic and general concepts and associated terms. *JCGM*, 2012.
- Christopher Jekeli. Alternative methods to smooth the earth’s gravity field. 1981.
- Maria Karbon, Johannes Böhm, Dudy D Wijaya, and Harald Schuh. Atmospheric effects on gravity space missions. In *Atmospheric Effects in Space Geodesy*, pages 159–180. Springer, 2013.
- R Kidd and E. Haas. Esa climate change initiative phase ii soil moisture - soil moisture ecv product user guide. 2015.
- Sonja Leitner. Response of soil respiration to repeated extreme events in a temperate beech forest in austria. In *2015 AGU Fall Meeting*. Agu, 2015.
- Yi Y Liu, RM Parinussa, Wouter A Dorigo, Richard AM De Jeu, Wolfgang Wagner, AIJM Van Dijk, Matthew F McCabe, and JP Evans. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrology and Earth System Sciences*, 15(2):425–436, 2011.
- YY Liu, Wouter A Dorigo, RM Parinussa, Richard AM de Jeu, Wolfgang Wagner, Matthew F McCabe, JP Evans, and AIJM Van Dijk. Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment*, 123:280–297, 2012.
- Torsten Mayer-Gürr et al. *Gravitationsfeldbestimmung aus der Analyse kurzer Bahnbögen am Beispiel der Satellitenmissionen CHAMP und GRACE*. PhD thesis, Universitäts-und Landesbibliothek Bonn, 2006.
- Carsten Pathe. *Surface soil moisture retrieval using Envisat ASAR global mode data*. PhD thesis, 2009. Wien, Techn. Univ., Diss., 2009.
- M Rodell and JS Famiglietti. An analysis of terrestrial water storage variations in

- Illinois with implications for the gravity recovery and climate experiment (grace). *Water Resources Research*, 37(5):1327–1339, 2001.
- M Rodell and JS Famiglietti. The potential for satellite-based monitoring of groundwater storage changes using grace: the high plains aquifer, central us. *Journal of Hydrology*, 263(1):245–256, 2002.
- BS Sheard, Gerhard Heinzl, Karsten Danzmann, DA Shaddock, WM Klipstein, and WM Folkner. Intersatellite laser ranging instrument for the grace follow-on mission. *Journal of Geodesy*, 86(12):1083–1095, 2012.
- S Siebert, V Henrich, K Frenken, and J Burke. Update of the global map of irrigation areas to version 5. *FAO: Rome, Italy*, 2013.
- Stefan Siebert, Jacob Burke, Jean-Marc Faures, Karen Frenken, Jippe Hoogeveen, Petra Döll, and Felix Theodor Portmann. Groundwater use for irrigation—a global inventory. *Hydrology and Earth System Sciences*, 14(10):1863–1880, 2010.
- Sean Swenson, Don Chambers, and John Wahr. Estimating geocenter variations from a combination of grace and ocean model output. *Journal of Geophysical Research: Solid Earth*, 113(B8), 2008.
- Byron D Tapley, S Bettadpur, M Watkins, and Ch Reigber. The gravity recovery and climate experiment: Mission overview and early results. *Geophysical Research Letters*, 31(9), 2004.
- W Wagner, Wouter Dorigo, Richard de Jeu, Diego Fernandez, Jerome Benveniste, Eva Haas, and Martin Ertl. Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture. In *Proceedings of the XXII International Society for Photogrammetry and Remote Sensing (ISPRS) Congress, Melbourne, Australia*, volume 25, 2012.
- John Wahr, Mery Molenaar, and Frank Bryan. Time variability of the earth’s gravity field: Hydrological and oceanic effects and their possible detection using grace. *Journal of Geophysical Research: Solid Earth*, 103(B12):30205–30229, 1998.

---

Cort J Willmott, Clinton M Rowe, and William D Philpot. Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *The American Cartographer*, 12(1):5–16, 1985.

Pingping Xie and Phillip A Arkin. Global monthly precipitation estimates from satellite-observed outgoing longwave radiation. *Journal of Climate*, 11(2):137–164, 1998.

# Appendix A

## Map of correlation coefficients

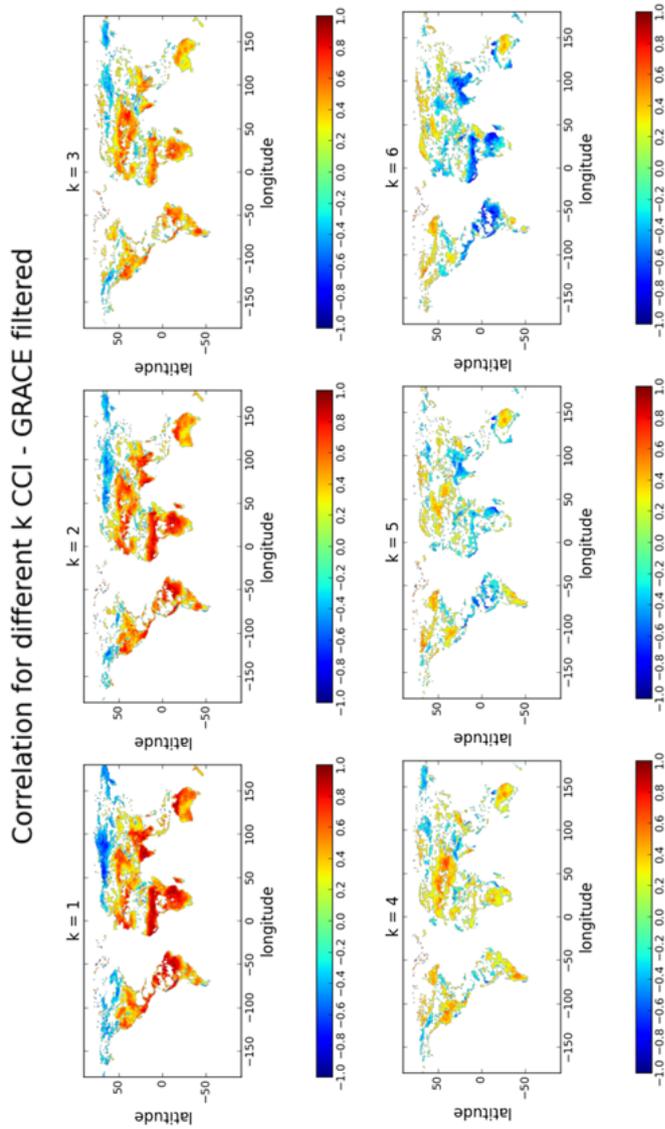


Figure A.1: The correlation coefficients between the CCI SM and GRACE Tellus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

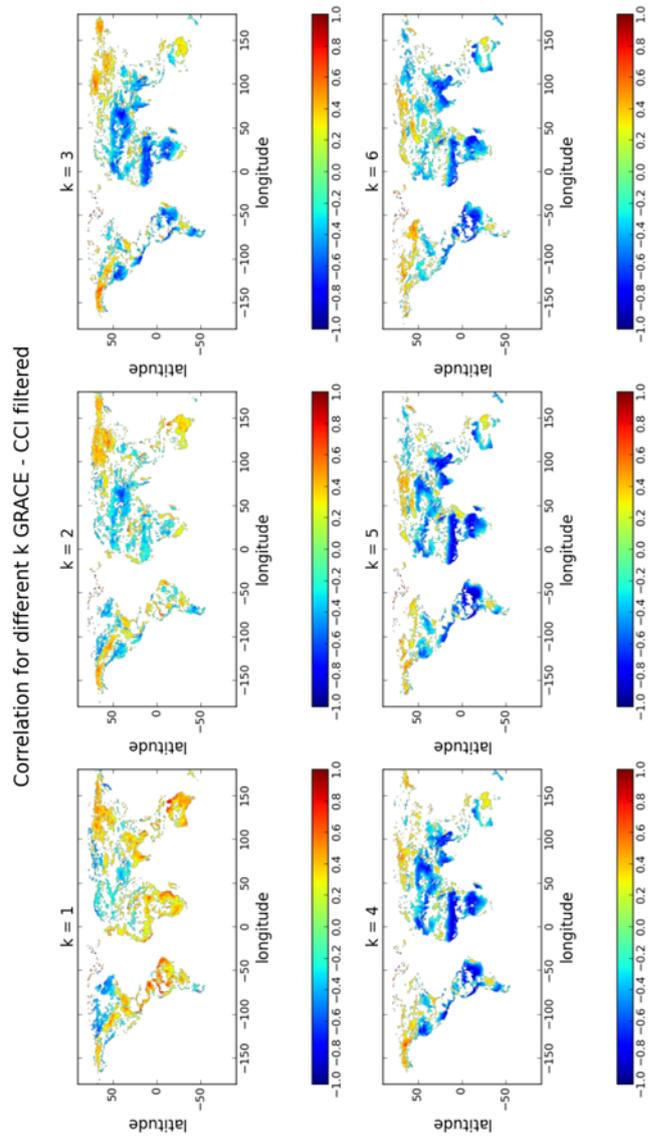


Figure A.2: The correlation coefficients between the CCI SM and GRACE  
Tellus datasets for different lags. The CCI SM dataset lags cor-  
responding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

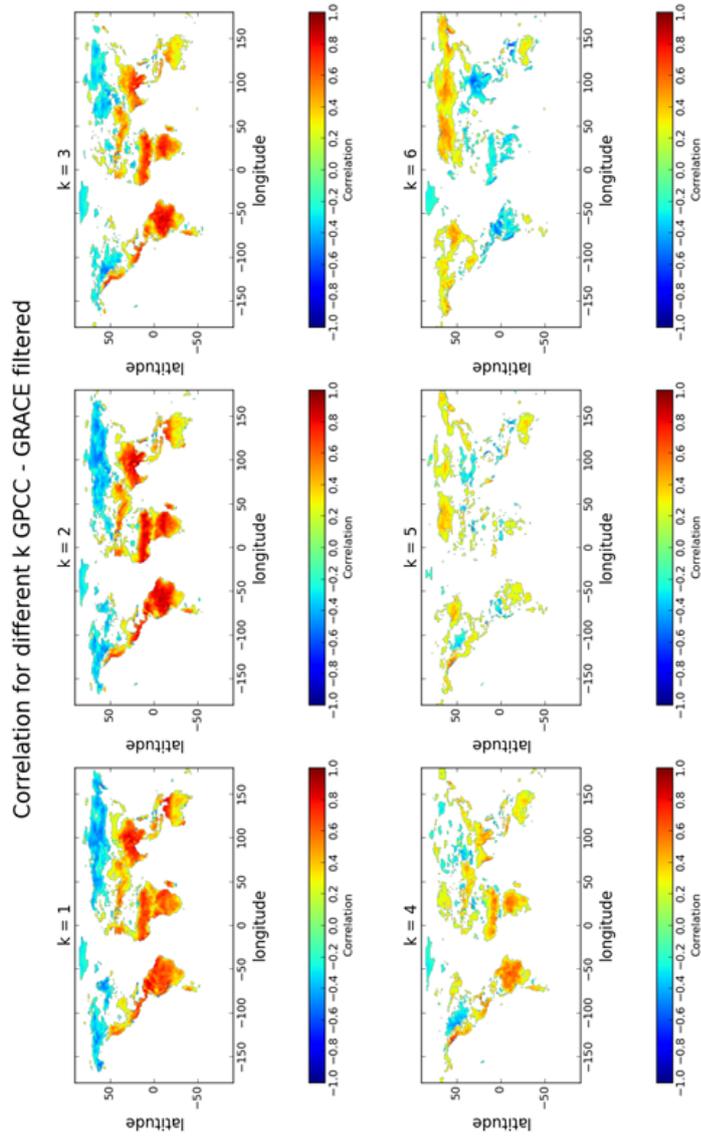


Figure A.3: The correlation coefficients between the GPCC and GRACE Tel- lus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

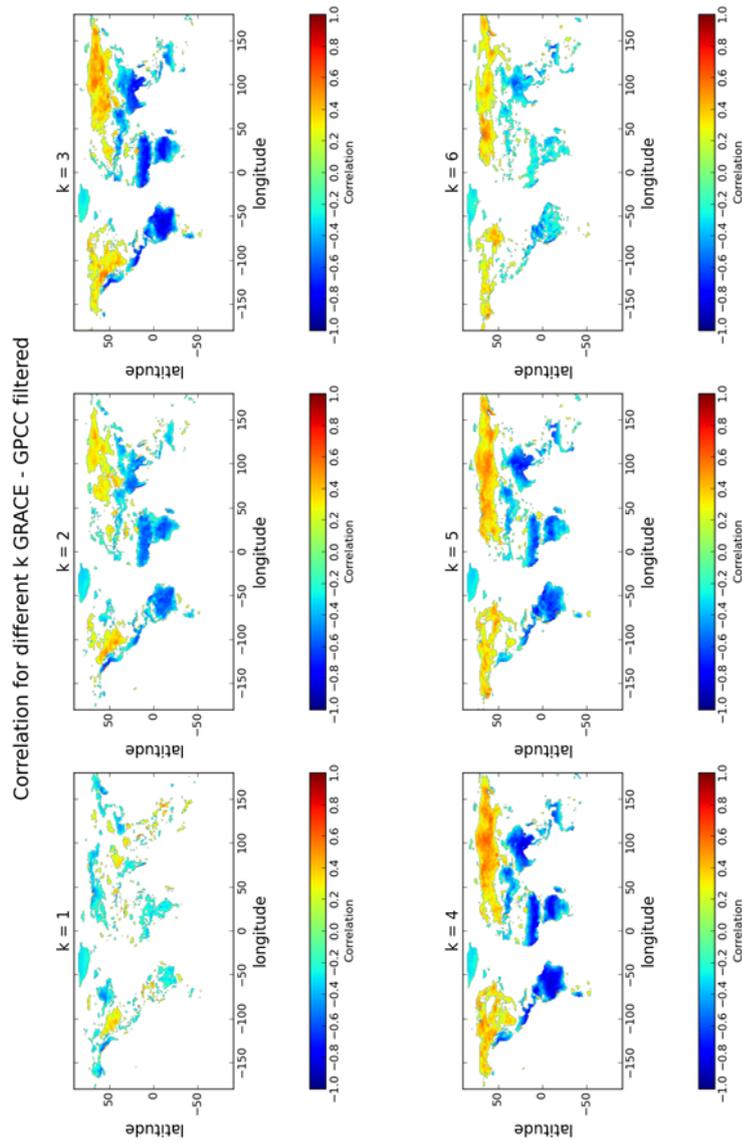


Figure A.4: The correlation coefficients between the GPCC and GRACE Telus datasets for different lags. The GPCC dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

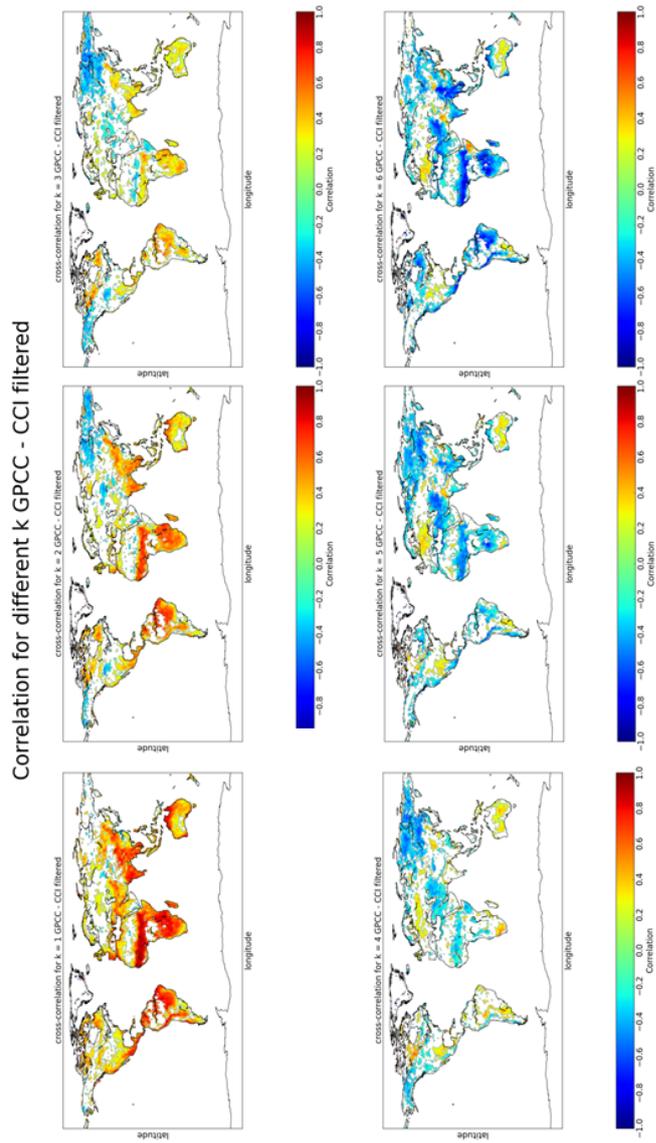


Figure A.5: The correlation coefficients between the GPCC and CCI SM datasets for different lags. The CCI SM dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

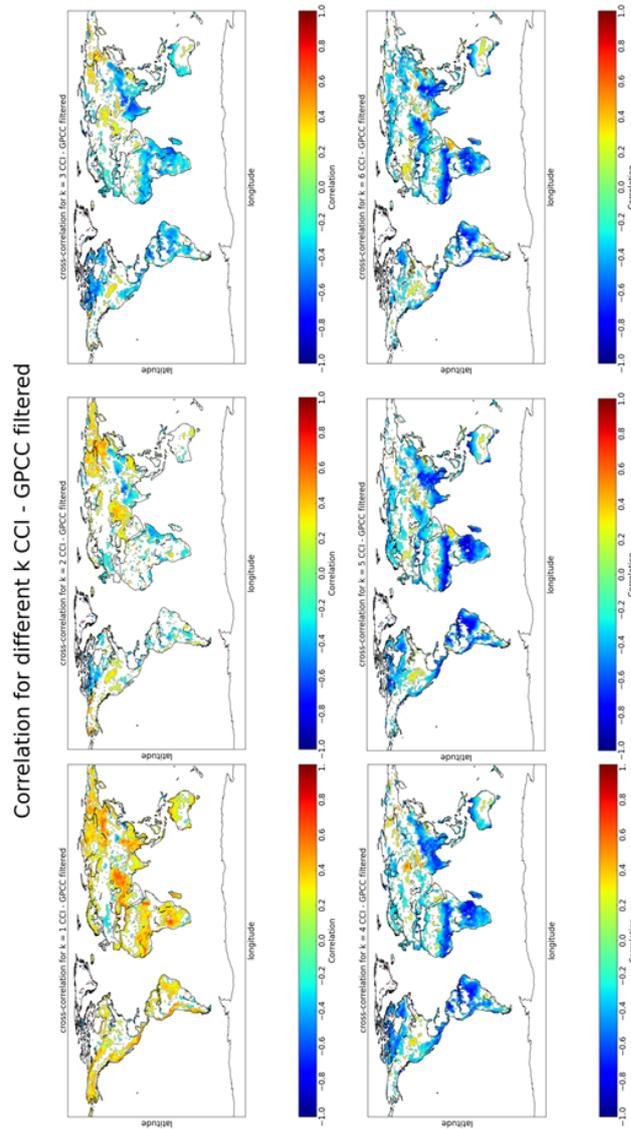


Figure A.6: The correlation coefficients between the GPCC and CCI SM datasets for different lags. The GPCC dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

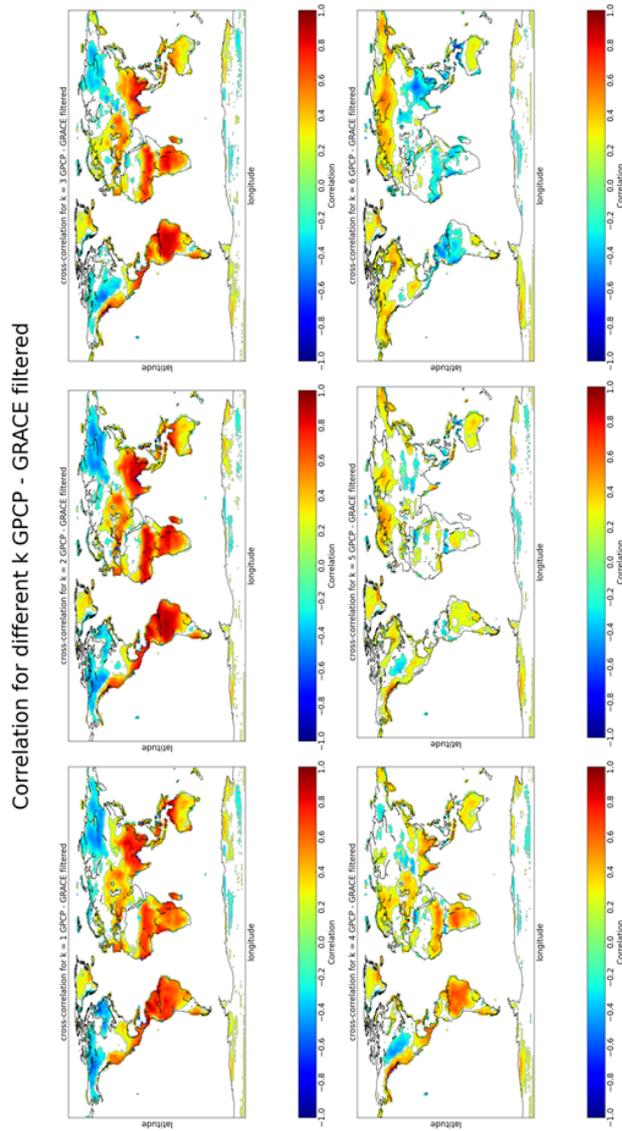


Figure A.7: The correlation coefficients between the GPCP and GRACE Tellus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

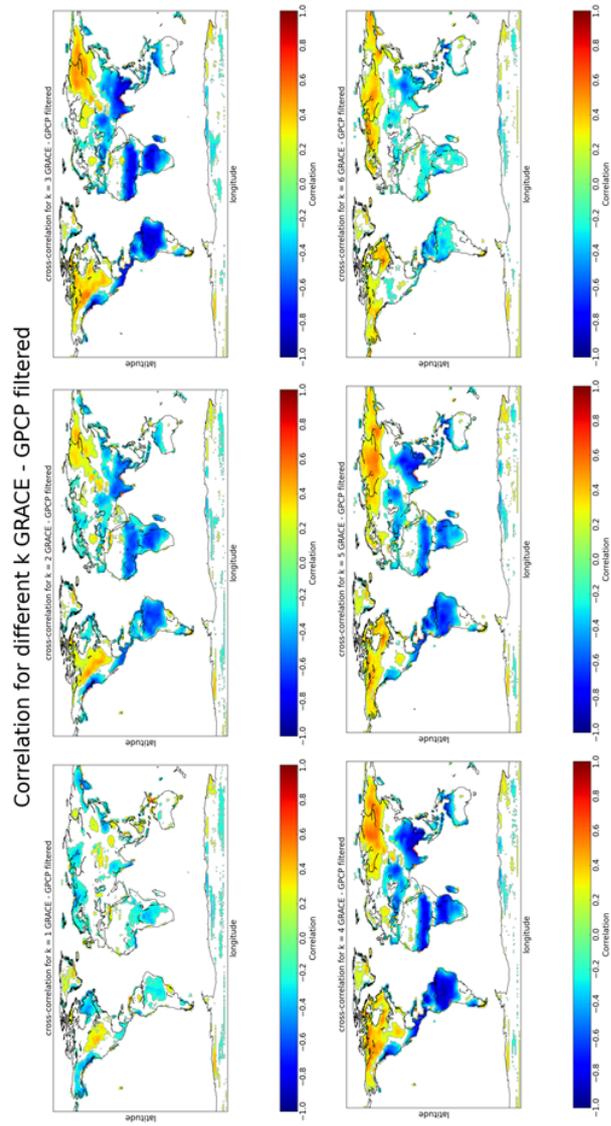


Figure A.8: The correlation coefficients between the GPCP and GRACE Tel-  
 lus datasets for different lags. The GPCP dataset lags corre-  
 sponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

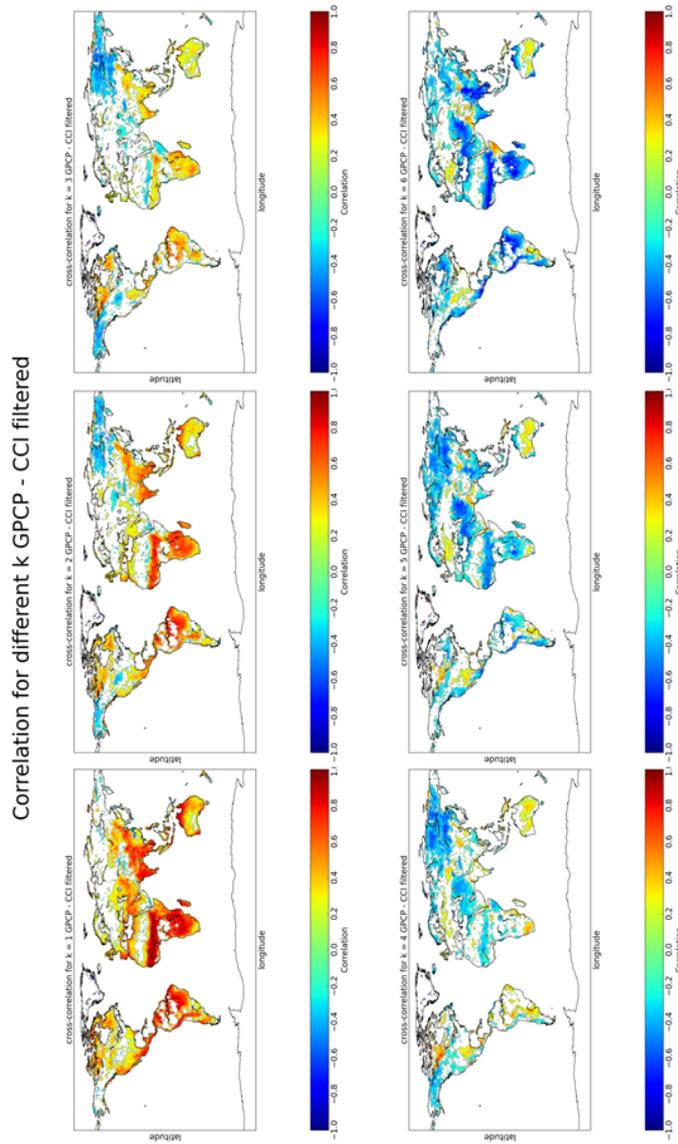


Figure A.9: The correlation coefficients between the GPCP and CCI SM datasets for different lags. The CCI SM dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

Correlation for different k CCI - GPCP filtered

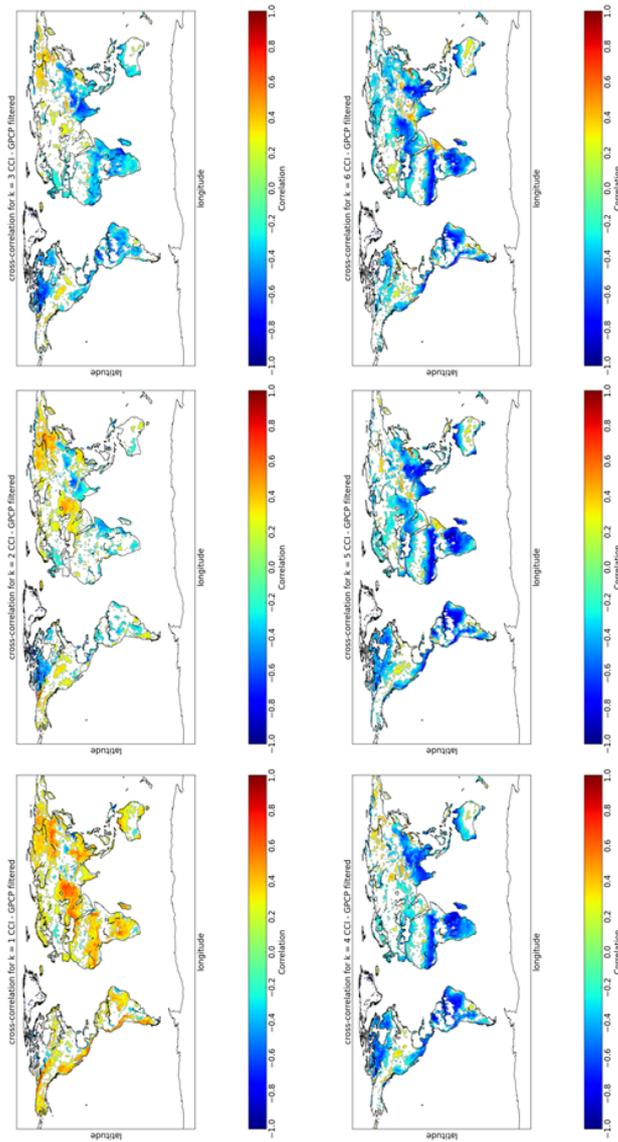


Figure A.10: The correlation coefficients between the GPCP and CCI SM datasets for different lags. The GPCP dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

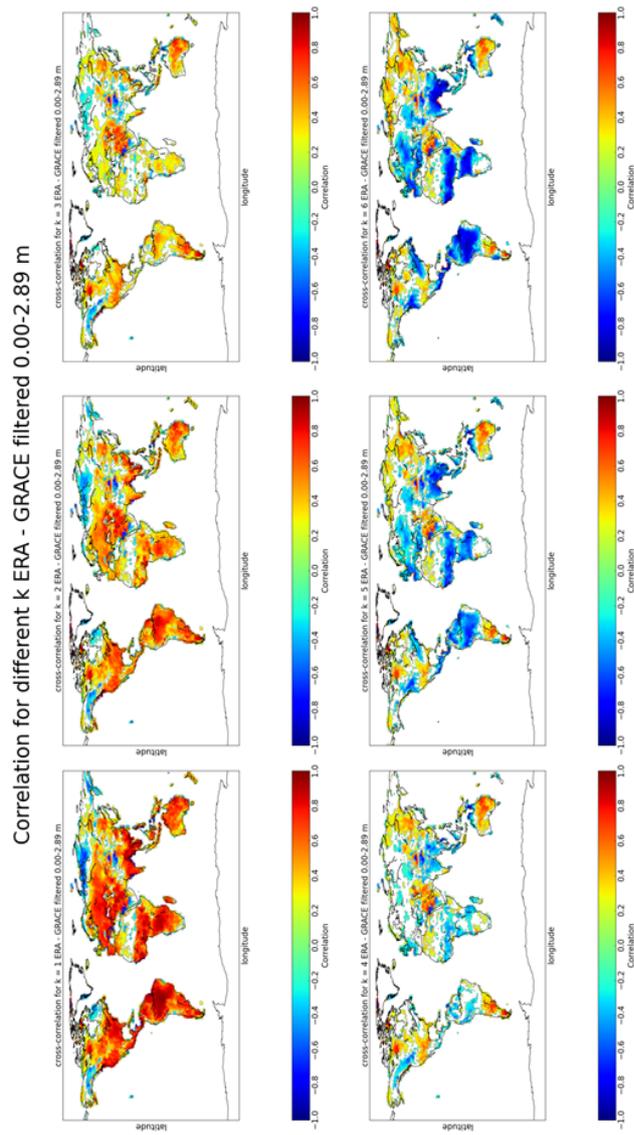


Figure A.11: The correlation coefficients between the ERA-Interim/Land all Layers and GRACE Tellus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

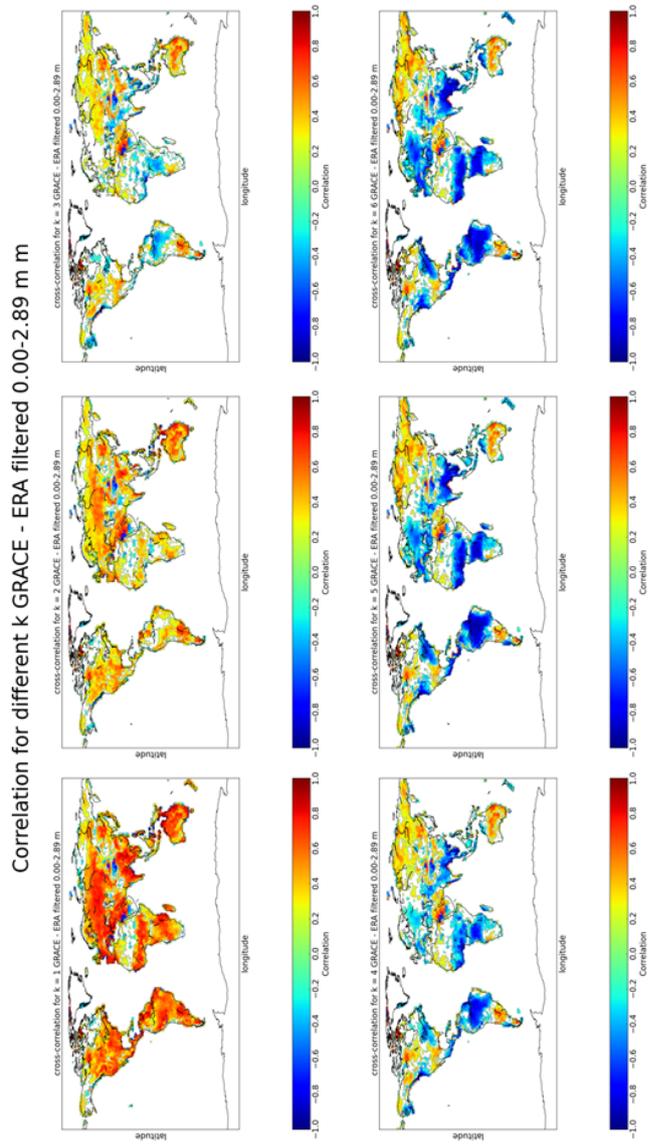


Figure A.12: The correlation coefficients between the ERA-Interim/Land all Layers and GRACE Tellus datasets for different lags. The ERA-Interim/Land dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

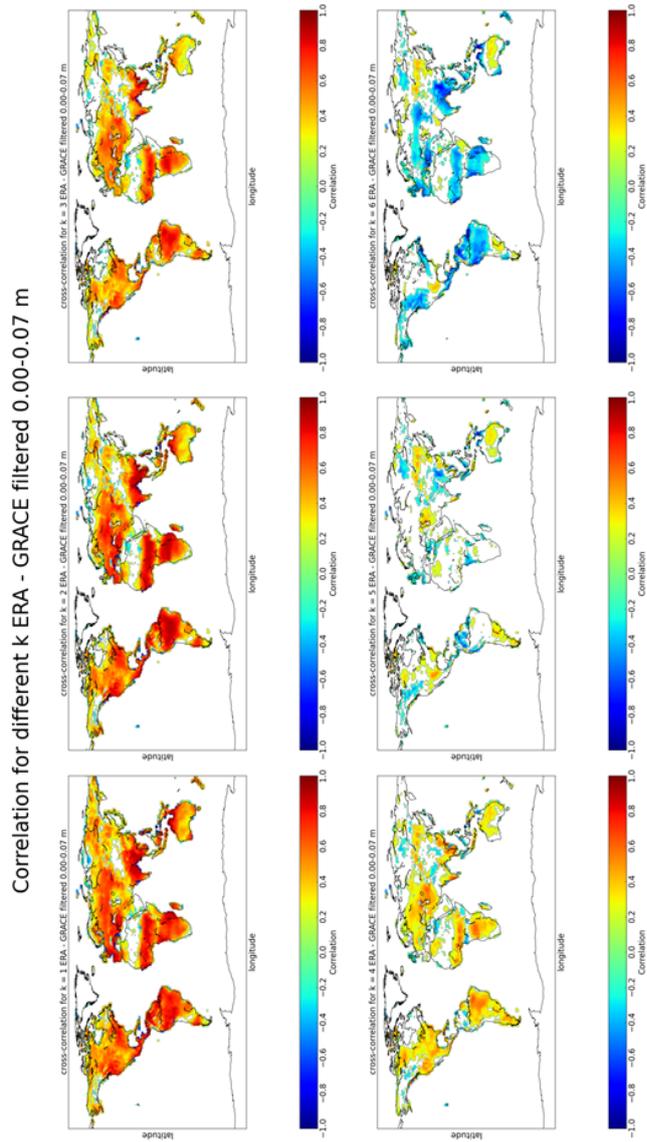


Figure A.13: The correlation coefficients between the ERA-Interim/Land Layer 1 and GRACE Tellus datasets for different lags. The GRACE Tellus dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

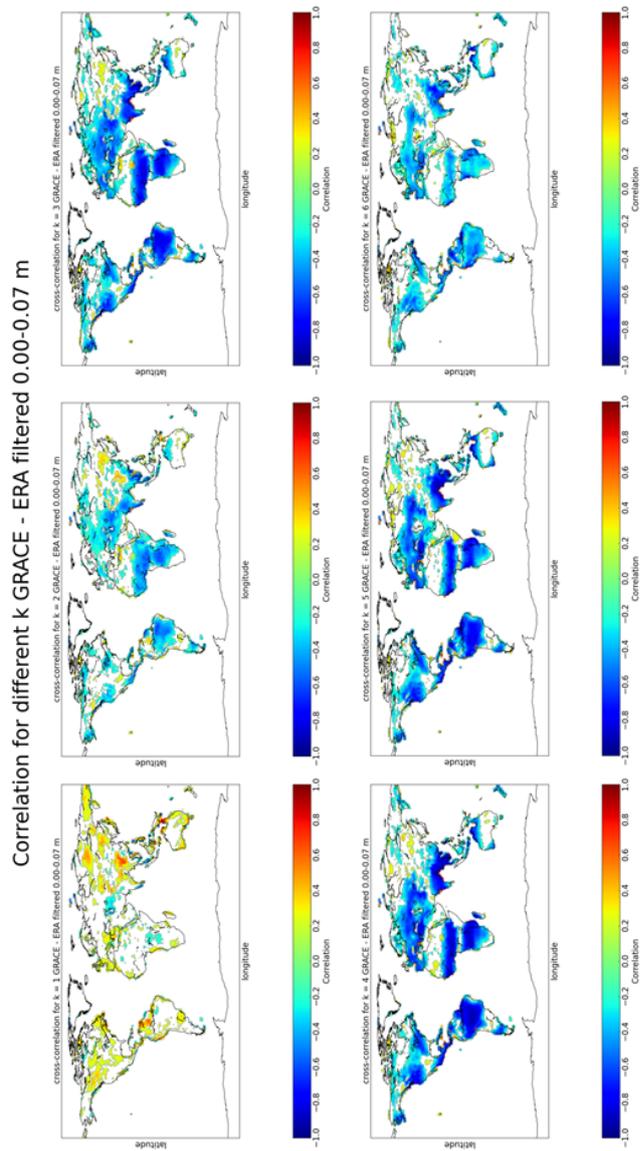


Figure A.14: The correlation coefficients between the ERA-Interim/Land Layer 1 and GRACE Tellus datasets for different lags. The ERA-Interim/Land dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

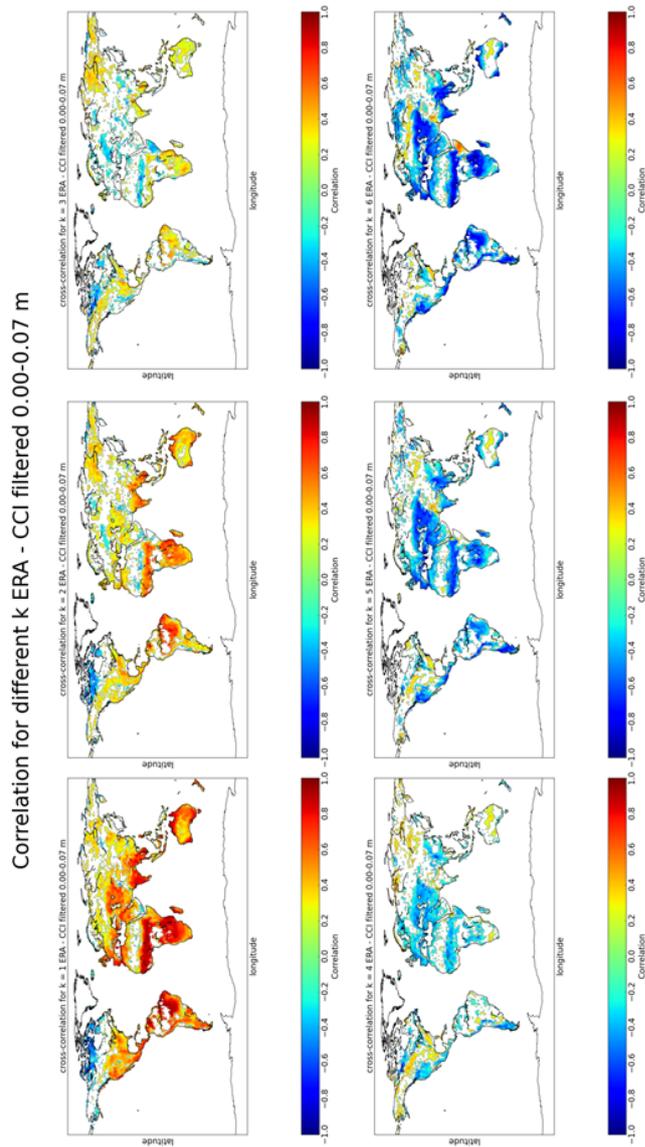


Figure A.15: The correlation coefficients between the ERA-Interim/Land Layer 1 and CCI SM datasets for different lags. The CCI SM dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

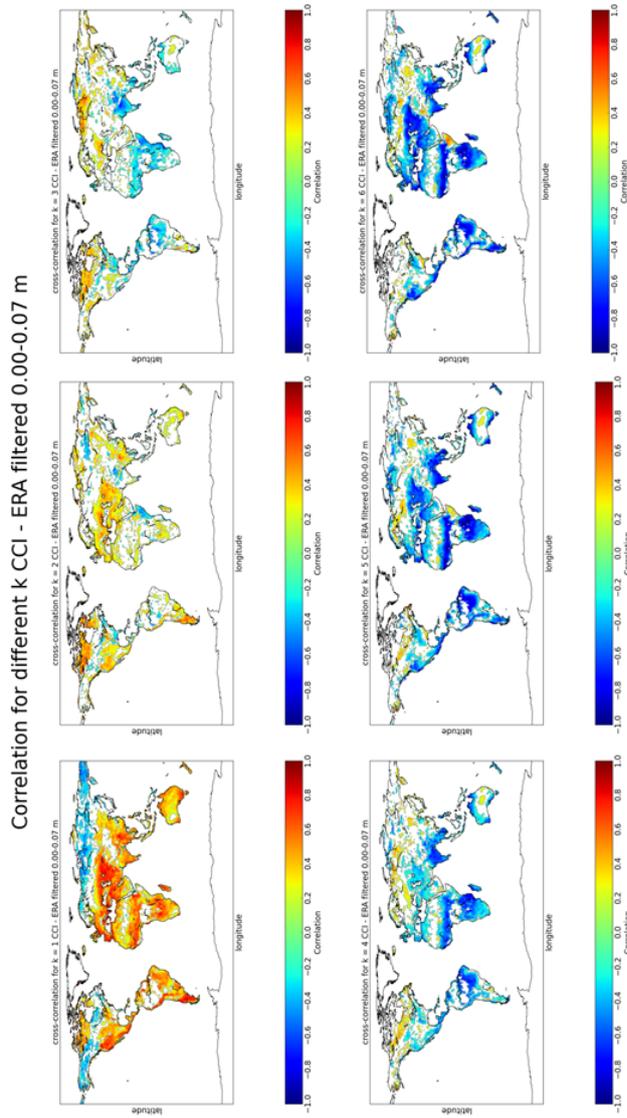


Figure A.16: The correlation coefficients between the ERA-Interim/Land Layer 1 and CCI SM datasets for different lags. The ERA-Interim/Land dataset lags corresponding to the given  $k$ -values ( $k = 1, 2, \dots, 6$ ).

# Appendix B

## Comparison of the datasets

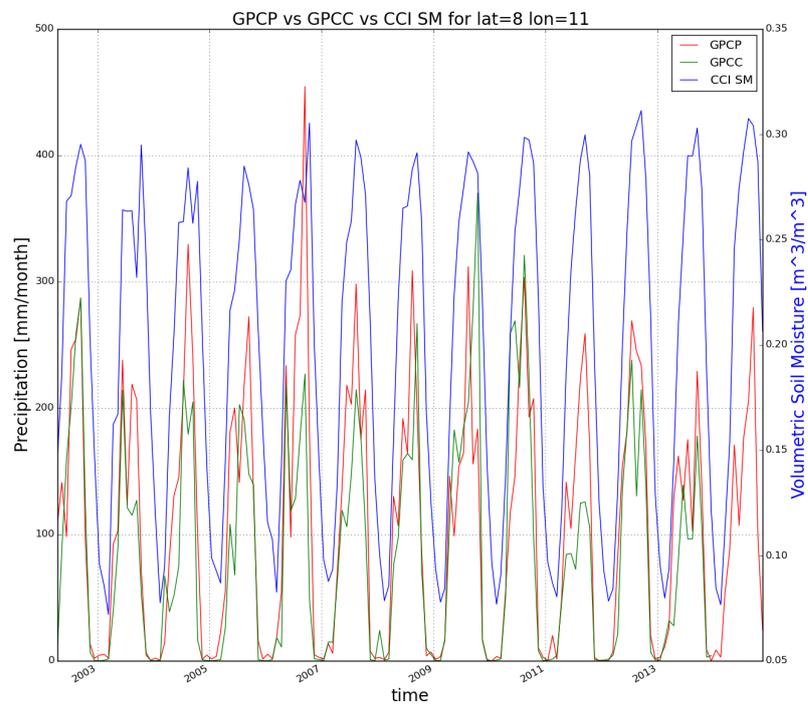


Figure B.1: A comparison of both precipitation datasets (GPCP and GPCC) and the CCI SM dataset. The site is located at a latitude of  $8^\circ$  and a longitude of  $11^\circ$ . The precipitation datasets are coloured in red (GPCP) and green (GPCC) in units of  $[mm/month]$ . The CCI SM dataset is shown in blue in  $[m_3/m_3]$ .

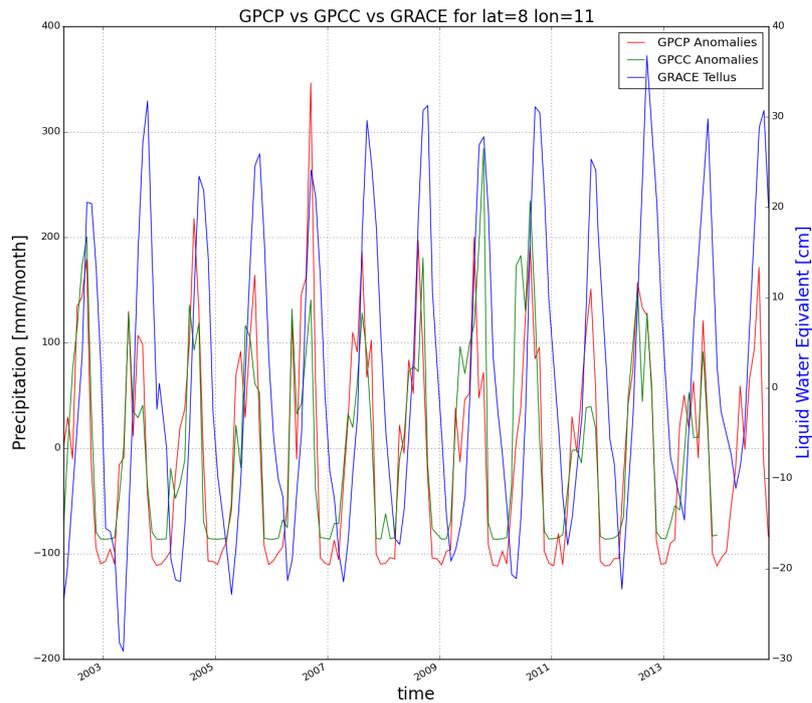


Figure B.2: A comparison of both precipitation datasets (GPCC and GPCP) and the GRACE Tellus dataset. The site is located at a latitude of  $8^\circ$  and a longitude of  $11^\circ$ . The precipitation datasets are coloured in red (GPCP) and green (GPCC) in units of  $[mm/month]$ . The GRACE Tellus dataset is shown in blue in  $[cm]$ .

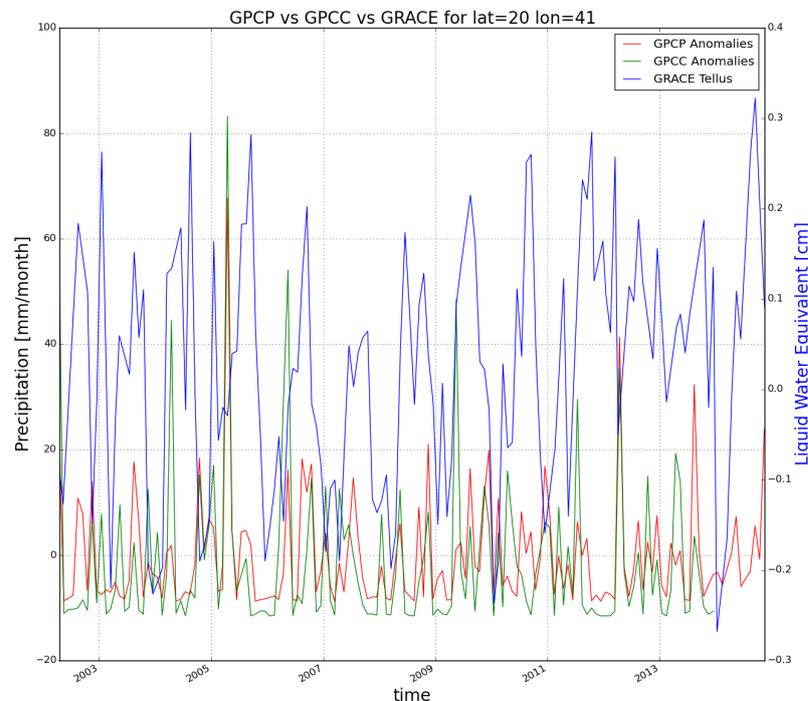


Figure B.3: A comparison of both precipitation datasets (GPCC and GPCP) and the GRACE Tellus dataset, for location in Saudi Arabia (latitude of  $20^\circ$  and longitude of  $41^\circ$ ). The anomalies of the precipitation datasets are coloured in red (GPCP) and green (GPCC) in units of  $[mm/month]$ . The GRACE Tellus dataset is shown in blue in  $[cm]$ .

# Appendix C

## Irrigation

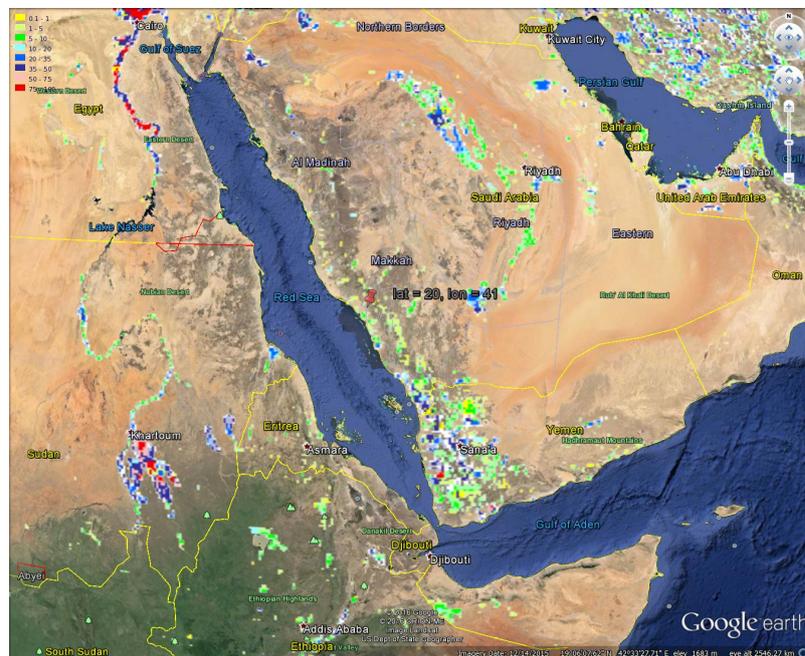


Figure C.1: The AEI for Saudi Arabia. The red marker indicates the site with a latitude of  $20^{\circ}$  and a longitude of  $41^{\circ}$ , which was chosen for the comparison of the different datasets (5). The figure was generated using Google Maps and GMIA. The AEI are the colored areas, mostly in the southern part of Saudi Arabia.

Irrigated crops		2006											
		Crop area as percentage of the full control actually irrigated area by month											
Area	1000 ha	J	F	M	A	M	J	J	A	S	O	N	D
Wheat	490	41	41	41	41	41	41						41
Barley	22	2	2	2	2	2							2
Sorghum	144						12	12	12	12	12	12	
Other cereals	18						2	2	2	2	2	2	
Vegetables	113						9	9	9	9	9	9	
Fruits	191	16	16	16	16	16	16	16	16	16	16	16	16
Citrus	11	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
Sesame	2						0.2	0.2	0.2	0.2	0.2	0.2	
Potatoes	15						1	1	1	1	1	1	
Fodder temporary	207	17	17	17	17	17							17
<b>Harvested irrigated crop area [AHI<sub>full</sub>]</b>	<b>1 214</b>												
<b>Area equipped for full control irrigation actually irrigated [AAI<sub>full</sub>]</b>	<b>1 191</b>	<b>77</b>	<b>77</b>	<b>77</b>	<b>77</b>	<b>77</b>	<b>77</b>	<b>58</b>	<b>41</b>	<b>41</b>	<b>41</b>	<b>41</b>	<b>77</b>
<b>Cropping intensity (%) = 100 x [AHI<sub>full</sub>]/[AAI<sub>full</sub>]</b>	<b>102</b>												
<b>Area equipped for full control irrigation [AEI<sub>full</sub>]</b>	<b>1 731</b>	*											
<b>% of full control equipped actually irrigated = 100 x [AAI<sub>full</sub>]/[AEI<sub>full</sub>]</b>	<b>69</b>												
<b>Total area equipped for irrigation [AEI<sub>total</sub>]</b>	<b>1 731</b>	*											

\* These areas refer to the year 2000

Figure C.2: The table represents the area actually irrigated per month in [%]. From January to May and in December the number is increasing correlated to the other month. This table can be found on the FAO website<sup>1</sup>.

<sup>1</sup> [http://www.fao.org/nr/water/aquastat/water\\_use.gr/index.stm](http://www.fao.org/nr/water/aquastat/water_use.gr/index.stm)

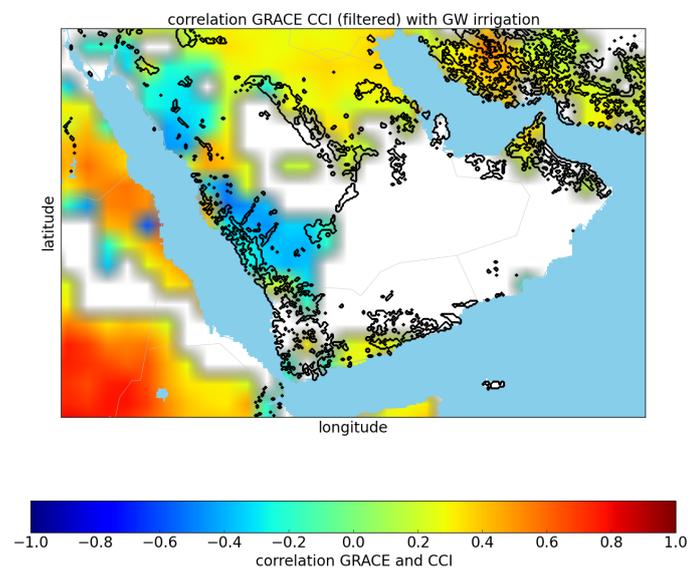


Figure C.3: The correlation coefficients between CCI SM and GRACE Tellus for Saudi Arabia are overlaid by a dataset showing areas using groundwater for irrigation.